NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

Northwest Atlantic Fisheries Organization

Serial No. N4384 **NAFO SCR Doc. 01/16**

SCIENTIFIC COUNCIL MEETING – JUNE 2001

Age Readings of *Sebastes marinus* Otoliths: Bias and Precision Between Readers and Otolith Preparation Methods

by

Christoph Stransky[1], Sif Guðmundsdóttir[2], Þorsteinn Sigurðsson[2], Svend Lemvig[3] and Kjell Nedreaas[3]

[1] Federal Research Centre for Fisheries, Institute for Sea Fisheries, Palmaille 9, D-22767 Hamburg, Germany
stransky.ish@bfa-fisch.de

[2] Marine Research Institute, Skúlagata 4, 101 Reykjavík, Iceland
sif@hafro.is steini@hafro.is

[3] Institute of Marine Research, P.O. Box 1870, Nordnes, N-5817 Bergen, Norway
svend.lemvig@imr.no kjell.nedreaas@imr.no

Abstract

This study presents a comparative age reading on *Sebastes marinus* from the Icelandic shelf within an otolith exchange program between institutes in Germany, Iceland and Norway. Out of a series of 212 otolith pairs, one otolith of each pair was prepared by the break-and-burn technique, while the other otolith was used for cross-sections. Age reading results are compared between readers and otolith preparation methods in terms of bias and precision, using a set of statistical tests and graphical methods. Significant bias was observed for both the comparison between readers and between methods, mainly caused by deviations between age scores in the higher ages (> 20 years). Precision estimates, involving the high longevity of redfish, were relatively good compared to previous age reading comparisons using other species. In contrast, the age dependent percent agreement was poor (< 30%) for a tolerance level of ± 0 years, particularly for the age range 21-30 years. A tolerance level of ± 3 years, however, lead to around 90% agreement for the age range up to 20 years. The fit of age reading scores with the von-Bertalanffy growth curve was relatively good, providing growth parameters comparable to *S. marinus* from the Norwegian shelf. The observed problems in bias and precision of age readings should to be improved by continuing with similar *Sebastes* otolith exchange programs and setting up a further age reading workshop to harmonise the interpretation of growth structures.

Keywords: redfish, *Sebastes marinus*; age reading, age determination, otolith preparation methods; bias, precision, percent agreement

Introduction

Age determinations provide essential input data for the stock assessment of marine fish stocks. The age-based stock assessment of redfish (*Sebastes* spp.) in the North Atlantic, however, proved to be difficult due to the lack of a sufficient amount of reliable age readings. The reliability of a set of hard body structures of the fish was addressed several times in the past. Various studies (e.g. Chilton and Beamish 1982, Nedreaas 1990) and workshops (e.g. ICES 1991, ICES 1996) have shown that the otoliths are the preferred structure for age readings on North Atlantic *Sebastes* species due to an underestimation of older ages using scales and difficulties in the interpretation of other structures such as fin rays or vertebrae. Most laboratories are still reluctant to implement routine age readings on otoliths of *Sebastes* since there are concerns about the error observed in age reading results.

Age reading error has two major elements: bias and precision. The **bias** of age readings is caused by a consistent deviation of reading results between readers and is skewed from the mean to one side or the other, while **precision** of age readings measures the closeness of repeated independent age estimates (Wilson *et al*. 1987, ICES 1996). Precision reflects the degree of agreement among readers and is not to be confused with accuracy which relates to the agreement with the true age of the fish.

Although there are routine testing systems and procedures for the assessment of bias and precision of age readings available (Kimura and Lyons 1991, ICES 1994, Campana *et al*. 1995, Hoenig *et al*. 1995), a broad-scale application of these methods in the laboratories carrying out redfish age readings is still missing. The 1995 ICES Workshop on Age Reading of *Sebastes* spp. (ICES 1996) showed a considerable bias between readers which was shown to be improved after discussion of general interpretation of growth structures on the sectioned otoliths. Thus, the need for further exchange of material and knowledge on age reading methods was stressed.

