Effects of sex, stock, and environment on the shape of known-age Atlantic cod (Gadus morhua) otoliths

Article *in* Canadian Journal of Fisheries and Aquatic Sciences · April 2011

DOI: 10.1139/f03-151 · Source: OAI

Massimiliano Cardinale
Swedish University of Agricultural Sciences

Peer Doering-Arjes
Springforelle

Henrik Mosegaard
Technical University of Denmark

Estimating population age structure using otolith morphometrics: a test with known-age Atlantic cod (*Gadus morhua*) individuals

P. Doering-Arjes, M. Cardinale, and H. Mosegaard

Abstract: Traditional age reading is a rather subjective method that lacks true reproducibility, producing ageing error that propagates up to stock assessment. One alternative is represented by the use of otolith morphometrics as a predictor of age. An important issue with such a method is that it requires known-age fish individuals. Here we used known-age Atlantic cod (*Gadus morhua*) from the Faroe Bank and Faroe Plateau stocks. Cod populations usually show quite large variation in growth rates and otolith shape. We showed that including otolith morphometrics into ageing processes has the potential to make ageing objective, accurate, and fast. Calibration analysis indicated that a known-age sample from the same population and environment is needed to obtain robust calibration; using a sample from a different stock more than doubles the error rate, even in the case of genetically highly related populations. The intercalibration method was successful but generalization from one stock to another remains problematic. The development of an otolith growth model is needed for generalization if an operational method for different populations is required in the future.

Résumé : La méthode courante de lecture des âges est plutôt subjective et elle manque de véritable reproductibilité; elle produit donc une erreur de lecture des âges qui se répercute jusque dans l'évaluation des stocks. Une méthode de rechange est l'utilisation des données morphométriques des otolithes comme variables de prédiction de l'âge. Un problème important associée à cette méthode est qu'elle requiert des poissons individuels d'âge connu. Nous utilisons ici des morues franches (*Gadus morhua*) d'âge connu des stocks du banc et du plateau des Féroé. Les populations de morue affichent généralement une variation assez importante des taux de croissance et de la forme des otolithes. Nous démontrons que l'inclusion de l'étude morphométrique des otolithes dans les processus de détermination de l'âge peut potentiellement permettre une estimation de l'âge qui soit objective, précise et rapide. Une analyse d'étalonnage indique qu'un échantillon de poissons d'âge connu provenant de la même population et du même milieu est nécessaire pour obtenir un étalonnage robuste; l'utilisation d'un échantillon d'une autre population plus que double le taux d'erreur, même lorsqu'il s'agit d'une population fortement apparentée du point de vue génétique. La méthode d'interétalonnage est efficace, mais les généralisations d'un stock à un autre restent problématiques. Il sera nécessaire d'élaborer un modèle de croissance des otolithes pour faire des généralisations, si on a besoin dans le futur d'une méthode opérationnelle applicable à différentes populations.

[Traduit par la Rédaction]

Introduction

Age-structured information is crucial for the assessment and management of fish populations in the North Atlantic (e.g., Campana 2001; Cardinale and Arrhenius 2004). For the majority of assessed fish species (International Council for the Exploration of the Sea (ICES) 2006*a*), otoliths have been the most reliable indicators of individual age because otoliths show annual growth structures and form a permanent record of life history events (Campana 1999). Also, otolith morphometric and otolith outer shape analyses are used to distinguish between stocks that mix in certain areas (i.e., Stransky et al. (2008) and references therein). However, traditional ageing methods rely on experienced age readers to obtain reliable individual age estimation. Because ageing criteria are only semiquantitative (reflecting relative changes in optical properties of the otoliths), age reading is a rather subjective method and thus traditional fish ageing procedures lack true reproducibility (Cardinale and Arrhenius 2004). Recurring calibration exercises are essential to standardize the criteria for assigning age to an individual fish (ICES 2006*b*). Also, the complexity in the otolith preparation for a precise age estimation in some species (see Bedford 1983) gives the traditional ageing techniques a low

Received 18 November 2007. Accepted 30 June 2008. Published on the NRC Research Press Web site at cjfas.nrc.ca on 20 October 2008.
J20273

P. Doering-Arjes.[1] Institute of Freshwater Research, Swedish Board of Fisheries, Stångholmsvägen 2, 178 93 Drottningholm, Sweden.
M. Cardinale.[2] Institute of Marine Research, Swedish Board of Fisheries, P.O. Box 4, 453 21 Lysekil, Sweden.
H. Mosegaard. Technical University of Denmark, National Institute for Aquatic Resources (DTU-Aqua), Charlottenlund Castle, 2920 Charlottenlund, Denmark.

[1]Present address: Institute of Inland Fisheries Potsdam-Sacrow, Im Königswald 2, 14469 Potsdam, Germany (e-mail: Pee.Doering-Arjes@ifb-potsdam.de).
[2]Corresponding author (e-mail: massimiliano.cardinale@fiskeriverket.se).

 doi:10.1139/F08-143

benefit-to-cost ratio when a large number of individuals has to be aged for stock assessment purposes (Cardinale and Arrhenius 2004). In that context, the use of otolith weight as a proxy for age has been advocated (e.g., Boehlert 1985; Worthington et al. 1995*a*; Cardinale et al. 2000), and it has been shown that otolith weight usually explains between 70% and 90% of variability in fish age. Moreover, because otoliths grow continuously from simple ellipsoidal bodies in early life stages to complicated species-specific three-dimensional growth structures in adults, shape and other morphometric features of the otolith may also be used to estimate the age of an individual (Doering and Ludwig 1990; Cardinale et al. 2004).

