What makes fisheries data informative?

Arni Magnusson[1,2] & *Ray Hilborn*[1]

[1]School of Aquatic and Fishery Sciences, University of Washington, Box 35520, Seattle, WA 98195, USA; [2]Marine Research Institute, Skulagata 4, PO Box 1390, 121 Reykjavik, Iceland

Abstract

Informative data in fisheries stock assessment are those that lead to accurate estimates of abundance and reference points. In practice, the accuracy of estimated abundance is unknown and it is often unclear which features of the data make them informative or uninformative. Neither is it obvious which model assumptions will improve estimation performance, given a particular data set. In this simulation study, 10 hypotheses are addressed using multiple scenarios, estimation models, and reference points. The simulated data scenarios all share the same biological and fleet characteristics, but vary in terms of the fishing history. The estimation models are based on a common statistical catch-at-age framework, but estimate different parameters and have different parts of the data available to them. Among the findings is that a 'one-way trip' scenario, where harvest rate gradually increases while abundance decreases, proved no less informative than a contrasted catch history. Models that excluded either abundance index or catch at age performed surprisingly well, compared to models that included both data types. Natural mortality rate, M, was estimated with some reliability when age-composition data were available from before major catches were removed. Stock-recruitment steepness, h, was estimated with some reliability when abundance-index or age-composition data were available from years of very low abundance. Understanding what makes fisheries data informative or uninformative enables scientists to identify fisheries for which stock assessment models are likely to be biased or imprecise. Managers can also benefit from guidelines on how to distribute funding and manpower among different data collection programmes to gather the most information.

Keywords abundance index, catch at age, informative data, reference points, stock assessment

Correspondence:
Arni Magnusson, Marine Research Institute, Skulagata 4, PO Box 1390, 121 Reykjavik, Iceland
Tel.: +354-575-2000
Fax: +354-575-2001
E-mail: arnima@hafro.is

Received 28 Nov 2006
Accepted 29 Aug 2007

Introduction

Stock assessment and informative data

Fisheries management relies on stock assessment models to provide estimates of population abundance, and to shed light on the underlying dynamics of the resources being managed. It is necessary to quantify and understand the uncertainty about model parameters and reference points to evaluate the consequences of alternative management actions. The uncertainty about estimated quantities reflects the information contained in the available data, but also depends on the choice of model and implicit assumptions that are made when the assessment is conducted. Despite theoretical and practical advances in the field of stock assessment, our ability to answer some key questions remains limited. This study focuses on one such question: What kinds of data are particularly informative in stock assessments, and how is this influenced by model assumptions?

Understanding what makes fisheries data informative or uninformative has obvious value for fisheries management, enabling us to identify fisheries for which stock assessment models are likely to be biased or imprecise. Managers can also benefit from guidelines on how to distribute funding and manpower among different data collection programmes to gather the most information. Moreover, adaptive management decisions can be taken today to make future data as informative as possible (Ludwig and Hilborn 1983; Walters 1986; Walters 2007).

Shepherd (1984) ranked types of fisheries data in terms of potential information provided by each type of data in isolation. Annual landings and age-specific abundance indices were ranked the highest, for example, while age-composition data alone was assigned a low score. Such statements are of course highly generalized, but nevertheless provide useful guidelines for planning data collection programmes. Shepherd (1984) also points out how different data types complement each other: landings provide information about the absolute scale of the fishery, age-composition data about the relative cohort size, and abundance-index data about the relative changes in abundance over time. Changes in growth or maturity can provide some information about changes in population density, confounded with other ecological and evolutionary factors (Rose *et al.* 2001). Less commonly used data types that provide information about stock status include tag recoveries and egg/larval surveys. In a Bayesian context, any prior information about estimated or derived parameters can also be seen as a type of data source (Gelman *et al.* 2004), where information from previous studies is expressed in the form of a probability distribution for estimated or derived parameters.

Many non-Bayesian estimation methods assume that specific parameters are known without error; the effect of such assumptions is similar to that of a highly informative Bayesian prior.

The general rule in statistical inference is that more data leads to less uncertainty. But other features of the data also play a role, e.g. the range of observed values and temporal patterns in time-series data. This is easy to show analytically for simple stock assessment techniques, such as depletion models and catch-curve analysis, as outlined below. When more complex models are used, it becomes less concrete what is meant when stock assessment modellers discuss the 'informative' data in a particular assessment, or perhaps more often, 'uninformative' data.

Before taking a closer look at what kinds of data are informative for a given model, it is helpful to begin with an overview of some commonly used models.

Models and assumptions in stock assessment

A variety of stock assessment models have been developed, as reviewed by Megrey (1989), Hilborn and Walters (1992), Quinn and Deriso (1999), Quinn (2003) and Smith and Addison (2003). This variety of models reflects the diversity of fisheries to which stock assessment techniques need to be applied, the data available for assessment purposes, and what is known or assumed about the fishery dynamics and stocks. There has been a move away from simple and restrictive assumptions (Schaefer 1954; Chapman and Robson 1960; Gulland 1965) towards more flexible models that incorporate all of the available data in a likelihood-based statistical framework.

A depletion model (Leslie and Davis 1939) can be used to estimate absolute abundance from a time series of landings and an index of relative abundance when a stock is fished down. Assuming a closed population where the impact of fishing mortality is much greater than those of recruitment, growth, and natural mortality, the biomass at the start of time step $t + 1$ equals the biomass at the start of time step t less the catch during time step t, that is $B_{t+1} = B_t - Y_t$, and the abundance index is proportional to the stock biomass, $I_t = qB_t$. The catchability coefficient, q, is assumed to be constant, although empirical data suggest that may not always be a justifiable assumption (Ricker 1958). The Schaefer (1954) biomass-dynamic model adds two parameters representing the maximum growth rate and carrying capacity, $B_{t+1} = B_t + rB_t(1-B_t/k) - Y_t$. When $r = 0$, it is identical to the depletion model. A slightly different approach is taken in the Kimura and Tagart (1982) stock-reduction model, where two parameters represent the natural mortality rate and recruitment, $B_{t+1} = B_t e^{-(Ft+M)} + R$. The fishing mortality rate F_t is evaluated from the landings, $Y_t = B_t(1-e^{-Ft-M})F_t/(F_t + M)$. As the above models do not distinguish between age groups, they can be formulated either in terms of numbers or biomass.

In catch-curve analysis (Chapman and Robson 1960), the total mortality rate of fully recruited fish in a given year can be estimated using the catch-at-age composition, assuming that recruitment variability is inconsequential. Catch curves can also be applied to individual cohorts (Hilborn and Walters 1992), relaxing the assumption that recruitment is the same across cohorts, $N_{t+1,a+1} = N_{t,a}e^{-(Ft+M)}$. In practice, M is often assumed to be known, and F can in turn be used to estimate absolute abundance if the annual landings are known. Related models include virtual population analysis (Gulland 1965), cohort analysis (Pope 1972), adaptive framework (Gavaris 1988) and extended survivors analysis (Shepherd 1999). The assumption of a known constant M is frequently challenged (Cotter *et al.* 2004), but restrictive assumptions about M and recruitment are often necessary to evaluate the consequences of alternative catch levels (Punt and Hilborn 1997). Even in fisheries where large quantities of data have been collected for decades, an age-structured assessment model can fit the data equally well when M is fixed at a very low value or a high value (Gavaris and Ianelli 2002).

The forward-projecting statistical catch-at-age model (Fournier and Archibald 1982; Deriso *et al.* 1985; Methot 1989) can be described conceptually as a stock-reduction model with variable recruitment, combined with catch-curve analysis of multiple cohorts. The basic framework can be tailored fairly easily to the specifics of the fishery to be modelled. The data types most commonly included in statistical catch-at-age analysis are: annual landed catch, catch at age and an index of relative abundance. The landed catch is often assumed to be measured without error, and model parameters are estimated by minimizing the difference between model predictions and the observed catch at age and abundance index. A statistical catch-at-age model can be fitted without any age data (Hilborn

1990) or without an index of abundance. However, Deriso *et al.* (1985) concluded that all three data types are required to estimate abundance and reference points reliably. Hilborn *et al.* (2003) describe how statistical catch-at-age models can incorporate sex-specific data from multiple fisheries with ageing error, catch-at-length data, and allow certain parameters to vary over time.

Parameters that are almost always estimated when fitting a statistical catch-at-age model include the age-structure of the population in the first year considered in the model, the selectivity curve for each fishery, the catchability coefficient for each abundance index and the annual recruitment. There remain, however, many decisions to fully specify a statistical catch-at-age model. For example, whether to estimate or fix the natural mortality rate M, how to model the relationship between spawning stock size and recruitment, whether to parametrize selectivity using an asymptotic or dome-shaped curve, and how to choose which parameters vary over time (Patterson *et al.* 2001; Gavaris and Ianelli 2002). M can be accurately estimated if the data include the catch age composition from a nearly unfished population, or if fishing effort is kept very low for some years (Beverton and Holt 1957), while the shape of the stock-recruitment curve can be estimated only if there is considerable contrast in stock size (Ricker 1958). Analysing a model that estimates M and right-hand selectivity, i.e. a shape parameter determining the selectivity of older age classes, Thompson (1994) noted that those parameters were confounded and recommended fixing either M or the selectivity shape parameter at an assumed value.

The simpler models (depletion, catch curves) can be emulated fairly adequately using statistical catch-at-age models, by fixing parameters, selecting specific functional forms for biological relationships, or excluding likelihood components from the objective function. The main difference between the depletion, biomass-dynamic, stock-reduction and delay-difference models is how recruitment, somatic growth and natural mortalities are handled. Schnute (1985) showed how the Schaefer (1954) biomass-dynamic model, the Deriso (1980) delay-difference model and Kimura and Tagart (1982) stock-reduction model are special cases of a generalized catch-effort model. Xiao (2000) showed further that the above models, along with the Leslie and Davis (1939) depletion model, Gulland's (1965) virtual population analysis and Fournier and Archibald's (1982) statistical catch-at-age models are all special cases of a generalized age-structured model. In light of their flexibility, superior performance (NRC 1998; Punt *et al.* 2002), and increasing usage, statistical catch-at-age models are used in this study to address the questions of interest.

Simulation studies and informative data scenarios

The term data scenario is used here to denote temporal patterns found in the data, regardless of the amount and types of data. These temporal patterns are impacted by how the fishery has been conducted historically, e.g. whether fishing effort and landings have been increasing or decreasing, held relatively steady, or perhaps the stock might be rebuilding after heavy depletion. One of the questions confronted in this study is which data scenarios are more informative than others.

A data scenario is informative when it enables a given model to estimate the status of a fishery with greater accuracy than most other data scenarios would. In the real world, a data scenario can be said to be informative if it resembles a scenario that has been shown to be informative, either analytically or in a simulation study. Analytical demonstration is only viable for the simplest of models, such as linear regression, but for more complex models, simulations are used to evaluate estimation accuracy. Simulation studies use an 'operating' model to generate artificial data similar to those used in stock assessment, except the true population parameters are known.

The depletion model can be expressed as a linear regression with the abundance index as the response variable and accumulated catch as the predictor, $\hat{I}_t = qB_{\text{init}} - q\sum_{i\leq t} Y_i$. As with any simple linear regression model, the uncertainty about the slope and intercept depends on (i) how closely the data points are aligned in a straight line, as residuals will be smaller when model assumptions are not violated substantially and when measurements are reasonably accurate, (ii) the range of values on the x-axis and (iii) the number of data points. The biomass before catches were removed, B_{init}, corresponds to the x-intercept. This intercept can be predicted more accurately when the y-values, relative abundance, are observed both at high and quite low values. Ricker (1958) noted that intense fishing effort that reduces the abundance considerably leads to informative data for the depletion model, and Pope (1972) found the same to be true for cohort analysis. Many fisheries have undergone a period of rapid removals and can therefore be expected to yield informative data, if scientific data were being collected at the time.

