ICES Journal of Marine Science, 60: 314–328. 2003
doi:10.1016/S1054–3139(03)00011-0

A simulation study of the implications of age-reading errors for stock assessment and management advice

Stuart A. Reeves

Reeves, S. A. 2003. A simulation study of the implications of age-reading errors for stock assessment and management advice. – ICES Journal of Marine Science, 60: 314–328.

Catch-at-age data are a crucial input to many stock assessments, so errors in age determination could have an adverse effect on the quality of the stock assessment and on the scientific advice based on that assessment. The results of simulation studies presented in this study are intended to quantify the effects of age-reading errors on the perception of stock trends and short-term management advice. The study is based on Eastern Baltic cod, in which problems with consistent interpretation of otolith structures result in the catch-at-age data being particularly problematic. The results indicate a clear distinction between the performance of the assessment, and the performance of catch forecasts and advice based on that assessment. The ageing error affected the absolute level of estimates of fishing mortality and spawning stock biomass from stock assessments, although overall trends are similar, and general conclusions about the state of the stock are likely to be broadly correct. Greater problems arose in catch forecasts and advice, for which ageing error led to discrepancies between the required and the effective fishing mortality, and a general tendency for ageing error to lead to advice on Total Allowable Catch that would be too optimistic and, therefore, less effective for stock conservation.

Keywords: age-reading errors, Baltic cod, fishery management advice, simulation, stock assessment.

Received 26 August 2002; accepted 13 January 2003.

S. A. Reeves: Department of Marine Fisheries, Danish Institute for Fisheries Research, Charlottenlund Slot, DK-2920 Charlottenlund, Denmark. Correspondence to S. A. Reeves: tel.: +45 3396 3467; fax: +45 3396 3333; e-mail: sar@dfu.min.dk.

Introduction

Catch-at-age data are a crucial input to the age-based assessments routinely used for many stocks in the North Atlantic. As a result, problems in age determination for a particular stock may cause problems for the assessment and management of that stock. Further, problems attributable to age-reading errors are not restricted just to catch-at-age data, because incorrect age reading is also likely to lead to errors in the estimation of other input data, including catch and stock weights at age, maturity at age, and any age-structured catch-per-unit-effort (cpue) indices. Hence, age-reading problems may influence virtually all the assessment inputs.

As a case study for the implications of age-reading problems for assessment and advice, this study uses the cod (*Gadus morhua*) stock of the Eastern Baltic. This stock has a long history of age-reading problems and hence makes a useful example. Simulation approaches based on the Baltic cod case are used in this work to investigate how age-reading errors influence both the results of stock assessments and the fishery management advice based on those assessments. The results are discussed in general terms and in terms of the specific consequences for the assessment and management of the stock.

Age-determination problems in the Eastern Baltic cod stock

The stock of cod in the Eastern Baltic (ICES subdivisions 25–32) has long had problems associated with age reading, problems addressed to a certain extent by ICES study groups. In the report of a recent meeting (ICES, 2000b), the Study Group on Baltic Cod Age Reading reviewed the history of age-reading problems for the stock. They noted that the first meeting on the subject took place in 1972, and commented, “After many successive meetings there are still systematic differences between age readers in the Baltic Sea.” The problem is discussed in detail in ICES (2000b), but broadly speaking, there is a difference in interpretation between readers from eastern and western countries around the Baltic, with eastern readers tending to allocate an older age to a given fish than their western

1054–3139/03/040314+15 $30.00

counterparts. There is no clear agreement on whether either of the two interpretations represents the true age of the fish. There are also problems with consistency of interpretation between readers from the same institute.

The Eastern Baltic cod stock is assessed annually by the ICES Baltic Fisheries Assessment Working Group. Recent assessments used Extended Survivors Analysis (XSA; Darby and Flatman, 1994), an age-based sequential population analysis model. The assessment is then carried forward into a deterministic short-term catch forecast, and management advice is formulated on the basis of the assessment and the catch forecast. The primary management tool is an annual Total Allowable Catch (TAC) based on the management advice and the short-term catch forecast. The approaches used in this study to simulate the stock assessment and advice process are based on those used for Eastern Baltic cod and many other ICES stocks.

Methods

Simulation approach

The approach used to investigate the effects of age-reading problems is essentially a subset of the management procedure approach as described, for instance, by Kell *et al.* (1999). An operating model is developed to describe essential elements of the fish population and of the fishery exploiting it. This operating model is used to generate assessment data subject to measurement error (in this case age-reading errors), which are then used in stock-assessment models, for which performance measures are generated. The operating model is run repeatedly to provide a number of different realizations of the dynamics of the simulated stock, and the performance measures are summarized over all the runs to evaluate the effects of the data problems on assessment performance.

In the current study, it is not the intention to evaluate the assessment model or the management approach, only to evaluate how one specific data problem will affect the stock assessment and the management advice. For this reason, the approach used is rather simpler than that used by Kell *et al.* (1999), with, for instance, no feedback to the underlying system. For similar reasons, the "true" system state in this study is represented by an assessment based on data without error. This is then used as a control against which the effects of age-reading errors can be evaluated.

Operating model

The operating model used in the simulations consists of two components, one representing the dynamics of the fish population, and the other the fishing fleets. The population model is parameterized such that the fish stock has broadly similar dynamics to that of Eastern Baltic cod in terms of growth rate and variability, stock and recruitment, size at age, etc. The fleet model is rather simpler, with two nations each with two fleets, and the effort in each fleet varying randomly. The intention in this study was not to provide a realistic representation of fleet development, but merely to ensure variability in fishing mortality and exploitation pattern, and in the proportion of the catch taken by each nation. The nature of the variation also ensures a wide range of outcomes in terms of the final state of the stock in each realization. The operating model is described in detail in the Appendix, but to summarize, the population data are generated using stochastic variation in recruitment, and growth and fleet efforts, and each realization uses a different starting population.

Age-reading errors and assessment data

The output data from the operating model include the true catches at age and length for each fleet/season combination. These can be used to calculate the true age–length keys (ALKs) for the catches, which can then be distorted to take account of age-reading errors. The approach used to distort the age readings follows that of Bradford (1991), matrices of ageing errors being used to represent the age-reading discrepancies. However, these matrices were applied in this study to the ALKs rather than directly to the age compositions. This allowed the latter to be estimated from the true length compositions. The approach used was, therefore, as outlined below.

If the age-error matrix is AEM, then AEM(i,j) is the proportion of fish of true age i allocated to age j. Similarly, ALK(l,a) is the true proportion of fish of length l that are of age a. To estimate the ALK that would result given the age-reading errors represented by AEM, ALK′, each individual element, ALK′(l,a) is given by

$$\mathrm{ALK}'(\mathrm{l},\mathrm{a}) = \frac{\sum_{\mathrm{i=A0}}^{\mathrm{i=An}} \mathrm{ALK}(\mathrm{l},\mathrm{i}) \cdot \mathrm{AEM}(\mathrm{a},\mathrm{i})}{\sum_{\mathrm{a=A0}}^{\mathrm{a=An}} \sum_{\mathrm{i=A0}}^{\mathrm{i=An}} \mathrm{ALK}(\mathrm{l},\mathrm{i}) \cdot \mathrm{AEM}(\mathrm{a},\mathrm{i})} \tag{1}$$

where A0–An is the age range of fish present in the catches, and the denominator is used to ensure that the elements at each length sum to one.

It is assumed in this study that the age-reading errors differ by nation, and hence the two "nations" used in the simulations are analogous to the two age-reading "schools" found among age readers of Baltic cod (ICES, 2000b). It has not been possible to estimate ageing-error matrices that correspond exactly to the Baltic cod case study. Although the Study Group on Baltic Cod Age Reading has done some comparative reading studies of otoliths (ICES, 2000b), it is not possible to use their data in this study, partly because too few comparative readings were made of otoliths from the relevant subdivisions, and partly because the true ages of the otolith samples are not known. Therefore, to investigate the possible scope of the problem, investigations have assumed that one nation's age readings are unbiased but made with some error, and that the other nation's readings are subject to either systematic under- or systematic over-ageing, as well as error. In addition, a run was made where the readings by both the nations were unbiased and subject only to noise. Although this situation does not apply to Baltic cod, it does allow investigation of the relative contributions of the noise and bias components to the overall problem. The ageing-error matrices assume that no catches are misallocated as 0-group because, in practice, it seems unlikely that there would be any significant allocation of fish to ages below the ones at which recruitment to the fishery is thought to occur. The ageing-error matrices used are given in Table 1, and the ageing-error scenarios are summarized in Table 2. In the absence of consensus on the consistent interpretation of Baltic cod otoliths, it is not clear which of these scenarios is closer to reality, but in practice they are likely to oversimplify the problem, both in the assumption that one nation's readings are essentially correct, and also in the assumption that the nature of the national age-determination errors has been consistent through time.

