Models and model selection uncertainty in estimating growth rates of endangered freshwater mussel populations

Yan Jiao, Richard Neves, and Jess Jones

Abstract: Appropriate inference of population status for endangered species is extremely important. Using a single model for estimating population growth rates is typically inadequate for assessing endangered species because inferences based on only one "best" model ignore model uncertainty. In this study, the endangered dromedary pearlymussel (*Dromus dromas*) in the Clinch and Powell rivers of eastern Tennessee, USA, was used as an example to demonstrate the importance of multiple models, with consideration of environmental noises for evaluating population growth. Our results showed that more than one model deserves consideration in making inferences of population growth rate. A Bayesian model averaging approach was used to make inferences by weighting each model using the deviance information criterion. To test the uncertainty resulting from model selection and the efficiency of the Bayesian averaging approach, a simulation study was conducted on the dromedary pearlymussel populations, which showed that model selection uncertainty is very high. The results of these tests lead us to recommend using Bayesian model averaging to assess population growth status for endangered species, by balancing goodness-of-fit and selection uncertainty among alternate models.

Résumé : Il est extrêmement important de déduire de façon appropriée le statut démographique des espèces menacées. L'utilisation d'un seul modèle pour estimer les taux de croissance démographique ne suffit généralement pas pour évaluer une espèce menacée parce que les déductions tirées d'un seul « meilleur » modèle ne tiennent pas compte de l'incertitude du modèle. Dans notre étude, nous utilisons la moule nacrée dromadaire (*Dromus dromas*), une espèce menacée, des rivières Clinch et Powell dans l'est du Tennessee, É.-U., comme exemple pour démontrer l'importance des modèles multiples, qui tiennent compte du bruit environnemental, pour évaluer la croissance de la population. Nos résultats montrent qu'il est nécessaire de tenir compte de plus d'un modèle lorsqu'on veut déduire le taux de croissance d'une population. Un modèle bayésien utilisant la méthode de la moyenne nous a servi à faire des déductions par pondération de chacun des modèles d'après le critère d'information de l'écart à la moyenne (« deviance information criterion »). Afin de vérifier l'incertitude associée à la sélection des modèles et l'efficacité de la méthode bayésienne par calcul des moyennes, nous avons fait des simulations des populations de moules nacrées dromadaires qui montrent que l'incertitude reliée à la sélection des modèles est très élevée. Les résultats de ces tests nous amènent à recommander l'utilisation du calcul de la moyenne des modèles bayésiens pour vérifier le statut de croissance des espèces menacées en équilibrant la précision de l'ajustement et l'incertitude de la sélection au sein des différents modèles de rechange.

[Traduit par la Rédaction]

Introduction

Freshwater mussels are long-lived species, with maximum life spans typically between 15 and 50 years depending on species. Many of the 18 species of freshwater mussels listed as endangered in the Clinch and Powell rivers in eastern Tennessee were once widespread and abundant throughout the Cumberland and Tennessee river drainages (Parmalee et al. 1980). However, these species experienced severe population declines during the 20th century; therefore, many of them have been under federal protection and management since 1976 (Terwilliger 1991). Concerns for current status of endangered mussel species are rising. For example, range reductions of some species may still be occurring even within the Clinch and Powell rivers (Ahlstedt and Tuberville 1997). A study of the population trends and growth rates of these endangered mussels is necessary to assist with their management and restoration.

The traditional exponential growth model, which analyzes the relationship between relative population abundance and

Received 10 December 2007. Accepted 3 May 2008. Published on the NRC Research Press Web site at cjfas.nrc.ca on 20 October 2008. J20314

Y. Jiao.[1] Department of Fisheries and Wildlife Sciences, 100 Cheatham Hall, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0321, USA.
R. Neves. US Geological Survey, Virginia Cooperative Fish and Wildlife Research Unit, Department of Fisheries and Wildlife Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0321, USA.
J. Jones. US Fish and Wildlife Service, Department of Fisheries and Wildlife Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0321, USA.

