Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

A simple method for comparing age−length keys reveals significant regional differences within a single stock of haddock (*Melanogrammus aeglefinus*)

Hans D. Gerritsen, David McGrath, and Colm Lordan

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 visualizing age−length keys (ALKs). An example of haddock (*Melanogrammus aeglefinus*) in ICES Division VIa (West of Scotland) is used to illustrate significant regional differences in the proportions of age-at-length. These differences are caused by regional variation in both length-at-age and relative abundance at age. As the length-at-age data are normally not 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 overestimate of recruitment of nearly 200%, and an underestimate of spawning-stock biomass of 15%. Comparing ALKs using this method has several practical applications.

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 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 age-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). A length-stratified sampling strategy ensures that fish from a wide range of sizes are represented in a relatively small age-sample.

All age-at-length data from an entire stock are often combined without weighting, under the assumption that differences between gear types and regions can be disregarded (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 age-length structure do have the potential to result in a biased ALK. These differences may be caused either by variation in length-at-age distributions or by variations in the relative abundance of age classes (age-at-length distributions). For example, fish of a certain age might have a larger mean length in one area than another as a consequence of differential growth rates or size-specific migration. Furthermore, in certain length classes, proportions of young fish might be higher in nursery areas than elsewhere, simply because they are locally more abundant than 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 application of these tests is straightforward, interpretation of the results is not, because there are as many *p*-values as the number of cells that are being considered. Additionally, any cells that do not contain sufficient 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 the 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 straightforward solution. Rindorf and Lewy (2001) applied multinomial models of continuation-ratio logits to age data. This approach has many advantages, but Rindorf and Lewy's model requires 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 length-at-age distributions, Rindorf and Lewy's method can be greatly simplified by removing the need for arbitrarily smooth functions.

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

The suggested approach allows for multinomial logistic models to be applied to test for differences between ALKs. In addition, the models can be used to predict missing values, to estimate uncertainty, and to help visualize ALKs. The method will be illustrated by examining the variability in ALKs of haddock West of Scotland (ICES Division VIa), through 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, and discarding and maturity ogives. In the case of ALKs, there are generally more than two overlapping age classes, so a multinomial logistic model is required to describe the proportions of age at length. Multinomial models can be fitted by maximizing 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()* that 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 generalized linear modelling (McCullagh and Nelder, 1989). Model selection allows identification of the factors that contribute significantly to the explanatory power of a 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 an overall length

frequency, no age samples are available. These gaps in the data need to be filled 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*) 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 (<75 m), medium ($75-125$ m) and deep (>125 m). Sampling targets of five age samples per 1 cm length class were set for each stratum, 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:

$$A \sim L + S + LS,$$

where A is the predicted age distribution at length L in stratum S. L was fitted as a continuous variable and S as a factor and LS 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 (v) using a Chi-squared test (Collett, 2003). For the current analysis, age classes of four year 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 (number) per unit effort (cpue), multiplied by the surface area of each stratum. The unit effort is a standard half-hour trawl, towed at 3 knots. 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 age 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, so 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. Length distributions were assumed to be known without error.

Results

A highly significant stratum effect was found for a model that contained data from all three strata ($\chi^2 = 133.3$; $v = 16$; $p < 0.001$). When the shallow stratum was omitted from the data set, the stratum effect was no longer significant ($\chi^2 = 9.2$; $v = 8$; $p = 0.32$). However, if one of the other strata was omitted, the stratum effect remained highly significant. This indicates that the ALK of the shallow stratum was significantly different from the ALKs of the other two 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. The main difference between the strata apparently lies in the proportions of 1-year-olds in length classes 25−35 cm, which were considerably greater in the shallow stratum than in the other strata.

In the medium and deep strata, 2-year-olds were by far the most common age class in the catch (Table 1). In the shallow stratum, 1-year-olds were most abundant relative to other age classes. In addition, mean length-at-age appeared to be higher for most age classes in the shallow stratum than in the others (Table 1). Combining all age data into an ALK without weighting resulted in estimated catch

numbers for 1-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 1). If these data were to be used as an absolute estimate of spawning-stock biomass (SSB), the unweighted estimate would result in an underestimate of SSB by 15%, assuming knife-edge maturity at age 2 (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 samples. Catch rates in the shallow stratum were around eight 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 with other age classes) and, on average, about 2 cm larger than in the other strata, the proportions of 1-year-olds at length were overestimated in many size classes of the unweighted ALK.