Table 1. Ageing-error matrices used in the simulations.

True	Allocated age									
age	1	2	3	4	5	6	7	8	9	10
	M1, control, no age-reading error									
1	1	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1
	M2, noise-only, unbiased (from Bradford, 1991)									
1	0.95	0.05	0	0	0	0	0	0	0	0
2	0.05	0.9	0.05	0	0	0	0	0	0	0
3	0.03	0.07	0.85	0.05	0	0	0	0	0	0
4	0	0.03	0.07	0.85	0.05	0	0	0	0	0
5	0	0	0.03	0.1	0.8	0.07	0	0	0	0
6	0	0	0	0.03	0.1	0.8	0.07	0	0	0
7	0	0	0	0	0.03	0.15	0.75	0.07	0	0
8	0	0	0	0	0	0.03	0.15	0.75	0.07	0
9	0	0	0	0	0	0	0.08	0.15	0.7	0.07
10	0	0	0	0	0	0	0	0.1	0.2	0.7
	M3, systematic under-ageing (adapted from Bradford, 1991)									
1	1	0	0	0	0	0	0	0	0	0
2	0.95	0.05	0	0	0	0	0	0	0	0
3	0.05	0.9	0.05	0	0	0	0	0	0	0
4	0.03	0.07	0.85	0.05	0	0	0	0	0	0
5	0	0.03	0.07	0.85	0.05	0	0	0	0	0
6	0	0	0.03	0.1	0.8	0.07	0	0	0	0
7	0	0	0	0.03	0.1	0.8	0.07	0	0	0
8	0	0	0	0	0.03	0.15	0.75	0.07	0	0
9	0	0	0	0	0	0.03	0.15	0.75	0.07	0
10	0	0	0	0	0	0	0.08	0.15	0.7	0.07
	M4, systematic over-ageing (adapted from Bradford, 1991)									
1	0.05	0.9	0.05	0	0	0	0	0	0	0
2	0.03	0.07	0.85	0.05	0	0	0	0	0	0
3	0	0.03	0.07	0.85	0.05	0	0	0	0	0
4	0	0	0.03	0.1	0.8	0.07	0	0	0	0
5	0	0	0	0.03	0.1	0.8	0.07	0	0	0
6	0	0	0	0	0.03	0.15	0.75	0.07	0	0
7	0	0	0	0	0	0.03	0.15	0.75	0.07	0
8	0	0	0	0	0	0	0.08	0.15	0.7	0.07
9	0	0	0	0	0	0	0	0.1	0.2	0.7
10	0	0	0	0	0	0	0	0	0	1

In practice, the total catch-at-age data used in a stock assessment are estimated by summing the catch data across all the national fleets, and catch weights at age are estimated in an analogous manner, using a mean across fleets, weighted by the numbers caught at age in each fleet. Hence, to estimate assessment data in the current study, the length compositions of the total catch for each fleet and season were first allocated to age using the distorted national ALKs, and the catches were then aggregated across the fleet and season to give annual catch-at-age data. The same data, disaggregated by fleet, were also used together with the fleet effort data as cpue series for tuning, so the tuning data also reflected national ageing errors. A similar approach was used for catch weights at age, where the mean length at each age was estimated from the true length composition and the distorted ALK for each fleet/season. This was converted to a weight at age using a fixed length–weight relationship, and an overall mean weight at age is estimated from the weighted average of the fleet/season estimates using the estimated catch for that fleet/season as weighting factor.

The extent to which the estimates of maturity at age and stock weight at age may be influenced by age-reading error is dependent on the source of the data. If fixed values are assumed for these biological parameters, then age determination may not be a problem, but if values are estimated from data from an international research vessel survey, then the age-reading errors may be similar to those in the catch data, owing to combining data across different age-reading "schools". For the purposes of the current study, it has been assumed that unbiased samples of the numbers and maturity at length in the population are available, as from an international survey. Further, it is assumed that the samples are divided equally between the two nations, so that national estimates of stock weight and maturity at age are obtained from combining the true length data with

Table 2. Details of ageing-error scenarios used to generate assessment data and results.

	Ageing-error matrix used for:	
Ageing-error scenario	Nation A	Nation B
C, control	M1	M1
N, noise-only	M2	M2
O, over-ageing	M2	M4
U, under-ageing	M2	M3

ALKs distorted according to the national age-reading characteristics, as previously described. The national estimates are then combined to give overall mean estimates analogous to the catch-weight-at-age data.

Performance measures for stock assessment and short-term forecasts

In order to define performance measures for stock assessments, it is necessary to define what the stock assessment is required to do. In this study, the primary concern is with the use of the stock assessment as the basis for scientific advice on fishery management, so it is on this basis that the performance measures are defined.

In formulating management advice, a stock assessment is first used to identify the current state of the stock in relation to defined management objectives. For many stocks in the ICES area, management objectives are defined in terms of precautionary reference points, with management intended to maintain spawning biomass above a defined precautionary level (known as B_{PA}, where the subscript PA stands for Precautionary Approach; ICES, 2000a), and to maintain fishing mortality below the defined precautionary level F_{PA}. In many cases, ICES has also proposed limit reference points that are the values associated with unknown stock dynamics or possible stock collapse. Therefore, the first stage in formulating management advice is to categorize the current state of the stock with respect to precautionary reference points. For this reason, this study has used the ability to categorize correctly the state of the stock with respect to precautionary reference points as a simple and pragmatic performance measure for stock assessments. By comparing the categorization with the "true" categorization (from a control assessment using data with no age-reading errors), and enumerating the proportion of simulations in which the state of the stock is misallocated, the effects of age-reading errors on this aspect of assessment performance can be evaluated.

Once the state of the stock has been determined, the next stage of the process is to specify management advice based on the state of stock categorization. There will typically be some correspondence between the stock categorization and the advice, i.e. the advice will be different depending on whether the stock is categorized as being above or below B_{PA}, or whether fishing mortality is categorized as being above or below F_{PA}. Further, in a TAC-based system, although the advice may be phrased in terms of a relative change in F or fishing effort required to achieve the management objectives, it will still be necessary to translate this into an absolute catch in order that a TAC may be set. This process can be represented by first using a simple rule to derive the advice, which is expressed in terms of the advised relative change in fishing mortality (or fishing effort; see subsequent discussion). This effort multiplier is then used in a conventional deterministic catch forecast to derive a total catch that can be used as a TAC.

In the present study, the term "effort" is used in relative terms, and is taken to be directly proportional to fishing mortality. It is useful to distinguish between "advised" effort, "effective" effort, and "required" effort. Here, the relative effort resulting from the use of the advice rule with an assessment with age-reading problems is defined as the advised effort. This value is then used in a catch forecast based on the same assessment to obtain a catch corresponding to the advised effort. The effort that will actually result from taking that catch from the true population is then defined as the effective effort. The required effort is then the relative effort that results from applying the advice rule to the true population, and is estimated using the control assessment.

To measure the performance of a catch forecast, the key point is whether the specified catch will achieve the required effort. Hence, in the current study, the deviation between the effective and the required effort multipliers is used as a measure of performance of the forecasts. This measure is referred to as the effort deviation in subsequent discussion.

Assessment runs and precautionary reference points

All stock assessment runs used XSA (Darby and Flatman, 1994) as the catch-at-age model, reflecting current practice in the assessment of Eastern Baltic cod (ICES, 2001b). All default model settings were accepted. This assumption was made for convenience, but in practice, the model settings used in the assessment given in ICES (2001b) correspond closely with these defaults. The assessments were tuned with age-structured cpue data from the four simulated fishing fleets (two gears for each of the two nations).