Age-structured stock assessment models do not require estimates of individual age but only reliable estimates of age composition in the stock, and new approaches combining a cost-efficient use of age-related morphometric features of the otoliths as an objective criteria have been suggested (Francis and Campana 2004). This new objective approach of analysing age composition of a stock may be achieved by a two-step procedure: (*i*) age calibration with individual known-age material (training sample), and (*ii*) application to production samples with unknown ages (test sample) (for a definition, see Francis and Campana 2004). Obviously, calibration requires samples of individuals of known age with minimal inherent error. Several options for production of such material exist. Mark–recapture experiments stocking young individuals from the target population and catching them at different ages is an obvious choice (Neilson et al. 2003). However, several concerns associated with this approach could be foreseen because external structures like shape and other morphometric features of the otolith reflect phenotype expressions that are mainly influenced by sex, age, year class, stock, and environment (Casselmann et al. 1981; Begg and Brown 2000; Cardinale et al. 2004). It has also been suggested that samples aged by expert readers might be used as the basis for calibrations (Francis and Campana 2004). However, agreement among readers (i.e., precision) rarely exceeds 90% and is usually much lower (e.g., Kimura and Lyons 1991).

To our knowledge, no studies have been performed on the suitability of other methods for production of known-age material for morphometric age determination, e.g., conducting rearing experiments under controlled conditions or applying recaptured known-age individuals to other populations under similar conditions. Therefore, in the present study, we analysed the success of age assignment by different combinations of the above-mentioned options to optimize calibration samples for objective ageing. We used samples of known-age individuals of Atlantic cod (*Gadus morhua*) from the Faroese stocks (for further details, see Cardinale et al. 2004). We argue that the inclusion of objective criteria, i.e., measurable parameters, will make ageing faster and more precise, thereby resulting in a higher benefit-to-cost ratio; however, the application of this method for routine ageing strictly depends on reliable calibration methods.

Materials and methods

Fish sampling

The Faroese Fisheries Laboratory and the Aquaculture Research Station initiated an enhancement programme for Atlantic cod in 1991 because of a substantial stock decline and a subsequent fishery collapse in 1990 (Jákupsstovu and Reinert 1994; Fjallstein and Jákupsstovu 1999). Prespawning cod representing two discrete spawning stocks (Jákupsstovu and Reinert 1994) were caught using bottom trawls when the fish were assumed to be at or near their spawning grounds (fig. 16.1 in Fjallstein and Jákupsstovu 1999) on the Faroe Bank (1994) or the Faroe Plateau (1994). Trawl-captured cod were transported to the laboratory and held until they matured. After incubation, both eggs and newly hatched cod larvae were transferred to tanks and mesh bags. Cod juveniles from the two stocks were reared separately in tanks and in pens. When the juvenile cod reached about 1 year old, they were tagged (Anchor Tags FD-68bc; FLOY TAG Inc., Seattle, Washington). Juveniles were then divided into two groups, one for tag and release and one for continued rearing. Individuals to be released were tagged a couple of weeks before transportation to the respective stocking locations. Individuals kept for rearing from the Bank and Plateau stocks were tagged and mixed in three pens, with approximately 50% Bank and 50% Plateau in each pen. The pens, mounted with nets of similar mesh size (depth, ~4 m; diameter, 8 m), were positioned at the sea surface supported by floating rings. Subsamples were taken between January and April of each year.

A total of 8408 fish from the Faroe Plateau was released in 1995, and 3500 from the Faroe Bank and 3000 from the Faroe Plateau were held mixed in pens until the spring of 2000. Individuals from Faroe Bank pen (FBP) and Faroe Plateau pen (FPP) represented the two separate stocks that were reared at the same growing conditions (temperature and feeding conditions). Recaptured individuals from Faroe Plateau (FPR) represented the wild counterpart (recaptures from the Faroe Bank stock were very few and therefore not included in our analysis). For each fish, sex, total fish length (TL; to the nearest millimetre), and wet weight (to the nearest gram) were recorded. Both sagittal otoliths were removed from each individual sampled and stored in dry envelopes for shape analysis. Otoliths were embedded in black epoxy, and transverse sections to be used for traditional age reading were cut with a diamond saw in the dorsal–ventral direction near the otolith centre. In sections coded for blind reading, readers were given different levels of information about origin of the otoliths (e.g., cod length, sampling time, and general otolith size – cod size relationship).

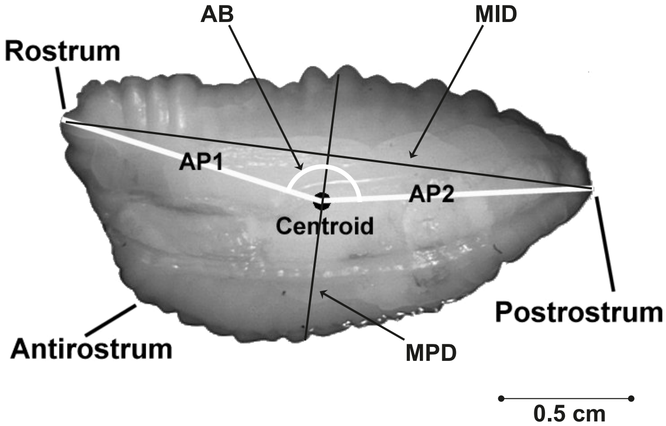
Otolith shape analysis

We recorded otolith morphometrics and normalised Fourier descriptors (NFDs) of otolith shape using IMAGIC (i.e., image analysis system; Van Heel et al. 1996) and procedures described in Cardinale et al. (2004).

Geometrical and nongeometrical otolith descriptors and their definition

Otolith orientation was standardized by positioning each otolith with the proximal otolith side facing up, the rostrum to the left, ventral edge away from the reader, and dorsal edge towards the reader (Fig. 1). The following variables were recorded for each otolith pair using the IMAGIC im-

Fig. 1. Left sagittal otolith of a 4-year-old Atlantic cod (*Gadus morhua*) from Faroe Bank. MID, length measured as the length of the maximum internal distance; MPD, width measured as the length of the longest axis perpendicular to MID; AP1, distance between the centroid (the centroid is determined as the centre of mass of the polygon defined by the contour points) of the otolith and the rostrum; AP2, distance between the centroid of the otolith and the postrostrum; AB, angle between AP1 and AP2.



age analysis procedures: area (OA), contour length (OC), length (MID), width (MPD), the distance between the centroid (the centroid is determined as the centre of mass of the two-dimensional (2D) polygon defined by the contour points) of the otolith and the rostrum (major apex, AP1), the distance between the centroid of the otolith and the postrostrum (minor apex, AP2), and the angle between AP1 and AP2 (AB). Otolith length was measured as the length of the maximum internal distance (MID), and otolith width was measured as the length of the longest axis perpendicular to the maximum internal distance (MPD). Each otolith was weighed to the nearest 0.1 mg. Finally, we estimated the ratio between otolith weight (OW) and fish total length (TL), i.e., otolith weight per unit of fish length ($OW \cdot TL^{-1}$).

