Model-based estimation of effective sample size in Stock Synthesis using the Dirichlet-Multinomial distribution

James T. Thorson1, Kelli F. Johnson2, Richard D. Methot1, Ian G. Taylor1

1Fisheries Resource Assessment and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd. East, Seattle, WA 98112, USA

2School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA

Theoretical considerations and applied examples suggest that stock assessments are highly sensitive to the weighting of different data sources whenever data sources conflict regarding parameter estimates. Previous iterative reweighting approaches to weighting compositional data are generally ad hoc, do not propogate uncertainty about data-weighting when calculating uncertainty intervals, and often are not re-adjusted when conducting sensitivity or retrospective analyses. We therefore incorporate the Dirichlet-Multinomial (DM) distribution into Stock Synthesis, and propose it as a model-based method for estimating effective sample size. This distribution incorporates one additional parameter per survey (with the option of mirroring its value among fleets), and we show that this parameter represents the ratio of nominal (“input”) and effective (“output”) sample size. We demonstrate this approach using data for Pacific hake, where DM and McAllister-Ianelli reweighting approaches give similar and plausible results. We also use simulation testing to demonstrate the estimation properties of this new estimator, and conclude by recommending further research of computationally efficient estimators of effective sample size based on alternative, *a priori* consideration of sampling theory is desirable rather than continuing current ad hoc practices.

Keywords: Dirichlet-Multinomial; integrated assessment; multinomial; statistical catch-at-age

**Introduction**

Data on commercially exploited marine fish populations typically come from a variety of sources and contain considerable amounts of variability. Theoretical considerations and applied examples suggest that integrated statistical stock assessments are highly sensitive to the weighting of different data sources whenever sources conflict regarding parameter estimates. Consequently, the status of the stock has the potential to be highly dependent on the method and values used to weight data sources included in stock assessments (Francis, 2011; Hulson et al., 2012).

Within integrated statistical stock assessments, the by variance or sample size assigned to likelihood components included in the objective function dictates the weight of each data source, where weight is inversely related to uncertainty. Four sources of error can affect the level of uncertainty inherent in the data: (i) measurement error, (ii) observation error, (iii) process error, and (iv) model-missspecification error. Assigning a given level of uncertainty to a data source becomes more complicated when unknown process error and model-misspecification exists because of their inability to be quantified.

For compositional data, which describe the distribution of ages or lengths in catches, the multinomial distribution is the most used. The multinomial distribution assumes samples are collected at random from the population. Therefore, employed sampling methods and fish behavior (e.g., schooling) will lead to overdispersion of the true uncertainty in the estimated proportions (citation).

* Assumptions and violations.
* Use of the multinomial requires determining an effective sample size.
  + Methods for data weighting should allow for correlations (Francis, 2011).

Alternatives to the multinomial, including ad hoc practices.

* McAllister-Ianelli (1997)
* Bootstrap (Stewart and Hamel, 2014)
* Dirichlet-Multinomial (DM)
* Logistic-normal (Francis, 2014)
* Adjusted log-normal (Legault, 2014)
* Multivariate-logistic (Schnute and Richards, 1995)

Given the inability of *ad hoc* methods to propagate uncertainty about data-weighting when calculating uncertainty intervals and their propensity to encourage ignoring data-weighting when conducting sensitivity or retrospective analyses, we incorporate the DM distribution into Stock Synthesis (SS) and propose it as a model-based method for estimating the effective sample size ().

**Methods**

*Dirichlet-Multinomial distribution*

We here use a Dirichlet-Multinomial distribution:

where is the proportion at age in the available data such that , *N* is the total number of samples in the available data (which is restricted to any positive real number), is the estimated proportion at age such that , and is the estimated variance inflation coefficient. The first term does not depend upon the parameters, but ensures that as , :

i.e., that the DM reduces to the multinomial likelihood in this circumstance.

*Effective sample size*:

We define the effective sample size of the DM distribution as the sample size of a multinomial distribution that has the same variance. We use a Dirichlet distribution:

where , is the true proportion at age, and is the effective sample size of the Dirichlet distribution:

Similarly, the variance of a single element from a multinomial distribution:

where *N* is the sample size, is:

Defining observed proportion , we see that:

Therefore the variance of the observed proportion at age for a DM distribution is approximately:

We therefore define the estimated effective sample size *Neff* as the sample size of a multinomial sample with identical variance:

i.e., that the effective sample size is the harmonic sum of *N* and . Preliminary results indicate that this approximation is very similar to estimates using Monte Carlo sampling to numerically generate samples from a DM distribution, and then this Monte Carlo variance to the variance of a multinomial distribution.