Precautionary reference points were estimated separately for each ageing-error scenario, based on the assumption that the ageing errors were a long-term problem that influenced all aspects of the assessment, including the estimation of reference points. The assessment run used to estimate reference points in each case omitted the five most recent years of data. This represents the situation where precautionary reference points were introduced five years before the present and were based on the assessment data available at that time. For each run, both precautionary and limit values were set for both fishing mortality (F) and spawning stock biomass (SSB). In addition, "extreme" values were set for both indicators, to allow additional resolution in cases where limit reference points were substantially exceeded. Fishing mortality reference points were based on the use of F_{med} (Jacobsen, 1993) as F_{PA}, with other reference points obtained by adjusting F_{PA} by a fixed multiplier. Similarly, biomass reference points were estimated by taking the lowest observed spawning stock (B_{loss}) as the limit SSB, B_{lim}, and were adjusted by a fixed multiplier to obtain the other biomass reference points. This procedure is summarized in Table 3. The use of F_{med} and

Table 3. Rules used for deriving precautionary reference points. B_{loss} is the lowest observed spawning stock in the time-series, and F_{med} is defined by Jacobsen (1993). Note also that 0.77 is the reciprocal of 1.3.

Parameter	PA	Limit	Extreme
Fishing mortality	$F_{PA} = F_{med}$	$F_{lim} = 1.3 \times F_{med}$	$F_{extreme} = 1.3 \times F_{lim}$
SSB	$B_{PA} = 1.3 \times B_{loss}$	$B_{lim} = B_{loss}$	$B_{extreme} = 0.77 \times B_{loss}$

B_{loss} as reference points in this study reflects their use as the basis of precautionary reference points defined for many ICES stocks (ICES, 2001a), and also means that ageing error in the assessment is incorporated in the reference points.

To estimate the current state of the stock, an assessment run was made using the full time-series of data. "Current" SSB was taken as the estimated SSB in the final year of the assessment run, and current F was estimated by first averaging the fishing mortalities at age over the last three years of the assessment run to estimate the current exploitation pattern (also used in the calculation of F_{med}) and then calculating an average fishing mortality (over ages 3–7) for this exploitation pattern.

Management advice and short-term forecasts

Once the current state of the stock has been established, a simple rule is used to derive advice based on the state of stock categorization. If current SSB is above B_{PA} then the advice would be to reduce F to F_{PA} unless F is already below F_{PA}, in which case the advice would be to maintain current F. At lower SSBs, there is a reduction in the maximum advised F according to the rule illustrated in Figure 1. At levels of SSB below B_{lim}, the rule used here is to advise an arbitrary low level of fishing mortality. In practice, this value might be set to zero, but here the rule specifies an advised F of no more than $0.2F_{PA}$ under these circumstances.

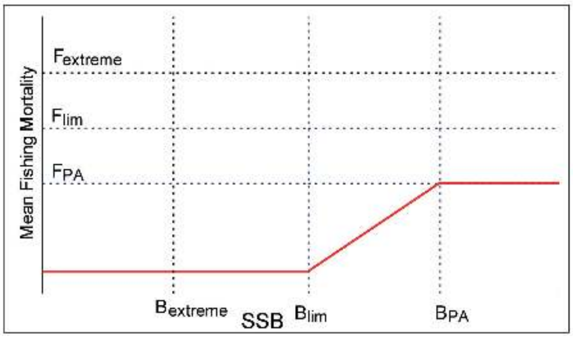

Figure 1. Diagrammatic representation of reference points and advice rules used in the simulations. The advised F is dependent on the current SSB and is indicated by the bold line. If current F is below this level, then the advice would be to maintain current F, i.e. an effort multiplier of 1.

To find the catch corresponding to the advised F, a conventional short-term catch forecast is used. It is assumed that the assessment is performed in year t, using data up to, and including, year t − 1, in order to provide advice for the fishery in year t + 1. In this case, the assessment provides estimates of the population numbers at the start of year t, so assumptions about fishing mortality and recruitment during year t are required to project the population numbers forward to the start of year t + 1. The exploitation pattern during year t is here assumed to be the mean over the years t − 3 to t − 1, and an effort multiplier of 1.0 is assumed. Recruitment during year t and subsequent years is taken as the geometric mean over years t − 10 to t − 1. Weights at age in the catch are taken as the means over years t − 3 to t − 1. The catch corresponding to the advised effort is then obtained by projecting the population numbers at the start of year t + 1, using Fs at age estimated by multiplying the recent mean exploitation pattern by the advised effort multiplier. Using these Fs and the population numbers in the conventional catch equation gives the estimated catches at age. These are then multiplied by the estimated weights at age, then summed across ages to give the estimated catch.

Once the advised catch has been estimated, the next stage is to derive the effective effort multiplier. This is done using a catch forecast based on the control assessment in reverse, i.e. setting a catch to estimate the effort multiplier, rather than *vice versa*. The performance measure used is then the effort deviation, defined as the difference between the effective and the required effort multipliers. Used in this way, a positive value means that assessment problems would lead to a higher F than required, and thus less conservative management, whereas a negative value indicates a lower F than required, i.e. an unnecessarily restrictive TAC.

Simulation algorithm

The individual components of the simulations are described in detail in the preceding discussion, but the overall procedure is described as:

For each stock realization from the operating model,

- generate assessment data for the required ageing-error scenario (N years),
- run assessment using first (N − 5) years of data,
- calculate PA reference points,
- run assessment using all N years of data,
- determine current (year N) state of stock with respect to PA reference points,
- derive advised effort multiplier from advice rule,
- find catch corresponding to advised effort from short-term forecast,

- find effective effort corresponding to advised catch from control forecast.

The performance measures for each realization are derived from a comparison of the estimated state of the stock with the state of stock in the control run, and from the difference between the effective and the required effort multipliers. This study used 100 realizations, each of 30 years duration. For each of these, data were generated representing the four ageing-error scenarios given in Table 2.

Results

As this is a simulation study, interpretation of results has focussed on identifying the general trends and features of the results rather than the detail of individual cases. Further, as some results may be dependent on the specific assumptions and parameters used to describe growth, mortality, and size selection, emphasis has been placed on understanding why the results have occurred rather than on the results themselves. In this way, it is hoped that the study will have wider application than just the Baltic cod example that prompted it.

Stock trends

Preliminary examination of the stock trends resulting from the three ageing-error scenarios indicated that the general result was for SSB and mean F to show similar trends in each of the scenarios, but with different absolute values, although the differences in general were not large (Figure 2). It was not the intention of the study to analyse the differences fully, but in order to give a coarse summary of the relative differences, a pairwise comparison of the trends was made for each of the 100 realizations. The comparison involved calculating the difference between the values for each year for a pair of scenarios. A comparison was classified as showing a consistent difference in level between two scenarios if 80% or more of the differences (24+ years out of 30) showed the same sign. The results from these comparisons are summarized in Table 4. There was a general tendency for age-reading errors to lead to lower estimates of SSB, because in some 60% of cases the control estimate was consistently higher (on the criteria used in this study) than that resulting from runs where under- or over-ageing error was applied to the data. For the noise-only ageing scenario, the estimated SSB was

