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

James T. Thorson1, Kelli F. Johnson2, Richard D. Methot1

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 Ianelli-McAllister 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; Multinomial; stock assessment

**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 stock assessments are highly sensitive to the weighting of different data sources whenever data 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 sets included in stock assessments.

Methods for data weighting should allow for correlations (Francis, 2011).

**Methods**

*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 Stock Synthesis. 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 testing and application*

**Results**

**Discussion**

**Acknowledgements**

**References**

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences. 68: 1124-1138.

Taylor, I. G., Grandin, C. Hicks, A. C., Taylor, N., and 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 p.

**Tables**

**Figures**