Data analysis

All measured otolith morphometric variables were tested for assumptions of normality and homogeneous variance using the Shapiro–Wilk W test (Shapiro et al. 1968) and Hartley's F_{max} statistic (Sokal and Rohlf 1995), respectively. Tests were computed for each selected variable among age classes, stocks, and environments. We used analysis of variance (ANOVA) to analyse shape differences between right and left otolith, as well as sexes, for samples from the same stock, environment, and age class by using the otolith length corrected otolith morphometric variables (for details, see Cardinale et al. 2004) and the NFDs. One-way ANOVA (unbalanced design, fixed effect) was used to inspect each variable. Tukey's honestly significant difference (HSD) test for unequal sample size was used for a posteriori comparisons of each significant variable. Each variable in at least one age class was not normally distributed (Shapiro–Wilk W test; $p < 0.05$), and none of the variables, except the ratio of otolith weight to fish length, had equal variance within all age classes (Hartley's F_{max} statistic, $p < 0.05$). Because parametric assumptions were not satisfied, these variables were natural logarithm transformed prior to statistical analysis. Following logarithmic transformation, all variables conformed to parametric statistical assumptions.

Stepwise discriminant analysis was used to find a small subset of significant variables separating cod of different ages. The significance of a variable staying in the model was set at $p = 0.01$ for the F test. In all analyses, a stepwise discriminant analysis preceded a linear discriminant analyses and was conducted for each location separately because the shape of the cod otoliths is strongly influenced by both stock and environment of the fish (Cardinale et al. 2004). In this way, any confounding variation attributable to differences in stock and environment of the samples was removed (Begg and Brown 2000). In the stepwise selection procedure of variables, only one variable can be entered into the model in each step. The procedure does not account for the relationships between variables that have not yet been selected, and important variables may therefore be excluded in the process (SAS Institute Inc. 2003). To ensure the inclusion of otolith weight in the model, this variable was always included a priori. Following stepwise discriminant analysis, we used linear discriminant analysis to estimate the discriminant function and the corresponding canonical variates (CanVar) for each analysis and thus represented the optimal combination of morphometric variables and NFDs providing the best overall discrimination between age groups (Begg et al. 2000). Jackknifed cross-validation procedures were used to calculate an unbiased estimation of individual classification success. The significance level was set at 0.05 for all statistical tests used, and results of multiple comparisons were evaluated on a case-by-case basis to judge their reliability (Perneger 1998). The following software was used for the statistical analysis: SPSS Inc. (1999) for ANOVA and Tukey's test; SAS Institute Inc. (2003) for discriminant analysis with jackknifed cross-validation procedures.

Comparisons between readers with different levels of information

When known-age material is not available for calibration then two options could be tested. It might be possible to (*i*) rely on similarities to a neighbouring stock with known-age material that could act as the calibration sample or (*ii*) let age readers define the baseline ages for the calibration in the unknown stock. We evaluated these two options. Moreover, an exercise following the traditional age reading method of the Faeroese Fisheries Laboratory was carried out using subsamples of the otoliths from the FPP and FPR sources (Table 1). Four readers (one from each of the Faroes (FA), Denmark (DK1 and DK2), and Sweden (SE)) were given the following different amounts of information: (*i*) low, only very general information about the experiment (Faeroe cod not identified, but two possible treatments, either rearing in captivity or recapture of marked fish); (*ii*) intermediate, the same as (*i*) but additional information about fish length and catch season; and (*iii*) high, the same as (*ii*) but additional information about stock, treatment, and fish weight, a plot of otolith sizes at different ages, and hints about techniques used by Faroe age readers.

For each otolith, the modal age was calculated for each selection of readings (with low, intermediate, or high levels of information). Readers examined otoliths several times to

Table 1. Available known-age cod otoliths from the three sources Faroe Bank pen (FBP), Faroe Plateau pen (FPP), and Faroe Plateau recaptures (FPR) used in the present analysis.

	Age (years)						
Stock	1	2	3	4	5	6	Total
FBP all		44	44	39	50	33	210
FPP subsample traditionally read		27	30	30	30	29	146
FPP subsample unread		19	17	21	18	14	89
FPR subsample traditionally read	12	2	4	11			29
FPR subsample unread	33	115	78	11			237
Total	45	207	173	112	98	76	711

provide an average estimate. If a single mode did not exist, e.g., because of equal number of estimates in each estimated age category, the median was chosen instead (Table 2).

Comparison of classification results was performed by two different methods: (*i*) based purely on true ages and (*ii*) based on age reader's expert opinion of true ages; all otoliths were still from the known-age material. In the first step of method *i*, each of the three sources (FBP, FPP, and FPR) was analysed with respect to misclassification of known-age otoliths by applying a separate linear discriminant analysis (DA) on the stepwise selected morphometrics from the entire otolith material from each source (Fig. 2*a*). The results are reported as the percentages of otoliths misclassified into other categories by known-age group. The cross-validated parameter estimates from the three DAs were then used in the second step of method *i* (generalization) to investigate to what degree the discriminant functions were applicable in age classification of other sources or populations (test samples) (Fig. 2*a*). The ability to use a discriminant function from one population to assign ages to individuals in another population would be useful in the very common case in which known-age material is not at hand (as an example, see ICES 2006*b*). The estimated discriminant functions (from the calibration samples) were applied to each of the other two sources (test samples). Results from the six possible comparisons, based on the selected morphometrics from each entire otolith sample, were reported as percentage misclassification by known-age group.