*Two alternative parameterizations*

[Explain why I re-define ]

Parameterization #1

Parameterization #2

Both parameterizations of the DM distribution were incorporated into SS (version 3.?), an integrated age-structured stock assessment framework frequently used to conduct assessments in many parts of the world (Methot and Wetzel, 2013).

*Case study: Pacfic hake*

Pacific hake (*Merluccius productus*) is a semi-pelagic schooling species of commercial importance to fisheries off of the US West Coast and Western Canada. Recent management includes an international treaty informed by annual stock assessments conducted using SS. Data used in the assessment includes catches from 1966 to 2014, an intermittent acoustic survey conducted between 1995 and 2013, 10 years of survey age-composition samples, and ‘empirical’ fishery weight-at-age data, which are assumed to be known without error (Taylor et al., 2015).

*Model application*

Four assessment models were fit to data for Pacific hake, where each model used a different approach to data-weighting: (i) unweighted, (ii) tuned using McAllister-Ianelli (1997), (iii) DM, and (iv) weight of zero for the age-composition data. Specify the McAllister-Ianelli approach.

*Simulation testing*

The performance of the DM distribution inside SS was explored using simulated data. Specify the simulation scenarios, operating models, and estimation methods.

*Model evaluation*

Estimation procedures were evaluated by comparing estimated parameters and derived quantities of interest to management to their true values from as defined in the operating model. Estimation error was quantified using relative error (, where and are estimated and true parameter values respectively).

Explain ESS across models compared to the true.

**Results**

Effective sample size

Explain linear (parameterization #1) versus asymptotic (parameterization #2) relationship between input sample size and (Fig. 1).

Pacific hake application

The method used to weight the fishery age-composition data had little effect on estimates of spawning stock biomass during years for which the model had the most data (Fig. 2). Conversely, estimates of spawning stock biomass and *F* were most different across methods of data weighting in the early years for which the model had relatively little data (Fig. 2).

Simulation

**Discussion**

**Acknowledgements**

Partial support for KFJ was provided by the National Marine Fisheries-Sea Grant Population Dynamics Fellowship (?). This publication was partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement No. ?, Contribution ?.

**References**

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124-1138.

Francis, R.I.C.C. 2014. Replacing the multinomial in stock assessment models: a first step. Fish. Res. 151, 70-84.

Hulson, P.-J.F., Hanselman, D.H., Quinn, T.J., II. 2012. Determining effective sample size in integrated age-structured assessment models. ICES J. Mar. Sci. 69, 281-292.

Legault, C.M. 2014. The ability of two age composition error distributions to estimate selectivity and spawning stock biomass in simulated stock assessments. Fish. Res. 158, 172-180.

McAllister, M.K., Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data nd the sampling – importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54, 284-300.

Methot, R.D., Wetzel, C.R. 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86-99.

Schnute, J. T., Richards, L. R. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. 52, 2063-2077.

Stewart, I.J., Hamel, O.S. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. Can. J. Fish. Aquat. Sci. 71, 581-588.

Taylor, I.G., Grandin, C., Hicks, A.C., Taylor, N., Cox, S. 2015. Status of the Pacific Hake (whiting) stock in U.S. and Canadian waters in 2015. Prepared by the Joint Technical Committee of the U.S. and Canada Packfic Hake/Whiting Agreement; National Marine Fishery Service; Canada Department of Fisheries and Oceans. 159 pp.

Table 1. Life-history, fishery, and modelling parameters used for the simulation.

Table 2. Combinations of operating and estimating models executed for the simulation.

