CSIRO PUBLISHING

www.publish.csiro.au/journals/mfr *Marine and Freshwater Research*, 2005, **56**, 735–743

Use of otolith weight in length-mediated estimation of proportions at age

R. I. C. Chris Francis[A,E], *Shelton J. Harley*[B], *Steven E. Campana*[C] *and Peer Doering-Arjes*[D]

[A]National Institute of Water and Atmospheric Research, Private Bag 14901, Wellington, New Zealand.
[B]Inter-American Tropical Tuna Commission, 8604 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA.
[C]Marine Fish Division, Bedford Institute of Oceanography, P.O. Box 1006, Dartmouth, Nova Scotia, Canada B2Y 4A2.
[D]Institute of Freshwater Research, Stångholmsvägen 2, SE-178 93 Drottningholm, Sweden.
[E]Corresponding author. Email: c.francis@niwa.co.nz

Abstract. Each year almost a million fish are aged from otoliths, primarily to estimate proportions at age for use in stock assessments. The preparation and reading of otoliths is time-consuming and thus expensive. Two techniques have been proposed to reduce costs. The first is length-mediated estimation, in which the length distribution from a large sample of fish is converted to an age distribution, using information (usually in the form of an age–length key) from a smaller sample containing length and age data. The second is to infer age from otolith weight (and/or other otolith measurements). These two cost-saving ideas are combined in a new method, length-mediated mixture analysis. It requires three samples – one with lengths only, one with lengths and otolith measurements, and one with lengths, otolith measurements and ages – and estimation is by maximum likelihood. The use of this method, which can be thought of as a generalisation of three established methods of age inference, is illustrated in two simulation experiments in a cost-benefit framework.

Introduction

Estimates of proportions at age, in either commercial or research catches, are an important and widely used input to stock assessments for fisheries in many parts of the world. It is primarily to obtain these estimates that almost a million fish are aged each year using otoliths (Campana and Thorrold 2001). The estimates are generated in one of two ways. The simplest is the direct method, in which a random sample of fish is aged, and the proportions at age in the sample are taken as estimates of those in the population being sampled. This is an expensive method because many otoliths are required to achieve acceptable precision and the collection, preparation and reading of otoliths is costly.

An alternative, and less costly, method is what we call length-mediated estimation. This requires measurements of fish length in a large sample, and ages and lengths from a small sample. The length distribution from the large sample is then converted to an age distribution using information from the small sample. Several length-mediated methods have been proposed. The first, and most popular, uses an age–length key (a matrix containing estimated proportions at age for each length class) constructed from the small sample to make the conversion from length distribution to age distribution (Fridriksson 1934). We call this the ALK (age–length key) method. It requires that the length distribution of the large sample be representative of the catch (which is achieved either by taking a simple random (SR) sample from the whole catch, or the weighted sum of simple random samples from various components of the catch). The requirement for the small sample is less stringent. Here, all that is necessary is that all fish of the same length should have the same probability of being selected. This requirement is, of course, met by SR sampling. However, it is more common to stratify the small sample by length, which we call length-stratified (LS) sampling. Ketchen (1950) proposed doing this by taking the same number of fish from each length class (or, for the extreme length classes where this number of fish is not available, as many as possible). This approach seems to be widely used, though some studies have attempted more optimal length-stratified designs (e.g. Baird 1983; Lai 1987, 1993; Horppila and Peltonen 1992; Oeberst 2000).

Other length-mediated methods of estimation have been proposed. Smith and Sedransk (1982) described a Bayesian method, which is equivalent to the ALK method if uninformative prior distributions are assumed. This requires no additional assumptions except for a specific form for the prior distributions (Dirichlet). Martin and Cook (1990) proposed a maximum-likelihood method based on the assumption that

 10.1071/MF04127 1323-1650/05/050735

length is normally distributed for each age. There is also an extensive literature on length-mediated estimation when there are no age data (e.g. see MacDonald and Green 1988; Fournier *et al*. 1990, and references therein).

Another approach to reducing costs in otolith-based age estimation was suggested by Boehlert (1985), who proposed estimating age from otolith measurements using multiple regression. The costs associated with measuring an otolith are much less than those for obtaining annulus counts (which often involve a substantial otolith preparation cost) and some otolith measurements (particularly weight) are highly correlated with fish age in many species (e.g. Boehlert 1985; Radtke and Hourigan 1990; Fletcher 1991; Anderson *et al*. 1992; Ferreira and Russ 1994; Worthington *et al*. 1995; Cardinale *et al*. 2000; Labropoulou and Papaconstantinou 2000; Luckhurst *et al*. 2000). Two samples are used: a small calibration sample, for which annulus counts and otolith measurements are recorded; and a large production sample, for which only otolith measurements are recorded. The calibration sample is used to calculate the regression relationship, and then this is applied to assign an age to each fish in the production sample. Subsequent work in this area has shown not only that weight is the most promising of otolith measurements for age inference, but that there are several alternatives to Boehlert's method of inference (see review by Francis and Campana 2004).

