A simple method for comparing age-length keys reveals significant regional differences within one stock of haddock (*Melanogrammus aeglefinus* L.)

Hans D. Gerritsen, David McGrath and Colm Lordan

Gerritsen, H.D., McGrath, D. and Lordan, C. 2006. A simple method for comparing age-length keys reveals significant regional differences within one stock of haddock (*Melanogrammus aeglefinus* L.). – ICES Journal of Marine Science, XXX:000-000

A multinomial logistic model is presented as a tool for comparing two or more age-length keys. The model provides an objective way to fill in missing values and can be used for estimating uncertainty and visualising age-length keys (ALKs). An example of haddock (*Melanogrammus aeglefinus* L.) in ICES Division VIa (West of Scotland) is used to illustrate that significant regional differences in the proportions of age classes-at-length can exist on a small spatial scale. These differences were caused by regional variation in both length-at-age and relative abundance-at-age. As the length-at-age data are not normally weighted by the local catch rate (abundance), the ALK of the combined age data can result in strongly biased estimates of numbers-at-age. In the present case, the use of unweighted age data would have resulted in an over-estimate of recruitment of nearly 200% and an under-estimate of the spawning stock biomass of 15%. Comparing ALKs using this method will have several applications in fisheries science.

Key words: age-length key; multinomial logistic model; sampling design; haddock.

H. D. Gerritsen and C. Lordan. Fisheries Science Services, Marine Institute, Rinville, Oranmore, Co Galway, Ireland. D. McGrath. Commercial Fisheries Research Group, Galway-Mayo Institute of Technology, Dublin Road, Galway, Ireland. Correspondence to: H.D. Gerritsen: tel: +353 91 780368; fax: +353 91 730470; e-mail: hans.gerritsen@marine.ie.

Introduction

Most fisheries stock assessments are based on estimates of numbers of fish per age class. Sampling for age data generally takes place on a non-random (length-stratified) basis where sampling targets are set by length class. Additionally, a larger random sample is taken to obtain the length frequency of the catch or landings. To estimate numbers at age, the aged sample is usually raised to the total length frequency using an Age-Length Key (ALK), which consists of the proportions at age for each length class (Fridriksson 1934). The length-stratified sampling strategy ensures that fish from a wide range of sizes are represented in a relatively small aged sample.

All age-at-length data from an entire stock are often combined without weighting under the assumption that differences between gear types or regions can be disregarded (e.g. ICES 2005). Differences in size selectivity among gears should not influence the proportions of age classes at a given length, assuming that within each length class the probability of capture is independent of age. However, regional differences in the length-at-age distributions do have the potential to result in a biased ALK. These differences might be caused either by variation in length-at-age distributions or by variation in relative abundance of the age classes. For example, fish of a certain age might have a larger mean length in one area than another due to differential growth rates or size-specific migration. Additionally, in certain length classes, proportions of young fish might be higher in nursery areas than elsewhere, simply because they are locally more abundant, relative to other age-classes.

Various methods have been applied to test for differences between ALKs. Hayes (1993) and Horbowy (1998) both suggested comparing individual cells of the ALKs using multiple Fisher's or Chi squared tests. Although the application of these tests is straightforward, the interpretation of the results is not, as there are as many p-values as the number of age and length classes that are considered. Additionally, any cells that do not contain enough data, have to be omitted, so the tests can only be applied to large data sets. Dwyer et al. (2004) took a different approach and suggested applying a two-dimensional Kolmogorov-Smirnov test. This approach only requires a single test to compare two ALKs. However, the two-dimensional Kolmogorov-Smirnov test is not widely available in statistical packages and it does not appear to be the most parsimonious solution. Rindorf & Lewy (2001) applied multinomial models of continuation-ratio logits to aged data. This approach has many advantages, however Rindorf & Lewy's model required a polynomial function to be defined to allow every possible type of distribution to be modelled. However if one makes the assumption of normality in the length-at-age distributions, Rindorf & Lewy's method can be greatly simplified by removing the need for arbitrary smooth functions.

The assumption of normality in length-at-age distributions is routinely being made, either with constant variance over the age groups or variance proportional to mean length (e.g. Schnute and Fournier 1980; Labonté 1983; Rosenberg and Reddington 1988). However, in contrast to these studies, the current assumption of normality is a weak one and applies only to the population from which sample was drawn, not the aged sample itself (which is non-random) or even the catch (which is often size selective).

The suggested approach allows for multinomial logistic models to be applied testing for differences between ALKs. In addition, the models can be used to predict missing values, estimate uncertainty and help visualise ALKs. The method will be illustrated by examining the variability in ALKs of haddock to the west of Scotland (ICES Division VIa) through the application of multinomial models to age-at-length data from the 2004 Irish Groundfish Survey.