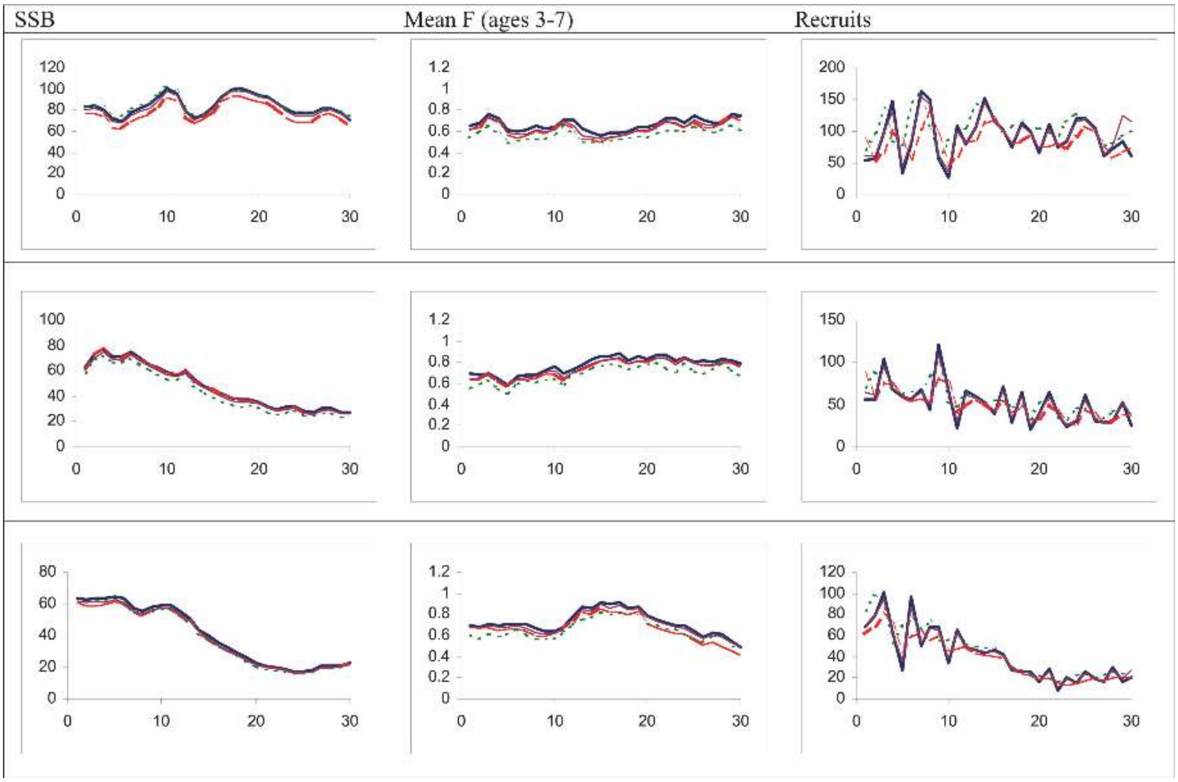

Figure 2. Example stock trends from three randomly chosen realizations. The four lines on each plot are the assessment results from runs using the control data (solid thick line), and runs where ageing error is applied as noise-only (solid thin line), under-ageing (dashed line), and over-ageing (dotted line). Each row gives the plots for a different realization. For reasons of clarity, the y-axis scales used for SSB and recruitment plots differ between realizations.

Table 4. Results of pairwise comparisons of trends in SSB and F from different ageing-error scenarios. The numbers give the number of realizations (out of 100) in which the given result was obtained. The results are given as follows: $A < B$: results from ageing-error scenario A are consistently less than those from B; $A > B$: results from ageing-error scenario A are consistently greater than those from B; $A \sim B$: no consistent difference obtained. To give an example, the table shows that estimates of SSB obtained from assessments using data with under-ageing applied were consistently lower than those from the control run in 57 out of 100 realizations. The criterion used to define a consistent difference in level is given in the text.

Comparison	SSB	Mean F(3–7)
Under-ageing (U) vs. control (C)		
$U < C$	57	99
$U \sim C$	35	1
$U > C$	8	0
Over-ageing (O) vs. control (C)		
$O < C$	63	100
$O \sim C$	24	0
$O > C$	13	0
Over-ageing (O) vs. under-ageing (U)		
$O < U$	41	90
$O \sim U$	22	10
$O > U$	37	0
Noise only (N) vs. control (C)		
$N < C$	95	100
$N \sim C$	5	0
$N > C$	0	0
Noise only (N) vs. under-ageing (U)		
$N < U$	36	0
$N \sim U$	25	65
$N > U$	39	35
Noise only (N) vs. over-ageing (O)		
$N < O$	33	0
$N \sim O$	23	1
$N > O$	34	99

consistently lower than the control estimate in 95% of the realizations. There are no consistent differences between the results for over-ageing, under-ageing, and noise only. For fishing mortality, it is striking that in all cases, age-reading error led to consistently lower estimates of mean F. In addition, in most cases, under-ageing and noise only led to a higher estimate of mean F than over-ageing.

Recruitment trends are not considered in the previously mentioned analysis because the differences in trend are more complex than those for SSB and F. The main effect of age-reading error is to distribute catches from 1-year class into neighbouring year classes. As a result, estimates of year-class strength (related to the accumulated catch of a given year class) are smoothed so that the recruitment time-series shows less extreme values, with increased autocorrelation. With the age-reading biases investigated in the current study, there is also a tendency for the apparent recruitment to be offset by 1 year relative to the true recruitment, the direction of the offset being dependent on whether the bias is under- or over-ageing.

State of stock

To summarize the effects of age-reading error on the stock assessment performance measure used in this study (i.e. the ability of the assessment to allocate the stock to the correct category with respect to precautionary reference points), the first concern is with the true categorization for each realization, and then with the extent and the nature of any incorrect categorizations that may result from age-reading errors. Table 5 summarizes the final state of stock categorization for the control run, i.e. the correct allocations for each of the 100 realizations. For each of the other two ageing-error scenarios, Table 6 summarizes the cases where the ageing error resulted in incorrect categorization of the state of stock, and Figure 3 indicates the magnitude and direction of these misallocations.

The results shown in Table 5 indicate that the random variation incorporated in the simulations resulted in a reasonably wide spread of outcomes across the possible categories. The higher frequency of outcomes around the leading diagonal of the table reflects the interdependence of fishing mortality and SSB, i.e. a high F tends to lead to a low SSB. The misallocations summarized in Table 6 indicate that, in 69–75% of cases, the state of the stock was correctly categorized despite the ageing error. Where the categorization was incorrect, this was more likely to be due to F than SSB being misallocated, with only a few realizations being misallocated on both indicators. The results were similar for all ageing-error scenarios. Figure 3 also indicates that most realizations were correctly allocated despite age-reading error, but that misallocations attributable to under-ageing mostly took the form of a slightly over-optimistic categorization of fishing mortality

Table 5. Final state of stock categorization with respect to PA reference points for control runs. Each cell indicates the number of realizations allocated to the category (out of 100).

State	$F < F_{PA}$	$F_{PA} \leq F < F_{lim}$	$F_{lim} \leq F < F_{extreme}$	$F \leq F_{extreme}$	Total
$SSB > B_{PA}$	14	12	7	0	33
$B_{PA} \geq SSB > B_{lim}$	0	10	7	1	18
$B_{lim} \geq SSB > B_{extreme}$	0	1	18	2	21
$SSB \geq B_{extreme}$	0	0	14	14	28
Total	14	23	46	17	100

Table 6. The number of realizations (out of 100) in which the terminal state of the stock was misallocated with respect to precautionary reference points.

	Misallocation on:			
Ageing-error scenario	SSB only	Mean F only	SSB and mean F	Total
Noise only	8	18	1	27
Under-ageing	10	18	3	31
Over-ageing	9	12	4	25

(i.e. F indicated to be below F_{PA} when it was actually above F_{PA}, but below F_{lim}). Correspondingly, over-ageing was most likely to lead to F being allocated to the next most pessimistic category. There was also a slight tendency for over-ageing to lead to an over-optimistic categorization of SSB. The results for the noise-only scenario were intermediate, with only a slight tendency for over-optimistic categorization of F. In all the cases, the misallocations were relatively small, leading only to allocation into an adjacent category.

Catch forecasts

The effort deviations for each ageing-error scenario are summarized by the true SSB and F categories in Figure 4. When summarized by SSB category, the highest deviations tended to be at SSBs above B_{PA}, with a general tendency for the deviations to be smaller at lower stock sizes. When summarized by F category, the highest deviations were when F was between F_{PA} and F_{lim}, and then tended to decrease in magnitude with increasing F. In all cases, there was a general tendency for the largest deviations to be positive, and in the most extreme cases, the result is an effective level of effort almost twice that required.

The ageing errors investigated can be considered to introduce a bias to the forecast if the mean of the effort deviations for a particular category is significantly different from zero. Paired sample t-tests were performed for each ageing-error scenario to compare the effective and required effort multipliers. As Figure 4 indicated different trends within each F and SSB category, the t-tests were run separately for data in each category. The results are summarized in Table 7. These indicate that, in some cases, the age-reading error led to bias in the forecasts, and in most cases, the bias was in a positive direction, i.e. the effective effort was consistently higher than the required effort.

Detailed examination of some assessment and projection runs indicated that the highest absolute values for effort deviation were present in cases where the ageing error resulted in failure to detect recent recruitment adequately. For example, in some cases, the most recent year classes were very weak, but this was not reflected in the estimated stock numbers, so the resultant catch forecasts were highly over-optimistic. This also happened in the other direction: in some cases, a recent strong year class was not detected. For smaller absolute effort deviations, the effects were much more complex and arose from a combination of estimated stock numbers, fishing mortalities, and weights and maturities at age all contributing to varying extents. The advice rule also contributes some error in this respect because it is notable that the highest absolute effort deviations were at Fs between F_{PA} and F_{lim}, when the advice rule led to an F proportional to the SSB rather than a fixed value.