Hilborn (1979) demonstrated why and how certain data scenarios are informative, using a simplified Schaefer biomass-dynamic model that has a closed-form solution. He concluded that contrast is needed in both abundance and harvest rate to obtain unbiased and precise parameter estimates. Specifically, Hilborn (1979) identified the most informative data scenario as one that includes a period of quite heavy exploitation, followed by a period where the stock is allowed to rebuild to an intermediate level, after which the exploitation rate increases again.

The parameter of main interest in catch-curve analysis is the fishing mortality rate in each year, F_t, frequently used in fisheries management. This parameter is confounded with the natural mortality rate, as cohorts decline at an exponential rate $Z_t = F_t+M$. If little or no fishing has taken place in previous years, Z corresponds to the rate of natural mortality M, which is otherwise a very difficult parameter to estimate. More generally, catch-at-age data that contain years with high and low fishing effort are informative to bound possible values of M, and therefore F_t (Beverton and Holt 1957). Variation in fishing effort has also been found to be informative for more complex age-structured models to separate natural and fishing mortalities (Hilborn and Walters 1992). Of course, uncertainty about the estimated parameters will also decrease if large numbers of fish are sampled at random and measured with negligible ageing error, and if M varies only slightly between years, without a consistent increasing or decreasing trend (Beverton and Holt 1957).

Several simulation studies have explored the behaviour of statistical catch-at-age models. Bence *et al.* (1993) found that current abundance is estimated more reliably when harvest rate has been high, and when the true survey selectivity curve is asymptotic rather than dome-shaped. The study of Sampson and Yin (1998), later updated by Yin and Sampson (2004), showed how low natural mortality M, high recruitment variability and small changes in the harvest rate all lead to unreliable estimates. They also concluded that for the U.S. West Coast groundfish fishery, it would be more cost-effective to gain information by increasing sampling for age composition than by improving the precision of the survey on which abundance indices are based, at least from a single-species perspective. Ianelli (2002) found that reference points are overestimated when the true steepness h of the stock-recruitment curve is low, and underestimated when the true value of h is high. In their simulation study, Punt *et al.* (2002) showed how depletion level, defined as current abundance compared with average virgin abundance, was estimated more reliably than other reference points. They also found that the statistical catch-at-age model performed substantially worse when age-composition data were not available.

This study

The goal of this study is to improve our understanding of how uncertainty about the status of a fishery resource depends on data, models and assumptions. An 'informative' data scenario is one that enables a given model to estimate the status of a fishery with greater accuracy than most other data scenarios would. The hypotheses that will be addressed are:

H_1 Fisheries data are most informative when they span a period where the population was fished down to a low level.

H_2 Fisheries data are most informative when they span a period where the population was fished down to a low level and then allowed to rebuild for some time.

H_3 The level of stock depletion is estimated more reliably than other reference points.

H_4 A data set that includes both an index of relative abundance and catch-at-age data is much more informative than a data set that includes only one of these two types of data.

H_5 Not knowing M, h and right-hand selectivity leads to inaccurate estimates of stock abundance and reference points.

H_6 Models estimating M perform about as well as models estimating h.

H_7 M can be estimated reliably if age-composition data are available from when the population was unfished.

H_8 M can be estimated reliably from the rate of population increase if the stock is allowed to rebuild from a low level.

H_9 h can be estimated reliably from catch-at-age data and an index of relative abundance when the data cover a period in which abundance varies substantially.

H_{10} Right-hand selectivity can only be estimated reliably when M is known.

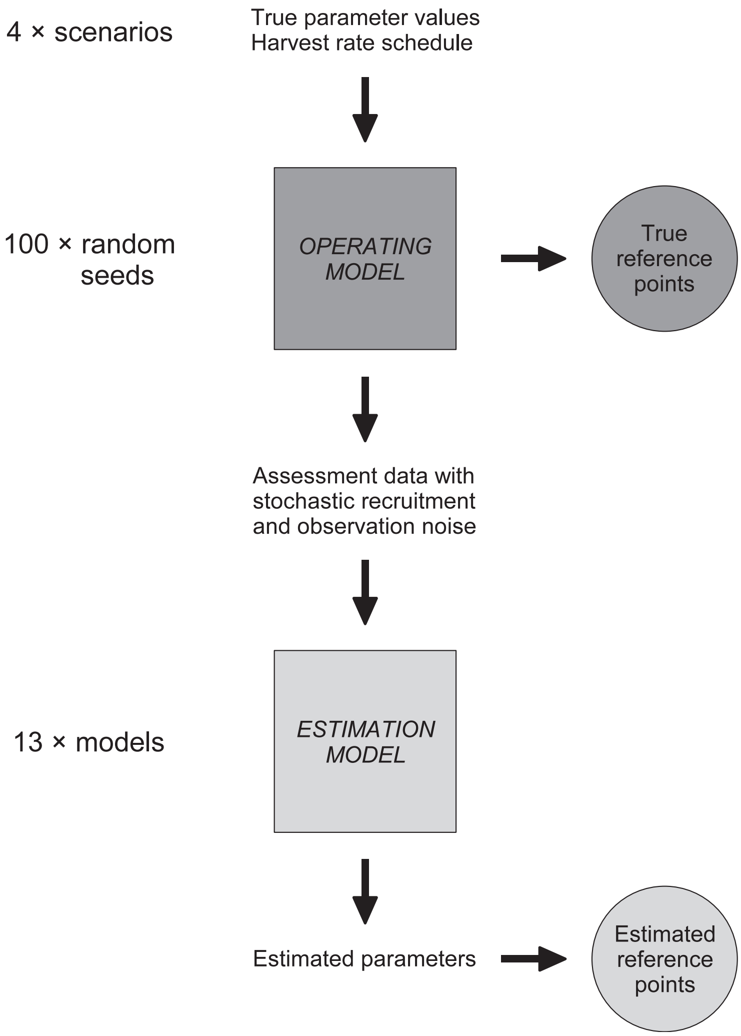

Figure 1 The stimulation procedure. Arrows and boxes indicate the workflow for a single run, and multiplications describe how the study consists of multiple runs.

Methods

First, we define four fishing history scenarios and generate stochastic data sets using an 'operating model,' based on an age-structured population dynamics model. The performance of a suite of estimation models is then evaluated, with respect to how well they estimate the values of six reference points. The simulation procedure, outlined in Fig. 1, is repeated for each scenario, random seed and estimation model. A scenario consists of chosen parameter values and a harvest rate schedule, described in more detail below. The operating model first applies stochastic recruitment and outputs the resulting reference point values. It then applies random observation noise and outputs the assessment data that are used as input for the estimation models. Finally, the estimated reference points are derived from the parameter estimates, and compared with the 'true' reference points that were not subject to observation noise.

Scenarios

Four fishing history scenarios are simulated in the analysis: (A) *one-way trip* where harvest rate is gradually increased while the abundance decreases, (B) *no change* where abundance is steady at a constant and somewhat low harvest rate, (C) *good contrast* where the stock is fished down to less than half its initial size and then allowed to rebuild and (D) *rebuild only* where the stock begins at a very low abundance and is allowed to rebuild under low fishing pressure. The fishing history scenarios are designed specifically to address hypotheses 1–2 and 7–9, in terms of harvest rate and the expected value

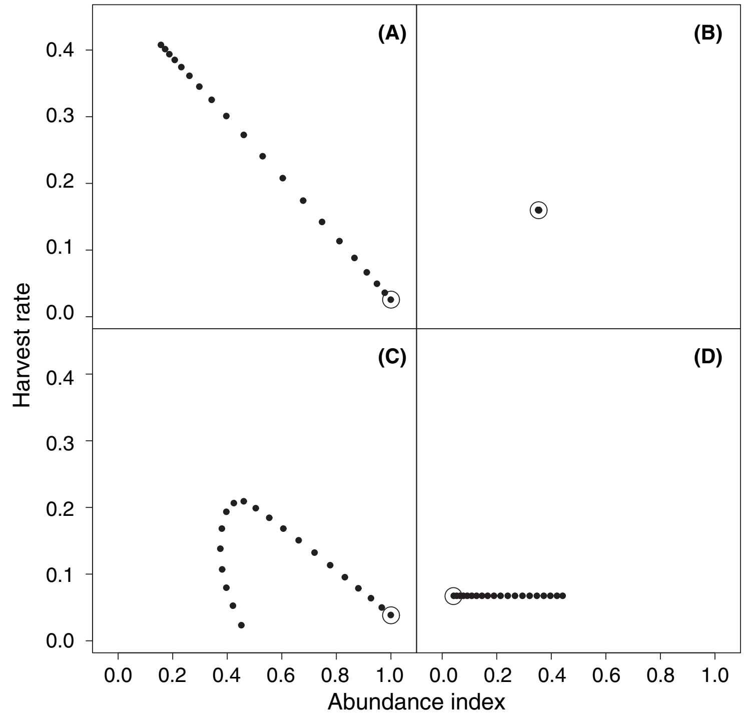

Figure 2 The four fishing history scenarios considered in this study, in terms of the relationship between the harvest rate and the expected value of the abundance index. Circles represent the status of the fishery in the first year. (A) One-way trip, (B) no change, (C) good contrast, (D) rebuild only.

of the abundance index (Fig. 2). Time trajectories offer a more traditional view of the same data (Fig. 3).

Operating model

The biological component

The operating model is a statistical catch-at-age model (Fournier and Archibald 1982) with biological characteristics (Table 1) and parameter values (Table 2) based on Atlantic cod (*Gadus morhua*, Gadidae). It follows the parametrization of the Coleraine statistical catch-at-age software (Hilborn *et al.* 2003), which is used to implement the estimation models.

The population dynamics are governed by the equation:

$$N_{t+1,a+1} = N_{t,a}e^{-M}(1 - {}_C S_a u_t) \tag{1}$$

where $N_{t,a}$ is population size at time t and age a, M is the rate of natural mortality, ${}_C S$ is the selectivity of the commercial fishery and u is harvest rate. The oldest age group, age A, is treated as a plus group:

$$N_{t+1,A} = N_{t,A-1}e^{-M}(1 - {}_C S_A u_t) + N_{t,A}e^{-M}(1 - {}_C S_A u_t) \tag{2}$$

Selectivity is an asymmetric normal curve determined by three shape parameters,

$$S_a = \begin{cases} \exp\left(\frac{-(a-S_{\text{full}})^2}{\exp(S_{\text{left}})}\right), & a \leq S_{\text{full}} \\ \exp\left(\frac{-(a-S_{\text{full}})^2}{\exp(S_{\text{right}})}\right), & a > S_{\text{full}} \end{cases} \tag{3}$$

where S_{full} is the age at full selectivity, S_{left} describes the left-hand slope and ${}_S S_{\text{right}}$ the right hand slope of the curve. The survey selectivity curve has a high ${}_S S_{\text{right}} = 15$ (Table 2) so the oldest fish are fully selected, but the commercial selectivity has an intermediate ${}_C S_{\text{right}} = 6$, resulting in a slightly dome-shaped curve (Fig. 4). Harvest rate is defined as the fraction removed from the vulnerable biomass in the middle of the fishing year, $u_t = Y_t / \sum_a ({}_C S_a N_{t,a} w_a) e^{-M/2}$, where Y is catch and w is body weight.