In the present study, we describe and investigate a new method of combining these two cost-saving ideas. We will call this length-mediated mixture analysis (LMMA). Its aim is to estimate proportions at age using three types of data: fish lengths, otolith annulus counts and otolith measurements (e.g. weight). Like the ALK method, it is length-mediated and involves measuring the lengths of many fish, but taking otoliths from relatively few. As with the method of Boehlert (1985), it requires measuring all otoliths, but counting annuli in only some of them. However, it differs from these methods in not using an age–length key (or anything similar), not assigning ages to any individual fish (other than those with annulus counts) and in requiring more statistical assumptions. The new method uses three samples: one with lengths only, one with lengths and otolith measurements and one with lengths, otolith measurements and annulus counts.

We describe LMMA, show that it is a generalisation of several established methods of age inference, and then illustrate its use with two simulation experiments based on data from known-age Faroese cod. These experiments show that, for this species, the use of otolith weight in estimating proportions at age is cost effective, but provides a much greater cost saving with direct estimation than it does with length-mediated estimation.

The length-mediated mixture-analysis (LMMA) method

As mentioned above, the aim of the method is to estimate proportions at age, p_A, in some population, and it uses three samples whose characteristics are described in Table 1. The *LOA* sample contains lengths, otolith measurements and ages; the *LO* sample contains lengths and otolith measurements and the *L* sample contains just lengths. The last sample must be SR; the other two samples may be either SR or LS. Some data types (and some samples) are optional (see next section). We need to make some assumptions about the distribution, for a given age A, of the vector $\mathbf{X}$ (containing length and otolith measurement(s)). The simplest assumption, which we will use in our simulation experiment below, is that this distribution is multivariate normal. More generally, we assume that for any given age, A, the distribution of $\mathbf{X}$ is described by the (known) density function $\mathrm{g}(\mathbf{X}; \boldsymbol{\theta}_A)$ for some unknown vector of parameters, $\boldsymbol{\theta}_A$. In other words, the distribution of $\mathbf{X}$ is a finite mixture with mixture proportions p_A. When g is multivariate normal, $\boldsymbol{\theta}_A = (\mathbf{M}_A, \mathbf{V}_A)$, where $\mathbf{M}_A$ is the vector of mean length and mean otolith measurements at age A, $\mathbf{V}_A$ is the associated covariance matrix, and

$$\mathrm{g}(\mathbf{X}; \boldsymbol{\theta}_A) = |\mathbf{V}_A|^{-0.5} \exp[(\mathbf{X} - \mathbf{M}_A)' \mathbf{V}_A^{-1} (\mathbf{X} - \mathbf{M}_A)] \quad (1)$$

With these assumptions we can write the log-likelihood for each sample as

$$\lambda_{LOA} = \sum\nolimits_i \log\left[\frac{p_{A_i}\, \mathrm{g}(\mathbf{X}_i; \boldsymbol{\theta}_{A_i})}{\mathrm{f}(L_i)}\right] \quad (2)$$

$$\lambda_{LO} = \sum\nolimits_j \log\left[\frac{\sum_A p_A\, \mathrm{g}\,(\mathbf{X}_j; \boldsymbol{\theta}_A)}{\mathrm{f}(L_j)}\right] \quad (3)$$

Table 1. The three samples used by LMMA (length-mediated mixture analysis)

Sample name	Sample type	Data	Notation
LOA	SR or LS	(Lengths), (otolith measurement(s)), ages	$\mathbf{X}_i, A_i; i = 1, \ldots, N_{LOA}$
LO	SR or LS	(Lengths), otolith measurement(s)	$\mathbf{X}_j; j = 1, \ldots, N_{LO}$
L	SR	Lengths	$L_k; k = 1, \ldots, N_L$

The sample types are simple random (SR) and length-stratified (LS). *A*, age; *L*, fish length; **O**, a vector of one or more otolith measurements; **X**, the vector (*L*, **O**) combining length and otolith measurements. Data types in parentheses are optional for that sample.

$$\lambda_L = \sum\nolimits_k \log\left[\sum\nolimits_A p_A\, \mathrm{h}(L_k; \boldsymbol{\theta}_A)\right] \tag{4}$$

where $\mathrm{h}(L; \boldsymbol{\theta}_A)$ is the marginal distribution of L and

$$\mathrm{f}(L) = \begin{cases} \sum_A p_A\, \mathrm{h}(L; \boldsymbol{\theta}_A) & \text{if the sample is LS} \\ 1 & \text{if the sample is SR} \end{cases} \tag{5}$$

Estimation is by maximum likelihood. That is, our estimates of the parameters $\{p_A, \boldsymbol{\theta}_A\}$ are those that maximise the total log-likelihood λ, given by $\lambda = \lambda_{LOA} + \lambda_{LO} + \lambda_L$. Note that our interest is just in the proportions at age, p_A, but we need also to estimate the other parameters, $\boldsymbol{\theta}_A$.

The above equations, and the simulation experiment below, assume, for simplicity, that the *LOA*, *LO* and *L* samples are quite separate. In practice it would often be convenient to nest these samples. Otoliths would be collected and measured from a subsample of the *L* sample. Some of these otoliths would then be aged (the *LOA* sample) and some would not be (the *LO* sample). To modify the likelihoods for nested samples with SR sampling is straightforward. All that is needed is to restrict the summation in Eqn (4) to those fish in the *L* sample that are not also in the *LO* or *LOA* samples. Modification for nested LS samples is more complicated and would be worthwhile only in the unusual situation in which the distribution of fish lengths within each length stratum contains information about the model parameters.