Methods

Logistic models with a binomial error distribution are widely used in fisheries science to describe the relative proportions of two overlapping distributions. Examples include size selection ogives for fishing gear, discarding ogives and maturity ogives. In the case of ALKs, there are mostly more than two overlapping age-classes and therefore a multinomial logistic model is required to describe the proportions of age-at-length. Multinomial models can be fitted by maximising the product of the conditional binomial trials simultaneously (Beare and McKenzie 1999; Rindorf and Lewy 2001). Alternatively the S-PLUS® and R packages provide the function *multinom()* which fits multinomial log-linear models via neural networks (Venables and Ripley 1994).

Multinomial model selection, testing and estimation can be carried out in a similar way to generalised linear modelling (McCullagh and Nelder 1989). Model selection allows one to identify which factors contribute significantly to the explanatory power of the model and to test for differences between regions, gear types etc. Model estimation can be used to interpolate missing values. It is a regular occurrence that, for certain length classes in the total length frequency, no aged samples are available.

These gaps in the data need to be filled in to allocate numbers-at-age for the relevant length classes. The multinomial logistic model provides an objective way to do so.

Here, ALKs of haddock (*Melanogrammus aeglefinus* L.) were obtained from the Irish Groundfish Survey, carried out by the Marine Institute in October and November 2004 on RV "Celtic Explorer". Data from ICES Division VIa (West of Scotland) were selected to illustrate the method. The area was divided into three depth strata: shallow (<75m), medium (75m-125m) and deep (>125m). Sampling targets of five age samples per cm length class were set for each of the strata, so a separate ALK was available for each stratum. Fish ages were determined by sectioning the sagittal otoliths through the nucleus and counting the number of hyaline rings.

Multinomial logistic models of the following form were fitted:

$$A \sim L + S + L.S$$

where A is the predicted age distribution at length L in stratum S. L was fitted as a continuous variable, S as a factor and L.S is an interaction term. The significance of the factor stratum in the model was tested by comparing the initial model to a model without that factor. The difference in residual variance of these nested models was tested against the difference in the model degrees of freedom (υ) using the Chi-square test (Collett 2003). For the current analysis, age classes of 4-year-olds and older were combined into a single plus-group. As catches of 0-group fish were scarce and did not overlap in size with the other age classes, they were omitted from the analysis.

All haddock from VIa are considered to be a single stock and for the purposes of stock-assessment it is common practice to use a single ALK to obtain numbers-at-age without weighting the age data in any way (ICES 2005). For the present study, numbers-at-age in the survey catches were estimated in two ways: firstly by combining all age data without weighting and secondly by weighting the age data by the relative abundance in each stratum. The relative abundance in each stratum was estimated from the catch numbers per unit effort (CPUE), multiplied by the surface area of each stratum. The unit effort is a standard half-hour trawl, towed at 3kn. The length frequency data were expressed as CPUE and weighted by stratum surface area in all cases to obtain an unbiased length frequency for the combined strata.

Standard errors for the numbers-at-age estimates were obtained using a bootstrapping routine (Efron and Tibshirani 1993). The individual fish in the aged sample were treated as independent sampling units and re-sampled 500 times. This approach, as opposed to re-sampling within length classes, can result in length classes without data, therefore a multinomial model was fitted to the data for each bootstrap iteration. Standard errors were estimated from the standard deviation of the bootstrapped estimates from the modelled data. The length distributions were assumed to be known without error.

Results

A very highly significant stratum effect was found for a model that contained data from all three strata (χ^2=133.3; υ=16; p<0.001). When the shallow stratum was omitted from the dataset, the stratum effect was no longer significant (χ^2=9.2; υ=8; p=0.32). However, if either one of the other strata were omitted, the stratum effect remained highly significant. This indicates that the ALK of the shallow stratum was significantly different from the ALKs of two other strata and that the ALKs of the deep and medium strata were not significantly different from each other. Figure 1 shows the observed and modelled proportions at age and length distributions. The figure indicates that the main difference between the strata lies in the proportions of one-year-olds in length classes 25-35cm, which were considerably higher in the shallow stratum than in the other strata.

In the medium and deep strata, two-year-olds were by far the most common age class in the catches (Table 1). In the shallow stratum, one-year-olds were most abundant, relative to other age classes. In addition, the mean length-at-age appeared to be higher for most age classes in the shallow stratum than in the others (Table 2). Combining all aged data into an ALK without weighting, resulted in estimated catch numbers for one-year-olds that were nearly twice as high (88 fish per unit effort) as the estimate using age data weighted by abundance (47 per unit effort; Table 2). If the present data were used as an absolute estimate of the spawning stock biomass, the unweighted estimate would have resulted in an under-estimate of the spawning stock biomass by 15%, assuming a knife-edge maturity at age two (ICES 2005).

