



From sequential to integrated Bayesian analyses: Exploring the continuum with a Pacific salmon spawner-recruit model



Benjamin A. Staton^{a,*}, Matthew J. Catalano^a, Steven J. Fleischman^b

^a School of Fisheries, Aquaculture, and Aquatic Sciences, Auburn University, 203 Swingle Hall, Auburn, AL 36849, United States

^b Division of Sport Fish, Alaska Department of Fish and Game, 333 Raspberry Road Anchorage, AK 99518, United States

ARTICLE INFO

Article history:

Received 3 May 2016

Received in revised form 31 August 2016

Accepted 2 September 2016

Handled by A.E. Punt

ABSTRACT

Stock assessment scientists are faced with decisions regarding how to incorporate fishery information into models. One primary decision revolves around how estimates that are summaries of raw data should be treated (e.g., abundance estimates derived from relative indices). The choice in this case is to either use estimates from a sequence of models as data in a final model (i.e., the model used for setting management goals) or to integrate the raw data into a more complex final model. Each approach has advantages and disadvantages that constitute a suite of trade-offs. These trade-offs are investigated here by comparing two sequential analyses (one that ignores measurement error and one that incorporates it) to an integrated analysis for a stock assessment of Pacific salmon using simulation-estimation, and the Kuskokwim River Chinook salmon stock of western Alaska as a case study. The major difference between approaches was that an abundance reconstruction was estimated separately from the spawner-recruit analysis in the sequential approaches, whereas the integrated approach did so in a single model. Primary findings showed that approaches that addressed the measurement error in the raw data returned very similar estimates of abundance, population dynamics parameters, and management reference points, both in terms of point estimates and uncertainty. When measurement error was ignored, similar point estimates were returned. However, this approach underestimated uncertainty in the spawner-recruit analysis but resulted in more uncertainty in the abundance reconstruction. These findings were consistent for both the Kuskokwim River case study and simulation-estimation analyses. The primary advantage of the integrated analysis was the added realism of sharing calendar year abundance data among brood years, but came at the cost of slow model run times. This exercise showed that while there is a trade-off between sequential and integrated analyses in terms of model complexity and realism, the benefits may not be large enough to warrant an integrated analysis in all cases, given that the terminal model carries forward uncertainty in the input estimates.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Stock assessment scientists are often faced with decisions regarding how to incorporate information into models used in setting management targets. This is particularly true when there are various types of information available, some of which are raw data and some are estimates from other models (e.g., unstandardized versus standardized catch-per-unit-effort; Gavaris, 1980). Additionally, these cases arise when two processes give rise to the observed data set, but only one is of interest, such as when data are produced by a biological process (i.e., population dynamics) and a

measurement process (i.e., sampling), but inference is desired only for the population process. In these cases, the practitioner has the option to either separate the information into a sequence of analyses (i.e., sequential analysis) or integrate the raw information into one analysis (i.e., integrated analysis). This choice may at first seem arbitrary and inconsequential, but may prove to have substantial ramifications on the results and interpretation of the assessment model (Brooks and Deroba, 2015).

Under the sequential analysis approach, raw data are pre-processed in some fashion to produce estimates of interest that are then passed to a final model, which forms the basis for inference on the population dynamics and management recommendations. In the next model in the sequence, these pre-processed estimates are either treated as perfectly known observations or assigned some uncertainty due to measurement and estimation error in the like-

* Corresponding author.

E-mail address: bas0041@auburn.edu (B.A. Staton).

likelihood or through informative priors (Michielsens et al., 2008). A common example of this sequential analysis practice in stock assessment modeling is the estimation of spawning stock biomass and a recruitment time series from a virtual population analysis-type model that are then passed to a spawner-recruit (or similar) model to estimate parameters or determine variables that govern the population dynamics (e.g., Wada and Jacobson, 1998; Gabriel et al., 1989). The major advantage to the sequential approach is that it allows for a simpler final model, which is likely to converge well and have relatively short run times (Michielsens et al., 2008). However, as noted by Maunder (2001) and Maunder and Punt (2013), the sequential approach has some potential disadvantages including loss of information in the raw data that could be exploited by the final analysis if not already summarized, inadequate treatment of uncertainty in the observation level-data, and reduced diagnostic ability. To address the issue of inadequate treatment of uncertainty, stock assessment practitioners have options that allow for uncertainty to transfer from one step to the next, such as hierarchical and/or Bayesian methods (Thorson and Minto, 2014; Punt and Hilborn, 1997).

An alternative approach to this problem is to undertake a more fully integrated analysis. Integrated analyses attempt to incorporate the raw data (as fully as appropriate) into a single analysis (Fournier and Archibald, 1982). An integrated analysis is often conducted with the goal of preserving the complete information content of the data and adequately treating uncertainty in the analysis due to measurement error and process variation. These types of integrated population models are gaining popularity in ecological assessments in both applied and academic settings (e.g., Schaub and Abadi, 2010; Royale and Dorazio, 2008). However, due to the inherent complexity of integrated analyses, they present unique problems such as difficulties with formulating the joint likelihood, convergence issues, and computational complexities (Maunder and Punt, 2013). With recent advances in computing power and statistical approaches to fitting complex non-linear models, particularly Bayesian estimation using Markov chain Monte Carlo (MCMC) methods, these barriers are becoming less troublesome for the applied stock assessment practitioner.