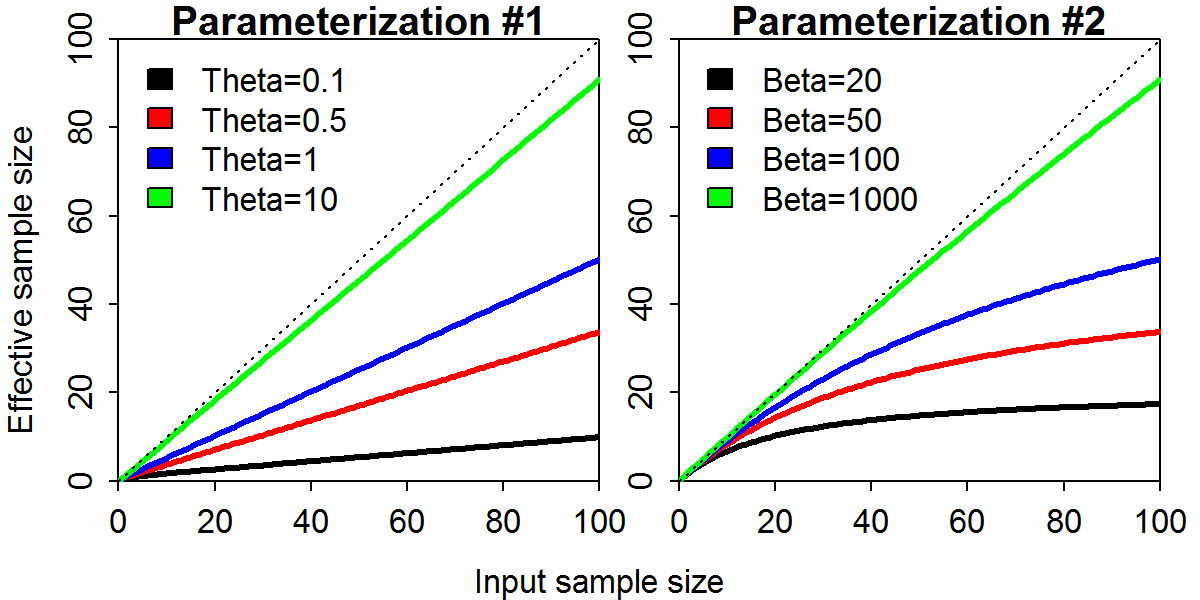
****

Fig. 1. Input sample size (x-axis) and effective sample size (; y-axis) for two paramaterizations of the Dirichlet-Multinomial (DM) distribution across varying values for the DM parameter specific to each parameterization. The dashed line represents the 1:1 line where the input sample size is the same as the .

Fig. 2 –

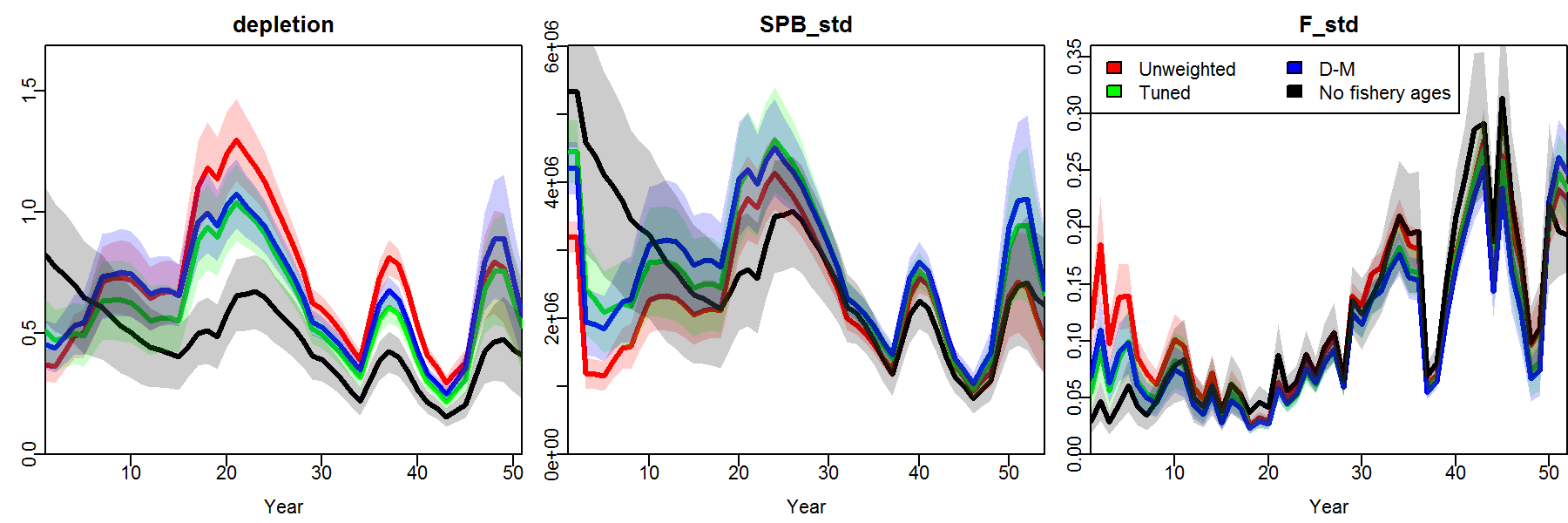


Fig. 2. Comparison of depletion (left), spawning stock biomass (SPB; middle), and fishing intensity (FSPR; right) for the Pacific hake assessment given four alternative methods of weighting the age-composition data: (i) unweighted (red), (ii) tuned (green); (iii) Dirichlet-Multinomial distribution (blue); and (iv) weight of zero for the age-composition data (black).