[1] Corresponding author (e-mail: yjiao@vt.edu).

doi:10.1139/F08-141

Table 1. Density of the dromedary pearlymussel (*Dromus dromas*; numbers per square metre) under survey by the US Geological Survey.

			Year					
Site (Tennessee)	River mile	Sample size	1979	1983	1988	1994	1999	2004
Clinch River (Swan Island)	172.2	40	0.1		0.1	0.1	0.8	1.5
Clinch River (Brooks Island)	183.8	26	0		0	0	0.46	0
Clinch River (Kyles Ford)	189.6	41	0		0	0	0.39	0.2
Total Clinch (sample weighted average)			0.0374		0.0374	0.0374	0.5603	0.6374
Powell River (Buchanan Ford)	99.2	40	0	0.1	0	0	0.1	0.1
Powell River (McDowell Schoal)	106.7	40	0.1	0.1	0	0	0	0.1
Powell River (Bales Ford)	111.8	20	0.2	0	0.2	0.2	0.2	0
Powell River (Fletcher Ford)	117.3	42	0	0	0	0	0.1	0
Total Powell (sample weighted average)			0.0564	0.0564	0.0282	0.0282	0.0859	0.0563

time for estimating population growth rates, can be inadequate for evaluating population status of endangered species because this model assumes that population growth rate is constant over time and that noise is stationary. That assumption can be violated by large environmental changes or strong and weak yearly recruitment, which are often observed for aquatic invertebrate species. Previous studies of aquatic invertebrate populations have often selected only one "best" model to interpret the observed data. That approach either does not consider other possible models to explain the data or ignores the uncertainty in model selection, leading to overconfident inferences and decisions that are more risky than expected (Draper 1995; Burnham and Anderson 2002).

Many aquatic species show strong and weak periods of recruitment that may relate to patterns of regime shift and recruitment autocorrelations (Glantz 1992; Beamish et al. 1999). Recent research has related these phenomena with the noise in nature. The classification of noise by spectral density is given "color" terminology, with different types named after different colors. White noise means a constant spectral density and temporally uncorrelated noise signal. For colored noise, the spectral density changes with changing frequency, and the noise signal is autocorrelated (Halley 1996; Petchey 2000). Noises in nature may be colored rather than white in many cases (Caswell and Cohen 1995; Halley 1996; Vasseur and Yodzis 2004). Colored noises have more serious implications for endangered species (Halley and Kunin 1999; Morales 1999; Schwager et al. 2006).

We based our analysis on the commonly used exponential growth model, which implies white noise (also called regression model by many freshwater mussel biologists), and then developed four additional models to simulate the colored environmental noises and explore the range of model uncertainty in relation to the population dynamics of an endangered mussel species. The four models were as follows: (1) residual autoregressive model, the residuals of the exponential growth model are assumed to be autocorrelated to simulate the colored environmental noise (Morales 1999; Schwager et al. 2006); (2) population growth autoregressive model, the population growth rates in the exponential growth model are assumed to be autocorrelated to simulate the colored environmental noise (Morales 1999; Schwager et al. 2006); (3) population growth random-walk model, this is a special case of the autoregressive model with fewer parameters (Peterman et al. 2003); and (4) hierarchical exponential population growth model, population growth rates are assumed to follow a multilevel structure to simulate the hierarchy of growth rates, which has been discussed relative to regime shifts, changes of productivity regimes, etc. (Beamish et al. 1999; Clark 2003).

