G Model
FISH-4592; No. of Pages 19

ARTICLE IN PRESS

Fisheries Research xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Full length article

Estimating multinomial effective sample size in catch-at-age and catch-at-size models

Samuel B. Truesdell[a,*], James R. Bence[a], John M. Syslo[a], Mark P. Ebener[b]

[a] Quantitative Fisheries Center, Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI, 48824, USA
[b] Chippewa/Ottawa Resource Authority, 179 W. Three Mile Road, Sault Ste. Marie, MI, 49783, USA

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 3 November 2016
Accepted 7 November 2016
Available online xxx

Keywords:
Fisheries
Stock assessment
Catch-at-age
Effective sample size
Coregonus clupeaformis
Great lakes

ABSTRACT

Catch-at-age or catch-at-size stock assessment models require specification of an effective sample size (ESS) as a weighting component for multinomial composition data. ESS weights these data relative to other data that are fit, and is not an estimable parameter within a model that uses a multinomial likelihood. The ESS is typically less than the actual sample size (the number of fish) because of factors such as sampling groups of fish (clusters) that are caught together. A common approach for specifying ESS is to iteratively re-fit the model, estimating ESS after each fit so that the standardized residual variance is "correct," until ESS converges. We survey iterative methods for determining ESS for a multinomial likelihood and apply them to two Great Lakes whitefish stocks. We also propose an extension of such methods: (the Generalized Mean Approach – GMA) for the case where ESS is based on mean age (or length) to account for correlation structures among proportions. Our extension allows for greater flexibility in the relationship between ESS and sampling intensity. Our results show that the choice of ESS estimation method can impact assessment model results. Simulations (in the absence of correlation structures) showed that all the approaches to calculating effective sample size could provide reasonable results on average, however methods that estimated annual ESS independently across years were highly imprecise. In our simulations and application, methods: that did account for correlation structure in annual proportions produced lower ESS than those that did not and suggested that these methods are adjusting for a deviation from the multinomial correlation structure. We recommend using methods that adjust for correlation structures in the proportions, and either assuming a constant ESS or, when there is substantial inter-annual variation in sampling levels, assuming ESS is related to sampling intensity and using the GMA or a similar approach to estimate that relationship.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Catch-at-age and catch-at-size models are commonly used tools in stock assessment (e.g., Legault and Restrepo, 1998; Methot and Wetzel, 2013; Punt et al., 2013). These models use observations of cohorts through time to estimate population parameters. Because cohort size is a fundamental component, an accurate implementation of the relative abundance of age or size classes is critical to model accuracy. In a model's likelihood function, observations of the relative abundance of class size (expressed as proportions) are frequently compared to model-produced estimates during fitting using the multinomial likelihood (Francis, 2014). The influence of the proportions-at-age or at-size on the fit of the likelihood function is determined by the multinomial's effective sample size parameter (ESS), which defines the expected amount of variability from a simple random sample of fish ages or sizes (Folmer and Pennington, 2000; Methot and Wetzel, 2013). Determining ESS is important because this weighting factor can impact the model output quantities used by managers such as population size and fishing mortality rates (Francis, 2011).

The observed population composition data may be more variable than or have a correlation structure that differs from that of a multinomial sample of the observed number of fish. Two causes are the spatial behavior of the fish and the spatial grouping of the sampling method (e.g., a trawl catches many fish together). This amounts to cluster samples (Cochran, 1977), which carry less information than the number of individuals actually aged or measured (McAllister and Ianelli, 1997; Folmer and Pennington, 2000; Stewart et al., 2014), so ESS is typically smaller than the

* Corresponding author at: Quantitative Fisheries Center, 375 Wilson Rd. Room 101, East Lansing, MI 48824, USA.
E-mail address: truesd16@msu.edu (S.B. Truesdell).

http://dx.doi.org/10.1016/j.fishres.2016.11.003
0165-7836/© 2016 Elsevier B.V. All rights reserved.

Please cite this article in press as: Truesdell, S.B., et al., Estimating multinomial effective sample size in catch-at-age and catch-at-size models. Fish. Res. (2016), http://dx.doi.org/10.1016/j.fishres.2016.11.003

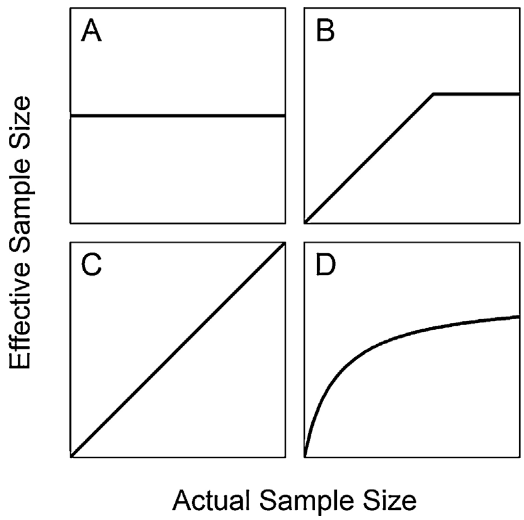

Fig. 1. Options for relating ESS to sampling intensity in catch-at-age or catch-at-size models: (A) a set ESS no matter the measured sample size, (B) proportional relationship between ESS and measured sample size up to a maximum, (C) proportional relationship between ESS and measured sample size, and (D) asymptotic relationship between ESS and measured sample size. In principle, relationships between ESS and actual sample size could apply to other measures of sampling effort, such as the number of trips sampled rather than number of fish aged or measured.

number of individuals processed. A third cause, applicable to length-structured models, is the potential for large recruitment events to impact multiple adjacent length bins, producing such correlations. Further complicating the issue, age compositions are often calculated based on both a length composition and an age-length key. Due to this complex data structure, ESS cannot be determined directly from the number of fish aged or measured, although in some cases it can be estimated based on sampling theory (e.g., Crone and Sampson, 1998; Pennington et al., 2002) or using an approach such as bootstrapping (e.g., Stewart et al., 2014); however it has been suggested that these data should not be weighted independently of an assessment model because much of the composition error may result from model process error rather than observation error (Francis, 2016). ESS also cannot be included as a parameter in models that use multinomial likelihoods for composition data because it is not estimable in the multinomial likelihood function.