It may be impractical to integrate all raw data into a single analysis, which leaves the investigator the task of deciding which data to include in raw form *versus* which information should be summarized before inclusion in the analysis. As such, it is more appropriate to view the contrast between sequential and integrated analyses as a continuum rather than a dichotomy. On the extreme sequential side of the continuum, all raw data are summarized or combined into estimates prior to being included in the final model. On the extreme integrated side of the continuum, every raw data point is included as an observation of the system and is used for inference. There are many intermediate scenarios between these extremes for a particular analysis where some data may be summarized pre-analysis and some data may be included in raw form. As one moves the analysis in either direction along this continuum, the relative advantages and disadvantages of either approach should become apparent in the output of the whole analysis.

Assessment approaches for Pacific salmon *Oncorhynchus* spp. typically involve collecting observations of annual spawner abundance and harvest followed by combining this information into a brood table to obtain the number of brood year recruits based on the year and age at which the fish returned (e.g., Clark et al., 2009; Bue et al., 2012). Note that only the calendar year processes of escapement and harvest can be observed directly, which are then used to obtain brood year recruitment that cannot be directly observed. Spawner abundance and harvest are usually not complete counts and must be estimated or expanded (e.g., *via* run reconstruction; Shotwell and Adkison, 2004). These estimates are then passed to a spawner-recruit analysis to estimate productiv-

ity, carrying capacity, and recruitment variation for the derivation of biological reference points used in setting escapement goals (Clark et al., 2009). This approach constitutes a sequential analysis, and could be conducted by either treating the reconstructed spawner and recruit abundances as known perfectly, or by treating the uncertainty in these abundances as measurement error in either a maximum likelihood (MLE) or a Bayesian estimation framework. Alternatively, in some cases it is possible to integrate the run reconstruction model directly into the spawner-recruit analysis whereby brood year spawner and recruit abundances are reconstructed from observational-level data and a spawner-recruit relationship is simultaneously fit to these internally reconstructed estimates (Fleischman et al., 2013). This approach may have the benefits of more fully addressing uncertainty due to the interaction between the raw data (*i.e.*, the observation submodel) with the spawner-recruit analysis (*i.e.*, the process submodel) and the sharing of information between calendar years due to the intrinsic link between calendar years *via* the spawner-recruit analysis (Maunder, 2001).

Here we present the application of an analysis that moves a stock assessment along the continuum from a sequential MLE analysis that ignores measurement error, to a sequential Bayesian analysis, to a more fully integrated Bayesian analysis to investigate trade-offs between these approaches to stock assessment. Potential trade-offs of moving along this continuum are investigated by independently running sequential MLE, sequential Bayesian, and integrated Bayesian versions of the same analysis and comparing relevant quantities of interest. The final model in the two Bayesian cases is fit using a state-space framework to allow for simultaneous incorporation of both measurement error in the input information and process variation in the spawner-recruit relationship (Fleischman et al., 2013), whereas measurement error is ignored entirely in the sequential MLE analysis. These analyses are conducted using both simulated data and actual data for the Kuskokwim River Chinook salmon *O. tshawytscha* stock, located in western Alaska. The objectives of this study are to (1) determine the consistency in point estimates and uncertainty among the three assessment approaches, (2) determine if the sequential Bayesian analysis constitutes an adequate treatment of uncertainty or if the more fully integrated model is justified, and (3) verify that findings resulting from the use of the actual data from the Kuskokwim River hold up when compared to many simulated data sets designed to model this system. The primary *a priori* expectations are (1) point estimates between Bayesian assessment approaches should be similar, (2) there should be greater estimated uncertainty in quantities of interest (e.g., abundance and spawner-recruit parameters) in the sequential Bayesian analysis than the sequential MLE analysis, and (3) the integrated analysis should show even more uncertainty than the sequential Bayesian analysis due to direct interaction between raw data and the spawner-recruit analysis.

2. Methods

2.1. Study system

The Kuskokwim River is the second largest drainage system in the state of Alaska (~130,000 km²), with the main stem flowing approximately 1500 km from its headwaters in the Kuskokwim Mountains to the southwestern coast where it empties into the Kuskokwim Bay of the Bering Sea (Hamazaki et al., 2012). Historically, the Kuskokwim River has provided one of the largest subsistence fisheries for Chinook salmon in the state (39 year average ~70,000 fish/year), but has seen low runs in recent years leading to substantial conservation measures including fishery closures for Chinook salmon and other species. While the directed

Download English Version:

<https://daneshyari.com/en/article/4542618>

Download Persian Version:

<https://daneshyari.com/article/4542618>

[Daneshyari.com](https://daneshyari.com)