As part of the four-year research project “Population structure, reproductive strategies and demography of redfish (Genus *Sebastes*) in the Irminger Sea and adjacent waters (ICES V, XII and XIV; NAFO 1)”, funded by the European Union (QLK5-CT1999-01222), an otolith exchange between redfish age reading experts of the participating institutions is being carried out to evaluate differences in age readings between readers and otolith preparation methods. In 2000, the otolith exchange was based on *Sebastes marinus* from the Icelandic shelf (ICES Sub-area Va), collected in March 1997. *S. marinus* from ICES Va are of special interest since two year classes (probably 1985 and 1990) have shown up as distinct modes in the length distributions and could provide information for the validation of age reading results (*e.g.* Mayo *et al*. 1981, Beamish and MacFarlane 1983, Nedreaas 1990). The age reading results of this otolith exchange are presented in this study, with regard to bias and precision between readers.

Before the otoliths can be used for age reading, they have to be prepared to be able to clearly identify growth structures. While Canada, Norway and Spain were mainly using the ‘break (and burn)’ method in the past (Chilton and Beamish 1982, Nedreaas 1990, Saborido-Rey 1995, MacLellan 1997), Germany, Iceland, Russia and the USA were using (cross-)sections of otoliths (ICES 1984, Gifford and Crawford 1988). Only few comparisons have been carried out to assess the variability between both methodologies with regard to *Sebastes* species (Boehlert and Yoklavich 1984, Stanley 1986). In this study, age readings of otoliths prepared with both methods were compared for bias and precision.

Materials and Methods

The otoliths used in this study were collected from *Sebastes marinus*, caught during the Icelandic groundfish survey in 1997, onboard M/V “Brettingur NS”, one of the participating vessels. The otoliths were taken from five hauls on the Icelandic shelf (ICES sub-area Va) on 14 March 1997. 212 otolith pairs were selected for age determination, covering fish of 10-54 cm in total length. From each pair, one otolith was prepared for age reading using the ‘break-and-burn’ technique (Christensen 1964, Chilton and Beamish 1982, MacLellan 1997), while the other otolith was thin-sectioned based on the technique described by Bedford (1983).

The preparation by break-and-burn was carried out at the Marine Research Institute (MRI) in Reykjavik/Iceland. The rings in the otoliths were counted using different microscope magnification (maximum of 100x). The light was coming from above and a drip of glycerine was put on the otolith before counting the rings. The angle of the light was about 30-45 degrees. For the comparison of otolith preparation methods, the cross-sections were read by the same person at MRI who prepared and read the corresponding other otoliths by break-and-burn.

The thin-sections were performed at the Institute for Sea Fisheries of the Federal Research Centre for Fisheries in Hamburg/Germany. A diamond-covered saw blade of 0.3 mm thickness and 100 mm diameter, rotating at 6000 rotations/min, was used to cut cross-sections of 0.5 mm thickness. The cross-sections were embedded onto glass plates with translucent polyester resin and read with a magnification of 150-200x using polarised transmitted light.

At the Institute of Marine Research (IMR) in Bergen/Norway, the otolith cross-section plates were read trough transmitted light using a magnification of 50x.

For the comparison of bias and precision between readers and methods, a set of statistical tests and graphical methods were applied.

Bias estimates were based on simple linear regression analysis, the parametric paired *t*-test and the nonparametric Wilcoxon matched-pairs rank test (Conover 1998, Hollander and Wolfe 1999). The slope and intercept of simple linear regressions are tested for significant differences (α = 0.05) from 1.0 and 0, respectively. The parametric paired *t*-test and the nonparametric Wilcoxon matched pairs rank test are used to detect significant differences from a paired difference of 0. Error terms are 95% confidence limits. Age bias plots (ICES 1994, Campana *et al*. 1995, Eltink 1997) visualise the deviation of the age scores of two readers or methods from the 1:1 equivalence line and also allow the detection of non-linear bias patterns, e.g. the underestimation of ages by one reader in one part of the age range and overestimation in another part of the age range. The mean age assigned by one reader for all fish assigned a given age by the second reader is presented including the standard deviation around the mean.

Various measures for **precision** were suggested for the comparisons of age readings. One of the more common indices is the percent agreement, which compares the percentage of age determinations that are in agreement within a specified number of years. This index, however, does not evaluate the degree of precision equally for all species. If for example 95% of the age readings agree within a range of ± 1 year for cod (*Gadus morhua*), this could be a very poor precision since there just few year-classes in the fishery. For *S. marinus*, a 95% agreement within a tolerance range of ± 4 years could represent a good precision given a 50-year longevity and 20-30 age groups present in the fishery. Beamish and Fournier (1981), therefore, suggested an average percent error, which is dependent on the average age of the fish species observed. Chang (1982) modified this index to a coefficient of variation, substituting the absolute deviation by the standard deviation from the mean age. Besides these indices, the correlation coefficient r^2 is given to evaluate the fraction of variation explained by the linear relationship between readers or otolith preparation methods.