The population size at the start of the first year is

$$\begin{aligned} N_{1,1} &= R_0 R_{\text{init}} \times \exp({}_R\varepsilon_{1,1} - \sigma_R^2/2) \\ N_{1,a} &= R_0 R_{\text{init}} e^{-(a-1)M} \sum_{i=1}^{a-1} (1 - {}_C S_i u_{\text{init}}) \\ &\quad \times \exp({}_R\varepsilon_{1,a} - \sigma_R^2/2) \\ N_{1,A} &= R_0 R_{\text{init}} e^{-(A-1)M} \\ &\quad \times \sum_{i=1}^{A-1} (1 - {}_C S_i u_{\text{init}}) / [1 - e^{-M}(1 - {}_C S_A u_{\text{init}})] \\ &\quad \times R_{\text{plus}} \end{aligned} \tag{4}$$

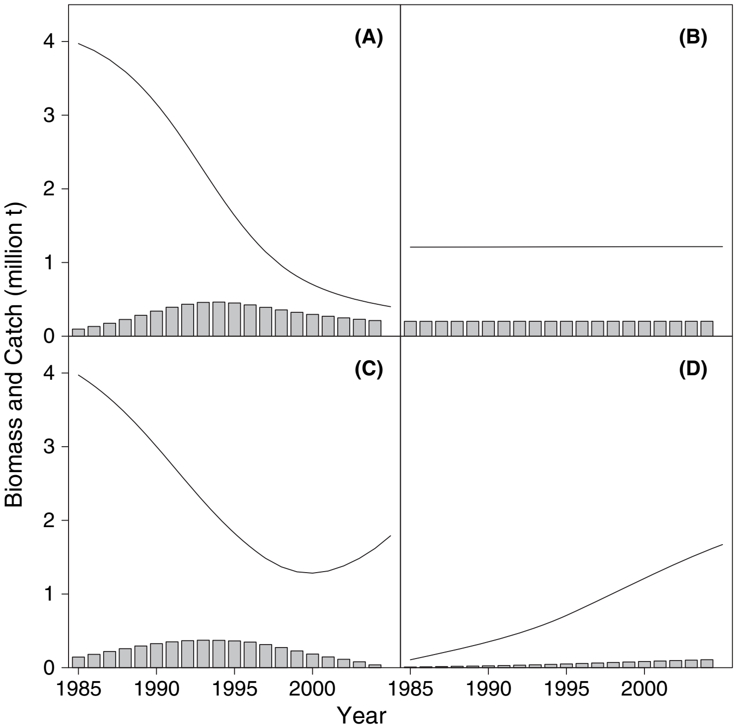

Figure 3 The four fishing history scenarios considered in this study, in terms of spawning biomass (line) and landed catch (bars). (A) One-way trip, (B) no change, (C) good contrast, (D) rebuild only.

Table 1 Age-specific weight (kg) and maturity (proportion) used in the operating and estimation model.

Age	1	2	3	4	5	6	7	8	9	10+
Weight (kg)	0.3	0.7	1.3	1.8	2.6	3.6	4.9	6.3	7.7	10.1
Maturity	0.0	0.0	0.1	0.4	0.6	0.8	0.9	1.0	1.0	1.0

for 1-year olds, intermediate ages and the plus group. R_0 is average virgin recruitment, R_{init} scales the initial population size across all ages and u_{init} is the initial harvest rate. The ${}_R\varepsilon$ elements are random recruitment deviates generated from the normal distribution, ${}_R\varepsilon \sim \text{Norm}(0,\sigma_R)$, where σ_R is recruitment variability. The R_{plus} term scales the initial plus group and is not drawn from the same distribution as the ${}_R\varepsilon$ recruitment deviates for the younger ages. Instead, a large number of initial ages are generated, up to 100 years old, and then ages 10 and over are aggregated in a plus group.

Recruitment is stochastic around a Beverton-Holt stock-recruitment function, reparametrized according to Francis (1992):

$$N_{t+1,1} = \frac{4hR_0(B_t/B_0)}{1-h+(5h-1)(B_t/B_0)} \times \exp({}_R\varepsilon_{t+1,1} - \sigma_R^2/2) \quad (5)$$

where $B_t = \Sigma_a N_{t,a}\Phi_a w_a$ is spawning biomass,

$$B_0 = \sum_{a=1}^{A-1} R_0 e^{-(a-1)M}\Phi_a w_a + R_0 e^{-(A-1)M}\Phi_A w_A/(1-e^{-M}) \quad (6)$$

is average virgin spawning biomass, h is steepness of the stock-recruitment curve, and Φ is proportion mature.

Generating the simulated data sets

One hundred data sets are generated for each fishing history scenario. These data sets vary in terms of landings, survey abundance index and commercial catch-at-age. The harvest rate is always the same in each scenario, but the resulting landings change as population size changes with stochastic recruitment. There are 10 age classes and 20 years of data, nominally referred to as 1985–2004. The landings are assumed to be known exactly, but the catch at age and abundance index are subject to random observation error. When stochastic recruitment and observation noise is

Table 2 Parameter values and harvest rate schedules for the four fishing history scenarios.

Scenario:	A One way	B No change	C Contrast	D Rebuild
Parameters				
R_0	250 000	*	*	*
h	0.7	*	*	*
M	0.2	*	*	*
R_{init}	1	0.8	1	0.2
u_{init}	0	0.16	0	0.4
R_{plus}	1	*	*	*
${}_{C}S_{full}$	5	*	*	*
${}_{C}S_{left}$	1	*	*	*
${}_{C}S_{right}$	6	*	*	*
${}_{S}S_{full}$	4	*	*	*
${}_{S}S_{left}$	1	*	*	*
${}_{S}S_{right}$	15	*	*	*
q	2.5×10^{-7}	*	*	*
Harvest rate				
1985	0.026	0.160	0.039	0.067
1986	0.036	0.160	0.050	0.067
1987	0.050	0.160	0.064	0.067
1988	0.067	0.160	0.079	0.067
1989	0.088	0.160	0.096	0.067
1990	0.113	0.160	0.114	0.067
1991	0.142	0.160	0.132	0.067
1992	0.174	0.160	0.151	0.067
1993	0.208	0.160	0.168	0.067
1994	0.241	0.160	0.184	0.067
1995	0.273	0.160	0.199	0.067
1996	0.301	0.160	0.209	0.067
1997	0.325	0.160	0.207	0.067
1998	0.345	0.160	0.193	0.067
1999	0.361	0.160	0.168	0.067
2000	0.374	0.160	0.138	0.067
2001	0.385	0.160	0.107	0.067
2002	0.394	0.160	0.080	0.067
2003	0.401	0.160	0.053	0.067
2004	0.408	0.160	0.023	0.067

An asterisk indicates that the same parameter value applies across all scenarios.

added to the original templates from Fig. 2, the observed abundance index shows random fluctuations, but the overall fishing history is still recognizable (Fig. 5).

Even though the harvest rates in Table 2 are followed precisely, the resulting landings vary among the data sets because of stochastic recruitment. The level of recruitment variability ($\sigma_R = 0.6$), observation noise for the abundance index ($\sigma_I = 0.2$) and observation noise for the commercial catch at age ($n = 50$) are similar to those used in recent assessments of the Icelandic cod stock (ICES 2003).

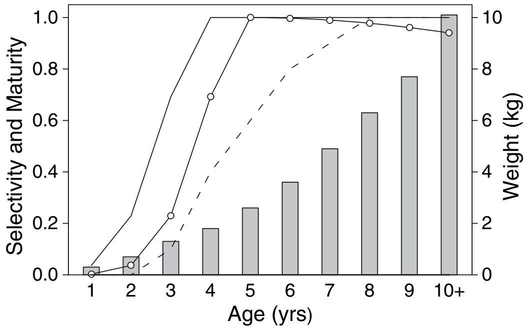

Figure 4 Age-specific characteristics of the operating model: survey selectivity (plain line), commercial selectivity (line with circles), maturity (dashed line) and weight (bars).

The survey abundance index is proportional to the biomass vulnerable to the survey in the middle of the fishing year:

$$I_t = q \sum_a {}_{S}S_a N_{t,a} w_a e^{-M/2} \times \exp({}_{I}\varepsilon_t) \tag{7}$$

where I is the observed abundance index, q is the catchability coefficient, ${}_{S}S$ is survey selectivity and ${}_{I}\varepsilon \sim \mathrm{Norm}(0, \sigma_I)$ is random observation noise. The commercial catch-at-age data are provided to the assessment model in the form of proportions at age. These proportions are generated assuming that the sampling is multinomial:

$$P_{t,a} \sim \mathrm{Multinom}\left(n, \frac{{}_{C}S_a N_{t,a}}{\sum_a {}_{C}S_a N_{t,a}}\right) \Big/ n \tag{8}$$

where P is the observed catch at age and n is the sample size used to generate observation noise.

Survey catch-at-age data are not used in this study, to keep the analysis and interpretation as simple as possible. The survey abundance index and the commercial catch at age are independent sources of information, one about changes in relative abundance, the other about relative cohort sizes and mortality rates. Data are assumed to be available for each year and the landings are output without observation error.

Estimation models

Thirteen estimation models are fitted to the simulated data. They have the same parametrization as the operating model (Equations 1–8) and are implemented with the Coleraine statistical catch-at-age software (Hilborn *et al.* 2003). The models differ in terms of which data types are included in the objective function and which parameters are

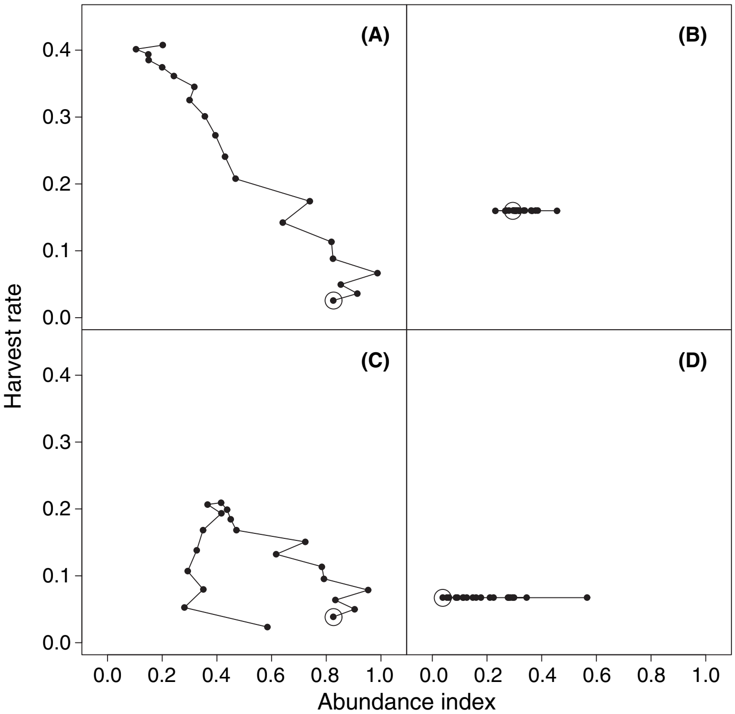

Figure 5 Examples of stochastic data sets (random seed = 100), in terms of the relationship between the harvest rate and the expected value of the abundance index. Circles represent the status of the fishery in the first year. (A) One-way trip, (B) no change, (C) good contrast, (D) rebuild only.

estimated (Table 3). The models are designed specifically to address hypotheses 4–10.

