Journal of Great Lakes Research 40 Supplement 3 (2014) 41–48

Contents lists available at ScienceDirect

Journal of Great Lakes Research

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

Comparing Bayesian and frequentist methods of fisheries models: Hierarchical catch curves

Jason C. Doll *, Thomas E. Lauer

Aquatic Biology and Fisheries Center, Department of Biology, Ball State University, Muncie, IN 47306, USA

ARTICLE INFO

Article history:
Received 4 October 2013
Accepted 13 June 2014
Available online 25 July 2014

Communicated by George Arhonditsis

Index words:
JAGS
Hierarchical model
Simulated data
Yellow perch

ABSTRACT

Bayesian inference is an emerging statistical paradigm and is becoming an increasingly used alternative to frequentist inference. Unfortunately, little is known about the efficacy of Bayesian inference and how it relates to the historical methodology of evaluating fisheries related models. Mortality information is routinely used in fisheries management to describe fish population abundance over time and has been historically estimated using catch curves and frequentist inference (i.e., maximum likelihood estimation). The objective of this study was to compare frequentist and Bayesian inference approaches to estimate instantaneous mortality (Z) from a hierarchical catch curve model. The data used in the comparison were from a long term monitoring program of yellow perch *Perca flavescens* from southern Lake Michigan in addition to a simulated dataset where parameter estimates were compared to known values. Point estimates of Z were similar among both methods. Similarly, Bayesian inference 95% credible intervals were concordant with frequentist 95% confidence intervals. However, the root mean squared error of frequentist inference increased at a higher rate than Bayesian inference with increasing variability in the simulated dataset. Our study builds on the literature that seeks to compare results between these two paradigms to assist managers to make the best decision possible when deciding what statistical paradigm to employ.

Introduction

Bayesian inference is a statistical inference paradigm that describes probability differently than traditional frequentist inference (Elison, 1996). Frequentist inference treats parameters as fixed unknowns, the data as random, and inference is drawn from 95% confidence intervals that are based on hypothetical replicates. In contrast, with Bayesian inference, parameters are random, data are considered fixed, and inference is drawn from posterior distributions of a parameter given the data, the model, and the prior belief about the parameter. Because Bayesian inference relies on specification of prior distributions, they are often considered subjective. Fortunately, much has been devoted to this subject, including non-informative priors (Berger, 2006). For example, non-informative prior distributions result in posterior distributions that are largely due to the likelihood function and thus are generally similar to the results of frequentist inference when evaluating simple models (e.g., linear regression). For a thorough review of Bayesian inference, see Gelman et al. (2004), Carlin and Louis (2008), and Kruschke (2011).

Bayesian inference is rapidly becoming an increasingly used method to address a variety of environmental science problems (Buckley et al., 2010; Fitzpatrick et al., 2010; Hilley and Young, 2008; Lee, 2008; Reckhow, 1990; Vivó-Truyols, 2012). Within ecology and fisheries, Bayesian inference has been used to disentangle the relationship between catch and detectability (McCarthy et al., 2013), assess the relationship between long-term fish assemblage structure with habitat and niche breadth (Jacquemin and Doll, 2013), evaluate the relationship of body size and geographic range with variation in abundance while incorporating phylogenetic relationships (Jacquemin and Doll, 2014), estimate abundance of fish from mark-recapture experiments (Rivot and Prévost, 2002), fit stock-recruitment models (Su and Peterman, 2012), estimate maturity parameters from a logistic regression model (Doll and Lauer, 2013), estimate mortality parameters (Bunnell et al., 2012), and determine efficiency and selectivity of gill nets (Askey et al., 2007). In limnology, Bayesian networks have been used to describe the processes involved in eutrophication in an estuary of North Carolina (Borsuk et al., 2004). The Bayesian framework has also been applied to the revaluation of phosphorus loads (Cheng et al., 2010) and predicting hypoxic volume in the Chesapeake Bay (Liu et al., 2011). Despite these examples, Bayesian inference is not in common use, nor has it been extended to cover the plethora of mathematical and statistical models associated with contemporary environmental research. Further, direct comparisons of results from models fit using frequentist and Bayesian inference have not been conducted on many ecological models, including the mortality processes of fish, despite both frequentist (Dutterer et al., 2012; Olsen et al., 2004) and Bayesian (Bunnell et al., 2012; Hall et al., 2004) methods being

* Corresponding author.
E-mail address: jcdoll@bsu.edu (J.C. Doll).

http://dx.doi.org/10.1016/j.jglr.2014.07.006
0380-1330/

applied to such models. Such a comparison is needed to identify discrepancies, particularly as these results are used for management recommendations.

While the advantages of taking a Bayesian over frequentist approach to statistical inference is well documented from many perspectives (Austin et al., 2001; Broomhall et al., 2010; Kruschke, 2010, 2013; Wagner et al., 2013), acknowledging specific differences in model fitting with small datasets is warranted. With limited information, Bayesian treatment using informative priors, a more ecologically realistic model, or sharing information across groups in a hierarchical framework minimizes biased parameter estimates when compared to frequentist models. For example using the Bayesian framework, phylogenetic accuracy increased creating a more ecologically realistic model (Alfaro et al., 2003), bias in spawning stock biomass from fisheries stock assessment decreased using uniform priors on all parameters (Nielsen and Lewy, 2002), and bias in parameter estimates was reduced in a multilevel model by sharing information across groups (Stegmuller, 2013). Using small or limited datasets in a hierarchical framework, Bayesian inference provides an advantage as the models converge more often (Doll and Lauer, 2013). Finally, frequentist inference of catch curves has resulted in missing results due to small sample size (Jackson and Noble, 2013) while Bayesian inference in a hierarchical framework has been used to estimate parameters of catch curves due to the known limitations (Bunnell et al., 2012). These shortcomings of frequentist inference methods highlight the need to evaluate and compare Bayesian inference as an alternative method of statistical inference.

