Fisheries Research 125–126 (2012) 173–183

Contents lists available at SciVerse ScienceDirect

Fisheries Research

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

Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics

Kotaro Ono[a,*], André E. Punt[a], Etienne Rivot[b]

[a] School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA
[b] Agrocampus Ouest, UMR Agrocampus Ouest, INRA, 985 ESE Ecologie et Santé des Ecosystèmes, Laboratoire Ecologie Halieutique, CS 84 215, 35 042 Rennes Cedex, France

A R T I C L E I N F O

Article history:
Received 19 September 2011
Received in revised form 30 January 2012
Accepted 21 February 2012

Keywords:
Bayesian
Model performance
State-space model
Octopus
OpenBUGS

A B S T R A C T

Bayesian observation error (OEM), process error (PEM) and state-space (SSM) implementations of a Fox biomass dynamics model are compared using a simulation–estimation approach and by applying them to data for the octopus fishery off Mauritania. Estimation performance is evaluated in terms of bias, precision, and reliability measured by the extreme tail-area probability and the mean highest posterior density interval. The PEM generally performs poorest of the three methods in terms of the these performance metrics. In contrast, the OEM is precise, but under-represents uncertainty. The OEM is outperformed by the SSM in terms of its ability to provide posterior distributions which adequately capture parameter uncertainty. It is key to consider the above four metrics when comparing estimation performance in a Bayesian context. Finally, although model performance measures are useful, there is still a need to examine goodness of fit statistics in actual applications.

1. Introduction

Biomass dynamics models have been used widely in fishery stock assessment, not only because of their algebraic simplicity, but also because they provide estimates of the management reference points, MSY, B_{MSY} and F_{MSY} (Fox, 1970; Hilborn and Walters, 1992; Schaefer, 1954) and require few data for parameter estimation (usually just a time-series of catches and an index of relative abundance). Although no longer the method of choice for stock assessment, they are still used in several jurisdictions when there are few data except for an index of relative abundance and catches (e.g., sea urchin or octopus in Mexico (Jurado-Molina et al., 2009; Jurado-Molina, 2010), prawns off northern Australia (Zhou et al., 2009), sea cucumbers off British Columbia, Canada (Hajas et al., 2011), and octopus fisheries off NW Africa (Ould Mahmoud et al., 2006)).

Biomass dynamics models are most commonly expressed in the form of a discrete (typically on a yearly basis) equation defining the biomass at time step $t+1$ as the biomass at time step t plus the production during the interval $[t,t+1]$ less the catches during this interval. These models are fitted to one or more time-series of observed abundance indices, which are assumed proportional to biomass with a proportionality constant (often denoted q) interpreted as catchability when the abundance indices are catch-rates.

Several methods have been developed to fit biomass dynamics models to observations to estimate key parameters and infer the unknown biomass trajectories. Random errors can be considered in the process equation (due to fluctuations in the size of population because of variation in recruitment, natural mortality, and growth), typically through multiplicative log-normal errors (Meyer and Millar, 1999; Polacheck et al., 1993; Punt, 1989, 2003), and/or in the observations (due to sampling variability and variation in catchability), again usually assumed to be log-normal.

Classical fitting approaches include equilibrium methods or methods that take account of the dynamics of the population (Hilborn and Walters, 1992; Polacheck et al., 1993). The latter methods usually assume that all of the error is due to process error or due to observation error. Comparisons of process and observation error estimators using simulation have suggested that the latter generally perform better than the former (more precise and less biased) (Polacheck et al., 1993). Methods have also been developed to simultaneously account for both process and observation error in a state-space model framework. Both maximum likelihood-based (e.g., de Valpine, 2002; de Valpine and Hastings, 2002; Freeman and Kirkwood, 1995) and Bayesian (Clark, 2003; Hammond and Trenkel, 2005; McAllister and Kirkwood, 1998; Meyer and Millar, 1999; Millar and Meyer, 2000; Zhou et al., 2009; Jiao et al., 2011) approaches have been used to implement biomass dynamics models within a state-space modelling framework. Conventionally, approaches which assume process error only are considered to be most appropriate for short-lived species which show large fluctuations in abundance, apparently unrelated to fishing pressure,

* Corresponding author. Tel.: +1 206 543 4270; fax: +1 206 685 7471.
E-mail address: kotarono@uw.edu (K. Ono).

0165-7836/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2012.02.022

while approaches which assume observation error are considered to most appropriate for species such as whales for which large natural fluctuations are rare.

The focus of this paper is on Bayesian estimation methods. Bayesian methods have received considerable attention in the ecological literature during the last 20 years, in particular in the fishery sciences, because they provide ready quantification of the uncertainty of parameters and model outputs, and hence provide the information needed to conduct probabilistic decision analyses (Harwood and Stokes, 2003; McAllister and Kirkwood, 1998; Punt and Hilborn, 1997; Rosenberg and Restrepo, 1994) or to form the basis of operating models in Management Strategy Evaluation (Smith et al., 1999). In addition, Monte Carlo (MC) estimation methods are now efficiently implemented in several software packages, including the free program OpenBUGS (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs, Lunn et al. (2000, 2009)) and enable posterior distributions for complex models (including non-linear non-Gaussian state-space models) to be represented fairly straightforwardly (Meyer and Millar, 1999; Rivot et al., 2004; Buckland et al., 2007; Zhou et al., 2009). In contrast, most frequentist methods either involve assumptions which are likely unrealistic (such as linearity – the extended Kalman filter – Freeman and Kirkwood, 1995; Punt, 2003) or have high computational and algebraic demands and hence are difficult to test by means of simulation. ADMB-RE (Pedersen et al., 2011), in contrast, provides a computational efficient way to implement state-space models. MC estimation methods have been used in many case studies in the fisheries sciences (e.g., Fernández et al., 2010; Ibaibarriaga et al., 2008; Michielsens et al., 2006; Rivot et al., 2004; Swain et al., 2009). However, relatively few papers (Clark, 2003; Hammond and Trenkel, 2005; Jiao et al., 2011; Robert et al., 2010 or Zhang et al., 2009) have assessed the performance of these estimation methods using a simulation–estimation (SE) approach.