The 13 models consist of three 'families,' indicated by the first digit of the abbreviation used to identify the model: family 1 uses only landings and abundance index, family 2 uses only landings and catch at age, family 3 uses all three data types. Thus, models from family 1 are akin to biomass-dynamic models (with R_0 scaling the absolute size of the population instead of K, and h and M determining the intrinsic growth rate instead of r), models from family 2 resemble catch-curve analysis based on multiple cohorts, and those from family 3 are several variants of statistical catch-at-age analysis.

Although it is possible to examine the implications of estimating every combination of parameters, the focus of this study is on three key parameters: the steepness of the stock-recruitment relationship (h), the natural mortality rate (M) and the right-hand selectivity shape parameter ($_{C}S_{\mathrm{right}}$) for the commercial fishery. Models that estimate these parameters have 'h,' 'M' or 'r' in their abbreviations. When parameters are not estimated, they are fixed at the true value, as is done for the survey selectivity parameters.

The Coleraine software requires that all estimated parameters be bounded. Wide bounds (Table 4) are assigned to all parameters so as not to impose any major constraints on the values for the parameters.

Table 3 The 13 estimation models in terms of data types used and parameters estimated.

Model	1	1h	1m	2	2h	2m	2r	3	3h	3m	3r	3mr	3hmr
Data													
Catch	x	x	x	x	x	x	x	x	x	x	x	x	x
Index	x	x	x					x	x	x	x	x	x
CA				x	x	x	x	x	x	x	x	x	x
Estimated													
R_0	x	x	x	x	x	x	x	x	x	x	x	x	x
h		x			x				x				x
M			x			x				x		x	x
R_{init}	x	x	x	x	x	x	x	x	x	x	x	x	x
u_{init}	x	x	x	x	x	x	x	x	x	x	x	x	x
R_{plus}				x	x	x	x	x	x	x	x	x	x
$_{C}S_{\mathrm{full}}$				x	x	x	x	x	x	x	x	x	x
$_{C}S_{\mathrm{left}}$				x	x	x	x	x	x	x	x	x	x
$_{C}S_{\mathrm{right}}$							x				x	x	x
q		x	x					x	x	x	x	x	x
R^{ε}				x	x	x	x	x	x	x	x	x	x

Catch stands for landings, index for survey abundance index and CA for commercial catch at age.

The objective function for the estimation models is the sum of three components. The first two relate

Table 4 Bounds on estimated parameters, along with the true values from the operating model.

Parameter	True value	Lower bound	Upper bound
R_0	250 000	1000	10 000 000
h	0.7	0.2	1
M	0.2	0	0.5
R_{init}	0.2–1	0	5
u_{init}	0–0.4	0	1
R_{plus}	1	0	2
${}_CS_{\mathrm{full}}$	5	3	10
${}_CS_{\mathrm{left}}$	1	−2	5
${}_CS_{\mathrm{right}}$	6	−2	15
$\log q$	−15.2	−30	0
${}_R\varepsilon$	*	−15	15

The true value of R_{init} and u_{init} varies among scenarios (see Table 2).
*Initial age-structure and annual recruitment varies between the simulated data sets.

to data included in the analysis and the last is a penalty on recruitment deviations from the stock-recruitment relationship:

$$f = -\log L_I - \log L_C + \mathrm{Pen} \tag{9}$$

The abundance-index likelihood component is lognormal:

$$-\log L_I = \sum_t \frac{(\log I_t - \log \hat{I}_t)^2}{2\sigma_I^2} \tag{10}$$

where I and $\hat{I}$ are the observed and model-predicted abundance indices. The robust normal likelihood for proportions (Fournier *et al.* 1990) is assumed for the catch-at-age data:

$$-\log L_C = \sum_t \sum_a \log\left[\exp\left(\frac{(P_{t,a} - \hat{P}_{t,a})^2}{2[P_{t,a}(1-P_{t,a}) + 0.1/A]n^{-1}}\right) + 0.01\right] \tag{11}$$

where P and $\hat{P}$ are observed and the model-predicted catch proportions at age. Finally, recruitment deviates are penalized under the assumption of lognormality:

$$\mathrm{Pen} = \sum_{a=2}^{A-1} \frac{{}_R\varepsilon_{t,a}^2}{2\sigma_R^2} + \sum_{t=2}^{t_{\max}-1} \frac{{}_R\varepsilon_{t,1}^2}{2\sigma_R^2}, \tag{12}$$

where ${}_R\varepsilon_{l,a}$ and ${}_R\varepsilon_{t,1}$ are recruitment deviates in the initial year and subsequent years, and σ_R is a measure of the extent of recruitment variability. The estimation models are given the correct (i.e. the operating model) values for $\sigma_I = 0.2$, the effective sample size $n = 50$ for the catch-at-age data and recruitment variability $\sigma_R = 0.6$.

Reference points

Six reference points are evaluated as potential management quantities of interest: B_{current}(current biomass), u_{current} (current harvest rate), Depletion (current depletion level), MSY (maximum sustainable yield), $B_{\mathrm{current}}/B_{\mathrm{MSY}}$ (current biomass relative to B_{MSY}) and Surplus (current surplus production). These reference points are chosen because they are commonly used in fisheries management. B_{current}, u_{current} and Depletion are calculated using the equations:

$$B_{\mathrm{current}} = \sum_a N_{2005,a} \Phi_a w_a \tag{13}$$

$$u_{\mathrm{current}} = Y_{2004} / \sum_a ({}_cS_a N_{2004,a} w_a) e^{-M/2} \tag{14}$$

$$\mathrm{Depletion} = B_{\mathrm{current}}/B_0 \tag{15}$$

The maximum sustainable yield, MSY, is defined as the long-term average catch when the harvest rate is set to an optimal value, u_{MSY}. The average catch at a given harvest rate can be calculated in closed form, by combining methods from Lawson and Hilborn (1985) and Francis (1992). First, the equilibrium age composition is standardized so that the number of 1-year olds equals 1:

$$n_a^* = \begin{cases} e^{-(a-1)M} \prod_i^{a-1} 1 - {}_CS_i u, & a < A \\ \dfrac{e^{-(A-1)M} \prod_i^{A-1} 1 - {}_CS_i u}{e^{-M}(1 - {}_CS_A u)}, & a = A \end{cases} \tag{16}$$

At this harvest rate, the average recruitment is $R^* = (SBPR^* - \alpha)/(\beta SBPR^*)$, where $\alpha = SBPR_0(1-h)/(4h)$, $\beta = (5h-1)/(4hR_0)$, $SBPR^* = \Sigma_a n^*_a \Phi_a w_a$, and $SBPR_0$ is calculated in the same way as $SBPR^*$ except that $u = 0$. The average long-term catch for a given harvest rate is

Table 5 True reference point values from the operating model, given deterministic recruitment.

Scenario:	A One way	B No change	C Contrast	D Rebuild
B_{current}	400	1216	1791	1672
u_{current}	0.408	0.160	0.023	0.067
Depletion	0.101	0.306	0.451	0.421
MSY	203	203	203	203
$B_{\text{current}}/B_{\text{MSY}}$	0.315	0.957	1.410	1.316
Surplus	170	203	195	185

B_{current}, MSY and Surplus are shown in thousands of tonnes.

$$Y^* = uR^* e^{-M/2} \sum_a n_a^* {}_C S_a w_a \tag{17}$$

and the corresponding spawning biomass is

$$B^* = R^* \times SBPR^* \tag{18}$$

MSY and B_{MSY} are calculated by searching iteratively for the u that maximizes Y^*. Finally, current surplus production is defined as the last year's catch, plus the resulting change in vulnerable biomass:

$$\text{Surplus} = Y_{2004} + \sum_a {}_C S_a w_a (N_{2005,a} - N_{2004,a}) e^{-M/2} \tag{19}$$

The true reference point values from the operating model vary due to stochastic recruitment, except u_{current} which is pre-defined in each scenario (Table 2) and MSY which depends only on R_0, h, M and commercial selectivity. The true MSY value is in all cases 203 thousand tonnes, with harvest rate $u_{\text{MSY}} = 0.154$ and spawning biomass $B_{\text{MSY}} =$ 1270 thousand tonnes. Table 5 gives an idea of the approximate values of the reference points, using the special case of deterministic recruitment (all ${}_R\varepsilon_{t,a} = 0$ and $\sigma_R = 0$) as an example.

Performance measures

The performance of an estimation model is quantified by comparing the estimates from the 100 data sets with the true values from the operating model, using two performance measures. One performance indicator is the bias of estimators:

$$\text{Median bias} = \text{median}\left(\frac{\hat{\theta} - \theta}{\theta}\right) \tag{20}$$

where $\hat{\theta}$ is the estimated value of a reference point and θ is the true value. The median bias is used rather than the mean, to make the performance indicator more robust to outlying estimates of the management quantities. The other performance indicator is the proportion of estimates that are less than half or greater than twice the true value:

$$\text{Failure rate} = \Pr(\hat{\theta}/\theta < 0.5 \cup \hat{\theta}/\theta > 2) \times 100 \tag{21}$$

where 0.5 and 2 bound an arbitrarily chosen range of 'acceptable' error. The failure rate is a robust measure of accuracy, capturing both bias and imprecision, while the median bias is better at detecting relatively small but consistent bias. Median bias has a possible range from −1 to ∞, and failure rate is between 0 and 100. An estimation model that performs well has median bias close to 0 and failure rate close to 0. The performance is also presented graphically using Tukey's boxplots, where a solid box shows the inner quartiles, and whiskers extend from the box to the outermost data point within 1.5 times the interquartile range (Tukey 1977).

Results

A total of 5200 model runs are analysed: 100 data sets for each of the four scenarios and 13 estimation models. In the first part of the results, we look at how well the models estimate the reference points, and the second part focuses on selected model parameters.

Reference points

To facilitate comparison, the distribution of the estimated reference points (Fig. 6) are expressed as ratios of the true values known from the operating model. The multipanel boxplot allows one to visually evaluate the estimation performance for each reference point across data scenarios and estimation models. For example, the top left panel shows how well each model estimates current spawning biomass when the data are simulated based on scenario A (one-way trip). In this panel the boxplot medians are not far from 1, indicating that the models estimate current abundance with relatively small bias. However, the uncertainty of the estimates is considerably greater for model families 1 and 2 than for model family 3. This is understandable, because model families 1 and 2 ignore the catch-at-age and abundance-index information, respectively, while model family 3 uses all of the

