G Model
FISH-4429; No. of Pages 12
ARTICLE IN PRESS
Fisheries Research xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Fisheries Research
journal homepage: www.elsevier.com/locate/fishres

Full length article

Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets

Mark N. Maunder[a,b,*], Kevin R. Piner[c]

[a] Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA
[b] Center for the Advancement of Population Assessment Methodology, Scripps Institution of Oceanography, La Jolla, CA, USA
[c] NOAA Fisheries, Southwest Fisheries Science Center, 8901 La Jolla Shores Dr, La Jolla, CA 92037-1508, USA

ARTICLE INFO

Article history:
Received 27 April 2016
Accepted 28 April 2016
Handled by A.E. Punt
Available online xxx

Keywords:
Fisheries
Maximum likelihood
Population dynamics
State-space model
Statistical inference
Stock assessment

ABSTRACT

Contemporary fisheries stock assessments often use multiple diverse data sets to extract as much information as possible about biological and fishery processes. However, models are, by definition, simplifications of reality and, therefore, misspecified. Model misspecification can cause degradation of results when multiple data sets are analyzed simultaneously. The process, observation, and sampling components of the model must all be, at least, approximately correct to minimize bias. Unfortunately, even the basic processes that are usually considered well understood (e.g., growth and selectivity) are misspecified in most, if not all, stock assessments. These misspecified processes, in combination with use of composition data, result in biased estimates of absolute abundance and abundance trends, which are often evident as “data conflicts.” This is compounded by over-weighting of composition data in many assessments owing to misuse of data-weighting approaches. The ‘law of conflicting data’ states that since data are facts, conflicting data implies model misspecification, but must be interpreted in the context of random sampling error. Down-weighting (or dropping) conflicting data is not necessarily appropriate because it may not resolve the model misspecification. Model misspecification and process variation can be accounted for in the variance parameters of the likelihoods (sampling error), but it is unclear when, or even if, this is appropriate. The appropriate method to deal with data conflicts depends on whether it is caused by random sampling error, process variation, observation model misspecification, or misspecification of the system (dynamics) model. Diagnostic approaches are urgently needed to evaluate goodness of fit and to identify model misspecification. We recommend external estimation of the sampling error variance in likelihood functions, modelling process variation in integrated models, and internal estimation of the standard deviation of the process variation. The required statistical framework is computationally intensive, but practical approximations are available, computational algorithms are being improved, and computer power is increasing. We provide a framework for model development that identifies and corrects model misspecification and illustrate the framework, using simulated data.

1. Introduction

The overarching goals of fisheries management have been optimization of yield and sustainability of stocks. These goals have led to the concept of maximum sustainable yield (MSY) and its proliferation throughout the fisheries literature (Larkin, 1977; Punt and Smith, 2001). Contemporary fisheries management objectives are much more complex (Hilborn and Walters, 1992), but often the main objectives relate to optimizing benefits to humanity (e.g., yield or profit) on a sustainable basis while mitigating adverse effects (e.g., bycatch), which parallels the concept of MSY to some extent. Attainment of these goals have been evaluated, using ever more complex quantitative analysis methods.

Stock assessments based on population dynamic models (Hilborn and Walters, 1992; Quinn and Deriso, 1999; Haddon, 2001) have been the gold standard for estimating MSY, and are conducted for most economically-valuable species. With increasing computer power and the popularization of integrated stock assessment modeling (Maunder et al., 2009), the complexity of modern

* Corresponding author at: Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Dr., La Jolla, CA 92037-1508, USA.
E-mail address: mmaunder@iattc.org (M.N. Maunder).

http://dx.doi.org/10.1016/j.fishres.2016.04.022
0165-7836/© 2017 Elsevier B.V. All rights reserved.

stock assessment modeling is rapidly increasing (e.g., Hampton and Fournier, 2001; Methot and Wetzel, 2013; Bull et al., 2005; Begley and Howell, 2004). Considerable research is directed toward improving the methods used to make model predictions (Quinn, 2003; Maunder and Piner, 2014) and to statistically compare predictions to data (Punt and Hilborn, 1997; Maunder, 2011; Francis, 2014; Thorson et al., this volume) and for parameter estimation (e.g., Skaug and Fournier, 2006; Fournier et al., 2012; Kristensen et al., 2014, 2015). The complexity of modern stock assessment models has been driven in part by the desire to incorporate a broad range of data types collected using different sampling techniques (Maunder and Punt, 2013; Punt et al., 2013). This is because we lack direct data on absolute stock abundance and fishing mortality, but we know that information about these quantities are contained in the available data. The use of integrated modelling is predicated on the "hope" that a "synthetic" solution will emerge.

A major issue in attempting to integrate disparate data series is apparent as conflicting signals from different data sets, given the model's structure. Although, conflicts can and do occur among all data types, conflicts among indices of relative abundance and composition data is particularly prevalent and concerning (Francis, 2011; Lee et al., 2014). Conflicts among data sets, which are often a symptom of model misspecification and evident as model misfit, can affect the estimates of important parameters and derived quantities. The current solution to data conflicts often is to eliminate one of the conflicting data sources, or, nearly equivalently, reduce its weight when fitting the model (e.g., Sharma et al., 2014), but this is dealing with the symptoms rather than the underlying cause of data conflicts (Wang et al., 2015).