In this study, we investigated the population growth rate of the endangered dromedary pearlymussel (*Dromus dromas*) in the Clinch and Powell rivers from 1979 to 2004. *Dromus dromas* inhabits shoals of coarse gravel and sand in medium to large rivers of the Cumberland Plateau and Southern Appalachian Mountains (Ortmann 1918). It has a maximum age of about 25–30 years (Jones et al. 2004). Historically, this species was widespread and abundant throughout the Cumberland and Tennessee river drainages and was one of the most abundant species in aboriginal shell middens along these rivers (Morrison 1942; Parmalee et al. 1982; Parmalee and Bogan 1998). It experienced severe declines in population abundance and distribution during the 20th century and was listed as an endangered species by the US Fish and Wildlife Service in 1976 (Terwilliger 1991). Currently, only 158 km of the Clinch River and 79 km of the Powell River, less than 10% of the historic range, contain reproducing populations of *D. dromas* (Parmalee et al. 1980; US Fish and Wildlife Service (USFWS) 1983; Jones et al. 2004).

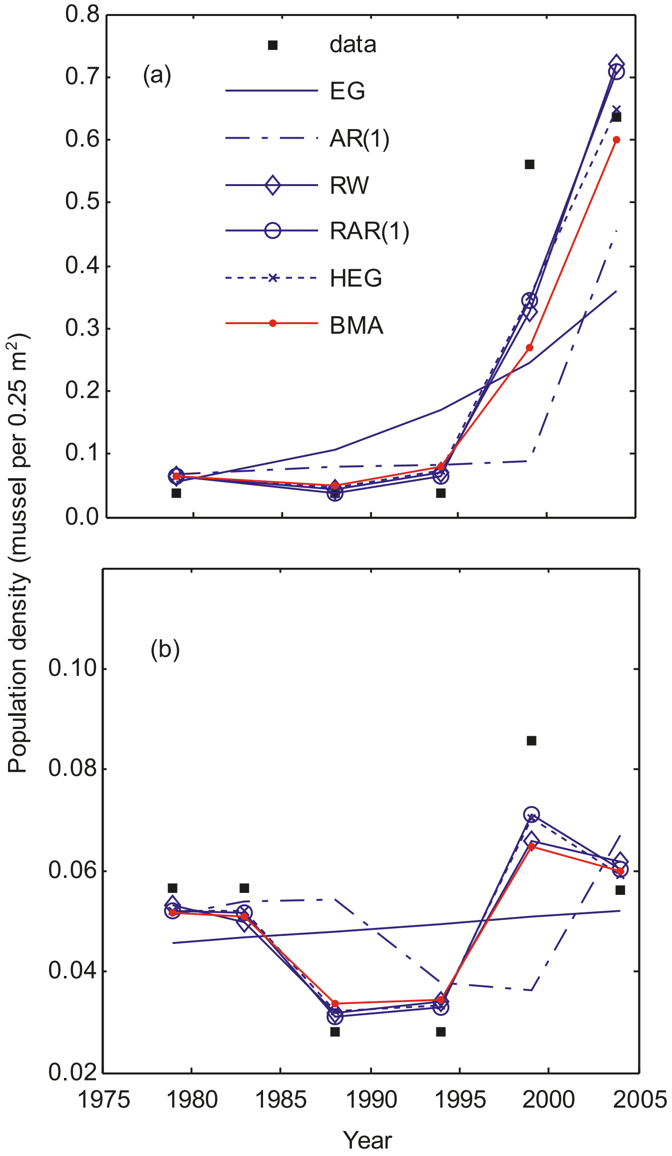
We used a Bayesian approach to solve the different time series models for population growth rates. A Bayesian approach has advantages in dealing with time series models, using observations to update prior models of process noise, measurement noise, and state variables. A deviance information criterion (DIC) was used to compare the different models and was further applied to weight the different models to provide a predictive posterior distribution of population growth rates (Hoeting et al. 1999; Spiegelhalter et al. 2004).

To test the efficiency of the Bayesian averaging approach and the uncertainty resulting from model selection, a further simulation study was done based on the endangered dromedary pearlymussel. Our ultimate goal was to find an appropriate method to assess population growth status of this and other endangered mussel species. The framework developed here can be used to estimate population growth rates of other rare species.