Relationship to other methods of inferring age

There are three published methods of inferring age that may be seen as special cases of LMMA. We create special cases by imposing additional assumptions or restrictions. One way to do this is to drop one, or even two, of our three samples. We can do this without violating any statistical principles; all that happens is that the corresponding term drops out of the log-likelihood. Another possible restriction is to drop either the otolith measurements or the fish lengths from the *LOA* and *LO* samples. That is, we set $\mathbf{X} = L$ or $\mathbf{X} = \mathbf{O}$, rather than $\mathbf{X} = (L, \mathbf{O})$.

The method of Martin and Cook (1990) is essentially the same as LMMA if we make the following restrictions: omit the otolith data (i.e. set $\mathbf{X} = L$); drop the *LO* sample; assume the *LOA* sample is LS; and assume that our density g is normal. The only remaining difference is that Martin and Cook (1990) assumed that their length observations were binned into length classes, which gives a different form to their log-likelihood. This difference is of no practical significance (as long as we can assume that the binning is not so coarse as to allow only a few length bins per age class – such coarse binning would be unlikely to occur in practice because it would restrict our ability to estimate proportions at age).

It is the *L* sample that makes our method length-mediated. If we drop this we move into the realm of direct methods of estimating proportions at age. If we also treat the length measurements in these samples as optional, then LMMA becomes the same as the mixture-analysis method of Francis and Campana (2004) (what they named the calibration and production samples are our *LOA* and *LO* samples, respectively).

Dropping both the *LOA* and *LO* samples and assuming normality produces the modal-analysis method similar to that implemented in the computer programme MIX (MacDonald and Green 1988).

Two simulation experiments

We illustrate the LMMA method with two simulation experiments based on a set of known-age Faroese cod data (Cardinale *et al*. 2004). Our aim in these experiments is to answer the following question for this cod stock: 'Is it worthwhile to use otolith weight when estimating proportions at age?'. A graphical examination of the data suggests that the answer may be 'Yes' because, for example, there is much less overlap between fish of ages 3 and 4 in the bivariate plot than in the length distributions (Fig. 1) . A sensible way of addressing this question is via a cost-benefit analysis (Francis and Campana 2004). Given a fixed sampling cost, we want to know whether we get more accurate estimates of proportions at age when we include otolith weights than when we exclude them. We did two experiments: one for length-mediated estimation, using LMMA; and the other for direct estimation, using the original mixture-analysis (MA) method of Francis and Campana (2004). We will see that these experiments produce quite different results.

Each experiment consisted of a series of scenarios, each of which was defined by: (*i*) the true proportions at age, p_A; (*ii*) the sample sizes, N_{LOA}, N_{LO}, N_L; (*iii*) what sampling method (SR or LS) was used in the *LOA* and *LO* samples; (*iv*) what measurements were used in the *LOA* and *LO* samples (length only, length and otolith weight or otolith weight only); and (*v*) which estimation method was used. For each scenario, 500 datasets were simulated, and estimates of the p_A were obtained for each dataset. Our performance measure for each scenario was based on the root-mean-square error (rmse). For a single age, A, the rmse associated with p_A is given by $\mathrm{rmse}_A = \left[(1/500)\sum_l (\hat{p}_{Al} - p_{A0})^2\right]^{0.5}$ where p_{A0} is the true proportion at age A and $\hat{p}_{Al}$ is an estimate of it from the lth dataset. Our performance measure was the 'total' rmse across ages, denoted as rmsetot (in analogy to the similar performance measure Vartot used by Kimura 1977) and defined as $\left[\sum_A \mathrm{rmse}_A^2\right]^{0.5}$. The smaller rmsetot is, the more accurately we have estimated the proportions at age. An approximate 95% confidence interval was calculated for rmsetot by bootstrap resampling of the $\hat{p}_{Al}$.

When comparing two scenarios, we will say that the estimation of proportions at age is better in the second than in the first if $\mathrm{rmsetot}_2 < \mathrm{rmsetot}_1$ and will define the percentage gain in efficiency of estimation in the second, compared