Various methods have been employed for fixing and estimating multinomial ESS (Francis, 2011; Maunder, 2011), and these include ad-hoc and iterative approaches. To recognize that the information content of the samples is less than the actual number of fish observed, ad-hoc methods may set a fixed ESS (e.g., Fournier and Archibald, 1982; Fig. 1A) or treat the annual number of observations as the ESS up to a maximum value, and use this maximum when the number of observations exceeds the threshold (e.g., Fournier et al., 1998; Caroffino and Lenart, 2010; Fig. 1B). These ad-hoc approaches can be based on estimation of actual variances in other fisheries if formal sampling designs permit this (Crone and Sampson, 1998), informal consideration of the observed variation in age compositions relative to what would be expected from a multinomial, or other forms of professional judgement.

A variety of iterative approaches have been advanced (e.g., McAllister and Ianelli, 1997; Francis, 2011; Maunder, 2011). Francis (2011) argued that decisions regarding weighting (variances) for other data should be made first, followed by tuning ESS using iterative approaches. Most approaches determine how variable data are about the model predictions, relative to how variable they are expected to be given the assumed ESS, and then refit the stock assessment model repeatedly, adjusting the ESS at each iteration to be consistent with the variation seen at the last iteration until ESS is stable.

These iterative methods were classified by Francis (2011) based on whether they accounted for correlation structures or not, and their assumptions about 'process error' (which in this case can be viewed as over-dispersion relative to a multinomial distribution based on the number of fish aged or measured). Herein, correlation structure refers to a deviation from the weak negative correlation in proportions between all pairs of bins that arises from the multinomial distribution and the constraint that proportions sum to 1.0. Our expectation is that such structure will generally involve the strongest positive correlations in observed proportions from proximal bins (e.g., ages 5 and 6) with positive correlations weakening and eventually becoming negative between proportions in bins that are farther apart (e.g., 4 and 9). Methods that do not allow for correlation structures generally seek to set ESS to match variation in the proportions at age or length versus what would be expected from a multinomial distribution. This includes McAllister and Ianelli's (1997) commonly used approach (e.g., Wilberg et al., 2005; Campana et al., 2010; Berger et al., 2012). Methods that can account for correlation structure seek to set ESS to match variation in mean age or mean length that would be expected if the composition data arose from a multinomial distribution. As originally implemented by McAllister and Ianelli, their iterative approach calculated an ESS for each year (for a data type), and then averaged these and used the same ESS for each year in the next iteration of the assessment model. Thus they assumed that information content was constant over years and unrelated to any variation in sampling effort (Fig. 1A). Francis (2011) proposed two hypotheses that account for overdispersion, based on the idea that the adjustment of ESS from the number of samples should either be multiplicative or additive. For the multiplicative case, if a particular composition sample was based on $\tilde{N}$ observations, then its information content (ESS) is $\ddot{N} = w\tilde{N}$, where $\ddot{N}$ is the ESS and w is a multiplicative scaling factor (Fig. 1C). For the additive case, $\frac{1}{\ddot{N}} = \frac{1}{\tilde{N}} + \frac{1}{N_{MAX}}$, the information content initially increases directly with sample number but approaches an asymptote, N_{MAX} (Fig. 1D).

The hypothesized direct proportionality between ESS and sampling intensity arising from multiplicative error could apply to other measures of sampling intensity such as number of trips rather than number of fish aged or measured. Iterative methods that do not account for correlation structure and use the observed variation in proportions along with the variability expected in a multinomial sample can at least theoretically calculate an ESS for each year. Maunder (2011) suggested that in such cases rather than using these directly one could fit a statistical model relating these nominal effective sample sizes to observed sampling intensities, and use the predictions from the statistical model as the ESS in the next iteration. This allows for consideration of more general relationships between ESS and sampling intensity than those arising from multiplicative or additive error acting alone. For example one might hypothesize that information content of composition samples increases to an asymptote as a function of the number of trips sampled, rather than number of fish aged, but there would be no reason to assume an initial slope of 1.0. Even when using number of ages or lengths as the predictor, an initial slope of less than 1.0 seems possible, i.e., both multiplicative and additive error could operate together. This approach is not directly applicable to the methods that allow for correlation structures, as only one deviation between observed and expected means is available for each year. Thus, for those methods, Francis (2011) indicated that either $w$for the multiplicative hypothesis or N_{MAX} for the additive hypothesis is adjusted so the resulting variation is matched exactly for each iteration.

Please cite this article in press as: Truesdell, S.B., et al., Estimating multinomial effective sample size in catch-at-age and catch-at-size models. Fish. Res. (2016), http://dx.doi.org/10.1016/j.fishres.2016.11.003

Download English Version:

https://daneshyari.com/en/article/5765460

Download Persian Version:

https://daneshyari.com/article/5765460

Daneshyari.com