Materials and methods

Data were obtained from mussel surveys taken in 1979,

Fig. 1. Estimated population density: (*a*) Clinch River population, (*b*) Powell River population. EG, exponential growth model; AR(1), first-order residual autoregressive model; RW, random-walk model; RAR(1), growth rate first-order residual autoregressive model; HEG, hierarchical exponential growth model; BMA, Bayesian model averaging.



1983, 1988, 1994, 1999, and 2004 (Ahlstedt et al. 2005). A 0.25 m^2 quadrat was used to collect samples at designated locations in the Clinch and Powell rivers. The densities and sampled numbers of mussels per 0.25 m^2 in different survey locations were recorded, and the sample-weighted mean population density in the Clinch and Powell rivers was calculated (Table 1; Fig. 1) according to the following equation:

$$(1)\qquad \text{weighted mean} = \frac{\sum_i \text{mean}_i \times \text{sample number}_i}{\sum_i \text{sample number}_i}$$

Only six time points were collected over a 26-year period, but each time interval is a fraction of life span (25–30 years) for this species (Ahlstedt et al. 2005). Thus, it is reasonable to analyze population trends through estimates of absolute population growth and assume that population growth rates are constant between survey time intervals.

Models used

The different models, including the Bayesian model averaging approach, were each implemented to detect the absolute population growth rate over time and to assess population status. In other words, was the population increasing or decreasing over the 26 years from 1979 to 2004? In all of the models, we did not consider a density-dependency effect because of the low densities observed over the 26-year period.

The first model used was the exponential growth model (EG):

$$(2)\qquad \begin{aligned} &N_{t+T_t} = N_t \lambda^{T_t} \mathrm{e}^{\varepsilon_1} \quad \text{or} \\ &\mathrm{Ln}(N_{t+T_t}) = T_t \mathrm{Ln}(\lambda) + \mathrm{Ln}(N_t) + \varepsilon_1 \end{aligned}$$

where λ is the absolute population growth rate, N_t is the population size at the survey year t, T_t is the time interval between year t and year $t + T_t$, and error ε_1 is independent and normally distributed with mean 0 and variance $\sigma^2_{\varepsilon_1}$. This model assumed constant population growth rate over time and locations.

The second model used was the first-order residual autoregressive model (AR(1)):

$$(3)\qquad \begin{aligned} &\mathrm{Ln}(N_{t+T_t}) = T_t \mathrm{Ln}(\lambda) + \mathrm{Ln}(N_t) + u_t \\ &u_{t+T_t} = \phi_{T_t} u_t + \varepsilon_{2,T_t} \end{aligned}$$

In this model, residual error u_t is modeled as a first-order autoregressive process, ϕ is the autocorrelation coefficient. Because the time intervals were different among observations (4 to 6 years, with 3 of them 5 years), we used the equation $u_{t+k} = \phi_0^k u_t + \sum_{j=0}^{k-1} \phi_0^j \varepsilon_2'$ in our preliminary analysis (Houseman 2005). In the analysis, by using $\phi_{T_t=5} = \phi = \phi_0^5$ when the time interval is 5, $\phi_{T_t=6} = \phi^{6/5}$ when the interval is 6, and $\phi_{T_t=4} = \phi^{4/5}$ when the interval is 4. The error term $\sum_{j=0}^{k-1} \phi_0^j \varepsilon_2' = \varepsilon_2$ when time interval is 5. We treated the error term equal to ε_2 when the time intervals were 4 and 6 because of the limited differences between $\sum_{j=0}^{5} \phi_0^j \varepsilon_2'$ and $\sum_{j=0}^{4} \phi_0^j \varepsilon_2'$ and between $\sum_{j=0}^{3} \phi_0^j \varepsilon_2'$ and $\sum_{j=0}^{4} \phi_0^j \varepsilon_2'$. The changes in the parameter estimates were trivial when compared with the result by assuming $\phi_{T_t=4,5,6} = \phi$ under different time intervals. Therefore, we treated ϕ as equal when the time intervals varied among surveys in this study. Error ε_2 is independent and normally distributed with mean 0 and var-

Table 2. Parameter estimates (posterior mean and 95% CI) and corresponding DIC values of the dromedary pearlymussel (*Dromus dromas*) population in the Clinch and Powell rivers.