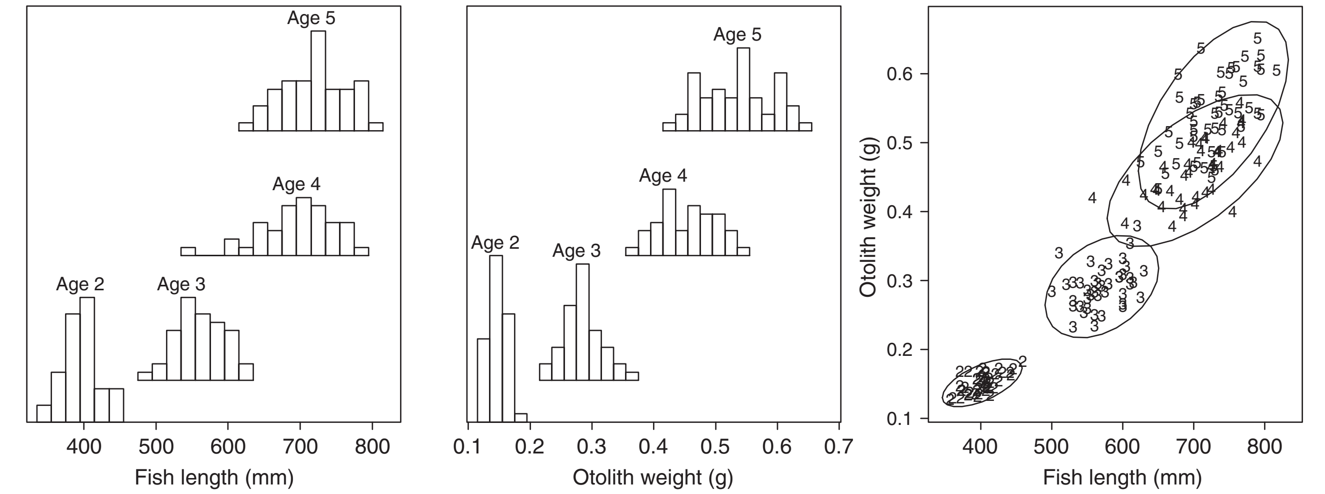

Fig. 1. The data from known-age Faroese cod that were used in the simulation experiments. In the right-hand panel, each point represents one fish (the plotting symbol identifies its age) and the ellipses are approximate 95% confidence regions (calculated by fitting bivariate normal distributions to the data for each age).

to the first, as $100(\text{rmsetot}_1^2 - \text{rmsetot}_2^2)/\text{rmsetot}_2^2$ (this formula is analogous to the common statistical practice of describing relative efficiency as a ratio of variances). A qualitative measure of how confident we can be that estimation is better in one scenario than in another is given by the degree of overlap between the 95% confidence intervals for rmsetot. The less overlap there is, the more confident we can be. To obtain a quantitative measure, we defined $\text{rmsetot}_l = \left[\sum_A (\hat{p}_{Al} - p_{A0})^2\right]^{0.5}$ to be the component of rmsetot associated with the lth simulated dataset and made a pairwise comparison of the rmsetot_l for the two scenarios. We counted how many datasets there were for which rmsetot_l was smaller for one scenario than for the other and determined significance using a two-sided sign test (Dixon and Massey 1969).

The 2-year old fish in our sample were not used in the simulations because they are so clearly separated (in both length and otolith weight) from older fish (Fig. 1) that they would serve no purpose in our experiment. In all scenarios we arbitrarily assumed that the true proportions at ages 3 to 5 were 0.2, 0.33 and 0.47, respectively.

The following procedure was used to simulate data, with SR sampling, from a catch in which the true age frequency is p_A.

(1) Define the weightings $w_s = p_{A_s}/m_{A_s}$, where A_s is the age of the sth fish and m_A is the number of fish of age A, in the cod dataset ($s = 1, \ldots, 131$).
(2) Simulate the *LOA* sample by selecting N_{LOA} fish at random (with replacement) from the dataset, where the probability of picking the sth fish is w_s, and recording fish length, otolith weight and age.
(3) Simulate the *LO* sample by selecting N_{LO} fish at random (with replacement) from the dataset, where the probability of picking the sth fish is w_s, and recording fish length and otolith weight.
(4) Simulate the *L* sample by selecting N_L fish at random (with replacement) from the dataset, where the probability of picking the sth fish is w_s, and recording fish length.

Table 2. Sampling costs used in the simulation experiments
For each estimation type the first cost covers travel and initial data entry; subsequent costs are additional to this; costs for annulus counts include otolith preparation

Estimation type	Data collected	Cost
Length-mediated estimation	250 fish lengths	\$68.64
	Additional lengths	\$0.083/fish
	Otoliths	\$0.416/otolith
	Otolith weights	\$0.08/otolith
	Annulus counts	\$2.14/otolith
Direct estimation	200 otoliths	\$131.00
	Additional otoliths	\$0.416/otolith
	Length measurements	\$0.083/fish
	Otolith weights	\$0.08/otolith
	Annulus counts	\$2.14/otolith

For the simulation of LS samples, six broad length classes of equal width were defined and equal numbers of fish were selected in each of these (the length-class boundaries were at 500, 553, 606, 659, 712, 765 and 818 mm).

To be able to calculate and compare the cost of collecting samples of varying sizes containing different measurements we need a detailed breakdown of sampling costs. The costs used in these experiments (Table 2) were based on experience in sampling cod landings in eastern Canada (S.E. Campana, unpublished data).

Table 3. Description of scenarios evaluated in the two simulation experiments
Each experiment evaluated two parallel sets of scenarios and the sampling costs were designed to be the same for all scenarios within an experiment

	Scenario number							
	0	1	2	3	4	5	6	7
(*a*) Length-mediated estimation (scenarios A0–A7 and B0–B7; sampling cost = \$601)								
N_{LOA}	200	200	194	180	170	160	140	120
N_{LO}	0	0	0	74	127	180	287	393
N_L	500	500	500	500	500	500	500	500
Estimation method	ALK	LMMA	LMMA	LMMA	LMMA	LMMA	LMMA	LMMA
Otolith weight?	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Fish length?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sampling								
A scenarios	LS	LS	LS	LS	LS	LS	LS	LS
B scenarios	SR	SR	SR	SR	SR	SR	SR	SR
(*b*) Direct estimation (scenarios C0–C6 and D1–D6; sampling cost = \$559)								
N_{LOA}	200	180	160	150	140	120	100	
N_{LO}								
C scenarios	0	38	132	178	225	319	413	
D scenarios	–*	74	180	233	287	393	499	
N_L	0	0	0	0	0	0	0	
Estimation method	Simp	MA	MA	MA	MA	MA	MA	
Otolith weight?	No	Yes	Yes	Yes	Yes	Yes	Yes	
Fish length?								
C scenarios	No	Yes	Yes	Yes	Yes	Yes	Yes	
D scenarios	–*	No	No	No	No	No	No	
Sampling	SR	SR	SR	SR	SR	SR	SR	

The estimation methods are ALK = age–length key method, MA = mixture-analysis method (Francis and Campana 2004), LMMA = length-mediated mixture analysis, Simp = simple estimation from sample proportions; the sampling methods are SR = simple random sampling, LS = length-stratified (these methods were specified for the *LOA* and *LO* samples only, the *L* sample was always SR). *There is no scenario D0.