The objective of this paper is to use simulations to compare the performance of Bayesian implementations of observation error (OEM), process error (PEM), and state-space (SSM) formulations of the biomass dynamics model in terms of the ability to accurately, precisely and reliably estimate some of the key quantities on which fisheries management advice is based. We extend the work of Punt (2003) by assessing the relative performance of the three estimation methods within a Bayesian framework and also introduce new metrics to assess estimation performance in a Bayesian context and show how a combination of these metrics can improve evaluation of the estimation performance. The analysis of catch and abundance data for the Mauritanian octopus (*Octopus vulgaris*) (Fig. 1) is used as an application example.

2. Materials and methods

2.1. State-space modeling of biomass dynamics models

The dynamic process equation was defined as the classical budget equation subject to multiplicative log-normal error with variance σ^2_{proc}. The Fox biomass dynamics model was used as the true model for the simulations because it formed the basis for past assessments of octopus off Mauritania (Chassot et al., 2010; Gascuel et al., 2010; Ould Mahmoud et al., 2006) (the operating model):

$$B_{t+1} = \left(B_t + rB_t\left(1 - \frac{\log(B_t)}{\log(K)}\right) - C_t\right)e^{\varepsilon_t - (\sigma^2_{proc}/2)}; \quad \varepsilon_t \sim N(0, \sigma^2_{proc}) \tag{1}$$

where B_t is the biomass at the start of the year t, r the intrinsic rate of growth, K the carrying capacity and C_t the (observed) catch during year t. The term $rB_t(1 - \log(B_t)/\log(K))$ is the assumed relationship between surplus production and biomass (Fox, 1970).

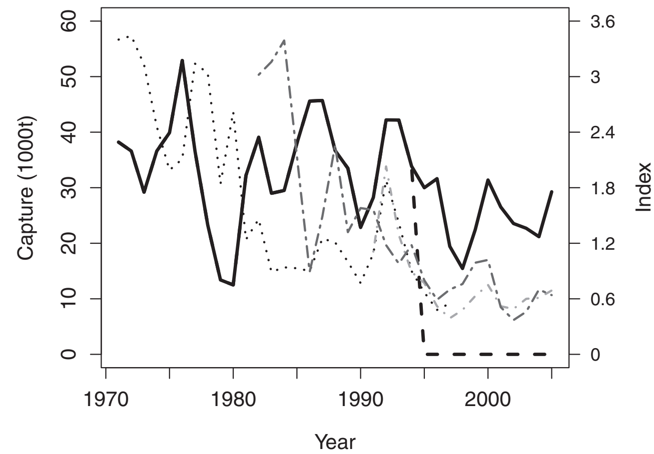

Fig. 1. Catch history for the Mauritanian octopus, 1971–2005 (solid line), the alternative catch scenario in which catches are zero for the 11 last years of the modeled period (dashed line), and the standardized (standardized by the mean abundance index from 1991 to 1997) index of abundance from the WG98 data (dotted line), "Survey" data (double-dashed line) and "Industry" data (dot-dashed line).

The observation equation(s) relates one or several time series of observed abundance indices I_t to biomass through a proportionality constant q. A log-normal observation error with variance σ^2_{obs} is used to capture sampling variability and variation in catchability:

$$I_t = qB_t e^{\tau_t - (\sigma^2_{obs}/2)} \quad \tau_t \sim N(0, \sigma^2_{obs}) \tag{2}$$

2.2. Simulation protocol

The first step of the SE approach involves simulating time-trajectories of biomass and an abundance index using Eqs. (1) and (2) with pre-specified parameters (Table 1) (assumed to be the true values for the analyses) and catches set equal to the observed catches of Mauritanian octopus during 1971–2005 (Fig. 1). The biomass at the start of the catch series (B_{1971}) was assumed to be equal to the carrying capacity K. In a second step, the simulated abundance indices and catches were used as "data" in the Bayesian analyses, and the performances of estimation methods based on OEM, PEM and SSM were evaluated.

The three estimation methods may behave differently depending on the available data and other hypotheses, so data were simulated for various scenarios, and the simulations were crossed

Table 1
Parameter specifications for the scenarios.

Scenario name[a]	Scenario no.	r	B_{2005}/K	σ_{proc}	λ
Baseline	1a	1.2	0.2	0.05	1
Variant1	1b	1.2	0.4	0.05	1
Variant2	1c	1.2	0.6	0.05	1
Higher r	2a	1.6	0.2	0.05	1
	2b	1.6	0.4	0.05	1
	2c	1.6	0.6	0.05	1
Lower r	3a	0.8	0.2	0.05	1
	3b	0.8	0.4	0.05	1
	3c	0.8	0.6	0.05	1
Lower catch, more informative data	4	1.2	0.4	0.05	1
Different variance ratio	5a	1.2	0.2	0.05	0.2
	5b	1.2	0.2	0.05	5
Different variance level	6a	1.2	0.2	0.01	1
	6b	1.2	0.2	0.25	1
Prior mis-specification	7a	0.8	0.2	0.05	1
	7b	1.6	0.2	0.05	1

[a] q was set to 10^{-6} for all scenarios.