Discussion and conclusions

Most age-based stock assessments assume that fish ages are measured without error, although this assumption is rarely met (Beamish and McFarlane, 1983; Richards *et al.*, 1992). Although there is extensive literature on the methods of age determination and their associated sources of error (e.g. Campana, 2001), there appear to be relatively few studies of the effects of these errors when catch-at-age data are used in a stock assessment. Indeed, Restrepo and Powers (1990) note that the problem of biases in age-based stock assessments attributable to systematic errors in ageing has been largely ignored in the literature. However, some progress was made by Lai and Gunderson (1987) and Tyler *et al.* (1989), who investigated the effects of age-reading problems on yield estimates, although those studies estimated yield by estimating growth parameters and then using them in a yield-per-recruit model, rather than using a catch-at-age analysis. Coggins and Quinn (1998) also investigated the effects of age-reading errors on yield estimates, and that study did use a catch-at-age analysis, with sustained yield estimates derived from an equilibrium approach. Bradford (1991) considered the effects of ageing error on estimates of recruitment from a catch-at-age analysis, but his study was concerned with the use of the estimates in the study of environmental effects on recruitment rather than in a fishery management context. Rivard (1989) considered the effects of ageing error on the estimation of recruitment as part of an overview of the source of error in catch-at-age analyses, and Restrepo and Powers (1990) used a simulation approach to investigate the effect that different methods of handling the plus-group have on results of a catch-at-age analysis in the presence of ageing errors.

The current study incorporates substantially more complexity than the studies summarized in preceding discussion because it attempts to represent the problems that may be found for stock assessment data in an international fishery where data are compiled across several national sources. In particular, the use of different ageing-error matrices for different nations, the variation in the proportion of the catch taken by each nation varying with time, and the incorporation of ageing error in the estimation of maturity at age and cpue data, are complexities that are not considered in

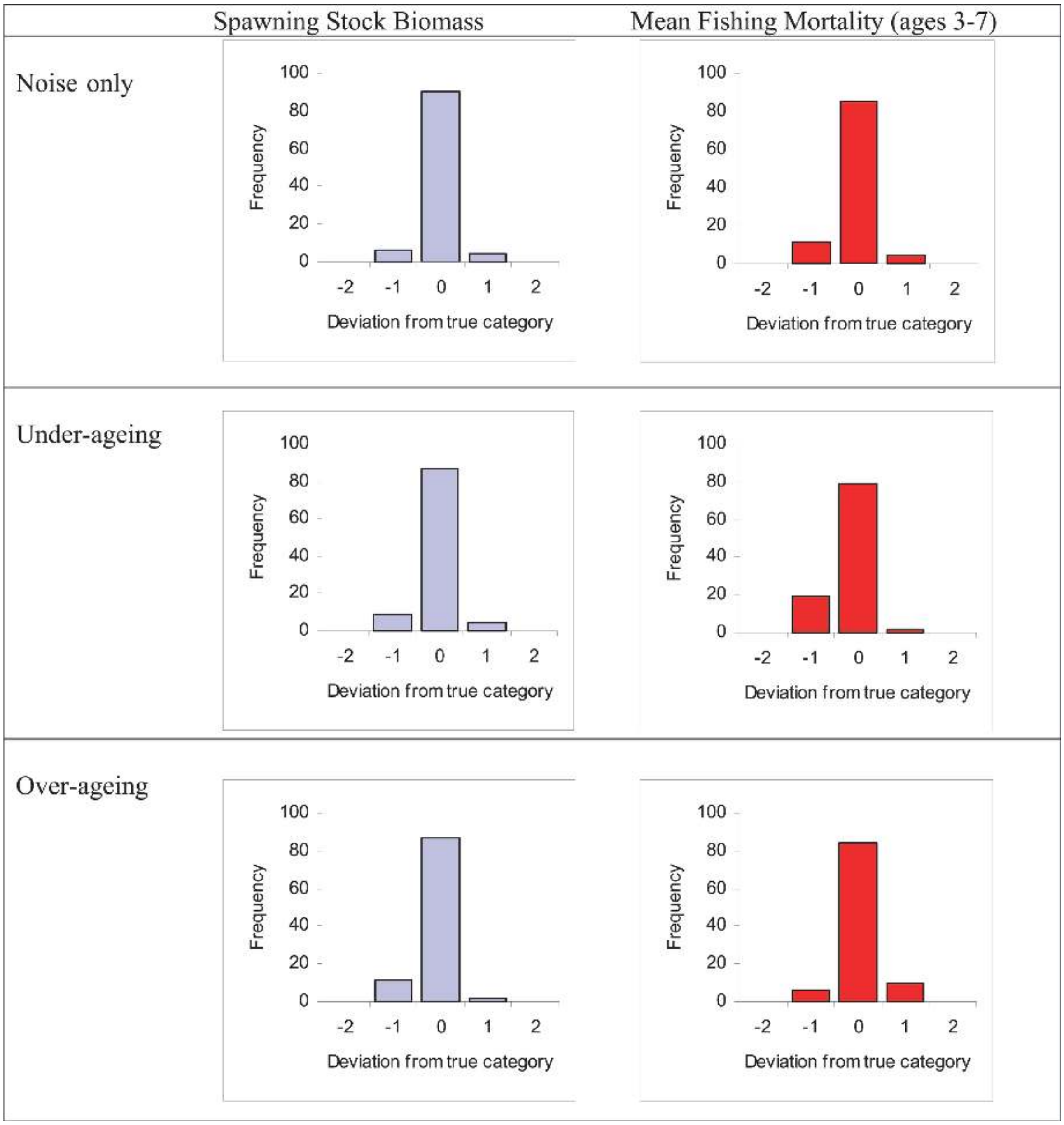

Figure 3. The extent and direction of misallocation for realizations, where the terminal state of the stock was incorrectly categorized with respect to precautionary reference points attributable to age-reading error. The x-axis indicates the deviation from the true category, so a deviation of 0 indicates that a realization has been allocated to the correct category, whereas a deviation of −1 indicates that a realization has been misallocated by one category in the optimistic direction, i.e. lower F or higher SSB. The columns refer to misallocations by SSB and F separately, and the rows refer to the different ageing-error scenarios.

the studies summarized previously. In addition, apart from describing the effects of ageing error on stock assessment results, this study emphasizes the practical implications of the effects on fisheries management advice.

One general result of the current study is that age-reading error tends to lead to estimates of F and SSB that show similar trends to the true values, but that differ in level. A similar result was found by Reeves (2001). For SSB, there is no clear tendency for ageing bias to lead to values consistently higher or lower than true ones. This is probably because SSB is estimated from the sum of products of numbers at age, maturity at age, and stock weight at age; estimates of all these quantities are affected to varying extents and directions by age-reading error. For instance,