Experiment using length-mediated estimation

In this experiment, two parallel sets of scenarios were investigated and these two sets differed only in the sampling strategy that was used for the *LOA* and *LO* samples (Table 3*a*). Scenarios A0–A7 used LS sampling, as is common when the ALK method is used; SR sampling was used in the other set of scenarios (B0–B7). Each set contained two baseline scenarios. The first (scenarios A0 and B0) used the ALK method with 500 fish lengths ($N_L = 500$) and 200 otoliths ($N_{LOA} = 200$). No otolith weights were used for this scenario and $N_{LO} = 0$. Exactly the same data were used for the second baseline scenario (A1 and B1), but the data were analysed using LMMA. The remaining scenarios used LMMA and all three samples. N_L was fixed at 500, a range of values was chosen for N_{LOA}, and N_{LO} was set so that the sampling cost was (almost) exactly the same for each scenario.

We illustrate the cost calculations for these scenarios with two examples. In all scenarios the cost of the *L* sample is \$89.39 (\$68.64 for the first 250 fish and 250 × \$0.083 for the remaining fish). For scenario A0 (or A1, B0 or B1), we have to collect and age 200 otoliths for the N_{LOA} sample, which costs 200 × (\$0.416 + \$2.14) = \$511.20, so the total cost is \$600.59 (= \$89.39 + \$511.20). For scenario A3 (or B3), it costs \$474.48 [= 180 × (\$0.416 + \$0.08 + \$2.14)] to collect, measure, and age 180 otoliths for the *LOA* sample and \$36.704 [= 74 × (\$0.416 + \$0.08)] for the *LO* sample, so the total cost is \$600.57 (= \$89.39 + \$474.48 + \$36.704). We assume in all these scenarios that there are no additional costs for length measurements for the *LOA* and *LO* samples because these are subsamples of the *L* sample.

We can draw three conclusions from the results of this experiment (Fig. 2*a*). First, SR sampling is better than LS (because rmsetot is lower for each B scenario than for the corresponding A scenario). This is not surprising when LMMA is used, because with SR sampling the proportions at age in the *LOA* sample provide direct additional information about the proportions at age in the population. It is not immediately obvious that SR should be better than LS for the ALK method, because this method has no way of using the additional information (the estimation algorithm is the same for SR and LS). However, the present results for the ALK method are consistent with those of Kimura (1977) in suggesting that estimation is more accurate with SR than with LS. Our second conclusion supports the finding of Martin and Cook (1990)

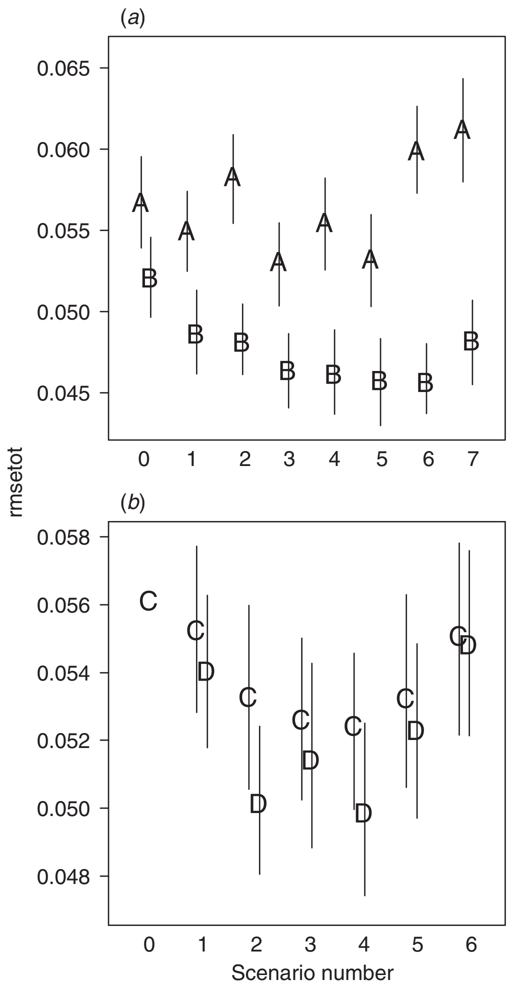

Fig. 2. Results of the simulation experiments for (*a*) length-mediated estimation, and (*b*) direct estimation, showing the estimated performance measure, rmsetot, for each scenario (with vertical bars representing approximate 95% confidence intervals). Individual scenarios are identified by the plotting symbol and horizontal position (e.g. the symbol 'A' above scenario number 6 represents the results from scenario A6).

that the ALK method is inferior to LMMA when applied to the same data (because rmsetot is lower for A1 than for A0, and lower for B1 than for B0). This result was significant for SR sampling ($P = 0.018$), but not for LS sampling.