The main reason for the bias in the unweighted ALK appears to be that fish from the shallow stratum were over-represented in the sampling. Catch rates in the shallow stratum were around 8 times lower than in the medium and deep strata, but the sample numbers for age were actually higher in the shallow stratum (Table 2). As the one-year-olds in the shallow stratum were relatively abundant (compared to other age classes) and, on average, about 2cm larger than in the other strata, the proportions of one-year-olds at length were over-estimated in many size classes of the unweighted ALK.

Discussion

The multinomial model used here, is a special case of the methodology presented by Rindorf & Lewy (2001). It eliminates the need to apply a polynomial function to the length classes, which improves the transparency and simplicity of the model. A model with A age classes only requires 2(A-1) model parameters; the apparently complex shape of the model (e.g. Figure 1) results from the added proportions of the various age classes.

The assumption of normality applies not to the aged data but only to the underlying population because the model uses proportions (age-at-length), not length-at-age distributions. This is most clearly demonstrated in the binomial logistic case, for example a discard ogive. The symmetric s-shaped curve that describes a discard ogive, results from the proportions of two overlapping distributions: one length distribution of discards and one of landings. If both distributions were strictly normal (at least in the area of overlap) with equal variance, a logistic binomial curve would describe the proportions-at-length exactly, regardless of any size selection in the sampling. For most binomial applications the assumption of normality cannot be made, however the proportions-at-length still tend to follow an s-shaped curve that is closely described by the logistic curve (McCullagh and Nelder 1989; Collett 2003). The multinomial case expands on the binomial model by describing the proportions of more than two overlapping distributions. Unlike many binomial applications, length-at-age distributions do tend to be approximately normally distributed with similar variances (e.g. Schnute and Fournier 1980; Labonté 1983; Rosenberg and Reddington 1988).

Sexual dimorphism in growth could result in bimodal, hence not normal, length-at-age distributions. In this case, it might be advisable to sample the sexes separately, as is feasible for some flatfish that can be sexed without dissection. Alternatively, one can apply an age-sex-length-key, which should restore the normal length-at-age distributions; the factor sex could then be added to the multinomial model.

The model appears to be a useful tool to detect significant differences between ALKs, although the likelihood of finding these differences will, of course, depend on the number of fish sampled. The model is also useful for obtaining confidence limits or variance estimates and it can deal with missing length classes: if no aged data exists for a certain length class, the model can predict the expected proportions of the age classes for that (or any other) length class. In the future, the model might be expanded to include seasonal changes, for example by fitting smooth curves through a time variable.

The current example shows that there can be a high degree of spatial variability in ALKs, which can result in strongly biased numbers-at-age estimates. This has many implications for the unit-stock and dynamic pool assumptions that underlie many age-based stock assessments. Many stocks are known to have nursery areas or age- or size-specific migration and will therefore have regional differences in the age structure. If the number of age samples is proportional to the local abundance of fish, the estimates will be unbiased, but otherwise the aged samples should be weighted by the abundance in each region before they are combined into an ALK to avoid bias. These considerations apply to survey data, as well as data from commercial sources, where data from many regions are often combined without weighting.

In the present case, the consequences of using an unweighted ALK would be a large bias in the estimated abundance of one-year-old fish. Many stock assessments use survey indices in a relative sense and this bias might be corrected by a catchability parameter. However if the bias changes from year-to-year due to year class effects, changes in survey design or other mechanisms, there will be implications for the assessment and management advice. If this survey were used in an absolute sense (e.g. Beare et al. 2005) the consequence of the bias would have been a nearly two-fold over-estimate of the 2003 year class and an underestimate of the spawning stock by 15%.

Acknowledgements

The authors would like to thank all the staff involved in the 2004 Irish Groundfish Survey and with the ageing of the fish. Also thanks to Rick Officer, Edd Codling and two referees for their valuable comments on the manuscript. Free R-software has been used for this work and we would like to thank the R development core team and all contributors to the R project (http://www.R-project.org).

References

Beare, D., McKenzie, E. 1999. The multinomial logit model: a new tool for exploring Continuous Plankton Recorder data. Fisheries Oceanography 8(suppl. 1): 25-39.

Beare, D. J., Needle, C. L., Burns, F., Reid, D. G. 2005. Using survey data independently from commercial data in stock assessment: an example using haddock in ICES Division VIa. ICES Journal of Marine Science. 62(5): 996-1005.

Collett, D. 2003 Modelling binary data. Chapman & Hall/CRC. Boca Raton, Florida.

Dwyer, K., Koen-Alonso, M., Walsh S. J. 2004. Finding the magical minimum sample size: a computer-intensive approach to minimize re-ageing effort to construct age-length keys for yellowtail flounder. NAFO SCR Doc 04/09: 1-4.