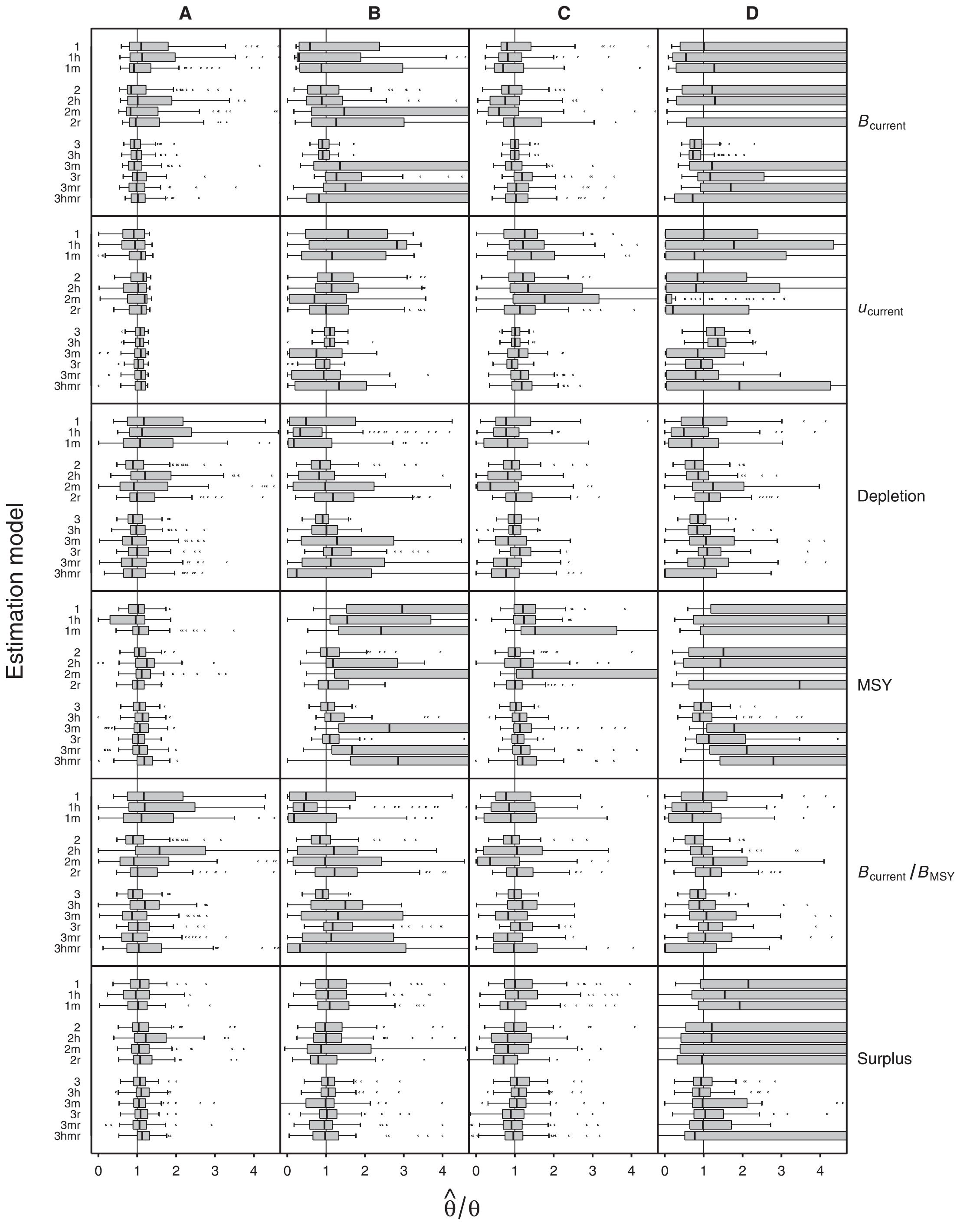

Figure 6 Distribution of estimated reference points. Panel columns correspond to fishing history scenarios A–D and panel rows are the six different reference points. Each Tukey boxplot shows the distribution of 100 estimates, divided by the true value of the reference point from the operating model. The *x*-axis is truncated to avoid loss of detail.

available data. The two performance measures, median bias and failure rate (Tables 6 and 7), summarize the information in Fig. 6.

$B_{current}$

When estimating current abundance (Fig. 6, top row of panels), the models exhibit only a small bias

Table 6 Bias of estimated reference points, by scenario and model.

	$B_{current}$	$u_{current}$	Depletion	MSY	$B_{current}/B_{MSY}$	Surplus
A1	+0.1	−0.1	+0.2		+0.2	+0.1
A1h	+0.1	−0.1	+0.1		+0.2	
A1m	−0.1	+0.1	+0.1		+0.1	
A2	−0.2	+0.2	−0.1		−0.1	
A2h			+0.2	+0.2	+0.6	+0.2
A2m	−0.2	+0.2	−0.1	+0.1	−0.1	
A2r		+0.1				+0.1
A3	−0.1	+0.1	−0.1	+0.1	−0.1	+0.1
A3h		+0.1		+0.1	+0.2	+0.1
A3m	−0.1	+0.1	−0.1	+0.1	−0.1	
A3r						+0.1
A3mr		+0.1	−0.1	+0.1	−0.1	+0.1
A3hmr		+0.1	−0.1	+0.2		+0.1
B1	−0.4	+0.6	−0.5	+2.0	−0.5	+0.1
B1h	−0.7	+1.8	−0.7	+0.5	−0.6	+0.1
B1m	−0.1	+0.2	−0.8	+1.4	−0.8	+0.1
B2	−0.1	+0.1	−0.2		−0.2	
B2h	−0.1	+0.1	−0.2	+0.2	+0.2	
B2m	+0.5	−0.3		+4.3		−0.1
B2r	+0.3		+0.2	+0.1	+0.2	−0.2
B3	−0.1	+0.1	−0.1		−0.1	
B3h	−0.1	+0.1		+0.1	+0.5	+0.1
B3m	+0.4	−0.2	+0.3	+1.6	+0.3	
B3r	+0.3		+0.2	+0.1	+0.2	
B3mr	+0.5	−0.1	+0.1	+0.7	+0.1	
B3hmr	−0.2	+0.3	−0.8	+1.9	−0.7	
C1	−0.2	+0.3	−0.2	+0.2	−0.2	
C1h	−0.2	+0.2	−0.2	+0.2	−0.1	+0.1
C1m	−0.3	+0.4	−0.2	+0.5	−0.1	−0.2
C2	−0.2	+0.2	−0.1		−0.1	
C2h	−0.3	+0.3	−0.2	+0.1	+0.1	−0.2
C2m	−0.4	+0.8	−0.6	+0.5	−0.6	−0.2
C2r		+0.1			+0.1	−0.3
C3						+0.1
C3h				+0.1	+0.2	+0.1
C3m	−0.1	+0.1	−0.2	+0.1	−0.2	
C3r	+0.2	−0.1	+0.1	+0.1	+0.1	−0.1
C3mr		+0.1	−0.2	+0.2	−0.2	−0.1
C3hmr		+0.2	−0.2	+0.2		
D1				+12.8		+1.2
D1h	−0.5	+0.8	−0.5	+3.2	−0.4	+0.5
D1m	+0.3	−0.2	−0.3	+12.1	−0.3	+0.9
D2	+0.2	−0.2	−0.2	+0.5	−0.2	+0.2
D2h	+0.3	−0.2	−0.1	+0.4	−0.1	+0.2
D2m	+26.9	−1.0	+0.2	+23.8	+0.2	+4.8
D2r	+4.3	−0.8	+0.1	+2.5	+0.2	
D3	−0.2	+0.3	−0.1	−0.1	−0.1	−0.1
D3h	−0.3	+0.4	−0.2	−0.1	−0.1	−0.1
D3m	+0.2	−0.2	+0.1	+0.8	+0.1	
D3r	+0.2	−0.1	+0.1	+0.1	+0.1	
D3mr	+0.7	−0.2		+1.1		
D3hmr	−0.3	+0.9	−1.0	+1.8	−1.0	−0.2

Blank entries denote negligible bias, between −0.05 and +0.05.

Table 7 Failure rates of estimated reference points, by scenario and model.

	$B_{current}$	$u_{current}$	Depletion	MSY	$B_{current}/B_{MSY}$	Surplus
A1	20	13	34	1	34	7
A1h	24	20	29	29	39	17
A1m	15	11	39	12	41	12
A2	9	3	10	0	10	5
A2h	24	15	25	6	45	20
A2m	16	10	39	7	40	7
A2r	19	4	16	1	17	5
A3	0	0	2	0	2	1
A3h	1	0	5	1	17	1
A3m	5	4	28	7	28	4
A3r	1	0	4	0	4	0
A3mr	2	1	26	7	26	4
A3hmr	2	1	25	4	27	1
B1	74	70	73	63	73	24
B1h	83	82	81	46	75	25
B1m	74	70	74	58	77	25
B2	34	27	22	12	22	21
B2h	45	38	41	28	49	28
B2m	65	64	67	61	69	50
B2r	45	39	29	21	30	30
B3	0	0	3	0	3	4
B3h	2	2	20	20	43	5
B3m	49	39	68	56	69	36
B3r	21	10	18	5	20	9
B3mr	45	37	62	45	65	34
B3hmr	60	58	81	69	88	29
C1	32	32	35	13	35	24
C1h	21	22	32	17	36	25
C1m	41	40	46	39	52	23
C2	16	14	8	4	8	11
C2h	36	36	33	31	47	38
C2m	51	52	62	35	62	39
C2r	21	16	8	6	9	37
C3	0	0	0	0	0	4
C3h	0	0	7	5	19	7
C3m	3	3	28	7	30	12
C3r	9	1	3	0	3	17
C3mr	8	8	30	8	31	22
C3hmr	8	8	35	15	42	22
D1	74	73	46	59	46	57
D1h	91	90	56	59	56	59
D1m	84	83	52	69	49	58
D2	69	66	26	54	26	61
D2h	77	76	24	73	25	77
D2m	92	90	44	88	48	87
D2r	83	85	21	78	23	77
D3	6	5	10	3	10	13
D3h	9	6	17	13	21	15
D3m	49	47	36	47	38	42
D3r	31	25	9	26	11	32
D3mr	53	46	33	52	35	41
D3hmr	85	89	64	57	65	55

in data scenario A (one-way trip), with models 2 and 2m exhibiting a negative bias of −0.2. The failure rate is also relatively low in scenario A, ranging from 0 for model 3, to 24 for models 1h and 2h. Most of the boxplots are wider in scenario B (no change), indicating that the data in this scenario are less informative about current abundance. Models 3 and 3h are exceptions from this general pattern, as their performance is comparable to scenario A. The considerably higher failure rate of models 3m, 3r, 3mr and 3hmr in scenario B shows how the uncertainty increases when the natural mortality rate and/or right-hand selectivity are unknown. The performance of the estimation models in scenario C (good contrast) is better than in scenario B and about as good as scenario A. The greatest bias in scenario C is −0.4 for model 2m, which also has a relatively high failure rate of 51, while models 3 and 3h have 0. Scenario D (rebuild only) is the least informative about current abundance. The lowest failure rates are 6 and 9 for models 3 and 3h, but the estimates from these models have a bias of −0.2 and −0.3. The other models have much higher failure rates in scenario D, including the highest of all cases, 92 for model 2m.

$u_{current}$

The current harvest rate (Fig. 6, second row of panels) is never greatly overestimated in scenario A. This is because the estimated fraction of the biomass caught in a year cannot be many times higher than the true value of 0.408 in this scenario (Table 5). Nevertheless, a small but consistent positive bias of *c.* +0.1 is shown by model family 3, but failure rates are quite low, 0 for models 3, 3h and 3r, up to 20 for model 1h. In scenario B, all models have high failure rates except for models 3, 3h and 3r, with 0, 2 and 10, respectively. Model 3r is unbiased, but 3 and 3h are positively biased by +0.1. The models that estimate natural mortality, 1m, 2m, 3m, 3mr and 3hmr all show high failure rates, between 37 and 70. Failure rates in scenario C are lower than in scenario B, but higher than in scenario A. The median bias ranges from 0 for models 3 and 3h, to +0.8 for model 2m, and failure rates are lower for model family 3 than the simpler models. Model performance in scenario D is considerably worse than in the other scenarios. Models 3 and 3h have low failure rates of 5 and 6, but consistently overestimate the harvest rate with a bias of +0.3 and +0.4. Model 3r has a smaller bias of −0.1, but a failure rate of 25, while model 2m shows an extreme −1.0 bias and a failure rate of 90.