In method *ii*, otolith ages in subsamples (S_{read}) from two of the sources (FPP and FPR) were estimated using traditional age reading (Fig. 2*b*). The traditionally estimated ages were used as baselines in the calibration of a discriminant analysis $\mathrm{DA}_{\mathrm{read}}$. The estimates were taken as the mode of five different readings by three of the readers (Table 2, Mode ($\mathrm{DK1}_{\mathrm{low}}$, $\mathrm{DK2}_{\mathrm{low}}$,$\mathrm{SE}_{\mathrm{low}}$, $\mathrm{DK1}_{\mathrm{intermediate}}$, $\mathrm{DK2}_{\mathrm{intermediate}}$)). The morphometric variables were selected by a stepwise procedure and the parameters of the linear discriminant function were cross-validated in the same way as when they were based on known ages. First, we analysed the success of cross-validated classification into age classes created by traditional reading. Secondly we compared the classified ages with the individual known ages. Finally, we investigated the efficiency of the estimated parameters from the $\mathrm{DA}_{\mathrm{read}}$ to classify ages of a different otolith subsample (S_{unread}) from the same source, thus the discriminant function from $\mathrm{DA}_{\mathrm{read}}$ was applied on the morphometric values in S_{unread}.

The classification efficiency based on otolith morphometrics calibrated by a discriminant function from a surrogate population (method *i*) was compared with the classification efficiency based on individual known ages with reader errors (method *ii*). For this comparison, only the subsample S_{read}, in which the otoliths were assigned by age readers, was used for discriminant analysis of the morphometrics (as calibration) to be applied on the remaining S_{unread} otoliths (as a test). First, cross-validated parameters were estimated from the FPR–S_{read} and applied to the FPR–S_{read} and FPR–S_{unread} subsamples; secondly, cross-validated parameters were estimated from the FPP–S_{read} and applied to the FPP–S_{read} and FPP–S_{unread} subsamples (Fig. 2*b*).

Results

Comparisons between readers with different levels of information

There were no consistent differences between the left and right otolith and between sexes (ANOVA, $p > 0.05$) for any of the variables investigated within different age classes, stocks, and environments. Therefore, we used the left otolith when it was available and the right otolith when the left otolith was not present or utilizable. The amount of discarded, damaged, or crystalline otoliths was less than 5% of all samples. Sexes were merged to increase the sample size used in the statistical analysis.

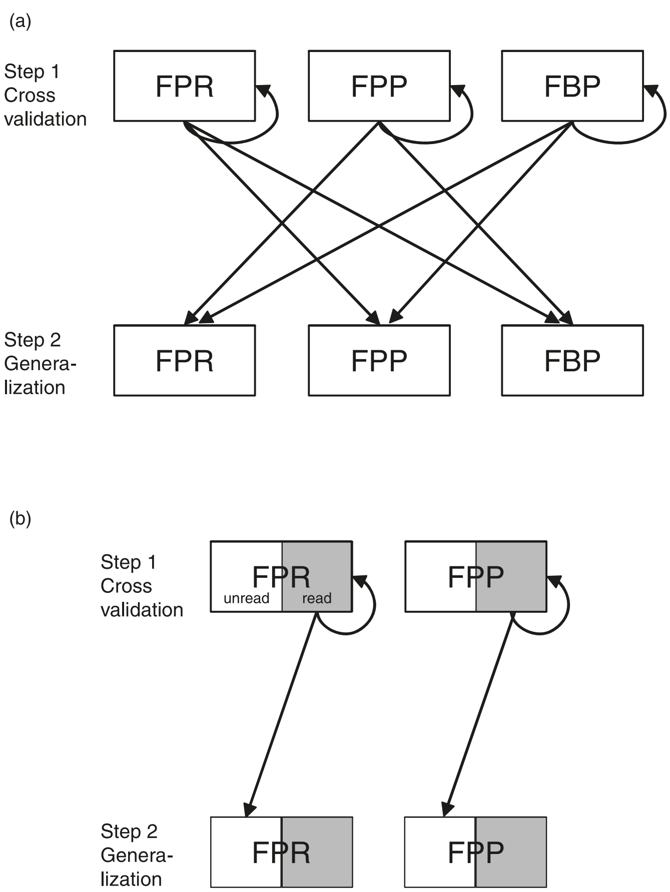
A successful estimation of otolith age from the two Faroe cod populations by traditional age reading seems to rely on the specific experience of the age reader and the degree of information about the fish. The Faroe reader (FA) exhibited 0% misclassification rate of pen-reared FPP but a high misclassification of recaptured FPR cod (87%) when no fish size and catch date information was given (Table 2). However, when the full information was available, misclassifications were as low as 1% and 5%, respectively. The two Danish readers (DK1 and DK2) and the single Swedish reader (SE) were unfamiliar with the Faeroe cod and generally had a low accuracy in their interpretations at all information levels tested. There was an overall better performance with the pen-reared cod otoliths (FPP mean misclassification = 47%), which had been sampled once a year as opposed to the year-round recaptures (FPR mean misclassification = 59%). However, there was no apparent improvement of readings with increasing information level or technique for any of the two sources of otoliths. Even the experienced reader failed for the continuously sampled fish. Combining the available five readings at the low and

Table 2. Classification error (in proportion of the total analysed otoliths) of traditional age reading with different levels of information on the analysed cod otolith.

Level of information	Reader	FPP (*n* = 146)	FPR (*n* = 29)
Low	DK1	0.53	0.92
Low	DK2	0.61	0.44
Low	SE	0.49	0.32
Low, but expert on the cod populations	FA	0.00	0.87
Intermediate	DK1	0.58	0.40
Intermediate	DK2	0.23	0.54
Intermediate, but expert on the cod populations	FA	0.01	0.05
High	DK1	0.63	0.82
High	DK2	0.25	0.70
Mode ($DK1_{low}$, $DK2_{low}$, SE_{low}, $DK1_{intermediate}$, $DK2_{intermediate}$)		0.22	0.41
Mode (all nine reading series)		0.05	0.48

Note: FPP, Faroe Plateau pen; FPR, Faroe Plateau recaptures.