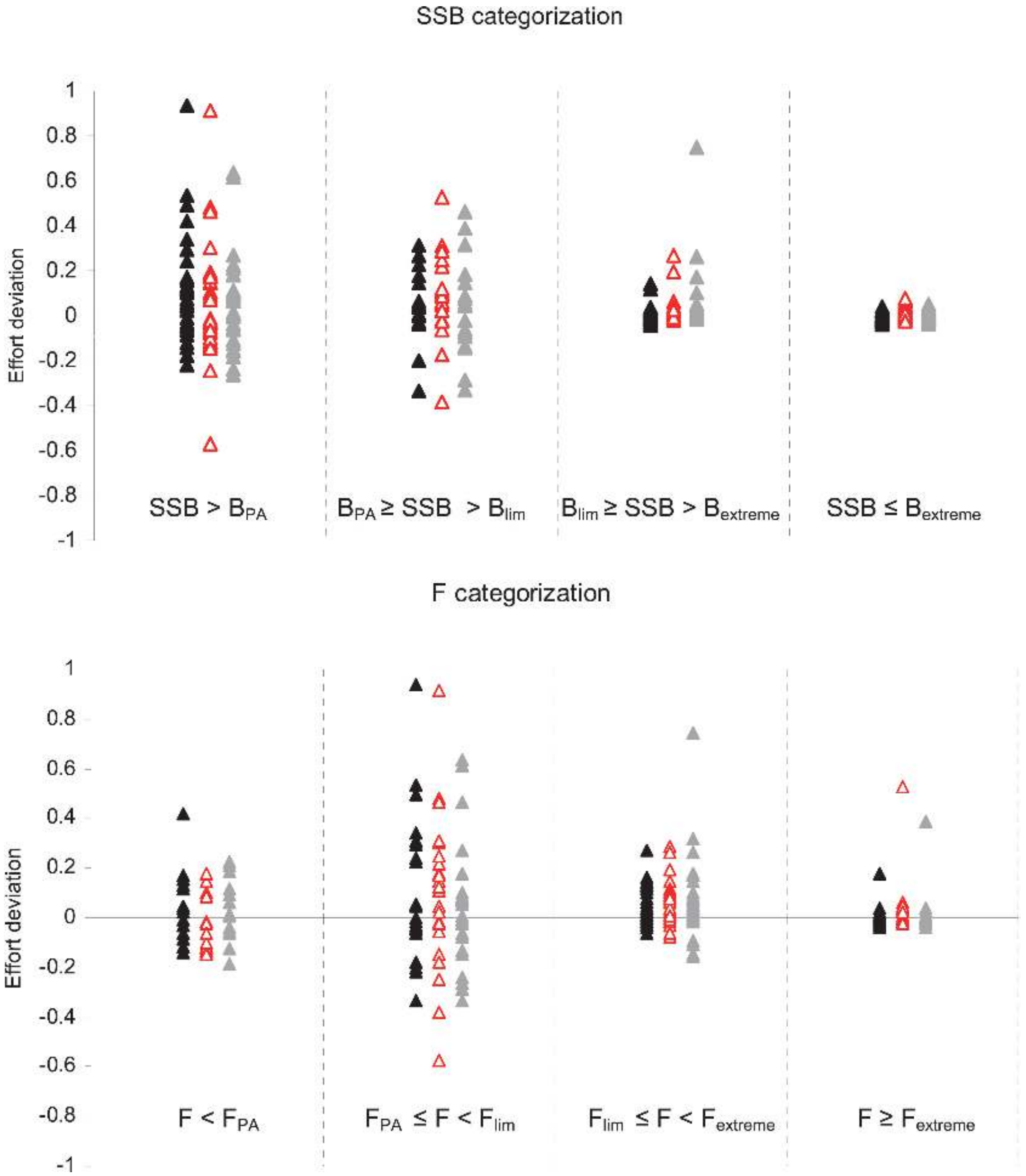

Figure 4. Effort deviations from individual realizations. The effort deviations are given for each ageing scenario (solid symbols, noise only; hollow symbols, under-ageing; grey symbols, over-ageing). They are summarized separately by the true SSB and F category for each realization.

under-ageing will tend to lead to lower estimates of recruitment (through higher F at younger ages and, therefore, faster convergence of the recruitment estimates), but with higher stock weights and faster maturity at the younger ages. Therefore, these effects will cancel each other out to some extent, resulting in a lack of a consistent bias.

In contrast to the ageing scenarios that included bias, the noise-only scenario led to estimates of SSB consistently lower than the control estimates, although the differences were not large. The effect of applying unbiased ageing error is to smooth differences in, for example, weight at age and maturity across ages, such that the estimated weights at age are slightly higher than the true weights at younger ages, but lower at the older ages. In the case investigated in this study, the net result of these effects has been SSB estimates slightly lower than the control estimates throughout. This result may be rather case-specific, and it seems likely that the result would be different under other assumptions about growth or

Table 7. Results of paired sample t-tests comparing required and effective effort multipliers from simulations for three age-reading scenarios (see Table 2). Results are summarized separately by true SSB category and true F category, as follows: ns, no significant difference; +/−, significant difference in positive or negative direction (e.g. + implies effective effort significantly greater than required effort); the number of symbols implies the degree of significance of the t-statistic, i.e. + is $0.01 < p \leq 0.05$; ++ is $0.05 < p \leq 0.01$; +++ is $p < 0.01$.

	Ageing-error scenario		
Category	Noise only	Under-ageing	Over-ageing
Spawning stock biomass			
$SSB > B_{PA}$	+++	ns	ns
$B_{PA} \geq SSB > B_{lim}$	ns	++	ns
$B_{lim} \geq SSB > B_{extreme}$	ns	+++	+++
$SSB \geq B_{extreme}$	−−	+++	ns
Fishing mortality			
$F < F_{PA}$	ns	ns	ns
$F_{PA} \leq F < F_{lim}$	++	ns	ns
$F_{lim} \leq F < F_{extreme}$	+++	+++	+++
$F \leq F_{extreme}$	ns	ns	ns

maturity. In general, all the results concerning SSB within this study will, to some extent, be driven by these assumptions, whereas the conclusions about F are likely to be less sensitive to the choice of biological parameters and should, therefore, be of more general application.

In contrast to SSB, ageing error resulted in a highly consistent tendency to underestimate the true mean fishing mortality. To an extent, this reflects the use of a simple mean over a fixed age range to represent overall annual fishing mortality. This is standard practice in, for example, ICES stocks (Darby and Flatman, 1994), with the age range intended to summarize the age classes most important in the catches. As such, catches from age classes within this age range are likely to dominate the catches, so misallocation of catches from these age classes into other age classes as a consequence of ageing error is likely to exceed misallocation in the other direction. This effect leads to lower estimates of F at age within these age classes, and hence a lower estimate of mean F. If the ageing error leads to a net allocation of catches from older age classes into younger ones, as in the under-ageing example used in this study, this leads to a tendency for an increase in the estimated accumulated F exerted on a cohort. This results from increased estimates of catch, and hence fishing mortality, on younger ages, so proportionately less of the cohort is lost to natural mortality. This effect accounts for the reason why under-ageing tends to result in a higher estimated mean F than over-ageing.

Any stock assessment model is conditioned on a number of assumptions, such as assuming a constant, known natural mortality (Quinn and Deriso, 1999; Patterson *et al.*, 2001). One consequence of this is that it is more appropriate to regard population trends from an assessment model as relative trends rather than absolute values. As such, the differences in level resulting from the ageing errors will not necessarily cause problems in the use of the assessment. Indeed, according to one of the performance criteria used in the present study, the ageing errors did not prove to be especially problematic, because the state of stock of around 70% of the realizations was correctly allocated despite the ageing error applied. These misallocations tended to be linked to F rather than to SSB. This is probably because the ageing error led to consistent underestimation of F, whereas there was no consistent bias in estimation of SSB.

Two measures of assessment performance were used in this study. The first was simply the ability of the assessment to categorize correctly the state of the stock with respect to precautionary reference points. The second was a measure of the ability of the assessment and the forecast to provide catch advice to achieve management objectives. This provides an integrated measure of the performance of the whole assessment/forecast/advice process. The results of the study indicate a clear distinction between the performance of the assessment, and the performance of catch forecasts and advice based on it. Considering first the stock assessment, the age-reading errors applied here do not cause serious problems for the assessment. In general, estimated stock trends are similar, and in most cases, the current state of the stock with respect to precautionary reference points is correctly allocated. However, when the assessment is carried forward into a catch forecast, these small discrepancies become more problematic. The observed effort deviations are such that a TAC set to achieve a specified fishing mortality is most likely to lead to a higher fishing mortality. In some cases, this discrepancy may be quite large. This result was obtained with ageing error as the only source of error in the assessment. In practice, as noted by Van Beek and Pastoors (1999) in their evaluation of ICES catch forecasts for a number of stocks, other problems, such as sampling coverage, discarding, and misreporting, will also add further error to the data, so it is likely that the effort deviations noted in this study will be underestimates of the effect such problems will have in practice. Even so, the contrast between relatively small problems in the stock assessment and the more substantial discrepancies that result in the catch forecast may go some way towards explaining the results observed by Van Beek and Pastoors (1999).

The three different ageing-error scenarios investigated in this study led to rather similar results; there was no clear distinction between the results from the noise-only scenario and those where bias was also included. This may imply that many of the problems attributable to age reading noted in this study result primarily from the noise component, with the bias component adding relatively little. However, this would imply that, for the assessment results, it might be more important to ensure consistency (i.e. precision) in age reading than for the age readings to be correct (i.e. unbiased). This is counter-intuitive, and further work would be required to elucidate the relative contributions of precision and bias in age reading to assessment problems.

This study was based on data for Eastern Baltic cod, so it is appropriate to review the current assessment for that stock in the light of the findings from this study. The current assessment indicates that the stock is at a very low level, with SSB well below B_{lim}, and F above F_{lim} (ICES, 2001b). The results of this study indicate that conclusions about the general state of the stock are robust to the age-reading problems, so notwithstanding other possible data problems, the conclusion that the stock is in a very poor state is likely to be correct.