An additional advantage of the Bayesian approach is the treatment of multiple comparisons and multilevel models. Under the frequentist paradigm, multiple comparisons are typically conducted by treating the variable as random. However, this treatment focuses on the overall, rather than the individual group effect. The only option of determining the individual group effect is to treat the variable as fixed. By doing so, unfortunately, the inherent multilevel structure of the data is neglected, which can result in biased parameter estimates (Gelman and Hill, 2007). Further, if the random effect was treated as a fixed effect, the post hoc multiple comparisons require the investigator to adjust their cut-off for significance for each comparison being made, typically using a Bonferroni correction or false discovery rates (Benjamini, 2010). Thus, the results would be influenced by the intention of the investigator (i.e., the cutoff for significance is different based on the number of comparisons the investigator intends to make). In contrast, Bayesian approaches to multiple comparisons are not subjected to these problems (Kruschke, 2011). With Bayesian inference, the posterior distribution of the parameters is the full joint posterior probability distribution (given the model, data and prior distribution) that specifies the probability the coefficient lies in a specific interval. There is no need to make adjustments to the decision criteria for significant differences based on the investigators' intentions of the study.

Mortality information is routinely used for successful management of fish populations and describes the rate at which population catch declines. Mortality rate is related to recruitment, growth, and harvest (Quinn and Deriso, 1999) and can be expressed as any decline in a group (e.g., year class or cohort) of fish over a constant time period (e.g., year). For example, the catch of the 2000 year class is estimated over time at age 0 (2000), age 1 (2001), age 2 (2002), etc: and the age structured data is used to describe the rate of decline. The instantaneous mortality rate (Z) is a metric that is typically used to describe the rate the population catch declines.

Catch curve models are typically used to estimate Z (Chapman and Robson, 1960; Robson and Chapman, 1961). Assumptions of the catch curve model include: constant recruitment over time, constant fishing mortality, constant natural mortality at age, and constant selectivity (Chapman and Robson, 1960). Unfortunately, these assumptions are rarely met and failure to meet these assumptions can result in biased and imprecise estimates. Combining samples from multiple years has been suggested to solve this problem (Ricker, 1975); yet, this technique imposes new assumptions, including each combined year class strength is similar. Further, without incorporating a single–multiple year hierarchical structure into the model, the investigator would not be able to determine trends of individual year classes. Attempts have been made to account for violations of some assumptions by incorporating age-specific natural mortality (Thorson and Prager, 2011) or adding a selectivity term (Cotter et al., 2007; Thorson and Prager, 2011). Despite these limitations, the catch curve model remains a widely used method to describe mortality processes (Dutterer et al., 2012; Newman et al., 2000; Olsen et al., 2004).

The objective of this study was to compare frequentist and Bayesian inference approaches to estimate instantaneous mortality (Z; Ricker, 1975) from a hierarchical catch curve model for two fisheries datasets: one long term and one simulated. Specifically, we compare the ability of the model to converge (i.e., estimate parameters), point estimates (frequentist mean parameter estimates and Bayesian median values of the posterior distribution), and precision of the estimates (frequentist 95% confidence intervals and Bayesian 95% credible intervals). We hypothesize that Bayesian inference will result in interpretable estimates for all year classes while frequentist inference will result in insignificant or non-calculable estimates when substantial noise is present or sample size is limited. We further hypothesize that frequentist mean estimate of Z will be similar to Bayesian median values of the posterior distribution. Although these point estimates will be similar, we hypothesize that Bayesian inference will provide more precise estimates of Z (i.e., narrower 95% credible intervals).

Methods

Yellow perch data

We used two datasets to compare statistical inference paradigms; including one long-term (26 years) and one simulated dataset. The first is from the Ball State University's long-term monitoring program of yellow perch *Perca flavescens* from southern Lake Michigan (Lauer and Doll, 2012). Yellow perch were sampled from up to three near shore sample zones between 1984 and 2009 semi-monthly using a semi-balloon bottom trawl. From 1984 to 1988, two sites were sampled and three sites were sampled from 1989 to 2009. Nighttime bottom trawling was conducted at the 5-m depth contour for a total of 6 h of effort at each site each year. This resulted in a total effort of 12 h from 1984 to 1988 and 18 h from 1989 to 2009. Only age-2 to age-9 fish were used to determine year class catch per unit effort (CPUE) at age by year, younger fish are not fully recruited to the trawl (Shroyer and McComish, 2000). Up to 10 fish per 10 mm length class per month were aged. Scales were used from 1984 to 1993 and opercular bones were used from 1994 to 2009 and aged independently by two readers. Aging methods were changed due to opercular bones having a lower coefficient of variation (Baker and McComish, 1998). Discrepancies were discussed between both readers until a consensus was reached. Un-aged fish were assigned an age using yearly and monthly specific age–length keys.

Simulated data

One hundred year classes of fish were simulated, each consisting of between 3 and 10 observations. The number of ages per year class was randomly assigned by independent draws from a uniform distribution. Survival for each year class was randomly drawn from a beta distribution and the slope was calculated as ln(survival). Intercepts for each year class were randomly drawn from a uniform distribution. Parameters of the age structure, beta distribution, and uniform distributions were selected to be consistent with estimates from the Ball State University yellow perch dataset. Linear predictors