Fig. 2. Scheme of the validation–generalization analysis: (*a*) discriminant analysis with complete samples; (*b*) discriminant analysis with subsamples.



intermediate information level for the Danish and Swedish readers by calculating the modal age reading improved the accuracy somewhat, especially for the FPP source (Table 2).

Exploring alternative calibration strategies

Effectiveness of otolith size and shape to calibrate age in the same population was analysed by linear discriminant analysis. The discrimination power among age classes was somewhat higher for the two pen-reared cod stocks than for the recaptured Plateau cod, with error rates 11%, 11%, and 14%, respectively, and when only comparing age classes 2–4, misclassifications were 5%, 6%, and 14%, respectively (Table 3*a*). The pattern of correctly classified and misclassified otoliths is apparent in the plots of the two first canonical variables (Figs. 3*a*–3*c*). Age classes 2 and 3 are well separated in Figs. 3*a* and 3*c*. A poorer separation of age classes is found in the recaptured FPR cod compared with the FBP and FPP pen-farmed cod.

For calibration on cod from pens but with individuals from a different stock, the misclassification rates were moderate (14% and 24% for ages 2–4 years; see Table 3) compared with when recaptured or pen cod are from the same stock (41% and 26% for ages 2–4 years; see Table 3). However, varying both treatment and stock gave unexpectedly low misclassification error (18% and 8% for ages 2–4 years; see Table 3). The misclassification rates increased considerably when older individuals were considered.

The strategy of using traditional age reading of a subsample as the baseline for calibrating another subsample from the same stock is explored in Table 3*b*. Here the result is quite dependent on the accuracy of the readings. The modal age of the FPP cod (S_{read}) included 22% incorrect ages (reader error; Table 2) in the calibration material and gave an overall error of the FPP validation material (S_{unread}) of only 30% for all ages. Although the cross validation of S_{read} indicated a 41% misclassification, the true error was only 22% (Table 3*b*).

In the case of FPR–S_{read}, the modal age included 41% incorrect ages (reader error) in the calibration material (Table 2). Although the cross validation of the S_{read} indicated 24% misclassification for all ages, the true error was as high as 47%, and the application on the test sample gave an overall misclassification of the FPR validation material (S_{unread}) of 54%.

Discussion

As shown by many authors (Boehlert and Yoklavich 1984; Kimura and Lyons 1991; Richards et al. 1992), errors caused by subjectivity in age determination are a common source of bias in otolith reading. Errors in ageing of fish are usually on the order of 5%–15% coefficient of variation (Kimura and Lyons 1991; Campana 2001). The interpretation of otoliths may change among replicates, as a reader becomes

Table 3. (*a*) Classification error (in proportion of analysed individuals) using known-age (method *i*) cod to calibrate discriminant analysis of otolith morphometrics to estimate cod individual age in the same or another sample. (*b*) Classification errors using known-age otoliths with reader error from a subsample S_{read} (method *ii*) to calibrate the discriminant analysis (DA) for prediction of ages in the same (S_{read}) or another (S_{unread}) subsample from the same population.

			Age (years)								
Row no.	DA from:	Used to predict ages for:	1	2	3	4	5	6	All	2–4	*n*
(*a*) Using known-age cod to estimate cod individual age in the same or another sample											
1	FPP	FPP	—	0.02	0.00	0.12	0.31	0.09	0.11	0.05	235
2	FPR	FPR	0.00	0.12	0.18	0.13	—	—	0.14	0.14	265
3	FBP	FBP	—	0.00	0.02	0.15	0.28	0.09	0.11	0.06	210
4	FPP	FBP	—	0.00	0.16	0.59	0.24	0.30	0.25	0.24	210
5	FPP	FPR	1.00	0.49	0.36	0.13	—	—	0.45	0.41	265
6	FBP	FPP	—	0.00	0.13	0.27	0.65	0.33	0.28	0.14	235
7	FBP	FPR	1.00	0.23	0.08	0.29	—	—	0.23	0.18	265
8	FPR	FPP	—	0.24	0.17	0.35	1.00	1.00	0.54	0.26	235
9	FPR	FBP	—	0.16	0.02	0.05	1.00	1.00	0.44	0.08	210
(*b*) Using known-age otoliths with reader error from S_{read} to predict ages in S_{read} or S_{unread} from the same population											
1	FPP–S_{read}	FPP–S_{read}	—	0.04	0.41	0.57	0.59	0.35	0.41	0.36	147
2	FPP–S_{read}	FPP–S_{read}	—	*0.07*	*0.20*	*0.33*	*0.27*	*0.21*	*0.22*	*0.20*	146
3	FPP–S_{read}	FPP–S_{unread}	—	0.05	0.18	0.48	0.56	0.15	0.30	0.25	88
4	FPR–S_{read}	FPR–S_{read}	0.40	0.20	0.50	0.13	0.00	1.00	0.24	0.20	29
5	FPR–S_{read}	FPR–S_{read}	*0.67*	*0.00*	*0.75*	*0.45*	—	—	*0.47*	*0.40*	29
6	FPR–S_{read}	FPR–S_{unread}	0.67	0.26	0.85	0.75	—	—	0.54	0.53	82

Note: FBP, Faroe Bank pen; FPP, Faroe Plateau pen; FPR, Faroe Plateau recaptures. In part *a*, for rows 1–3, classification errors are from cross validation of the same sample; for rows 4–9, classification errors are from using a surrogate population to calibrate the DA to predict ages in another population. In part *b*, the method assumes that the modal age of traditionally read otoliths is correct, whereas there is an actual reader error in each of the S_{read} subsamples: row 1: apparent classification errors using cross-validated DA parameters estimated from the FPP–S_{read} using reader-assigned ages with inherent 22% incorrect ages; row 2, the same classification as in row 1 but compared with the true ages (in italics); row 3, cross-validated parameters estimated in row 1 and applied to the FPP–S_{unread} subsample; row 4, apparent classification errors using cross-validated DA parameters estimated from the FPR–S_{read} using reader-assigned ages with inherent 41% incorrect ages; row 5, the same classification as in row 4 but compared with the true ages (in italics); and row 6, cross-validated parameters estimated in row 4 and applied to the FPR–S_{unread} subsample.