	Clinch River				Powell River		
		Estimates			Estimates		
Model	Parameters	Mean	95% CI	DIC	Mean	95% CI	DIC
EG model	λ	1.08	0.95, 1.19	–3.85	1.01	0.94, 1.06	–24.92
	$\sigma^2_{\varepsilon_1}$	2.05			0.47		
AR(1) model	λ	1.20	0.42, 1.94	–3.71	1.03	0.30, 1.88	–23.02
	ϕ	0.69			0.77		
	$\sigma^2_{\varepsilon_2}$	2.47			0.81		
RW model	λ_t (mean)	1.15	0.52, 1.84	–4.53	1.02	0.60, 1.53	–25.45
	$\sigma^2_{\varepsilon_3}$	1.57			0.51		
	$\sigma^2_{\varepsilon_4}$	0.37			0.09		
RAR(1) model	$\overline{\lambda}$	1.16	0.48, 1.87	–4.61	1.03	0.53, 1.64	–23.60
	φ	0.24			0.29		
	$\sigma^2_{\varepsilon_5}$	1.54			0.71		
	$\sigma^2_{\varepsilon_6}$	0.40			0.13		
HEG model	a	1.14	0.15, 1.90	–4.96	1.02	0.30, 1.76	–26.11
	b	0.25			0.07		
	$\sigma^2_{\varepsilon_7}$	1.52			0.49		
BMA	λ_t (mean)	1.14	0.28, 1.88		1.03	0.46, 1.67	

Note: CI, credible interval; DIC, deviance information criterion.

iance $\sigma^2_{\varepsilon_2}$. λ, the parameter in which we are interested, is very robust to this approximation.

The third model used was the random-walk model (RW):

$$\begin{aligned} &\mathrm{Ln}(N_{t+T_t}) = T_t\mathrm{Ln}(\lambda_t) + \mathrm{Ln}(N_t) + \varepsilon_3 \\ &\lambda_{t+T_t} = \lambda_t + \varepsilon_{4,T_t} \end{aligned} \tag{4}$$

where population growth rate λ_t is modeled as a random-walk process. The error term $\sum_{j=0}^{T_t-1} \varepsilon_4' = \varepsilon_4$ when time interval is 5, where ε_4' is the error when the time interval is 1 year. We treated the error term equal to $\sum_{j=0}^{4-1} \varepsilon_4'$ and $\sum_{j=0}^{6-1} \varepsilon_4'$ when the time intervals were 4 and 6, i.e., $\frac{4}{5}\varepsilon_4$ and $\frac{6}{5}\varepsilon_4$, respectively. Errors ε_3 and ε_4 are independent and normally distributed with mean 0 and variances $\sigma^2_{\varepsilon_3}$ and $\sigma^2_{\varepsilon_4}$, respectively.

The fourth model used was the population growth autoregressive model (RAR(1)):

$$\begin{aligned} &\mathrm{Ln}(N_{t+T_t}) = T_t\mathrm{Ln}(\lambda_t) + \mathrm{Ln}(N_t) + \varepsilon_5 \\ &\mathrm{Ln}(\lambda_{t+T_t}) = \mathrm{Ln}(\overline{\lambda}) + \varphi_{T_t}[\mathrm{Ln}(\lambda_t) - \mathrm{Ln}(\overline{\lambda})] + \varepsilon_{6,T_t} \end{aligned} \tag{5}$$

where population growth rate λ_t is modeled as a first-order autoregressive process, and φ_{T_t} is the autocorrelation coefficient when the time interval is T_t. We treated $\varphi_{T_t} = \varphi$ and $\varepsilon_{6,T_t} = \varepsilon_6$ the same under different time intervals because of the reasons explained in the AR(1) model. Error ε_6 is independent and normally distributed with mean 0 and variance $\sigma^2_{\varepsilon_6}$.