Our third conclusion is that there would be a slight advantage in using otolith weight for length-mediated estimation of proportions at age for this stock, particularly if SR sampling is to be used. Use of these data provide a 7% increase in efficiency with LS sampling (comparing A5 with A1, $P = 0.02$) and a 13% increase with SR sampling (comparing B5 with B1, $P = 0.002$).

Experiment using direct estimation

In this experiment, $N_L = 0$ for all scenarios (because estimation was direct, rather than length-mediated) and there was just one baseline scenario, labelled C0 (Table 3*b*), in which $N_{LOA} = 200$, $N_{LO} = 0$, sampling was SR and no fish length or otolith weight data were used. In this scenario, the estimated proportions at age in the population are simply those in the *LOA* sample. There was no need to simulate data for scenario C0 because we know from multinomial sampling theory that rmsetot is exactly equal to $[p_3(1-p_3)+p_4(1-p_4)+p_5(1-p_5)]^{0.5}/N_{LOA}^{0.5}$ (Stuart and Ord 1987), so we can simply calculate rmsetot for any sample size, N_{LOA}, rather than having to estimate it via a simulation experiment. We investigated two parallel sets of scenarios that differed only in whether fish length data were used (they were used in scenarios C1–C6, but not in D1–D6). These scenarios used MA estimation. As in the first experiment we picked a range of values for N_{LOA} and then set N_{LO} so that the sampling cost was (almost) exactly the same for each scenario.

Table 4. Estimates of rmsetot from three scenarios differing only in sample type

For all scenarios the sample sizes and composition were exactly as for C4 (i.e. $N_{LOA} = 140$, $N_{LO} = 220$, $N_L = 0$, estimation method = MA, samples contain both otolith weight and fish length)

Scenario	*LOA* sample	*LO* sample	rmsetot (95% confidence interval)
C4	SR	SR	0.0533 (0.0506, 0.0563)
C4.1	LS	SR	0.0637 (0.0605, 0.0665)
C4.2	SR	LS	0.0611 (0.0583, 0.0639)

Sample types are simple random (SR) and length-stratified (LS).

We can draw two conclusions from the results of this experiment (Fig. 2*b*). First, when otolith weight data are used it is slightly better not to measure fish lengths (because rmsetot is slightly less for each D scenario than for the corresponding C scenario). Thus the loss of information when fish lengths are not collected is more than compensated for by the gain from the bigger sample sizes that are possible for the same cost. The second conclusion is that there is a clear advantage in using otolith weight for direct estimation of proportions at age for this stock. The MA method using otolith weight (but not fish length) is 27% more efficient than the simple approach using annulus counts alone (comparing D4 with C0). To achieve the same rmsetot value as in scenario D4 with annulus counts alone (i.e. using the simple estimation method) would require $N_{LOA} = 253$, which would cost 24% more than scenario D4.

We also considered the effect of using LS sampling with direct age estimation. With this type of estimation there is less scope for LS sampling. For example, there would be no point in using LS sampling with the simple method because the data would then contain no information about the proportions at age in the population. The same would be true with the MA method if both the *LOA* and *LO* samples were LS. Also, we cannot construct a LS sample without measuring fish length, so we need not consider LS sampling in the absence of fish length data. However, LS sampling could be applied to just one of the two samples when the MA method was used, as long as fish length data were collected. Two additional scenarios, which were the same as C4 except for the sampling strategy, showed that this reduces efficiency (Table 4). Applying LS sampling to one sample reduces efficiency by about 30% in this case.

Discussion

We have shown how otolith measurements can be used in length-mediated estimation of proportions at age. Under the conditions considered in our first simulation experiment, this use was shown to be cost-effective. However, the increase in efficiency from using otolith weight was relatively small: 7% (for LS sampling) or 13% (for SR sampling). A much greater increase in efficiency (27%) was obtained from otolith weight when estimation was direct, rather than length-mediated.

The second simulation experiment may seem out of place in the present study. Our title refers to length-mediated estimation but this experiment concerns direct estimation. We included both experiments to illustrate a point that became clear to us only part way through this study. Both are intended to address the question 'Is otolith weight useful in estimating proportions at age?'. However, the meaning of this question changes when we change our method of estimation. With direct estimation we are concerned with how well we can estimate the age of a fish from the weight of its otolith. With length-mediated estimation the question is much more complex. We are concerned with correlations between three quantities: otolith weight, fish length and age. The relevant question seems to be 'Can otolith weight help us to refine our knowledge of the relationship between length and age so that our conversion from a length distribution to an age distribution is more accurate?'. Once we see that our two experiments are asking quite different questions it is less surprising that the answers they produce (in terms of the gain in efficiency from using otolith weight) are so different.