Depletion

Many of the models estimate current depletion (Fig. 6, third row of panels) in scenario A about as well as current abundance, but there are noteworthy exceptions. Specifically, the failure rate is consistently higher for current depletion compared to current abundance when the natural mortality rate is unknown, in models 1m, 2m, 3m, 3mr and 3hmr. Scenario B is again less informative than scenario A about current depletion, with median bias ranging from −0.8 for models 1m and 3hmr, to +0.3 for model 3m, and failure rate from 3 for model 3 to 81 for models 1h and 3hmr. Models 2, 2r, 3, 3h and 3r perform quite well in scenario C, with failure rates below 10, while model 2m is negatively biased and has a high failure rate of 62. Scenario D is clearly more informative about depletion than absolute abundance, with the models showing less extreme biases and failure rates. Nevertheless, models 1h, 1m and 3hmr provide negatively biased and inaccurate estimates of current depletion in scenario D.

MSY

All models in scenario A estimate the maximum sustainable yield (Fig. 6, fourth row of panels) with quite low failure rates, although the estimates are often slightly biased towards overestimation. Model 1h has the highest failure rate, 29, but the failure rates of the remaining models are between 0 and 12. Models 2, 2r, 3, 3h and 3r perform relatively well in scenario B, but the other models overestimate MSY considerably. Model performance in scenario C is again similar to that in scenario A, except that models 1m and 2m have larger bias and higher failure rates. Scenario D is highly uninformative about MSY for all models except 3 and 3h, which have failure rates 3 and 13, respectively. Other models in this scenario are positively biased with failure rates between 26 and 88.

$B_{\mathrm{current}}/B_{\mathrm{MSY}}$

The ability to estimate the ratio $B_{\mathrm{current}}/B_{\mathrm{MSY}}$ (Fig. 6, fifth row of panels) is largely similar to that for current depletion, the other ratio reference point, reflecting a strong correlation between the B_0, B_{current} and B_{MSY} parameter estimates. The failure rates in scenario A range from 2 for model 3, to 45 for model 2h. Scenario B is less informative about the stock status relative to B_{MSY}, although models 2, 3 and 3r perform relatively well. The estimation models in scenario C are subject to rather small biases, with the exception of −0.6 for model 2m. In scenario D, most of the models have failure rates below 50, but model 3hmr is strongly biased downwards, with median bias −1.0.

Surplus

All models estimate current surplus production (Fig. 6, bottom row of panels) quite accurately in scenario A, with failure rates ranging from 0 for model 3r, to 20 for model 2h, which also shows the greatest bias of +0.2. Scenario B is much more informative about surplus production than about other reference points, with a bias of −0.2 to +0.1, and failure rates from 8 for model 3, to 45 for model 2m. The estimation performance is also good in scenario C, with the greatest bias being −0.3 for model 2r, and failure rates ranging from 4 for model 3, to 39 for model 2m. In scenario D, failure rates are generally high, over 50 for all models in families 1 and 2. The only models that perform well here are 3 and 3h, with failure rates of 13 and 15, and a bias of −0.1. Thus, estimating surplus production reliably in scenario D requires that both catch-at-age data and an index of abundance are available, as well as perfect knowledge about the true natural mortality rate, recruitment steepness and right-hand selectivity.

Parameters

Steepness h and natural mortality M are estimated directly in the model, while selectivity at oldest age S_{10} is a derived parameter from Equation (3). In Fig. 7, the estimated parameter values are divided by the true values from the operating model, which are $h = 0.7$, $M = 0.2$ and $S_{10} = 0.94$.

Steepness is overestimated by all models in all scenarios, but relatively accurate estimates are seen in scenario D using models 2h and 3h. By definition, steepness has an upper bound of 1 (Francis 1992) and many estimates in the top row of panels in Fig. 7 run into this bound, where $\hat{h}/h = 1.0/0.7 = 1.43$, but less frequently in scenario D. Estimates of natural mortality rate are generally unreliable, especially using the 1m or 3hmr models, but relatively accurate estimates are seen in scenario A using the 3m model. When right-hand selectivity is estimated as well, in models 3mr and 3hmr, M becomes biased towards underestimation. In other words, when the estimated selectivity does not fully target older fish, the relatively high frequency of older fish in the catch can be fitted by increasing the natural mortality rate. Selectivity at oldest age is consistently underestimated (Fig. 7, bottom row of panels). This bias is partly due to the true value of $S_{10} = 0.94$ being so near the theoretical upper bound of 1, but the estimates are also inaccurate, in many cases less than half the true value. The performance does not differ much between models 2r, 3r, 3mr and 3hmr,

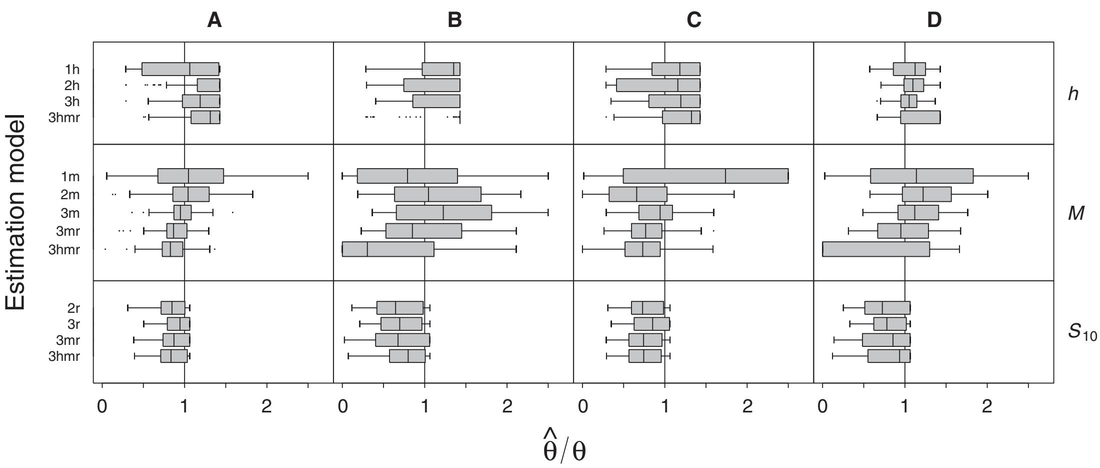

Figure 7 Distribution of estimated parameters h (steepness), M (natural mortality) and derived parameter S_{10} (selectivity at oldest age). Panel columns correspond to fishing history scenarios A–D and panel rows are the different parameters. Each Tukey boxplot shows the distribution of 100 estimates, divided by the true value of the parameter from the operating model.

but scenario A is slightly more informative than the others about selectivity at oldest age.

Summary and discussion

Below, the hypotheses are reviewed in light of the results, using average failure rate as a summary statistic. This is followed by a general discussion about implications and the strengths and weaknesses of the experimental design.

Hypotheses

H_1 Fisheries data are most informative when they span a period where the population was fished down to a low level.

H_2 Fisheries data are most informative when they span a period where the population was fished down to a low level and then allowed to rebuild for some time.

Table 8 shows the estimation performance in each scenario, where each average is based on 65 failure rates from Table 7, across models and reference points. There is a clear division, also noticeable in Fig. 6, where fishing histories A (one-way trip) and C (good contrast) provide more reliable data to estimate the reference points than B (no change) and D (rebuild only). The results provide slightly more support to hypothesis 1 (average failure rate in scenario A = 12.7) than hypothesis 2 (C = 21.0), but both of those scenarios are much more informative than B or D.

The results from the 'one-way trip' scenario imply that fisheries data spanning an early period of high abundance followed by low abundance are likely to be informative in age-structured stock assessment, even if the fishing history does not include subsequent rebuilding. This is in contrast to findings from simulation studies of biomass-dynamic models, where a rebuilding phase provides necessary information to estimate the population growth parameters (Hilborn 1979; Hilborn and Walters 1992). It is worth noting that those studies looked at how well the parameters of the Schaefer model were estimated, not just reference points.

H_3 The level of stock depletion is estimated more reliably than other reference points.

The reference point with the lowest overall failure rate is not the current depletion level, but surplus production (Table 9). But from Fig. 6 it is clear that Depletion is a more robust reference point across all scenarios. The reference points that are in absolute biomass units (B_{current}, MSY and Surplus) become highly unreliable in scenario D (rebuild only), particularly when natural mortality is unknown. The relative biomass estimates (Depletion and $B_{\mathrm{current}}/B_{\mathrm{MSY}}$) perform much better in those cases.

Punt *et al.* (2002) and other studies have shown that depletion is generally estimated more reliably than other reference points. The greater the correlation is between the B_{current} and B_0 parameter estimates, the smaller the variance around the estimated ratio of the two. The results from scenarios A through D indicate that depletion may be subject to slightly higher failure rates than some other reference points when the data are informative, but is a robust quantity to estimate in worst-case uninformative scenarios.

An interesting exception is how accurately current surplus production is estimated in scenario B. This is understandable, since if the abundance index and catch is constant over time, then the surplus production must be roughly equal to the catch.

Table 8 Average failure rate in each scenario, across all reference points and models.

Scenario A	Scenario B	Scenario C	Scenario D
12.7	41.8	21.0	49.1

A, one-way trip; B, no change' C, good contrast; D, rebuild only.

Table 9 Average failure rate for each reference point, across all models and scenarios.

B_{current}	u_{current}	Depletion	MSY	$B_{\mathrm{current}}/B_{\mathrm{MSY}}$	Surplus
34.3	31.4	32.4	27.2	35.4	26.1

B_{current}, current spawning biomass; u_{current}, current harvest rate; Depletion, current depletion level; MSY, maximum sustainable yield; $B_{\mathrm{current}}/B_{\mathrm{MSY}}$, current abundance relative to B_{MSY}; Surplus, current surplus production.

Table 10 Average failure rate for each estimation model family, across all reference points and scenarios.

Family 1	Family 2	Family 3
45.4	35.8	20.9

1, landings and abundance-index data only; 2, landings and catch-at-age data only; 3, all three data types.

H_4 A data set that includes both an index of relative abundance and catch-at-age data is much more informative than a data set that includes only one of these two types of data.

Estimation models of family 3 (all data types) perform better than family 1 (no age data) and family 2 (no abundance index), as shown in Table 10. This comes as no surprise and is in agreement with the recommendations by Deriso *et al.* (1985). Models similar to those of family 1 have been used by Hilborn (1990) and others, and are seen by many as a preferable alternative to traditional biomass-dynamic models (Maunder 2003). Their argument is that traditional biomass-dynamic models make implicit assumptions that offer limited freedom to explore different hypotheses about the fishery dynamics.

Model family 2 performs surprisingly well. Even in the absence of abundance-index data, the landings and age-composition data provide considerable information to estimate both absolute abundance and relative depletion. The common view is that estimation of these quantities requires either an abundance index or highly restrictive assumptions (Shepherd 1984; Deriso *et al.* 1985; Hilborn and Walters 1992), but here the assumptions of model family 2 are similar to families 1 and 3. Furthermore, convergence diagnostics indicated that models of family 2 were no less estimable than the other models. This behaviour may be due to the simplified nature of the simulation environment, but it is worth remembering that a precursor of the statistical catch-at-age model (Doubleday 1976) did not include abundance-index data. Statistical catch-at-age models combine the landings and the commercial catch-at-age data in a framework that provides more insight than analysing each cohort separately (Fournier and Archibald 1982). An additional element of information for model families 1–3 is the penalty (Equations 9 and 12) that allows the estimated recruitment to vary considerably (*c.* 20-fold difference between largest and smallest recruitment in scenario B with $\sigma_R = 0.6$) but not by many orders of magnitude, whereas recruitment is completely free in the model used by Deriso *et al.* (1985).

H_5 Not knowing *M*, *h*, and right-hand selectivity leads to inaccurate estimates of stock abundance and reference points.

As expected, the models perform better when parameters are fixed at the true value, than when they are estimated. Even so, the 3hmr model performs quite well in the informative scenarios A and C, especially estimating current abundance, harvest rate, MSY and surplus production (Fig. 6; Tables 6 and 7).