Results

Comparison of readers

Three age readers from partner institutes in Germany, Iceland and Norway were participating in the *S. marinus* otolith exchange in the year 2000 (Table 1).

As indicated by the age bias plots (Figure 1), all between-reader comparisons are subject to a certain degree of bias. In all three cases, the deviation from the 1:1 equivalence line is non-linear, most pronounced in the comparisons between readers 1 and 2, and between readers 2 and 3. In these two cases, the mean age assigned by one reader deviate considerably from the age assignments of the second reader, particularly in the age range 17-30 years. Table 2 presents the statistical tests applied to the comparison of readers in terms of bias. Regression analysis as well as the nonparametric Wilcoxon test and the parametric paired *t*-test show high significance levels, indicating bias between readers in all three comparisons. The overestimation of ages assigned by reader 2 compared to reader 1 in the older ages (deviation up to 10 years), as shown in Figure 1, results in a slope >1 and negative intercept of the linear regression. The overall bias between readers 1 and 2 is about 1 year, as indicated by the mean paired difference. Slopes of <1 and positive intercepts are present in the comparisons between readers 1 and 3 as well as between readers 2 and 3. In the latter case, an underestimation of the older ages assigned by one reader compared to the second reader was clearly detectable in Figure 1, contributing to the observed slope of <1 and positive intercept. In all three reader comparisons shown in Figure 1, a general trend in increasing standard deviations around the mean with increasing age is visible.

From the precision estimates between readers (Table 3), the correlation coefficient, the coefficient of variation and the average percent error show relatively good agreement, whereas the percent agreement value is comparatively low. As pointed out earlier, the percent agreement index does not take account of the mean age, as the average percent error does. A 24-28% agreement of age determinations between readers, as in our case, might be very low compared to values achieved for age readings of species with a shorter life-span. For redfish, however, we might have to consider a wider age range when comparing percent agreement with other species. If we increase the tolerance level of agreement between readers, as illustrated in Figure 3, a level of around 90% is reached with a tolerance of ± 3 years, considering the whole age range. In the age range 0-10 years, an agreement of 28-45% is reached with a tolerance of ± 0 years, and more than 95% agreement is reached applying ± 3 years tolerance. In contrast, the percent agreement for all pair-wise readers comparisons in the age range 21-30 years lies below 20% with a tolerance level of ± 0 years, and does not exceed 75% agreement with a tolerance of ± 5 years for the comparison of readers 1 and 2 (Figure 3). The comparison of precision indices between pairs of readers shows no major differences (Table 3), while the percent agreement plots (Figure 3) indicate a better agreement of readers 2 and 3 compared to the other two reader pairs, particularly in the age range 0-10 years.

Comparison of otolith preparation methods

The age bias plot for the comparison of otolith preparation methods (Figure 2) shows a generally slight underestimation of ages from >11 years onwards, using the break-and-burn technique. This observation is also indicated by the statistical test results given in Table 4. The slope of the linear regression is <1 with a positive intercept. The mean paired difference between both methods was around 0.8 years.

Although the regression explains about 93% of the observed variation, and the coefficient of variation and average percent error are relatively low, the percent agreement between otolith preparation methods is below 30% (Table 5). Generally, precision between methods was higher than precision between readers (Table 3). The percent agreement plot for the comparison of methods (Figure 4), however, also shows a relatively poor agreement in the age range 21-30 years.

Discussion

Comparison of readers

All between-reader comparisons in this otolith exchange showed a considerable bias, predominantly caused by an over- or underestimation of up to 10 years within the 21-30 years age range. Although maximum ages of over 40 years were reported for this species in the Northeast Atlantic (Nedreaas 1990), the observed differences in age determinations in the ages >20 years lead to relatively high error. As stated in the report of the most recent *Sebastes* age reading workshop (ICES 1996), there is further need for exchange of otoliths and knowledge on the interpretation of growth structures. Particularly the interpretation of growth zones around the nucleus and following the transition zone should be harmonised between readers. The best forum to do so would be a further age reading workshop in the frame of ICES, which is planned for 2002.