Discussion

The multinomial model used here is a special case of the methodology presented by Rindorf and Lewy (2001). It eliminates the need to apply a polynomial function to 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
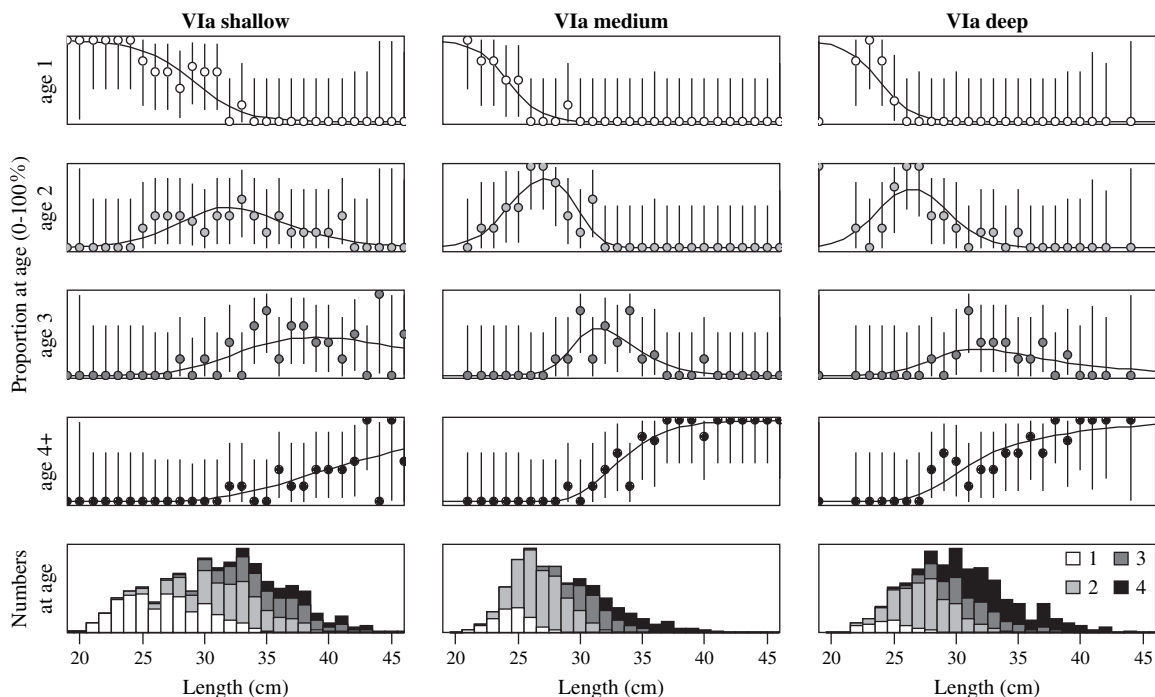

Figure 1. Proportions at length of age groups 1−4+ in the three depth strata, and the length-at-age distributions estimated by applying ALK models to the length frequencies of the total catch. Circles represent observed proportions with their individual 95% confidence intervals, and curves represent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

Table 1. Estimated numbers-at-age and mean lengths-at-age for individual strata and for the combined area after application of weighted and non-weighted ALKs. Weighting the ALKs resulted in considerably lower estimates for the number of 1-year-old fish. Standard errors were obtained by bootstrapping and are given in parenthesis. Numbers- and lengths-at-age were obtained by applying the ALK to the catch length frequency in each stratum.

Stratum	Values per age class					
	0	1	2	3	4	5
Numbers-at-age (per unit effort)						
VIa Shallow	1 (5)	21 (2)	18 (3)	13 (3)	3 (1)	4 (5)
VIa Medium	3 (0)	71 (19)	226 (24)	91 (17)	40 (10)	30 (7)
VIa Deep	2 (0)	34 (8)	160 (22)	36 (14)	93 (20)	129 (21)
Numbers-at-age for combined strata (per unit effort)						
VIa – non-weighted	2 (1)	88 (10)	121 (10)	46 (6)	34 (5)	36 (5)
VIa – weighted	2 (1)	47 (10)	143 (13)	53 (9)	39 (7)	43 (6)
Mean length-at-age (cm)						
VIa Shallow	14.2 (3.1)	26.7 (1.4)	32.1 (0.8)	34.8 (0.7)	36.0 (1.2)	37.4 (3.5)
VIa Medium	15.1 (0.3)	24.6 (0.5)	27.0 (0.2)	31.0 (0.4)	33.3 (0.7)	36.0 (0.9)
VIa Deep	13.8 (0.3)	24.7 (0.7)	27.7 (0.4)	31.7 (0.8)	31.9 (0.6)	33.2 (0.5)

of the model (Figure 1) results from the added proportions of the various age classes.

The assumption of normality applies not to the age data but only to the underlying population, because the model uses proportions (age at length), not length-at-age distributions. This can be illustrated using the binomial logistic case, e.g. 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, the proportions of length would be described by a logistic binomial curve regardless of size selection in the sampling. For most binomial applications the assumption of normality cannot be made, but 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 tend to be approximately normally distributed with similar variances (Schnute and

Fournier, 1980; Labonté, 1983; Rosenberg and Beddington, 1988).

Sexual dimorphism in growth could result in bimodal, and hence not normal, length-at-age distributions. In this case, it might be advisable to sample 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 age data exist for a certain length class, the model can predict the expected proportions of the age classes for that (or any other) length class. In 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 estimates of numbers-at-age. This has many implications for the unit-stock and dynamic pool assumptions that underlie many age-based stock assessments. Many stocks have nursery areas or age- or size-specific migration patterns and will therefore have regional differences in their 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 preclude bias. These considerations apply to survey data, as well as to data from

Table 2. Summary statistics of the three depth strata in ICES area VIa. Given for each stratum are the surface area, the number of stations; the catch in number per unit effort (cpue) and the number of fish sampled for ageing.

Stratum	Area (nautical miles^2)	Number of stations	cpue	Numbers aged
VIa Shallow	4 000	18	59	96
VIa Medium	5 400	17	460	62
VIa Deep	2 700	6	455	41