The fifth model used was the hierarchical exponential growth model (HEG):

$$\begin{aligned} &\mathrm{Ln}(N_{t+T_t}) = T_t\mathrm{Ln}(\lambda_t) + \mathrm{Ln}(N_t) + \varepsilon_7 \\ &\lambda_t \in N(a, b) \\ &a \in U(c, d) \end{aligned} \tag{6}$$

where error ε_7 is independent and normally distributed with mean 0 and variance $\sigma^2_{\varepsilon_7}$, λ_t follows a hierarchical distribution, i.e., a, the mean of λ, is uniformly distributed between c and d, and $N(a, b)$ is truncated to make sure that λ has positive values.

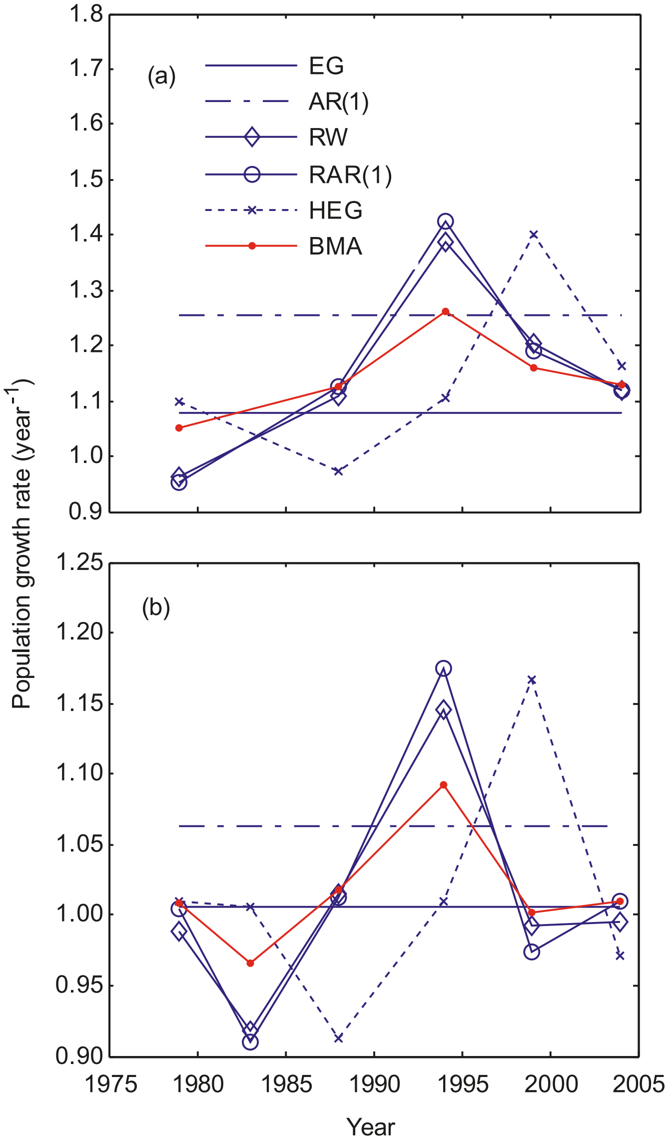
Population growth rates in both rivers were predicted using each of the five models separately, then using a Bayesian model averaging over the five models together, without a priori conditioning on model goodness-of-fit (Hoeting et al. 1999). Population sizes N were expressed as the mean population densities from the mussel surveys.

Bayesian method and priors

A Bayesian method with Metropolis–Hasting within Gibbs sampling algorithm was implemented in WinBUGS 1.4 software to estimate the model parameters (Gilks 1996). Bayesian implementation of the models requires specification of prior distributions on all unobserved quantities. Noninformative uniform distribution priors were used for parameters $\lambda \sim U(0.001, 2)$, $\phi \sim U(-1, 1)$, $\varphi \sim U(-1, 1)$, $a \sim U(0.001, 2)$, and for variances $\sigma^2_{\varepsilon} \sim U(0.001, 5)$. Uniform prior distributions work better as noninformative priors than inverse-gamma distributions for variance parameters when dealing with multilevel models (Gelman 2005).