Considering age reading comparisons of previous otolith exchanges (*e.g.* Villamor 1995, Eltink 1997, Bergstad *et al*. 1998), the age-dependent precision estimates (coefficient of variation, average percent error) applied in this study indicate relatively good reproducibility. The observed CV values point to medium precision, compared with other studies on long-lived species (*e.g.* Kimura and Lyons 1991, Stevenson and Secor 1999). In contrast, the percent agreement within a tolerance of ± 0 years does not exceed 30% which is relatively poor compared to age reading results of short-living species (*e.g.* Corten 1993). Only at a tolerance of ± 3 years, around 90% agreement are reached, looking at the whole age range. 100% agreement is only reached for all readers comparisons in the age range 0-10 years with a tolerance level of ± 4 years. This represents a relatively poor fit and should sought to be improved in future. It should be noted that there were just few data available for the age range of 0-5 years to allow a conclusive evaluation of the precision of age readings in the younger ages.

Comparison of otolith preparation methods

Although the precision of the comparison between otolith preparation methods was generally higher than that for the readers comparison, a significant bias between age readings based on different otolith preparation methods was observed. Age readings based on the 'broken and burnt' otoliths showed an underestimation of age in the higher ages, relative to the results obtained from cross-section readings. The number of age determinations achieved through break-and-burn (n=105), however, represents just about half of the age readings obtained from the cross-sectioned otoliths (n=199), pointing to a better readability of otoliths prepared by the latter method.

One disadvantage of the break-and-burn method is the reading variability introduced by different angles of the light applied to the broken surface.

Age-length relationship

While a relatively high degree of bias was observed between readers and otolith preparation methods, the age-length relationship (Figure 5) obtained from the combination of readers (a) and methods (b) shows a relatively good fit of age scores with the von-Bertalanffy growth function. The derived parameters L_{inf}, k and t_0 roughly coincide with the values described by Nedreaas (1990) for *S. marinus* from the Barents Sea (L_{inf} = 49.0 cm, k = 0.06, t_0 = -2.47 years).

Future otolith exchange

The otolith exchange in 2001 will be based on *S. mentella* otoliths, collected on a German commercial vessel in the Irminger Sea in July 1999. One otolith of each pair was thin-sectioned at the Institute for Sea Fisheries in Hamburg. The other otolith will be used for radiometric age validation (*e.g.* Campana *et al*. 1990).

References

BEAMISH, R. J. and D. A. FOURNIER. 1981. A method for comparing the precision of a set of age determinations. *Can. J. Fish. Aquat. Sci*., **38**: 982-983.

BEAMISH, R. J. and G. A. McFARLANE. 1983. The forgotten requirement for age validation in fisheries biology. *Trans. Am. Fish. Soc.*, **112**: 735-743.

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.

BERGSTAD, O. A. 1998. Intercalibration of age readings of ling (*Molva molva* L.), blue ling (*Molva dipterygia* Pennant, 1784) and tusk (*Brosme brosme* L.). *ICES J. Mar. Sci*., **55**: 309-318.

BOEHLERT, G. W. and M. M. Yoklavich. 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., K. C. ZWANENBURG, and J. N. SMITH. (1990): $^{210}Pb/^{226}Ra$ determination of longevity in redfish. *Can. J. Fish. Aquat. Sci*., **47**: 163-165.

CAMPANA, S. E., ANNAND, M. C., and J. I. McMILLAN. 1995. Graphical and statistical methods for determining the consistency of age determinations. *Trans. Am. Fish. Soc.*, **124**: 131-138.

CHANG, W. Y. B. 1982. A statistical method for evaluating the reproducibility of age determination. *Can. J. Fish. Aquat. Sci*., **39**: 1208-1210.

CHILTON, D. E. and R. J. BEAMISH. 1982. Age determination methods for fishes studied by the groundfish program at the Pacific Biological Station. *Can. Spec. Pub. Fish. Aquat. Sci*., **60**: 1-102.

CHRISTENSEN, J. M. 1964. Burning otoliths, a technique for age determination of soles and other fish. *J. Cons. Int. Explor. Mer*, **29**: 73-81.

CONOVER, W. J. 1998. *Practical nonparametric statistics.* 3rd ed., John Wiley & Sons, New York, 584 pp.

CORTEN, A. MS 1993. Results of a comparative age reading experiment on herring from the North Sea and adjacent waters. *ICES C.M.* 1993/H:16, 6 pp.

ELTINK, A. MS 1997. Horse mackerel otolith exchange in 1996. *ICES C.M.* 1997/HH:24, 30 pp.