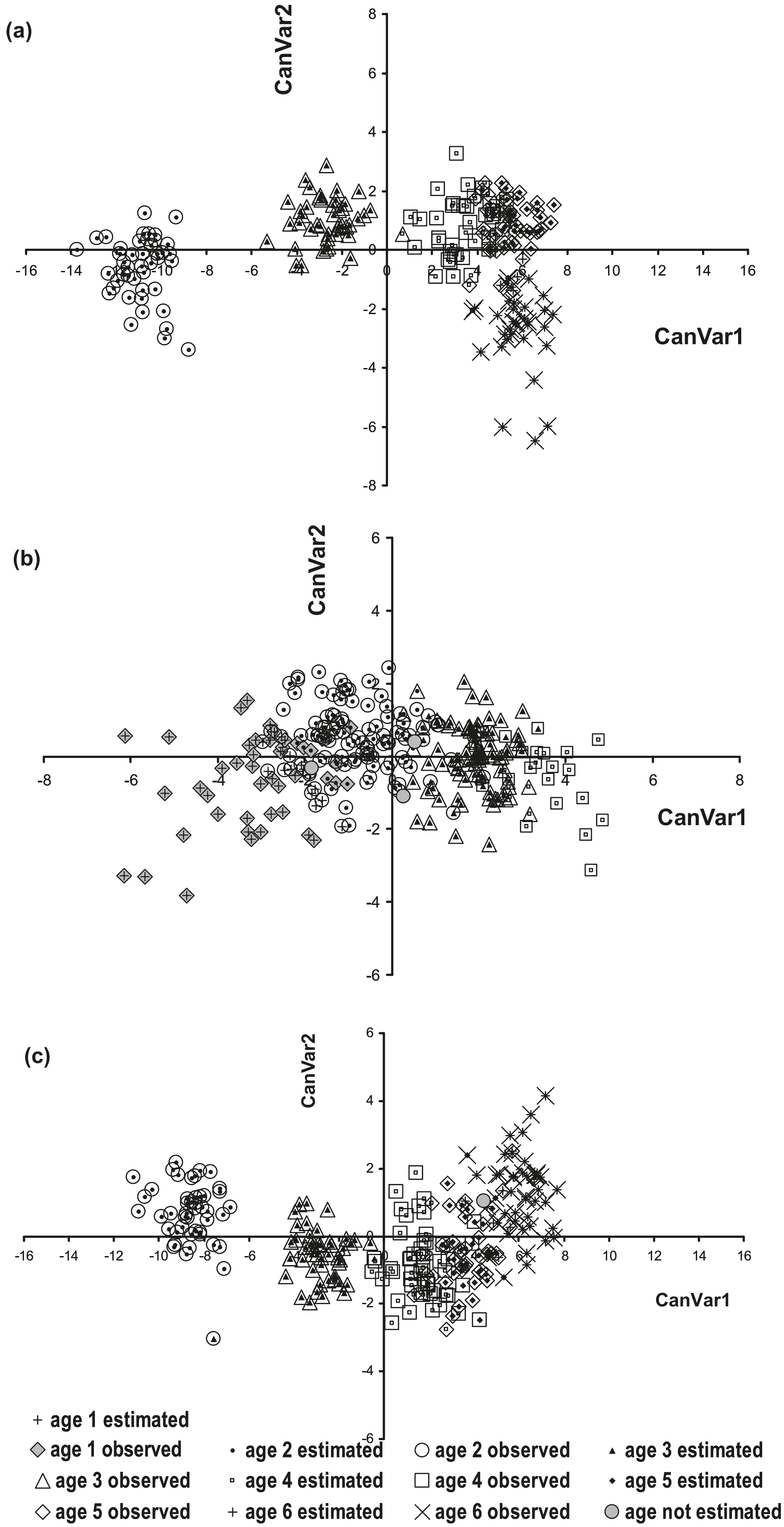
more experienced, or among different readers (Worthington et al. 1995*b*). Such errors are believed to largely affect the estimates derived from virtual population analysis (VPA) models (ICES 2006*b*). Different stocks usually show differences in otolith morphometrics, as both genetic and environmental influences on otolith shape are likely to exist (Cardinale et al. 2004; Ruzzante et al. 2006). Both otolith shape (Doering and Ludwig 1990) and otolith weight (Worthington et al. 1995*b*; Cardinale and Arrhenius 2004; Francis and Campana 2004) have been used as objective measures for ageing. The combination of both has the potential to increase the accuracy of predicting individual fish age.

Estimated proportions at age that are usually based on age–length keys may contain significant bias (Kimura 1977). Francis and Campana (2004) discussed four types of bias. Discriminant bias associated with the cutting rule for assigning an individual to an age class is one source of bias in a discriminant analysis. Francis and Campana (2004) and Francis et al. (2005) argued that it is possible to avoid such bias by applying a maximum likelihood based mixture analysis. The suggested method consists of two linked procedures: the calibration with known-age material or growth observations (learning sample) and the application of the calibrated parameters to full scale (test sample). In our study, we tried to avoid this discriminant bias in the analysis by having approximately balanced sample sizes at age for the objectively based calibration. The selection of otoliths from recaptures was not under our control so the resulting, somewhat unbalanced sample sizes may have resulted in the observed bias, where the proportion of the smallest age classes was overestimated. Nevertheless, this bias is estimated to be small and to have not affected the results of the analysis.

Otoliths are known to grow bigger with increasing fish age. However, when the exact relationship between otolith growth and age for a stock or population is unknown, it may be tempting to generalize from a population with a known relationship. Although we demonstrated that including otolith morphometrics in the ageing process makes the method more reliable, using surrogate material from a different source still remains problematic. Calibration analysis shows that a known-age sample from the same population (in the statistical sense) is needed to obtain a robust calibration and that using another population may more than double the age reading error. Even though the populations are highly genetically related, their use as calibration sample does not eliminate problems related with the generalization method. Here we show that the reader intercalibration method might work, but calibration with a surrogate source does not because it depends on stock-specific growth rates. The best basis for reader-independent calibration is when the age classes are well separated, as in the once-per-year samples from the pens.