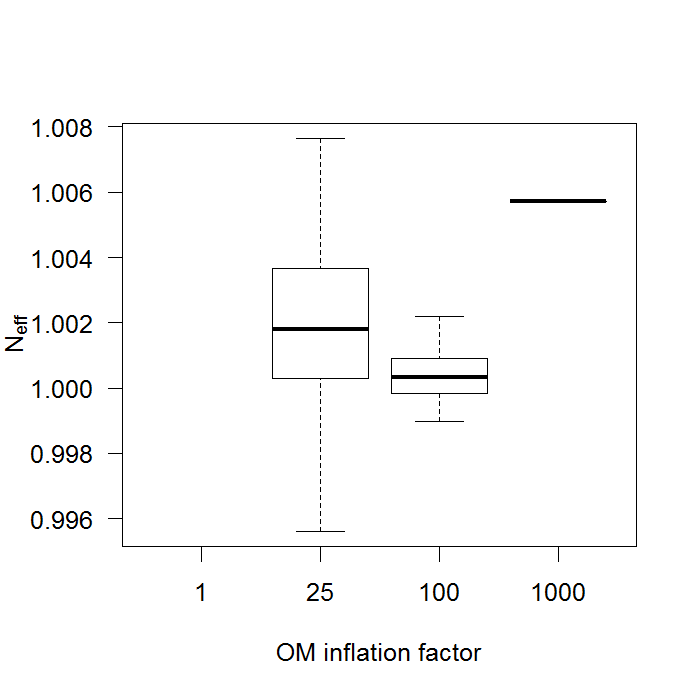


Fig. 3. Estimated effective sample size () from parameterization #2 of the Dirichlet-Multinomial (DM) distribution implemented in Stock Synthesis as the amount of information contained in the data decreased (i.e., operating model (OM) inflation factor).

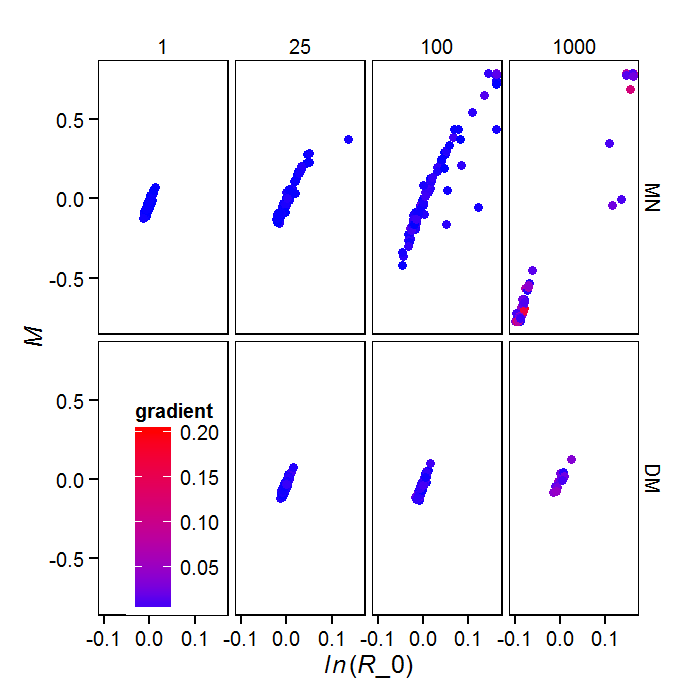


Fig. 4. Relative error in estimates of and *M* across estimation methods (rows; multinomial (MN) and Dirichlet-Multinomial (DM)) and levels of inflation for the fishery age-composition data in the operating model (columns). Gradients are shown for the estimation method with red indicating models that may not have properly converged. Lower left paneldoes not contain results because the DM estimation method was not used when the inflation factor was one.