GIFFORD, V. M. and M. K. CRAWFORD. 1988. Redfish *Sebastes fasciatus*. *In*: Age determination methods for Northwest Atlantic Species. J. Penttila and L. M. Dery. (eds.), pp. 93-96.

HOENIG, J. M., M. J. MORGAN, and C. A. BROWN. 1995. Analysing differences between two age determination methods by tests of symmetry. *Can. J. Fish. Aquat. Sci*., **52**: 364-368.

HOLLANDER, M. and D. A. WOLFE. 1999. *Nonparametric Statistical Methods*. 2nd ed., John Wiley & Sons, New York, 787 pp.

ICES 1984. Report of the workshop on ageing of redfish. *ICES C.M.* 1984/G:2, 9 pp.

ICES 1991. Report of the workshop on age determination of redfish. *ICES C.M.* 1991/G:79, 18 pp.

ICES 1994. Report of the workshop on sampling strategies for age and maturity. *ICES C.M.* 1994/D:1, 67 pp.

ICES 1996. Report of the workshop on age reading of *Sebastes* spp. *ICES C.M.* 1996/G:1, 32 pp.

KIMURA, D. K. and J. J. LYONS. 1991. Between-reader bias and variability in the age-determination process. *Fish. Bull*., **89**: 53-60.

MACLELLAN, S. E. 1997. How to age rockfish (*Sebastes*) using *S. alutus* as an example - the otolith burnt section technique. *Can. Tech. Rep. Fish. Aquat. Sci*., **2146**: 1-42.

MAYO, R. K., V. M. GIFFORD, and A. JEARLD Jr. 1981. Age validation of redfish, *Sebastes marinus* (L.), from the Gulf of Maine-Georges Bank region. *J. Northw. Atl. Fish. Sci*., **2**: 13-19.

NEDREAAS, K. 1990. Age determination of Northeast Atlantic *Sebastes* species. *J. Cons. Int. Explor. Mer*, **47**: 208-230.

SABORIDO-REY, F. MS 1995. Age and growth of redfish in Flemish Cap (Div. 3M). *NAFO SCR Doc.*, No.31. Serial No. N2540, 16 p.

STANLEY, R. D. 1986. A comparison of age estimates derived from the surface and cross-section methods of otolith reading for Pacific ocean perch (*Sebastes alutus*). *Proceedings of the International Rockfish Symposium*, Oct. 1986, Anchorage, Alaska, pp. 187-196.

STEVENSON, J. T. and D. H. SECOR. 1999. Age determination and growth of Hudson River Atlantic sturgeon, *Acipenser oxyrinchus*. *Fish. Bull*., **97**: 153-166.

VILLAMOR, B. and M. MEIXIDE. MS 1995. Results of the mackerel (*Scomber scombrus*, L.) otolith exchange programme in 1994. *ICES C.M.* 1995/H:12, 30 pp.

WILSON, C. W., R. J. BEAMISH, K. D. BROTHERS, K. D. CARLANDER, J. M. CASSELMAN, J. M. DEAN, A. JEARLD Jr., E. D. PRINCE, and A. WILD. 1987. Glossary. *In*: Age and growth of fish. Summerfelt, R.C. and G.E. Hall, G.E. (eds.), Iowa State University Press, Ames, Iowa, pp. 527-530.

Table 1. Participating age readers.

Reader ID	Name	Institute
1	Christoph Stransky	Federal Research Centre for Fisheries, Institute for Sea Fisheries, Hamburg, Germany
2	Sif Guðmundsdóttir	Marine Research Institute, Reykjavík, Iceland
3	Svend Lemvig	Institute of Marine Research, Bergen, Norway

Table 2. Statistical tests for the detection of **bias** for age readings of *S. marinus* between **readers**.

	Age reader		
Statistic	Reader 1 versus Reader 2 (N = 199)	Reader 1 versus Reader 3 (N = 212)	Reader 2 versus Reader 3 (N = 199)
		Regression	
Slope	1.157 ± 0.038	0.950 ± 0.034	0.783 ± 0.024
P	0.000	0.000	0.000
Intercept	-0.993 ± 0.510	1.221 ± 0.460	2.615 ± 0.357
P	0.053	0.009	0.000
		Wilcoxon test	
P	0.000	0.000	0.012
		Paired *t*-test	
Mean paired difference	-1.005 ± 0.320	-0.585 ± 0.274	-0.372 ± 0.296
P	0.000	0.000	0.014

Table 3. Measures of **precision** for age readings on *S. marinus* between **readers**.