We argue that the practice of data elimination does not address the more important issue highlighted by internal conflicts in models. Rather, data conflicts may be indicative of misspecification of the system (dynamics) model, which controls the population dynamics. Misspecification of important model processes will lead to biased estimates of information needed for management. We highlight how composition data are often the cause of conflicts and then discuss the various ways sampling error, unmodeled process variation, and model structure misspecification can lead to conflicts between composition data and indices of abundance. We offer recommendations on designing a stock assessment that systematically identifies root causes of data conflicts and how to solve them. We illustrate this framework by evaluating (blindly) misspecified models fit to simulated data.

2. Causes of data conflicts

Data conflicts occur when two or more data sets, given the model structure, provide information about a model state or process that disagree. For example, the data are in conflict about abundance when an index of relative abundance supports high abundance while catch-at-length data support low abundance. It has become common practice to simply conclude that one of the conflicting data sources is unrepresentative of the system and reduce the weight assigned to that data source. However, if a data set was considered good after initial analyses, it should be considered representative of the system and considered as "facts" to be used in the model.

Conflicts among data imply that either the model is misspecifed (Table 1) or that the precision of the data has been overstated (leading to a false impression of data conflicts). Precision of the data is related to sampling error, which is variation due to taking a sample, rather than a census. Sampling error can usually be reduced by increasing the sample size, which is important because the causes of data conflicts must be interpreted within the context of this error. Some data may have large sampling error and, therefore, low information content (e.g., a large confidence interval around an estimate) so the apparent conflicts in the data can be attributed solely to sampling error. Conflicts among data within the extent of sampling error are not indicative of data and model structure that are in disagreement about the underlying system dynamics. Therefore, it is important to get the assumptions about the sampling error and the distribution associated with that error correct to interpret apparent data conflicts. For example, data from an index of relative abundance may be inconsistent with the underlying population dynamics model if the error is assumed to be multiplicative, but not so if it is assumed to be additive (Fig. 1). In this case, the

Table 1
The law of conflicting data.

The law of conflicting data
Axiom
Data are facts
Implication
Conflicting data implies model misspecification
Caveat
Data conflict needs to be interpreted in the context of random sampling error
Significance
Down weighting or dropping conflicting data is not necessarily appropriate because it may not resolve the model misspecification

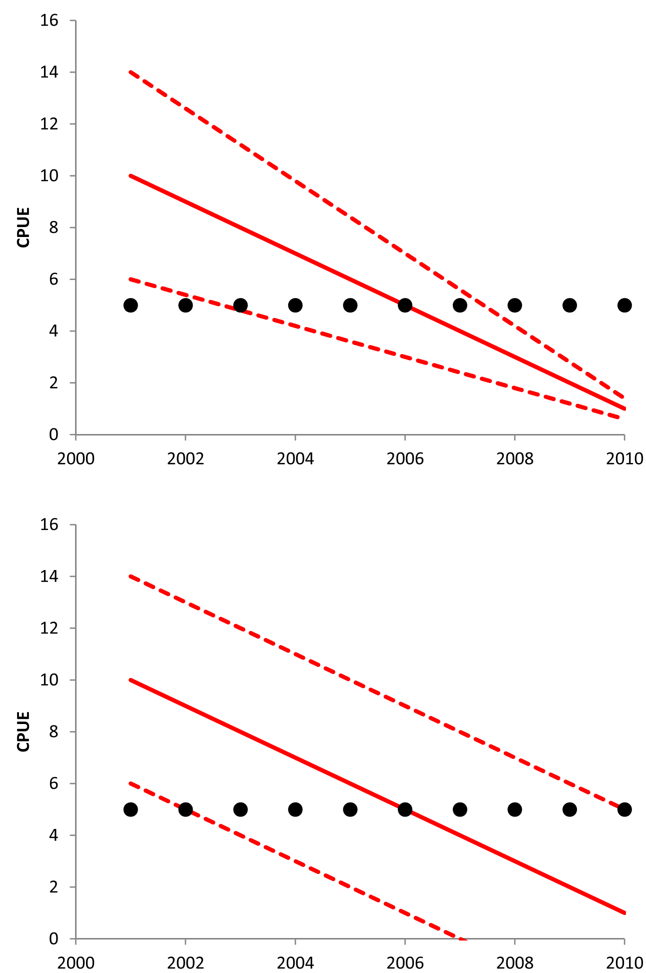

Fig. 1. Illustration of how apparent data conflicts in catch-per-unit-of-effort (CPUE), a relative index of abundance, (upper panel with multiplicative error) could simply be a consequence of large sampling error (lower panel with additive error) and that correct specification of sampling variation is necessary to understand the apparent data conflicts. Points = data, solid line = true relative abundance, dashed line = 95 percentiles of the assumed sampling distribution.