The results of the current study indicate that the estimate of mean F from the current assessment is likely to be an underestimate of true fishing mortality. Historically, assessment of the adjacent cod stock in the Western Baltic (ICES subdivisions 22–24) has shown rather higher estimates of fishing mortality than that in the Eastern Baltic (average mean F of ages 3–6 over the period 1991–2000 is 1.25 in the Western Baltic, 0.82 in the Eastern Baltic; ICES, 2001b), even though the fisheries in the two areas are similar, with many vessels fishing in both areas. The otoliths of cod in the Western Baltic are considered to be more straightforward to interpret than those of cod in the Eastern Baltic (H. Mosegaard, pers. comm.), and there is likely to be more consistency in age reading because most age determination is done by readers from Western Baltic nations, so there is less influence of the different age-reading schools. As a result, it seems likely that at least part of the East/West discrepancy in apparent fishing mortality is due to the age-reading problems of the Eastern Baltic stock.

While the broad conclusions about the state of the Eastern Baltic cod stock seem likely to be correct despite the ageing errors present in the data, it is desirable to correct the data problems if at all possible. The failure to achieve this despite almost 30 years of study (ICES, 2000b) suggests that it is not possible to do this through interpreting the otolith ring structure. More objective approaches, such as otolith weight (Boehlert, 1985) may offer a way to make progress. However, the results from the current study indicate that it will not be sufficient just to introduce an improved and consistent method for age determination without correcting the entire past data. If this is not done then the assessment will indicate an increase in fishing mortality, and probably also a change of uncertain direction in SSB corresponding to the change in age-determination methodology, rather than any real change in the stock. Correcting all past data, including survey indices, maturity at age, and weight at age, as well as catch at age, would be a substantial task, but this should lead to considerable improvements in understanding the stock dynamics.

This study concentrated on the use of stock assessments as the basis of management advice, and as a result, has placed particular emphasis on how age reading affects the relative perception of SSB and F, reflecting their use in describing the current state of the stock for management purposes. The effects of ageing error on recruitment estimates have only been mentioned briefly in this study, but as demonstrated by Rivard (1989) and Bradford (1991), they are rather more complex than just a change of level. The relationship between spawning stock and recruitment in Eastern Baltic cod has been a fruitful area for research studies, many of which have used either the assessment data or the results in some form (Sparholt, 1996; Cardinale and Arrhenius, 2000; Jarre-Teichman *et al.*, 2000). In relation to such studies, Bradford (1991) notes that “recruitment researchers should consider correcting their catch-at-age matrix for ageing errors to determine the robustness of their conclusions to this source of error”. While it would not be straightforward, this comment seems particularly relevant in the case of Eastern Baltic cod.

Acknowledgements

I thank Henrik Mosegaard and Yvonne Walther for discussion, and two anonymous referees for comments that helped improve the manuscript. The work was funded by the Danish Directorate for Food, Fisheries, and Agribusiness, as part of project “Bestandsvurdering”.

References

Beamish, R. J., and McFarlane, G. A. 1983. The forgotten requirement for age validation in fisheries biology. Transactions of the American Fisheries Society, 112: 735–743.

Boehlert, G. W. 1995. Using objective criteria and multiple regression models for age determination in fishes. Fishery Bulletin, US, 83: 103–117.

Bradford, M. J. 1991. Effects of ageing errors on recruitment time series estimates from sequential population analysis. Canadian Journal of Fisheries and Aquatic Sciences, 48: 555–558.

Campana, S. E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. Journal of Fish Biology, 59: 197–242.

Cardinale, M., and Arrhenius, F. 2000. The influence of stock structure and environmental conditions on the recruitment process of Baltic cod estimated using a generalised additive model. Canadian Journal of Fisheries and Aquatic Sciences, 57: 2402–2409.

Coggins, L. G., and Quinn, T. J. 1998. A simulation study of the effects of ageing error and sample size on sustained yield estimates. *In* Fishery Stock Assessment Models, pp. 955–975. Ed. by F. Funk, T. J. Quinn, J. Heifetz, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C-I. Chang. University of Alaska, Fairbanks, Alaska Sea Grant College Program Report No. AK-SG-98-01.

Cook, R. M., Kunzlik, P. A., Hislop, J. R. G., and Poulding, D. 1999. Models of growth and maturity for North Sea cod. Journal of Northwest Atlantic Fisheries Science, 25: 91–99.

Cook, R. M., and Reeves, S. A. 1996. A comparison of the selectivity parameters of Scottish demersal seines and trawls estimated from commercial catch data and designed experiments. ICES CM 1996/D: 10.

Darby, C. D., and Flatman, S. 1994. Virtual Population Analysis: Version 3.1 Windows/Dos User Guide. MAFF Information Technology Series, 1. Directorate of Fisheries, Lowestoft.

Hoydal, K., Rørvik, C. J., and Sparre, P. 1982. Estimation of effective mesh sizes and their utilization in assessments. Dana, 2: 69–95.

ICES. 2000a. Report of the ICES Advisory Committee on Fishery Management, 2000. ICES Cooperative Research Report, 242.

ICES. 2000. Report of the Study Group on Baltic Cod Age Reading. ICES CM 2000/H: 01.

ICES. 2001a. Report of the Study Group on the Further Development of the Precautionary Approach to Fishery Management. ICES CM 2001/ACFM: 11.

ICES. 2001b. Report of the Baltic Fisheries Assessment Working Group. ICES CM 2001/ACFM: 18.

Jacobsen, T. 1993. The behaviour of F_{low}, F_{med} and F_{high} in response to variation in parameters used for their estimation. *In* Risk Evaluation and Biological Reference Points for Fisheries Management, pp. 119–125. Ed. by S. J. Smith, J. J. Hunt, and D. Rivard. Canadian Special Publications in Fisheries and Aquatic Sciences, 120, National Research Council, Canada.

Jarre-Teichman, A., Wieland, K., Mackenzie, B., Hinrichsen, H-H., Plikshs, M., and Aro, E. 2000. Stock–recruitment relationships for cod (*Gadus morhua callarais* L.) in the central Baltic Sea incorporating environmental variability. Archives of Fisheries and Marine Science, 48(2): 97–123.

Jones, R. 1983. An investigation of North Sea haddock (*Melanogrammus aeglefinus*) length-at-age data, with particular reference to the effects on growth of the outstanding year-classes of 1962 and 1967. Journal du Conseil de l'Exploration de la Mer, 40: 50–62.

Kell, L. T., O'Brien, C. M., Smith, M. T., Stokes, T. K., and Rackham, B. D. 1999. An evaluation of management procedures for implementing a precautionary approach in the ICES context for North Sea plaice (*Pleuronectes platessa* L.). ICES Journal of Marine Science, 56: 834–845.

Lai, H. L., and Gunderson, D. R. 1987. Effects of ageing errors on estimates of growth, mortality and yield per recruit for walleye pollock (*Theragra chalcogramma*). Fisheries Research, 5: 287–302.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V., Skagen, D. W., and Steffánson, G. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries, 2: 125–157.

Quinn, T. J., and Deriso, R. B. 1999. Quantitative Fish Dynamics. Oxford University Press, New York. 542 pp.

Reeves, S. A. 2001. Age-reading problems and the assessment of Eastern Baltic Cod. Annexe 1 in Report of the Baltic Fisheries Assessment Working Group. ICES CM 2001/ACFM: 18.

Restrepo, V. R., and Powers, J. E. 1990. A comparison of three methods for handling the "plus" group in Virtual Population Analysis in the presence of ageing errors. Collective Volume of Scientific Papers, ICCAT, 35(2): 346–354.

Richards, L. J., Schnute, J. T., Kronlund, A. R., and Beamish, R. J. 1992. Statistical models for the analysis of ageing error. Canadian Journal of Fisheries and Aquatic Sciences, 49: 1801–1815.

Rivard, D. 1989. Overview of the systematic, structural, and sampling errors in cohort analysis. American Fisheries Society Symposia, 6: 49–65.

Schnute, J., and Fournier, D. 1980. A new approach to length-frequency analysis: growth structure. Canadian Journal of Fisheries and Aquatic Sciences, 37: 1337–1351.

Sparholt, H. 1996. Causal correlation between recruitment and spawning stock size for central Baltic cod? ICES Journal of Marine Science, 53: 771–779.