	Age reader		
Statistic or index	Reader 1 versus Reader 2 (N = 199)	Reader 1 versus Reader 3 (N = 212)	Reader 2 versus Reader 3 (N = 199)
Correlation coefficient (r^2)	0.824	0.787	0.840
Coefficient of variation (%)[a]	8.79	8.19	7.66
Average percent error[b]	6.21	6.17	5.42
Percent agreement	24.12	25.00	27.64

[a] from Chang (1982)
[b] from Beamish and Fournier (1981)

Table 4. Statistical tests for the detection of **bias** for age readings of *S. marinus* between **otolith preparation methods**.

Statistic	Cross-sections versus break-and-burn (N = 105)
Regression	
Slope	0.877 ± 0.024
P	0.000
Intercept	0.987 ± 0.368
P	0.009
Wilcoxon test	
P	0.000
Paired *t*-test	
Mean paired difference	-0.771 ± 0.315
P	0.000

Table 5. Measures of **precision** for age readings on *S. marinus* between **otolith preparation methods**.

Statistic or index	Cross-sections versus break-and-burn (N = 105)
Correlation coefficient (r^2)	0.930
Coefficient of variation (%)[a]	6.69
Average percent error[b]	2.49
Percent agreement	28.57

[a] from Chang (1982)
[b] from Beamish and Fournier (1981)