In analysing the internal consistency of our procedure in

Fig. 3. Two first canonical variables (CanVar) from (*a*) Faroe Bank pen (FBP), (*b*) Faroe Plateau recaptures (FPR), and (*c*) Faroe Plateau pen (FPP) cross validations.



estimating the correct age, we can state that the pen cod show an error rate of 11% for ages 2–6 years and ~5% for ages 2–4 years and the wild fish show an error rate of 14% for ages 2–4 years. There can be several explanations for this. Individual growth in the wild varies. There is variation in age in days: the wild fish are more or less continuously caught, whereas pen cod are sampled block-wise. Fisheries act selectively: the fast growers are caught first. The strong separation of age groups of reared cod (because they were sampled only once per year) and the procedure of cross validation to some degree corrected the classification towards a better correspondence with the true ages. For the recaptured cod, however, the larger inherent size variation by the continuous sampling over the year increased the classification errors.

Our results therefore indicate the importance of a sampling strategy that minimizes the effects of continuous otolith growth and different environmental influence for a robust calibration procedure if age classification based on otolith metrics should replace traditional age reading.

To be fair when comparing the two strategies, the same otolith material should be used. Because only subsamples were used for the reader-assigned baselines, the first strategy (using different stocks for calibration and testing) was also recalculated using the appropriate subsamples (instead of the entire material). Using FPR known-age material (from the subsample S_{read}) as calibration for the age estimation of FPP (subsample S_{unread}), the error was 39% (not tabulated) and in the same range as the whole material, where the error was 54% for all ages and 26% for ages 2–4 years. Using subsamples of FPP on FPR also gave a result quite similar to the earlier mentioned 41%–44% error rate for the entire material.

Thus, the discriminant analysis can separate age classes for the pen cod because they are sampled once per year, whereas the recaptures are continuously sampled and therefore not easily separated. This calls for developing an otolith growth model. On the other hand, because of variation in fish and otolith growth, there is an increasing overlap of otolith weight in successive age classes with increasing age within the pen-reared cod.

The main goal of this study was to estimate the age classification success at the individual fish level to find valid calibration procedures for objective ageing. Including otolith weight in the age estimation process makes the method reliable and fast. However, generalizing from one population to another remains problematic. To have an operational procedure with which to obtain a reliable age, three different approaches are possible: (*i*) obtain known-age material from the same stock and growth conditions, e.g., by tagging and recapturing fish, preferably combined with otolith marking (e.g., strontium) as comparing pen with wild fish is insufficient; (*ii*) use traditional ageing relying on accurate expert readers; or (*iii*) generalize from one population to another. The first procedure appears superior to the other two.

However, the results shown here demonstrate that the development of a growth model is needed for generalization if an operational method for a different population is required (Campana 2005). For good performance of traditional ageing as calibration, there is a need for sufficiently large material, inclusion of all possible age groups, many readers, and cross validation. Selecting otoliths with low reader variation could possibly refine the method and further criteria may be extracted from the discriminant analysis or related methods by the posterior probability of the individual scoring in relation to estimated age group.

In conclusion, the fish age estimation procedure may be optimized in relation to the most discriminating calibration sample by using otolith weight and otolith morphometrics. However, more work is still needed: the accuracy of age composition in the production sample has to be explored using different calibration strategies and the relative precision of the stock assessment output has to be investigated. This might be done using estimates of fishing mortality, spawning stock biomass, and the number of recruits by simulation studies with different calibration strategies. Finally, a cost–benefit analysis of the strategy has to be made in relation to total uncertainties.

Acknowledgements

We thank the Faroe Fisheries Laboratory for providing known-age cod otoliths and participating in the traditional age reading. Further, we acknowledge the funding received from the European Commission for the FP4 project FAbOSA (FAIR CT97 no. 3402) and the FP6-STP project Automated Fish Ageing (AFISA) (SSP8 no. 044132). We also thank the anonymous referees for their useful suggestions on an earlier version of the manuscript.

References

Bedford, B.C. 1983. A method for preparing sections of large numbers of otoliths embedded in black polyester resin. J. Cons. Int. Explor. Mer, **41**: 4–12.

Begg, G., and Brown, R. 2000. Stock identification of haddock *Melanogramus aeglefinus* on Georges Bank based on otolith shape analysis. Trans. Am. Fish. Soc. **129**: 935–945. doi:10.1577/1548-8659(2000)129<0935:SIOHMA>2.3.CO;2.

Begg, G., Overholtz, W.J., and Munroe, N.J. 2000. The use of internal otolith morphometrics for identification of haddock (*Melanogramus aeglefinus*) stocks on Georges Bank. Fish. Bull. (Washington, D.C.), **99**: 1–14.

Boehlert, G.W. 1985. Using objective criteria and multiple regression models for age determination in fishes. Fish. Bull. (Washington, D.C.), **83**: 103–117.

Boehlert, G.W., and Yoklavich, M.M. 1984. Variability in age estimates in *Sebastes* as a function of methodology, different readers, and different laboratories. Calif. Fish Game, **70**: 210–224.

Campana, S.E. 1999. Chemistry and composition of fish otoliths: pathways, mechanisms and applications. Mar. Ecol. Prog. Ser. **188**: 263–297. doi:10.3354/meps188263.

Campana, S.E. 2001. Accuracy, precision and quality control in age determination including a review of the use and abuse of age validation methods. J. Fish Biol. **59**: 197–242. doi:10.1111/j.1095-8649.2001.tb00127.x.

Campana, S.E. 2005. Otolith science entering the 21st century. Mar. Freshw. Res. **56**: 485–495. doi:10.1071/MF04147.

Cardinale, M., and Arrhenius, F. 2004. Application of the otolith weight–age relationship to estimate the age-structure of haddock (*Melanogramus aeglefinus*). J. Appl. Ichthyol. **20**: 470–475. doi:10.1111/j.1439-0426.2004.00576.x.