Tyler, A. V., Beamish, R. J., and McFarlane, G. A. 1989. Implications of age determination errors to yield estimates. Canadian Special Publications in Fisheries and Aquatic Sciences, 108: 27–35.

Van Beek, F. A., and Pastoors, M. A. 1999. Evaluating ICES catch forecasts: the relationships between implied and realised fishing mortality. ICES CM 1999/R: 04.

Appendix: Details of the operating model

All simulations were run on a seasonal basis, dividing the year into four quarters. Population numbers at age were modelled using traditional age-based equations with appropriate modifications for season. Numbers at age in the population were allocated across length classes using the growth model subsequently described. Gear selectivity and maturity were modelled as length-based processes. The individual components of the model are also described in the following discussion. The parameter values used in the population models were intended to simulate a population that is broadly similar in its dynamics to the Eastern Baltic cod stock, and are not intended to provide an exact representation of this stock. The parameter values used are given in Table A1.

Natural mortality

Natural mortality is assumed to be constant for all ages and years, with a value of 0.2 per year being assumed. The proportion of natural mortality occurring in each season is assumed to be equal (i.e. 0.05 per quarter-year).

Recruitment

Recruitment is generated from a Ricker stock–recruitment model with log–normal error, i.e. $\ln(\text{recruits}) = ln(SSB) + (\alpha - (\beta SSB)) + \varepsilon$, where ε is a random error term, $\sim N(0,\sigma_{SR})$. Spawning is assumed to take place at the start of the year, and fish to recruit to the population at age 1.

Maturity

A fixed maturity at length ogive is assumed for all years. If the proportion of fish mature at length l is PM_l, then the maturity ogive can be described by the model

$$PM_l = \frac{1}{3^{((LM_{50}-l)/(MR/2))} + 1} \tag{A1}$$

where, LM_{50} is the length at which 50% of fish are mature, and MR is $(LM_{75} - LM_{25})$, and hence describes the steepness of the maturity ogive. The parameters used were based on fitting a curve of this form to maturity-at-length data for female cod in the Eastern Baltic from a Danish survey in 2000.

Growth

Length at age is modelled using the approach of Schnute and Fournier (1980). Thus, if a_l and a_n are the minimum and maximum ages of fish in the population, and the mean lengths at these ages are given by l_{a1} and l_{an}, respectively, then the mean length at age a_i in year y is given by

Table A1. Parameter values used in the operating model.

Natural mortality	
M	0.2 per year
Stock–recruitment relationship	
α	0.2
β	0.0000002
σ_{SR}	0.4
Maturity	
LM_{50}	30.92 cm
MR	5.85 cm
Growth	
a_1	1
a_n	10
L_{a1}	10 cm
L_{an}	100 cm
S_{a1}	2.5 cm
S_{an}	25 cm
K	0.95
N_A	20 000
N_B	100 000
D_A	−0.05
D_B	0.05
σ_G	0.01
Weight at age	
A	0.000009
B	3
Gear selectivity	
Trawl al_{50}	35 cm
Trawl aSR	4 cm
Trawl bl_{50}	150 cm
Trawl aSR	50 cm
Gillnet al_{50}	40 cm
Gillnet aSR	4 cm
Gillnet bl_{50}	60 cm
Gillnet aSR	4 cm
Effort	
$E_{i,1}$	1.0
σ_E	0.1
Fishing mortality	
F_i^n	0.275

$$l_{ai,y} = l_{a1} + (l_{an} - l_{a1})\frac{(1 - k_c^{(a_i-1)})}{(1 - k_c^{(a_n-1)})} \tag{A2}$$

where c is a subscript indicating the year class spawned in year (y−a), and k_c is a parameter representing the growth rate of cohort c. Similarly, if the standard deviations of length at ages a_1 and a_n are S_{a1} and S_{an}, then the standard deviation of length at age a_i in year y is given by

$$S_{ai,y} = S_{a1} + (S_{an} - S_{a1})\frac{(1 - k_c^{(a_i-1)})}{(1 - k_c^{(a_n-1)})} \tag{A3}$$

Following Jones (1983) and Cook *et al.* (1999), growth rate is assumed to be cohort-specific. Here, growth rate is assumed to be density-dependent, with smaller year classes growing at a faster rate than larger ones. However, the effect is assumed only to take place over a limited range of year-class sizes, such that above and below specified such sizes, there is no additional variation in growth rate. If cohort c was spawned in year y, then its growth rate is estimated as

$$k_c = (1 + \delta_c)K \tag{A4}$$

where K is an overall average for the growth rate parameter. If $N_{1,y+1}$ is the number of recruits in year class c at age 1, N_A and N_B are the minimum and maximum year-class strengths over which density-dependence operates, and D_A and D_B are the corresponding bounds on δ, then δ_c is calculated as follows:

$$\begin{aligned}
&\text{If } (N_{1,y+1} < N_A),\ \text{then}\ \delta_c = D_A + \varepsilon\\
&\text{If } (N_A \geq N_{1,y+1} \geq N_B),\ \text{then}\\
&\quad \delta_c = ((D_A - D_B)/(N_A - N_B))(N_{1,y+1} - N_B) + D_B + \varepsilon\\
&\text{If } (N_{1,y+1} > N_B),\ \text{then}\ \delta_c = D_B + \varepsilon
\end{aligned} \tag{A5}$$

Note that ε is a random error term $\sim N(0,\sigma_G)$, and also that an increase in k_c implies a decrease in growth rate.

Weight at age

Lengths are converted to weights using a fixed relationship of the form Weight = A × Length^B.

Gear selectivity

The availability of fish to the fishery is assumed to be determined only by their size distribution in relation to the selectivity characteristics of the gears in use by the fishing fleets. For each fleet, the relative selectivity at length l, S_l, is described by a curve of the form

$$S_l = \frac{1}{(3^{((al_{50}-l)/(aSR))} + 1)(3^{((l-bl_{50})/(bSR))} + 1)} \tag{A6}$$

This is a combination of an ascending ogive, described by the parameters al_{50} and aSR as in a conventional selection curve for a towed gear, and a descending ogive described by the parameters bl_{50} and bSR, which describes the "deselection" or less efficient capture of larger fish (Hoydal *et al.*, 1982; Cook and Reeves, 1996). The parameters al_{50} and aSR are only approximately equivalent to the length at 50% selection and the selection range of a conventional trawl selection ogive, unless there is no overlap between the selection and deselection sections of the curve. However, in cases of limited overlap, the approximation is sufficiently close that estimates from selectivity trials could be used to parameterize the model. For the current study, it has been assumed that each of the two nations has two fleets, and that these correspond to a trawl fleet and a gillnet fleet, so parameter values have been assumed that approximate to these gear types. It has been assumed that the selectivity for a given gear type remains fixed with time and does not vary between nations. In use, the selectivity curve is rescaled by dividing throughout by its maximum value, in order that the maximum selectivity at length is 1.0.

Fleet effort

To model the relative effort by each fleet, effort is assumed to vary according to an unbiased random walk model, so the effort by fleet i in year y, $E_{i,y}$, is given by

$$E_{i,y} = E_{i,y-1} + d \tag{A7}$$

where d is a random term $\sim N(0,\sigma_E)$. All fleets are assumed to have the same starting value for effort, $E_{i,1}$, and effort is assumed to be distributed evenly across seasons within a year.

Fishing mortality

The numbers at age in a given year and season are allocated to length according to the growth model given above. The proportions of each age class at each length, $p_{a,l}$, are then obtained by dividing the numbers of age a at length l by the total number at age a in the relevant year/season. This is then used together with a nominal starting value for fishing mortality, along with the effort and selectivity curve of a given fleet, to obtain the fishing mortality for that fleet/season, $F_{a,y,q,i}$. Thus,

$$F_{a,y,q,i} = E_{i,y,q} \cdot F_i^n \cdot \sum_{l=L_1}^{l=L_n} \cdot p_{a,l,y,q} S_{l,i} \tag{A8}$$

where a, y, q, and i index age, year, season, and fleet respectively; F_i^n is the nominal starting value for fishing mortality for fleet i, and L_1 are L_n are the minimum and maximum lengths of fish encountered in the population.

Starting populations

The starting population for each realization is generated by first assuming fixed values for recruitment, growth rate, and fleet effort. These are used to generate an equilibrium population corresponding to these conditions. The population is then projected forward for a_n years using the stochastic variation in growth, fishing mortality, and recruitment detailed in the preceding discussion in order to generate initial population numbers for the realization. This approach means that every realization has a different starting population.