It would be quite wrong to generalise from the results of our very limited experiments, which are intended to be illustrative rather than definitive. In other species, with sampling costs different from ours, there may or may not be a gain in efficiency from using otolith weight (or some other otolith measurement), and if there is a gain, it may be smaller or larger than found here. Even with the same species and sampling costs, the gain in efficiency is likely to vary depending on the true proportions at age in the catch, the cost allocated to sampling and how that cost is structured. It is recommended that anyone contemplating using LMMA should carry out simulation experiments similar to those given here, but using data for their own species and appropriate sampling costs and covering a range of different values of proportions at age. We chose to carry out our cost-benefit analysis by comparing the accuracy of different methods using same-cost samples; other researchers may prefer to compare the costs required to achieve a target level of accuracy. In any set of sampling costs (like those in Table 2) it is the relative costs of the various parts of the sample that are important, rather than the absolute costs. We note that there is no clear best performance measure for an estimator of a vector of proportions at age. The one we have used, rmsetot, seems reasonable, but others are possible. Rmsetot is based on absolute errors, so, for example, an error of 0.01 in an estimate of p_A has the same impact whether the true value of p_A is 0.3 or 0.03. A possible alternative measure is $\left[(1/500)\sum_{A,l}(\hat{p}_{Al} - p_{A0})^2/p_{A0}^2\right]^{0.5}$, which is based on relative errors (so, e.g. an error of 0.01 when the true p_A is 0.3 is equivalent to an error of 0.001 with a true p_A of 0.03). This would put more emphasis on the accuracy of estimates of the small proportions, which may be desired.

We have assumed throughout that all samples are taken from the same catch. It is well known that what Clark (1981) colourfully referred to as the promiscuous application of age–length keys (applying an age–length key derived from a sample taken in one year to a length distribution collected in a different year) can produce significant bias in estimated proportions at age (Kimura 1977; Westrheim and Ricker 1978). The LMMA method would be similarly affected, as would the other length-mediated methods described in the introduction. However, LMMA is easily modified to avoid this bias. Suppose the *LOA* and *LO* samples were collected from a year in which the proportions at age were p'_A, whereas the *L* sample had proportions p_A. All that is required is to modify Eqns (2), (3) and (5) by replacing p_A by p'_A. This increases the set of parameters to be estimated from $\{p_A, \boldsymbol{\theta}_A\}$ to $\{p_A, p'_A, \boldsymbol{\theta}_A\}$. It remains to be seen how well this modified method performs compared to other length-mediated methods that have been developed to allow for samples from different years (Clark 1981; Bartoo and Parker 1983; Hoenig and Heisey 1987; Kimura and Chikuni 1987). It should be noted that all these methods require that the distributions of length at age do not vary from year to year (i.e. the parameters $\boldsymbol{\theta}_A$ do not vary with time).

Two other possible extensions of LMMA are reasonably straightforward. Extension to Bayesian estimation would not be difficult. This requires the user to supply prior distributions for every model parameter. Also, estimation is more complex because what is required is a joint posterior distribution for all parameters, rather than just a single value. However, there are several widely available software packages, such as BUGS (http://www.mrc-bsu.cam.ac.uk/bugs/, verified June 2005) and AD Model Builder (http://otter-rsch.com/admodel.htm, verified June 2005), that facilitate Bayesian estimation. The specification of prior distributions is often difficult. However, it would be simpler if LMMA were to be used year after year with the same fish stock. In this situation it would be sensible to use the posterior distributions for the $\boldsymbol{\theta}_A$ from one year as the prior distributions in the following year (but of course it would not be sensible to do this for the p_A). A second extension would be to allow for ageing error. This can be simply done as long as we can provide a misclassification matrix $\mathbf{M}$ to characterise this error (this can be constructed from replicate age estimates). The element $\mathbf{M}_{A,A'}$ of this matrix denotes the probability that a fish of true age A is given age A' (so the rows of this matrix must sum to one). The only change that is needed is to replace the term $p_{A_i}\,\mathrm{g}(\mathbf{X}_i;\,\boldsymbol{\theta}_{A_i})$ in Eqn (2) by $\sum_A p_A \mathbf{M}_{A,A_i}\,\mathrm{g}(\mathbf{X}_i;\,\boldsymbol{\theta}_A)$.

We close by acknowledging a practical problem. It is often not possible to select an SR sample from the entire catch from a fishery. With the ALK method, the common approach is to divide the catch into segments (each of which may be a landing or a catch from an individual tow) and then to collect samples from some or all segments. The large sample is made up of an SR sample from each sampled segment. The length distribution of the large sample is calculated as a weighted sum of these SR samples, where the weights represent the contribution of each segment to the total catch weight. The weights are necessary because it is accepted that the length distribution of the catch varies from segment to segment. However, no weighting is used for the small sample, which is simply the union of otolith samples from each segment. This requires the strong assumption that, for fish of a given length, the expected distribution of ages is the same in all segments. In other words, the selectivity factors that cause the length distribution to vary between segments (e.g. differences in fishing gear or spatio-temporal heterogeneity in the fished population) must depend only on fish length, and not on age. The ALK method is unaffected by this change in sample structure, but this is not true of LMMA or other statistically-based methods (e.g. the methods of Smith and Sedransk 1982; Martin and Cook 1990). In principle, these methods could be extended by treating the L sample as a (weighted) collection of SR samples and estimating a length-based selectivity curve for each. How easy this would be in practice remains to be seen.

Acknowledgments

We are grateful to Rick Fletcher, Ian Wright and an anonymous referee for reviewing an earlier version of the present paper.

References

Anderson, J. R., Morison, A. K., and Ray, D. J. (1992). Age and growth of Murray cod, *Maccullochella peelii* (Perciformes: Percichthyidae), in the lower Murray-Darling Basin, Australia, from thin-sectioned otoliths. *Australian Journal of Marine and Freshwater Research* **43**, 983–1013.

Baird, J. (1983). A method to select optimum numbers for aging in a stratified approach. *Canadian Special Publication of Fisheries and Aquatic Sciences* **66**, 161–164.

