Fisheries Research 192 (2017) 41–51

Contents lists available at ScienceDirect

Fisheries Research

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

Full length article

Is down-weighting composition data adequate for dealing with model misspecification, or do we need to fix the model?

Sheng-Ping Wang[a,*], Mark N. Maunder[b,c]

[a] Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Keelung 202, Taiwan
[b] Inter-American Tropical Tuna Commission, La Jolla, CA 92037, United States
[c] Center for the Advancement of Population Assessment Methodology, Scripps Institution of Oceanography, La Jolla, CA 92093, United States

ARTICLE INFO

Article history:
Received 17 July 2016
Received in revised form 1 December 2016
Accepted 3 December 2016
Handled by A.E. Punt
Available online 22 December 2016

Keywords:
Statistical catch-at-age
Stock assessment
Data weighting
Model misspecification
Monte carlo simulation
Summer flounder

ABSTRACT

Data weighting has become a major area of focus for fisheries stock assessment research. The sensitivity of results to data weighting is due mainly to model misspecification. We develop an age-structured population dynamics model for summer flounder (*Paralichthys dentatus*) in the U.S.A. mid-Atlantic region to explore the influence of model misspecification and data weighting on estimation performance using simulation. Our results confirm those of numerous previous studies that model misspecification can lead to substantial bias in estimates of quantities of management interest. We also show that data weighting generally cannot counter the effects of model misspecification, except by a small amount when composition data from a fishery for which the selectivity is misspecified are down-weighted. Additionally, estimating natural mortality was not useful as a way to address model misspecification, except when natural mortality itself was misspecified. In conclusion, down-weighting composition data is generally not adequate for dealing with model misspecification. Therefore, we recommend that more work be done to ensure that models are correctly specified. Unfortunately, there is still a lack of understanding and uncertainty in the fundamental biological and fishing processes of most, if not all, fish stocks, making the removal of model misspecification difficult.

1. Introduction

Integrated stock assessment models are frequently used to derive estimates of management quantities and their associated estimates of uncertainty (Maunder and Punt 2013; Punt et al., 2013; Stewart and Hamel, 2014). This methodology allows several sources of data to be included in a single analysis, including indices of abundance, age- or size-composition data, and biological information (Fournier and Archibald, 1982; Maunder and Punt, 2013; Punt et al., 2013; Stewart and Hamel, 2014). Integrated analysis, which is essentially the construction of a joint likelihood for the observed data, can be used in any likelihood-based framework (Maunder and Punt, 2013). However, fisheries stock assessment models can depend strongly on the relative weights assigned to each data set, which also affects all the usual tools of statistical inference that forms part of these assessments (Francis, 2011). This is particularly true when models are misspecified and, because models are a simplification of reality, they are all misspecified to some degree. Data weighting is consequently a common task that arises when developing stock assessment models.

The relative weighting of each likelihood component may be specified as the variance or sample size assigned to each observation (Stewart and Hamel, 2014). Thorson (2014) emphasized that composition data must be weighted by their "effective sample size" to avoid a situation in which composition data have a greater influence on estimated abundance trends than abundance index data, and effective sample sizes must be estimated from available data to down-weight data that are not statistically independent. However, there is no consensus among practitioners as to the best approach to data weighting (Francis, 2011; Francis, 2017; Maunder and Piner, 2017). Often, some data sets are arbitrarily down-weighted to reduce their influence on results due to concerns about model misspecification (Francis, 2011; Piner et al., 2011; Lee et al., 2014). However, it is not clear if down-weighting data deals adequately with any forms of model misspecification. Other methods to deal with model misspecification include correcting the misspecification and estimation of additional parameters.

Length-frequency data are often down-weighted to account for model misfit, presumably caused by model misspecification, as seen in residual patterns. For example, Aires-da-Silva and Maunder

* Corresponding author.
E-mail address: wsp@mail.ntou.edu.tw (S.-P. Wang).

http://dx.doi.org/10.1016/j.fishres.2016.12.005

(2014) down-weighted length-composition data to account for a residual pattern in recruitment, possibly caused by unmodeled spatial structure and Sharma et al. (2014) down-weighted length-frequency data to account for a pattern in length-composition residuals caused by unmodeled temporal variation in selectivity or sampling.

The appropriate weighting of data sets in stock assessment models is an important component of model development (Francis, 2017). To explore the importance of data weighting, diagnostics can be created by constructing a profile for each data component on a key parameter (e.g., unfished biomass, B_0, or recruitment, R_0) (Maunder and Starr 2003; Francis, 2011; Piner et al., 2011; Lee et al., 2014; Wang et al., 2014b; Ichinokawa et al., 2014). Francis (2011) proposed that the primacy of data weighting should be assigned to abundance data, while Lee et al. (2014) suggested that down-weighting composition data could lead to substantial uncertainty in the model estimates and could potentially result in degraded estimation performance. Wang et al. (2014b) also indicated that composition data can provide substantial information on estimates of absolute abundance when the model is correctly specified, but can substantially bias estimates when there is model-mis-specification.