By admitting uncertainty about *M*, *h*, and right-hand selectivity, model 3hmr represents the real task facing stock assessment scientists. These parameters are highly confounded, so model 3hmr cannot be expected to perform reliably when fitted to real fisheries data, that come from a much more complex system than the operating model used in this study. In practice, some or all of these parameters would be fixed at an assumed value, or be assigned an informative Bayesian prior probability distribution. The effect of fixing these parameters at values that are very different from the true dynamics has been explored by Thompson (1994), Clark (1999), Ianelli (2002) and others.

H_6 Models estimating *M* perform about as well as models estimating *h*.

Models estimating *h* perform better on the average than those estimating *M*, especially when the data include both catch at age and an index of abundance (Table 11). This means that uncertainty about the natural mortality rate is more important than the uncertainty about the shape of the stock-recruitment curve, when estimating the stock status. Model 3m shows particularly bad performance in scenarios B and D, so in those scenarios any external information about *M* would be valuable. This information could be used to construct a Bayesian prior for *M*, or to fix the parameter, which is analogous to an extremely narrow Bayesian prior (Gelman *et al.* 2004).

The highest overall failure rates were shown by model 2m in scenario D. As expected (Beverton and Holt 1957), it is simply not feasible to estimate harvest rate and natural mortality from catch-at-age, when harvest rate has been steady and low.

H_7 *M* can be estimated reliably if age-composition data are available from when the population was unfished.

Table 11 Average failure rate for estimation models 1h, 1m, 2h, 2m, 3h and 3m, across all reference points and scenarios.

	h	m
Model family 1	46.4	47.7
Model family 2	39.0	51.9
Model family 3	9.8	30.6

H_8 M can be estimated reliably from the rate of population increase if the stock is allowed to rebuild from a low level.

Model 3m estimates M with greater accuracy in scenarios A and C than in the other scenarios (Fig. 7). This was expected, since the age structure in the first few years of the fishery carries information about the natural mortality rate (Beverton and Holt 1957). After taking the individual cohort sizes into account, using data from all years, M can be inferred from the age composition in the first years. Importantly, the model was not 'told' that the stock was unfished in the first year in scenarios A and C, as the parameters R_{init} and u_{init} were estimated in all cases.

The estimation of M is less accurate in scenario D. One might have expected this scenario to be informative about the value of M, as the rate at which the stock rebuilds is dependent on this parameter. The other main factor determining the rebuilding rate is recruitment, so the variable recruitment ($\sigma_R = 0.6$) might explain why scenario D is not informative about M. Another reason could be that the observation noise ($\sigma_I = 0.2$) causes the observed abundance index to suggest random fluctuations instead of a steady growth.

H_9 h can be estimated reliably from catch-at-age data and an index of relative abundance when the data cover a period in which abundance varies substantially.

Although scenario A involves the widest range of abundance (Fig. 2), it is only in scenario D that model 3h estimates h reliably (Fig. 7). Recruitment success at low spawning stock levels is informative about the shape of the stock-recruitment curve (Ricker 1958), and scenario D includes a large number of cohorts spawned by a small parent stock. The initial stock status in scenario D (*c.* 3% of B_0) is also considerably lower than the last year's stock status in scenario A (*c.* 10% of B_0). When the data do not include years of very low abundance, the models tend to overestimate the steepness parameter (Fig. 7).

H_{10} Right-hand selectivity can only be estimated reliably when M is known.

The results from this study are not conclusive about the estimation of right-hand selectivity, as the only case considered is when the true selectivity curve is nearly asymptotic. Nonetheless, the results do suggest that when the true selectivity is nearly asymptotic, the reliability of estimating right-hand selectivity (Fig. 7, bottom panel) depends more on the scenario than on whether M is estimated or fixed. Thompson (1994) performed a more thorough analysis of the relationship between these parameters, concluding that right-hand selectivity can only be estimated reliably when M is known.

Implications

The results presented here show how the perceived uncertainty about stock status is not only affected by the available data, but also by the assumptions made in the estimation process.

The features of different fishing history scenarios determine how informative the data are about management quantities. The main feature of an informative fishing history is a large decrease in abundance, while other features, such as contrast in harvest rate, seem to be of secondary importance. Although strong depletion is to be avoided due to the ecological risk and economic cost, it does provide informative data. In the words of John G. Pope (personal communication), 'the more fish you catch, the better you know how many there were.'

An uninformative fishing history, commonly seen in practice, is when a relative index of abundance and age data are not available from the early years of the fishery. In these cases, depletion level tends to be more robust than other commonly used reference points, although surplus production can also be estimated accurately when the abundance remains stable over a long period. When the data are informative, other reference points can be expected to perform just as well, or better. Despite regular criticism, MSY remains a key concept in fisheries management, if not as a goal, then as an upper limit of a precautionary approach (Mace 2001; Punt and Smith 2001). MSY is independent of the current stock status, being a function of R_0, M, h, commercial selectivity, weight and maturity at age. When the estimation models are given the true value of most of those quantities, MSY will be estimated quite accurately. This can be seen from the performance of our models 1, 2 and 3, which generally estimated MSY with a lower failure rate than the other reference points. However, it is also important to note that MSY was more often overestimated than underestimated.

Catch-at-age data can provide information about the current stock status, even without a relative abundance index. When the true value of M is known, the total mortality rate of cohorts leads to an accurate estimate of annual harvest rate, which combined with known annual catches leads to an accurate estimate of vulnerable biomass. The assumption of a known constant M plays a central role here. This assumption is commonly made in practice, and the effects of its violations are largely understood (Mertz and Myers 1997; Clark 1999). It is an unrealistic assumption (Cotter *et al.* 2004), but a time-constant M, estimated or fixed, is seen as necessary to evaluate the consequences of alternative catch levels (Punt and Hilborn 1997), which is the central purpose of fisheries stock assessment.

The statistical catch-at-age model yields more information from catch-at-age data than earlier catch-curve methods, given that the added assumptions about recruitment are justifiable. Catch-at-age and abundance-index data become particularly informative when used together, as they provide complementary information about different aspects of the population dynamics, and are subject to different assumptions. It is known that the sampled and processed catch-at-age data do not necessarily reflect the population age-structure very well (Pope 1988), and empirical evidence also undermines the assumption about a constant linear relationship between the abundance index and population abundance (Harley *et al.* 2001). When these two data types tell a consistent story about the population trends, it indicates that the model assumptions are likely to be justifiable. This can be checked by fitting models that exclude data components, as was carried out in this study, or by changing the likelihood weights via the catch-at-age sample size and observation uncertainty σ_I about the abundance index. When the two data types provide contradicting information about the stock status, the validity of each data source needs to be examined (Schnute and Hilborn 1993).

When evaluating confidence bounds around estimated quantities, one should strive to incorporate all major sources of uncertainty. This means estimating parameters instead of fixing them, but this is not always statistically feasible. Confounded parameters like natural mortality rate M, stock-recruitment steepness h and declining right-hand selectivity can be estimated when the data carry information about these quantities. For M, this means catch-at-age data from the early years of the fishery, or at least a contrasted history of harvest rates (Beverton and Holt 1957) and for h it means catch-at-age data from a period of very low abundance (Ricker 1958), as verified in this study. With the constant harvest rates in scenarios B and D, there is no information to separate the total mortality rate between natural mortalities and fishing mortalities. The selectivity of older fish can be estimated when M is a fixed parameter (Thompson 1994). In a Bayesian model, one or more of these parameters can be assigned an informative prior distribution, perhaps from a meta-analysis of many related stocks (Myers *et al.* 1999), instead of estimating as a free parameter or fixing completely.

Overall, the estimation models showed considerably high failure rates, where management quantities were underestimated or overestimated by a factor of two or more. Bearing in mind the simple 'laboratory conditions' of this simulation study, stock assessment models can only be expected to have higher failure rates when fitted to real fisheries data. A retrospective look at fisheries assessments around the world shows that management quantities are not estimated as accurately as statistical theory suggests, due to violated assumptions and ignored sources of uncertainty (NRC 1998; Walters and Martell 2004).

Strengths and weaknesses

This study advances our understanding of fisheries stock assessment models, with respect to what kinds of data are informative or uninformative, and highlights the role of assumptions. Based on the experimental design and findings of Hilborn (1979), Hilborn and Walters (1992), NRC (1998), Gavaris and Ianelli (2002) and Punt *et al.* (2002), this simulation study uses up-to-date statistical methods that take advantage of the computing power available today. The scope is wide, addressing a variety of questions, and the conclusions can be used to support various decisions made in any fisheries stock assessment.

Compared to real fisheries, with their complex interaction between biological and human systems, the simulation approach is a simplified abstraction. Apart from stochastic recruitment, the parameters in the operating model are constant over time (natural mortality rate, catchability and selectivity), and the estimation models are specified without model error and given the true survey selectivity. These decisions were made deliberately to make the results as easy to interpret as possible. Excluding survey catch-at-age data from the study allowed a clear separation between two kinds of information: commercial catch at age reflecting the age distribution of the population over time and a survey abundance index reflecting relative changes in biomass over time.

The wide scope of this study comes at a cost, as the experimental design is not optimal for any one of the 10 hypotheses. Each of them could be tested more rigorously with a simulation study specifically designed for that purpose. Similarly, there are many more hypotheses that could be addressed using the same simulation framework, but applying other treatments than was done here. For example, examining the effect of fixing parameters at values that are substantially lower or higher than the true value. Model errors and violated assumptions are inevitable in stock assessments, but the combined experience from real fisheries and simulation studies will help making fisheries data as informative as possible.

Acknowledgements

This research was supported by the Icelandic Centre for Research and the New Zealand Seafood Industry Council. We thank André Punt, Jim Ianelli, Rick Methot, John Pope, and two anonymous reviewers for insightful comments, and Pamela Nelle for a thorough review of the manuscript.

References

Bence, J.R., Gordoa, A. and Hightower, J.E. (1993) Influence of age-selective surveys in the reliability of stock synthesis assessments. *Canadian Journal of Fisheries and Aquatic Sciences* **50**, 827–840.

Beverton, R.J.H. and Holt, S.J. (1957) *On the Dynamics of Exploited Fish Populations*. Her Majesty's Stationery Office, London.

Chapman, D.G. and Robson, D.S. (1960) The analysis of a catch curve. *Biometrics* **16**, 354–368.

Clark, W.G. (1999) Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 1721–1731.

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 and Fisheries* **5,** 235–254.

Deriso, R.B. (1980) Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences* **37**, 268–282.

Deriso, R.B., Quinn, T.J. II and Neal, P.R. (1985) Catch-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 815–824.

Doubleday, W.G. (1976) A least squares approach to analysing catch at age data. *ICNAF Research Bulletin* **12**, 69–81.

Fournier, D. and Archibald, C.P. (1982) A general theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 1195–1207.

Fournier, D.A., Sibert, J.R., Majkowski, J. and Hampton, J. (1990) MULTIFAN: a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 301–317.

Francis, R.I.C.C. (1992) Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. *Canadian Journal of Fisheries and Aquatic Sciences* **49**, 922–930.

Gavaris, S. (1988) *An Adaptive Framework for the Estimation of Population Size*. Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 88/29, 12 pp.

Gavaris, S. and Ianelli, J.N. (2002) Statistical issues in fisheries' stock assessments. *Scandinavian Journal of Statistics* **29**, 245–271.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2004) *Bayesian Data Analysis*, 2nd edn. Chapman and Hall, Boca Raton, FL.

Gulland, J.A. (1965) *Estimation of Mortality Rates*. ICES CM 1965 Gadoid Fisheries Committee Document 3, 9 pp.