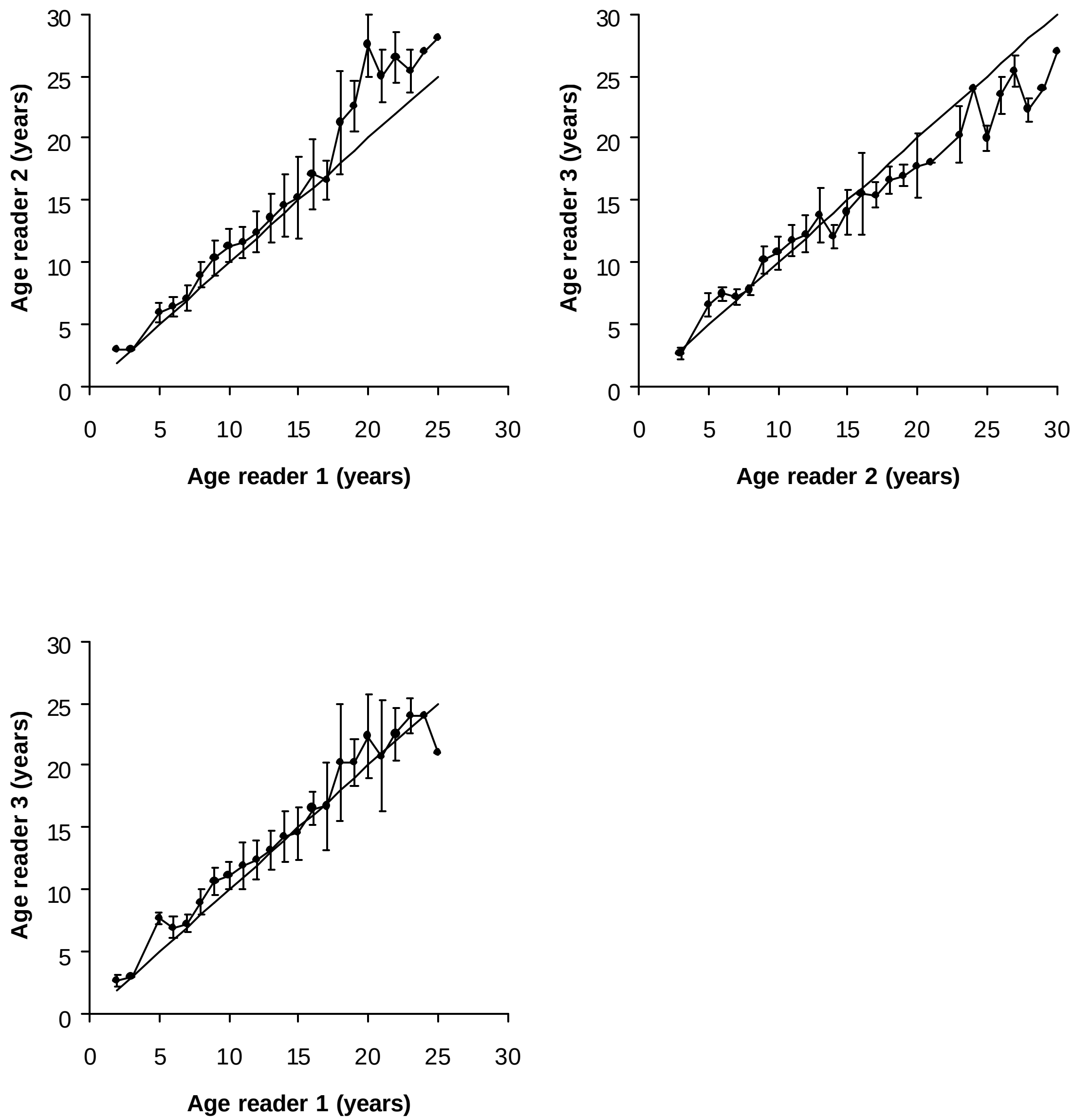


Fig. 1. Age bias plots for the **reader** comparisons given in Table 2. Each error bar represents the standard deviation around the mean age assigned by one reader for all fish assigned a given age by the second reader. The 1:1 equivalence (straight line) is also indicated.

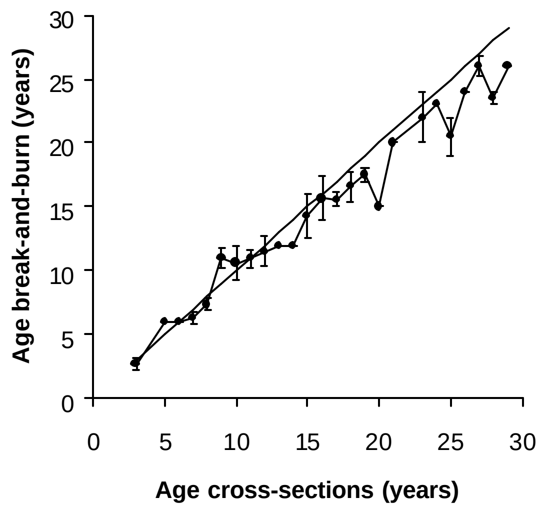


Fig. 2. Age bias plot for the comparison of **otolith preparation methods** given in Table 4. Each error bar represents the standard deviation around the mean age assigned by one reader for all fish assigned a given age by the second reader. The 1:1 equivalence (straight line) is also indicated.