The abundance information extracted from relative indices of abundance and age- or size-composition data are not expected to be informative about population dynamics processes without adequate knowledge of growth, recruitment, natural mortality, selectivity, and sampling processes, which are also important for determining fisheries management and reference points, but are poorly known for many, if not most, species. Growth is generally considered to be one of the most well-estimated processes due to the prevalence of ageing data, but there are inherent difficulties in reliably estimating the parameters of stock-recruitment relationship and natural mortality for most stocks. Many of the direct estimates of natural mortality are based on data that are used in contemporary integrated stock assessments (e.g., age composition) (Aanes et al., 2007; Brodziak et al., 2011; Lee et al., 2011; Maunder and Piner, 2015). Lee et al. (2011) indicated that the unrealistic model estimates are a result of model misspecification, rather than an inability to estimate natural mortality, while Piner et al. (2011) stated that unrealistic estimates of natural mortality are a good indication of severe model misspecification if a simulation analysis shows that natural mortality can be estimated with reasonable precision and accuracy. In addition, selectivity has traditionally been assumed to be well estimated in integrated stock assessment models when ample age- or size-composition data are available, but recent research has shown that misspecification of the selectivity curve can have a substantial impact on stock assessment results and management advice (Lee et al., 2014; Wang et al., 2014a,b; Maunder and Piner, 2015).

Using summer flounder (*Paralichthys dentatus*) in the U.S.A. mid-Atlantic region as an example stock, we use simulation to determine if: (1) data weighting influences assessment outcomes, (2) model misspecification (related to the stock-recruitment relationship (steepness parameter), selectivity, and natural mortality) influences assessment outcomes, (3) data weighting and model misspecification in combination influence assessment outcomes, (4) data weighting can correct for model misspecification, and (5) estimation of natural mortality can correct for model misspecification. In addition, the R_0 profile is used to explore the contribution of data components when an assessment model is either correctly or incorrectly specified. The results of this study will provide information on the appropriateness of using data weighting to deal with model misspecification.

2. Materials and methods

2.1. Summer flounder

Summer flounder is one of the most important commercial and recreational species on the U.S.A. Atlantic coast. Total landings peaked in 1983 at 26,100 mt, then declined substantially during the late 1980s reaching 6,500 mt by 1990 and have been relatively constant since the early 1990s (Maunder and Wong, 2011). The assessment and management of the summer flounder fishery has been very contentious since the implementation of the Fisheries Management Plan in 1989, given the estimated highly depleted status of the summer flounder stock (Terceiro, 2002). Information about the fishery, stock assessment, and management can be found in the reports of the Northeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, U.S.A. (e.g., Terceiro, 2015).

2.2. Simulation analysis

The stock assessment covers the years 1976–2008. Data used in previous assessments consist of catches from six fisheries, three relative indices of abundance from trawl surveys, age-composition data from six fisheries and three surveys, and mean body weight data from one fishery (see Maunder (2012) for the details of the data and the model assumptions). To explore the influence of data weighting on model estimation we simplified the stock assessment model for summer flounder to include (1) a single fishery that includes catch data combined from six fisheries (1976–2008); (2) a single survey index of abundance (1976–2007); (3) age-composition data from the fishery (1982–2006) and survey (1976–2007); and (4) a single sex. Life-history parameters used in this study are consistent with the findings of Maunder and Wong (2011) and Maunder (2012).

The simplified version of the stock assessment model is used as both the simulator and estimator. (Appendix A describes the basis for the population dynamics model.) The "true" values for natural mortality and the steepness of the stock-recruitment relationship are $M = 0.25\text{yr}^{-1}$ and $h = 0.9$, respectively, and selectivity is assumed to be dome-shaped for the fishery. The model is implemented in Stock Synthesis (SS, version 3.24f) (Methot and Wetzel, 2013), and the parametric bootstrap feature of SS was used to generate the simulated data. The fishing mortality rates estimated by fitting the model to the original data (Step 1 below) were halved to prevent the population from being reduced to a very low level, which led to convergence issues during exploratory simulation analyses. This assumption does not reduce the appropriateness of the analysis because the aims of the study relate to general concepts and are not meant to be specific to summer flounder. The simulation analysis involved the following steps:

(1) Fit the population dynamics model to the actual data to obtain estimates of its parameters (Table A1) and multiply the estimated fishing mortality rates by 0.5.
(2) Set the parameters of the population dynamic model to the values obtained from Step 1 and replace the recruitment deviations by random values generated from a normal distribution with mean 0 and standard deviation ($\sigma = 0.6$) and hence construct a "true" population – these are the "true" values for the replicate.
(3) Use the SS bootstrap procedure to generate a data set. Coefficients of variation (CVs) between 0.1–0.3 (index of abundance) and effective sample sizes of 200 (age-composition data), adopted previously by Maunder (2012), are used when generating data sets (Table S1).