Efron, B., Tibshirani, R. J. 1993. An introduction to the bootstrap. Chapman & Hall.CRC. Boca Raton, Florida.

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 Réunions. Conseil International pour l'Éxploration de la Mer. 86(6): 1-5.

Hayes, D. B. 1993. A statistical method for evaluating differences between age-length keys with application to Georges Bank haddock, Melanogrammus aeglefinus. Fishery Bulletin. 91(3): 550-557.

Horbowy, J. 1998. Comparison of age-length keys of Baltic cod derived from polish commercial and research data. Fisheries Research. 36: 257-266.

ICES 2005. Report of the Working Group on the Assessment of Northern Shelf Demersal Stocks (WGNSDS), Murmansk, Russia.

Labonté, S. S. M. 1983. Aging capelin: enhancement of age-length keys and importance of such enhancement. *In* Sampling commercial catches of marine fish and invertebrates Ed. by W. G. Doubleday, D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences. 66, pp 171-177.

McCullagh, P., Nelder, J. A. 1989. Generalized linear models. Chapman & Hall.CRC. Boca Raton, Florida.

Rindorf, A., Lewy, P. 2001. Analysis of length and age distributions using continuation-ratio logits. Canadian Journal of Fisheries and Aquatic Sciences. 58(6): 1141-1152.

Rosenberg, A. A., Reddington, J.R. 1988. Length-based methods of fish stock assessment. *In* Fish Population Dynamics (Second Edition). Ed. by J. A. Gulland. John Wiley & Sons Ltd., Chochester, pp 83-103.

Schnute, J., Fournier, D. 1980. A new approach to length-frequency analysis: growth structure. Canadian Journal of Fisheries and Aquatic Sciences. 37(9): 1337-1351.

Venables, W. N., Ripley, B. D. 1994. Modern Applied Statistics with S-PLUS. Springer, New York.

Table

Table 1. Summary statistics of the three depth strata in ICES area VIa. Given for each stratum are the surface area (in nautical miles); the number of stations; the catch per unit effort (CPUE) and the number of fish sampled for ageing.

	Area (nm)	Stations	CPUE	Nos aged
VIa Shallow	4000	18	59	96
VIa Medium	5400	17	460	62
VIa Deep	2700	6	455	41

Figure

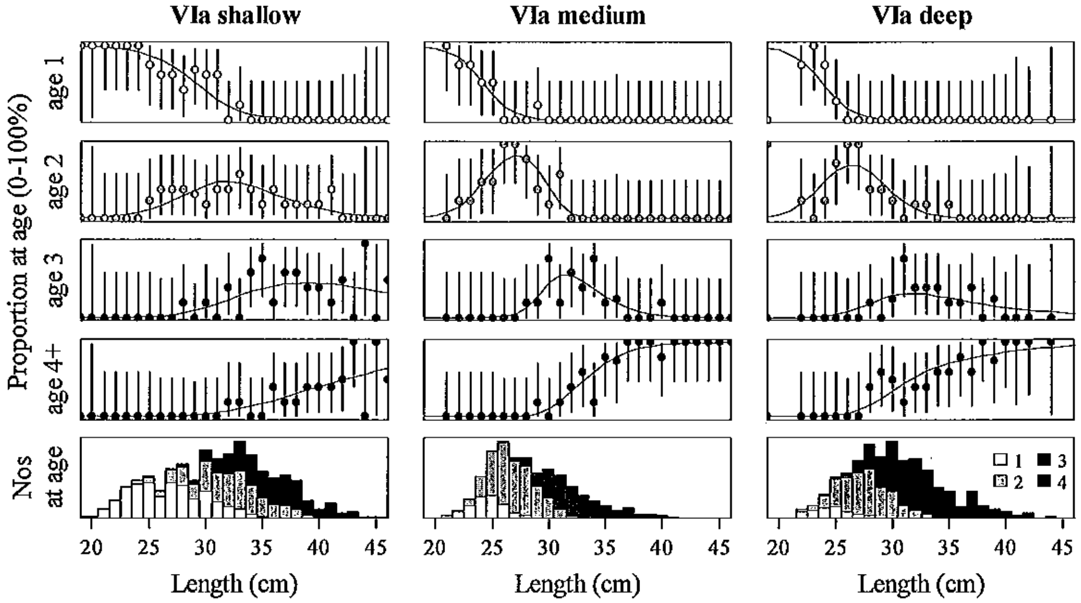

Figure 1. Proportions-at-length of age groups 1 to 4+ in the three depth strata and the length-at-age distributions estimated by applying ALK models to the total catch length frequencies. The circles represent the observed proportions with their individual 95% confidence intervals and the curves represent the predicted proportions from the multinomial models. The shades of the stacked bars correspond to the different age classes. The proportions of one-year-olds were higher in the shallow strata than in the other strata for length classes up to 35cm.