Bartoo, N. W., and Parker, K. R. (1983). Stochastic age-frequency estimation using the von Bertalanffy growth equation. *Fishery Bulletin* **81**, 91–96.

Boehlert, G. W. (1985). Using objective criteria and multiple regression models for age determination in fishes. *Fishery Bulletin* **83**, 103–117.

Campana, S. E., and Thorrold, S. R. (2001). Otoliths, increments, and elements: keys to a comprehensive understanding of fish populations? *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 30–38. doi:10.1139/CJFAS-58-1-30

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*). *Fisheries Research* **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. *Canadian Journal of Fisheries and Aquatic Sciences* **61**, 158–167.

Clark, W. G. (1981). Restricted least-squares estimates of age composition from length composition. *Canadian Journal of Fisheries and Aquatic Sciences* **38**, 297–307.

Dixon, W. J., and Massey, F. J. (1969). 'Introduction to Statistical Analysis.' (McGraw-Hill: New York.)

Ferreira, B. P., and Russ, G. R. (1994). Age validation and estimation of growth rate of the coral trout, *Plectropomus leopardus*, (Lacepede 1802) from Lizard Island, Northern Great Barrier Reef. *Fishery Bulletin* **92**, 46–57.

Fletcher, W. J. (1991). A test of the relationship between otolith weight and age for the pilchard *Sardinops neopilchardus. Canadian Journal of Fisheries and Aquatic Sciences* **48**, 35–38.

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

Francis, R. I. C. C., and Campana, S. E. (2004). Inferring age from otolith measurements: a review and a new approach. *Canadian Journal of Fisheries and Aquatic Sciences* **61**, 1269–1284. doi:10.1139/F04-063

Fridriksson, A. (1934). On the calculation of age-distribution within a stock of cod by means of relatively few age-determinations as a key to measurements on a large scale. *Rapports Et Proces-Verbaux Des Reunions, Conseil International Pour l'Exploration De La Mer* **86**, 1–5.

Hoenig, J. M., and Heisey, D. M. (1987). Use of a log-linear model with the EM algorithm to correct estimates of stock composition and to convert length to age. *Transactions of the American Fisheries Society* **116**, 232–243. doi:10.1577/1548-8659(1987)116<232:UOALMW>2.0.CO;2

Horppila, J., and Peltonen, H. (1992). Optimizing sampling from trawl catches: contemporaneous multistage sampling for age and length structures. *Canadian Journal of Fisheries and Aquatic Sciences* **49**, 1555–1559.

Ketchen, K. S. (1950). Stratified subsampling for determining age distributions. *Transactions of the American Fisheries Society* **79**, 205–212.

Kimura, D. K. (1977). Statistical assessment of the age-length key. *Journal of the Fisheries Research Board of Canada* **34**, 317–324.

Kimura, D. K., and Chikuni, S. (1987). Mixtures of empirical distributions: an iterative application of the age-length key. *Biometrics* **43**, 23–35.

Labropoulou, M., and Papaconstantinou, C. (2000). Comparison of otolith growth and somatic growth in two macrourid fishes. *Fisheries Research* **46**, 177–188. doi:10.1016/S0165-7836(00)00144-2

Lai, H.-L. (1987). Optimum allocation for estimating age composition using age-length key. *Fishery Bulletin* **85**, 179–185.

Lai, H.-L. (1993). Optimal sampling design for using the age-length key to estimate age composition of a fish population. *Fishery Bulletin* **92**, 382–388.

Luckhurst, B. E., Dean, J. M., and Reichert, M. (2000). Age, growth and reproduction of the lane snapper *Lutjanus synagris* (Pisces: Lutjanidae) at Bermuda. *Marine Ecology Progress Series* **203**, 255–261.

MacDonald, P. D. M., and Green, P. E. J. (1988). 'User's Guide to Program MIX: An Interactive Program for Fitting Mixtures of Distributions.' (Icthus Data Systems: Hamilton, Ontario.)

Martin, I., and Cook, R. M. (1990). Combined analysis of length and age-at-length data. *Journal Du Conseil International Pour L'exploration De La Mer* **46**, 178–186.

Oeberst, R. (2000). A universal cost function for the optimization of the number of age readings and length measurements for age-length-key tables. *Archive of Fishery and Marine Research* **48**, 43–60.

Radtke, R. L., and Hourigan, T. F. (1990). Age and growth of the Antarctic fish *Nototheniops nudifrons*. *Fishery Bulletin* **88**, 557–571.

Smith, P. J., and Sedransk, J. (1982). Bayesian optimization of the estimation of the age composition of a fish population. *Journal of the American Statistical Association* **77**, 707–713.

Stuart, A., and Ord, J. K. (1987). 'Kendall's Advanced Theory of Statistics, Volume 1, Distribution Theory.' (Charles Griffin and Company Limited: London.)

Westrheim, S. J., and Ricker, W. E. (1978). Bias in using an age-length key to estimate age-frequency distributions. *Journal of the Fisheries Research Board of Canada* **35**, 184–189.

Worthington, D. G., Doherty, P. J., and Fowler, A. J. (1995). Variation in the relationship between otolith weight and age: implications for the estimation of age of two tropical damselfish (*Pomacentrus moluccensis* and *P. Wardi*). *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 233–242.

Manuscript received 12 July 2004; revised 17 December 2004; and accepted 28 January 2005.