Cardinale, M., Arrhenius, F., and Johnsson, B. 2000. Potential use of otolith weight for the determination of age-structure of Baltic cod (*Gadus morhua*) and plaice (*Pleuronectes platessa*). Fish. Res. **45**: 239–252. doi:10.1016/S0165-7836(99)00122-8.

Cardinale, M., Doering-Arjes, P., Kastowsky, M., and Mosegaard, H. 2004. Effects of sex, stock, and environment on the shape of known-age Atlantic cod (*Gadus morhua*) otoliths. Can. J. Fish. Aquat. Sci. **61**: 158–167. doi:10.1139/f03-151.

Casselmann, J.M., Collins, J.J., Crossman, E.J., Ihssen, P.E., and Spangler, G.R. 1981. Lake whitefish (*Coregonus clupeiformis*) stocks of the Ontario waters of Lake Huron. Can. J. Fish. Aquat. Sci. **38**: 1772–1789. doi:10.1139/f81-225.

Doering, P., and Ludwig, J. 1990. Shape analysis of otoliths — a tool for indirect ageing of eel, *Anguilla anguilla* (L.)? Int. Rev. Gesamten Hydrobiol. **75**: 737–743. doi:10.1002/iroh.19900750607.

Fjallstein, I., and Jákupsstovu, S.H. 1999. Release, recapture and migration of reared Atlantic cod (*Gadus morhua* L.) on the Faroe Islands. *In* Stock enhancement and sea ranching. *Edited by* B.R. Howel, E. Moksnes, and T. Svåsand. Fishing News Books, Blackwell Science Ltd., Osney Mead, Oxford, UK. pp. 231–242.

Francis, R.I.C.C., and Campana, S.E. 2004. Inferring age from otolith measurements: a review and a new approach. Can. J. Fish. Aquat. Sci. **61**: 1269–1284. doi:10.1139/f04-063.

Francis, R.I.C.C., Harley, S.J., Campana, S.E., and Doering-Arjes, P. 2005. Use of otolith weight in length-mediated estimation of proportions at age. Mar. Freshw. Res. **56**: 735–743. doi:10.1071/MF04127.

International Council for the Exploration of the Sea. 2006*a*. Report of the ICES Advisory Committee on Fishery Management, Advisory Committee on the Marine Environment and Advisory Committee on Ecosystems, 2006. ICES Advice, Books 1–10, 2, Copenhagen, Denmark.

International Council for the Exploration of the Sea. 2006*b*. Report of the Study Group on Ageing Issues of Baltic Cod (SGABC), 16–19 May 2006, Gdynia, Poland. ICES CM 2006/BCC:08.

Jákupsstovu, S.H., and Reinert, J. 1994. Fluctuations in the Faroe Plateau cod stock. ICES Mar. Sci. Symp. **198**: 194–211.

Kimura, D.K. 1977. Statistical assessment of the age–length key. J. Fish. Res. Board Can. **34**: 317–324.

Kimura, D.K., and Lyons, J.J. 1991. Between reader bias and variability in the age-determination process. Fish. Bull. (Washington, D.C.), **89**: 53–60.

Neilson, J.D., Stobo, W.T., and Perley, P. 2003. Age and growth of Canadian east coast pollock: comparison of results from otolith examination and mark–recapture studies. Trans. Am. Fish. Soc. **132**: 536–545. doi:10.1577/1548-8659(2003)132<0536:AAGOCE>2.0.CO;2.

Perneger, T.V. 1998. What’s wrong with Bonferroni adjustments. BMJ, **316**: 1236–1238. PMID:9553006.

Richards, L.J., Schnute, J.T., Kronlund, A.R., and Beamish, R.J. 1992. Statistical models for the analysis of ageing error. Can. J. Fish. Aquat. Sci. **49**: 1801–1815. doi:10.1139/f92-200.

Ruzzante, D.E., Mariani, S., Bekkevold, D., André, C., Mosegaard, H., Clausen, L.A.W., Dahlgren, T.G., Hutchinson, W.F., Hatfield, E.M.C., Torstensen, E., Brigham, J., Simmonds, E.J., Laikre, L., Larsson, L.C., Stet, R.J.M., Ryman, N., and Carvalho, G.R. 2006. Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. Proc. R. Soc. Lond. B Biol. Sci. **273**: 1459–1464. doi:10.1098/rspb.2005.3463.

SAS Institute Inc. 2003. SAS Version (*r*) 9.1. SAS Institute Inc., Cary, N.C.

Shapiro, S.S., Wilk, M.B., and Chen, H.J. 1968. A comparative study of various tests of normality. J. Am. Stat. Assoc. **63**: 1343–1372. doi:10.2307/2285889.

Sokal, R.R., and Rohlf, F.J. (*Editors*). 1995. Biometry. 3rd ed. W.H. Freeman and Company, New York.

SPSS Inc. 1999. SYSTAT 10.0 statistics. SPSS Incorporated, Chicago, Ill.

Stransky, C., Baumann, H., Fevolden, S.-E., Harbitz, A., Høie, H., Nedreaas, K.H., Salberg, A.-B., and Skarstein, T.H. 2008. Separation of Norwegian coastal cod and Northeast Arctic cod by outer otolith shape analysis. Fish. Res. **90**: 26–35. doi:10.1016/j.fishres.2007.09.009.

Van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. 1996. A new generation of the IMAGIC Image Processing System. J. Struct. Biol. **116**: 17–24. doi:10.1006/jsbi.1996.0004. PMID:8742718.

Worthington, D.G., Fowler, A.J., and Doherty, P.J. 1995*a*. Variation in the relationship between otolith weight and age: implications for the estimation of age of two tropical damsel-fish (*Pomacentrus moluccensis* and *P. wardi*). Can. J. Fish. Aquat. Sci. **52**: 233–242. doi:10.1139/f95-224.

Worthington, D.G., Fowler, A.J., and Doherty, P.J. 1995*b*. Determining the most efficient method of age determination for estimating the age structure of a fish population. Can. J. Fish. Aquat. Sci. **52**: 2320–2326. doi:10.1139/f95-224.