Harley, S.J., Myers, R.A. and Dunn, A. (2001) Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1760–1772.

Hilborn, R. (1979) Comparison of fisheries control systems that utilize catch and effort data. *Journal of the Fisheries Research Board of Canada* **36**, 1477–1489.

Hilborn, R. (1990) Estimating the parameters of full age-structured models from catch and abundance data. *INPFC Bulletin* **50**, 207–213.

Hilborn, R. and Walters, C.J. (1992) *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Chapman and Hall, New York.

Hilborn, R., Maunder, M., Parma, A., Ernst, B., Payne, J. and Starr, P. (2003) *Coleraine: a Generalized Age-structured Stock Assessment Model*. User's Manual Version 2.0. University of Washington Report SAFS-UW 0116, 54 pp. Available at: http://fish.washington.edu/research/coleraine/coleraine.pdf, accessed on 11 June 2007.

Ianelli, J.N. (2002) Simulation analyses testing the robustness of productivity determinations from West Coast Pacific Ocean perch stock assessment data. *North American Journal of Fisheries Management* **22**, 301–310.

ICES (2003) *Report of the North-Western Working Group*. ICES CM 2003/ACFM 24, Icelandic cod, pp. 144–227.

Kimura, D.K. and Tagart, J.V. (1982) Stock reduction analysis: another solution to the catch equations. *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 1467–1472.

Lawson, T.A. and Hilborn, R. (1985) Equilibrium yields and yield isopleths from a general age-structured model of harvested populations. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 1766–1771.

Leslie, P.H. and Davis, D.H.S. (1939) An attempt to determine the absolute number of rats on a given area. *Journal of Animal Ecology* **8**, 94–113.

Ludwig, D. and Hilborn, R. (1983) Adaptive probing strategies for age-structured fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences* **40**, 559–569.

Mace, P.M. (2001) A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. *Fish and Fisheries* **2**, 2–32.

Maunder, M.N. (2003) Is it time to discard the Schaefer model from the stock assessment scientist's toolbox? *Fisheries Research* **61**, 145–149.

Megrey, B.A. (1989) Review and comparison of age-structured stock assessment models from theoretical and applied points of view. In: *Mathematical Analysis of Fish Stock Dynamics* (eds E.F. Edwards and B.A. Megrey). American Fisheries Society, Bethesda, MD, pp. 8–48.

Mertz, G. and Myers, R.A. (1997) Influence of errors in natural mortality estimates in cohort analysis. *Canadian Journal of Fisheries and Aquatic Sciences* **54**, 1608–1612.

Methot, R.D. (1989) Synthetic estimates of historical abundance and mortality for northern anchovy. In: *Mathematical Analysis of Fish Stock Dynamics* (eds E.F. Edwards and B.A. Megrey). American Fisheries Society, Bethesda, MD, pp. 66–82.

Myers, R.A., Bowen, K.G. and Barrowman, N.J. (1999) Maximum reproductive rate of fish at low population sizes. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 2404–2419.

NRC (1998) *Improving Fish Stock Assessments*. National Academy, Washington, DC.

Patterson, K., Cook, R., Darby, C. et al. (2001) Estimating uncertainty in fish stock assessment and forecasting. *Fish and Fisheries* **2**, 125–157.

Pope, J.G. (1972) An investigation of the accuracy of virtual population analysis using cohort analysis. *ICNAF Research Bulletin* **9**, 65–74.

Pope, J.G. (1988) Collecting fisheries assessment data. In: *Fish Population Dynamics* (ed. J.A. Gulland), 2nd edn. Wiley, Chichester, pp. 63–82.

Punt A.E. and Hilborn R. (1997) Fisheries stock assessment and decision analysis: The Bayesian approach. *Reviews in Fish Biology and Fisheries* **7**, 35–63.

Punt, A.E. and Smith, A.D.M. (2001) The gospel of maximum sustainable yield in fisheries management: birth, crucifixion and reincarnation. In: *Conservation of Exploited Species* (eds J.D. Reynolds, G.M. Mace, K.H. Redford and J.G. Robinson). Cambridge University Press, Cambridge, pp. 41–66.

Punt, A.E., Smith, A.D.M. and Cui, G. (2002) Evaluation of management tools for Australia's South East Fishery 2: how well can management quantities be estimated? *Marine and Freshwater Research* **53,** 631–644.

Quinn, T.J. II (2003) Ruminations on the development and future of population dynamics models in fisheries. *Natural Resource Modeling* **16**, 341–392.

Quinn, T.J. II and Deriso, R.B. (1999) *Quantitative Fish Dynamics*. Oxford University Press, New York.

Ricker, W.E. (1958) Handbook of computations for biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada* **119**, 300 pp.

Rose, K.A., Cowan, J.H. Jr, Winemiller, K.O., Myers, R.A. and Hilborn, R. (2001) Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. *Fish and Fisheries* **2**, 293–327.

Sampson, D.B. and Yin, Y. (1998) A Monte Carlo evaluation of the stock synthesis assessment program. In: *Fishery Stock Assessment Models* (eds F. Funk, T.J. Quinn II, J. Heifetz et al.). Alaska Sea Grant Program, Fairbanks, pp. 315–338.

Schaefer, M.B. (1954) Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *IATTC Bulletin* **1**, 27–56.

Schnute, J.T. (1985) A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 414–429.

Schnute, J.T. and Hilborn, R. (1993) Analysis of contradictory data sources in fish stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences* **50**, 1916–1923.

Shepherd, J.G. (1984) The availability and information content of fisheries data. In: *Exploitation of Marine Communities* (ed. R.M. May). Springer, Berlin, pp. 95–109.

Shepherd, J.G. (1999) Extended survivors analysis: an improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science* **56**, 584–591.

Smith, M.T. and Addison, J.T. (2003) Methods for stock assessment of crustacean fisheries. *Fisheries Research* **65**, 231–256.

Thompson, G.G. (1994) Confounding of gear selectivity and the natural mortality rate in cases where the former is a nonmonotone function of age. *Canadian Journal of Fisheries and Aquatic Sciences* **51**, 2654–2664.

Tukey, J.W. (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

Walters, C. (1986) *Adaptive Management of Renewable Resources*. Macmillan, New York.

Walters, C.J. (2007) Is adaptive management helping to solve fisheries problems? *Ambio* **36**, 304–307.

Walters, C.J. and Martell, S.J.D. (2004) *Fisheries Ecology and Management*. Princeton University Press, Princeton.

Xiao, Y. (2000) A general theory of fish stock assessment models. *Ecological Modelling* **128**, 165–180.

Yin, Y. and Sampson, D.B. (2004) Bias and precision of estimates from an age-structured stock assessment program in relation to stock and data characteristics. *North American Journal of Fisheries Management* **24**, 865–879.

FISH and FISHERIES

FISH and FISHERIES, 2012, **13**, 479–480

Erratum

Over the years, we have discovered typographical errors in the following equations of our paper (Magnusson and Hilborn 2007).

Equation 2

$$N_{t+1,A} = N_{t,A-1}e^{-M}(1 - {}_CS_Au_t) + N_{t,A}e^{-M}(1 - {}_CS_Au_t)$$

should be

$$N_{t+1,A} = N_{t,A-1}e^{-M}(1 - {}_CS_{A-1}u_t) + N_{t,A}e^{-M}(1 - {}_CS_Au_t)$$

changing the first ${}_CS_A$ to ${}_CS_{A-1}$.

Equation 4

$$\begin{aligned} N_{1,1} &= R_0R_{\text{init}} \times \exp({}_R\varepsilon_{1,1} - \sigma_R^2/2) \\ N_{1,a} &= R_0R_{\text{init}}e^{-(a-1)M}\sum_{i=1}^{a=1}(1 - {}_CS_iu_{\text{init}}) \times \exp({}_R\varepsilon_{1,a} - \sigma_R^2/2) \\ N_{1,A} &= R_0R_{\text{init}}e^{-(A-1)M}\sum_{i=1}^{A-1}(1 - {}_CS_iu_{\text{init}})/[1 - e^{-M}(1 - {}_CS_Au_{\text{init}})] \times R_{\text{plus}} \end{aligned}$$

should be

$$\begin{aligned} N_{1,1} &= R_0R_{\text{init}} \times \exp({}_R\varepsilon_{1,1} - \sigma_R^2/2) \\ N_{1,a} &= R_0R_{\text{init}}e^{-(a-1)M}\prod_{i=1}^{a-1}(1 - {}_CS_iu_{\text{init}}) \times \exp({}_R\varepsilon_{1,a} - \sigma_R^2/2) \\ N_{1,A} &= R_0R_{\text{init}}e^{-(A-1)M}\prod_{i=1}^{A-1}(1 - {}_CS_iu_{\text{init}})/[1 - e^{-M}(1 - {}_CS_Au_{\text{init}})] \times R_{\text{plus}} \end{aligned}$$

changing both $\sum$ to $\prod$.

Equation 11

$$-\log L_c = \sum_t\sum_a \log\left[\exp\left(\frac{(P_{t,a} - \hat{P}_{t.a})^2}{2[P_{t,a}(1 - P_{t,a}) + 0.1/A]n^{-1}}\right) + 0.01\right]$$

should be

$$-\log L_c = -\sum_t\sum_a \log\left[\exp\left(\frac{-(P_{t,a} - \hat{P}_{t.a})^2}{2[P_{t,a}(1 - P_{t,a}) + 0.1/A]n^{-1}}\right) + 0.01\right]$$

changing $\sum$ to $-\sum$

and $(P_{t,a} - \hat{P}_{t,A})^2$ to $-(P_{t,a} - \hat{P}_{t,a})^2$.

Equation 12

$$\text{Pen} = \sum_{a=2}^{A-1}\frac{{}_R\varepsilon_{t,a}^2}{2\sigma_R^2} + \sum_{t=2}^{t_{\max}-1}\frac{{}_R\varepsilon_{t,1}^2}{2\sigma_R^2}$$

should be

$$\text{Pen} = \sum_{a=2}^{A-1}\frac{{}_R\varepsilon_{1,a}^2}{2\sigma_R^2} + \sum_{t=2}^{t_{\max}-1}\frac{{}_R\varepsilon_{t,1}^2}{2\sigma_R^2}$$

changing ${}_R\varepsilon_{t,a}^2$ to ${}_R\varepsilon_{1,a}^2$.

DOI: 10.1111/j.1467-2979.2012.00479.x

Equation 16

$$n_a^* = \begin{cases} e^{-(a-1)M}\prod_{i}^{a-1} 1 - {}_cS_iu, & a < A \\ \dfrac{e^{-(A-1)M}\prod_{i}^{A-1} 1 - {}_cS_iu}{e^{-M}(1 - {}_cS_Au)}, & a = A \end{cases}$$

should be

$$n_a^* = \begin{cases} e^{-(a-1)M}\prod_{i=1}^{a-1}(1 - {}_cS_iu), & a < A \\ \dfrac{e^{-(A-1)M}\prod_{i=1}^{A-1}(1 - {}_cS_iu)}{1 - e^{-M}(1 -_c S_Au)}, & a = A \end{cases}$$

changing $\prod_{i}^{a-1}$ to $\prod_{i=1}^{a-1}$,

$e^{-M}(1 - {}_cS_Au)$ to $1 - e^{-M}(1 - {}_cS_Au)$,

and $1 - {}_cS_iu$ to $(1 - {}_cS_iu)$.

The corrected equations are those used in the original analysis.

Reference

Magnusson, A. and Hilborn, R. (2007) What makes fisheries data informative? *Fish and Fisheries* **8**, 337–358.