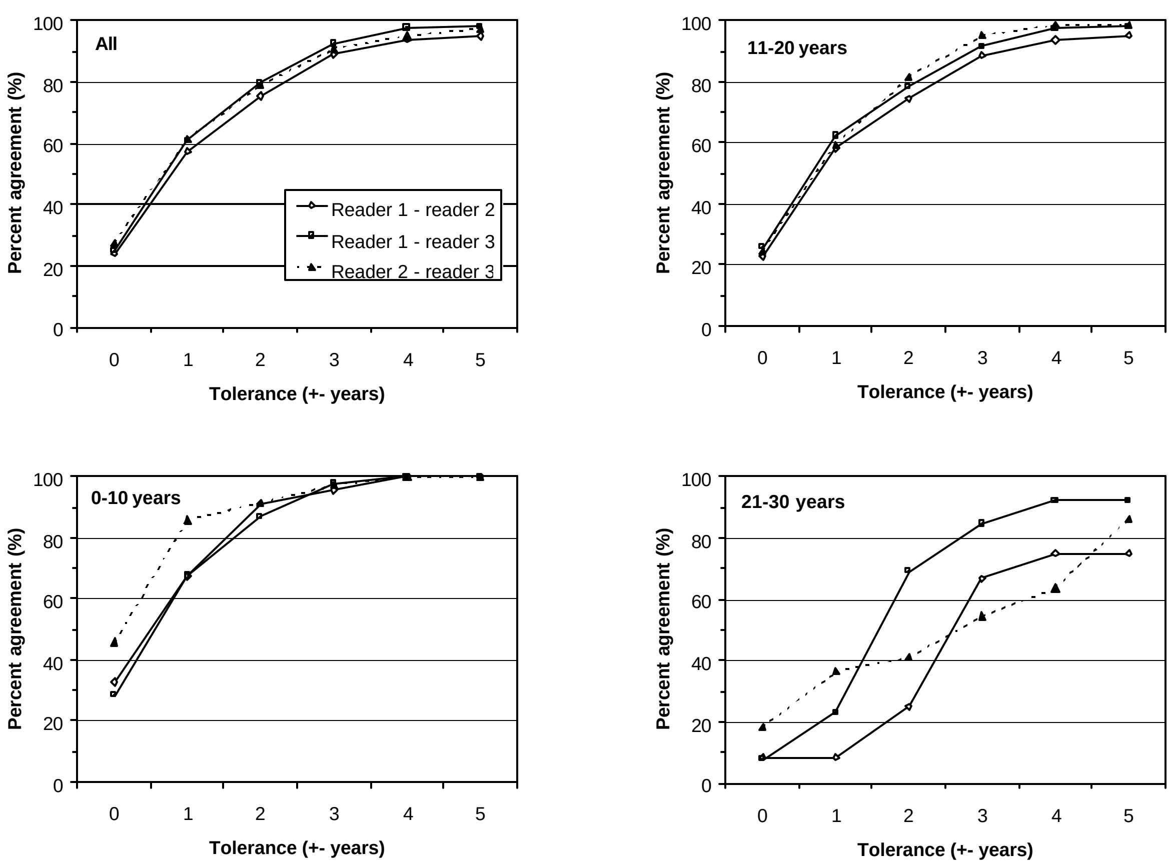


Fig. 3. Percent agreement for the **reader** comparisons given in Table 3 for a tolerance level (deviation of assigned ages between both readers) of ± 0 (total agreement) to ± 5 years, applied to all age groups and sub-sets of age ranges assigned by the first reader.

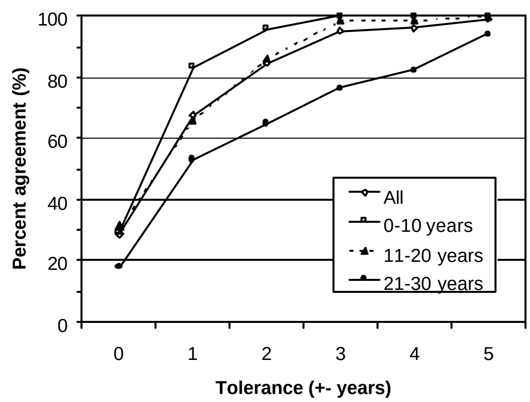
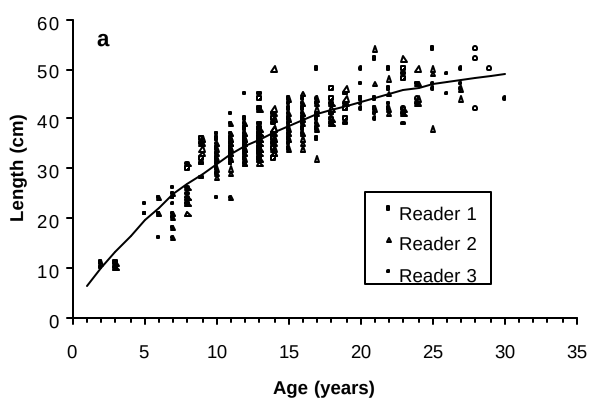


Fig. 4. Percent agreement for the comparison of **otolith preparation methods** given in Table 5 for a tolerance level (deviation of assigned ages between both readers) of ± 0 (total agreement) to ± 5 years, applied to all age groups and sub-sets of age ranges assigned by the first reader.



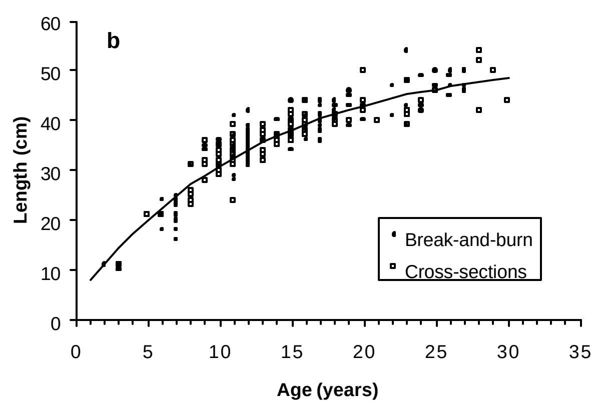


Fig. 5. Age-length relationship and fitted von-Bertalanffy growth curves of the comparison of readers (a) and otolith preparation methods (b). The fitted growth parameters are: (a) L_{inf} = 53.258 cm, k = 0.083, t_0 = -0.499 years and (b) L_{inf} = 53.239 cm, k = 0.079, t_0 = -1.047 years, respectively.