Because inverse-gamma is often used as a noninformative prior, we also performed analyses using IG(0.001, 0.001) for variances of $\sigma^2_{\varepsilon_1}$, $\sigma^2_{\varepsilon_2}$, … , $\sigma^2_{\varepsilon_7}$ in the Bayesian time series models to test whether the results are sensitive to the priors (Spiegelhalter et al. 2004; Gelman 2005).

Fig. 2. Population growth rate over time: (*a*) Clinch River population; (*b*) Powell River population. EG, exponential growth model; AR(1), first-order residual autoregressive model; RW, random-walk model; RAR(1), growth rate first-order residual autoregressive model; HEG, hierarchical exponential growth model; BMA, Bayesian model averaging.



Convergence diagnostics

A critical issue in using Markov chain Monte Carlo (MCMC), such as the Metropolis–Hasting within Gibbs sampling algorithm that we used here, is how to determine when random draws have converged to the posterior distribution. Here, three methods were considered: monitoring the trace, diagnosing the autocorrelation plot, and Gelman and Rubin statistics (Spiegelhalter et al. 2004). In this study, three chains were used. After several sets of analysis for each chain, the first 20 000 iterations with a thinning interval of three were discarded, and another 20 000 to ~50 000 iterations were used in the Bayesian analysis.

Bayesian model averaging (BMA)

The DIC was used as a model selection criterion to weight different models in this study. It was also used as the criterion to select the “best” model in the simulation study:

$$
\begin{aligned}
&\mathrm{DIC} = 2\overline{D} - \widehat{D} \text{ or } \overline{D} + p_D \\
(7)\quad &D(y, \theta) = -2\mathrm{loglikelihood}(y|\theta) \\
&p_D = \overline{D} - \widehat{D}
\end{aligned}
$$

where D is deviance, a measurement of prediction goodness for our models; p_D is the effective number of parameters in a Bayesian model. The DIC is a hierarchical modeling generalization of the AIC (Akaike information criterion) and BIC (Bayesian information criterion, also known as the Schwarz criterion). DIC is particularly useful in Bayesian model selection problems, where the posterior distributions of models have been obtained by MCMC simulation. As with AIC and BIC, it is an asymptotic approximation as the sample size becomes large (Spiegelhalter et al. 2002, 2004).

Following the Bayesian analysis of each model, a BMA was used to balance model goodness-of-fit and model selection uncertainty. The posterior distributions of estimated population growth rates from these models were further weighted based on their DIC to provide our predictive posterior distribution of population growth rates (Hoeting et al. 1999):

$$
\begin{aligned}
&\Delta_{\mathrm{DIC}_i} = \mathrm{DIC}_i - \min(\mathrm{DIC}) \\
(8)\quad &\mathrm{weight}_i = \frac{\mathrm{e}^{-2\Delta_{\mathrm{DIC}_i}}}{\sum_i \mathrm{e}^{-2\Delta_{\mathrm{DIC}_i}}}
\end{aligned}
$$

According to the weight of each model, we selected the corresponding proportion of the posterior MCMC runs randomly from each model and then combined all the selected posterior MCMC runs as the model-averaged MCMC. When weighting the models, models with ΔDIC larger than 10 were ruled out because of the extremely small weight (Spiegelhalter et al. 2002).

Simulation study

A simulation study was designed to test the performance of the proposed Bayesian model averaging approach and the model selection uncertainty. The following simulation algorithm was used: (*i*) estimate population growth rates from the models using the test species and treat these estimates as the true population growth rates; (*ii*) generate population density data from a Monte Carlo simulation with uncertainty levels equivalent to the uncertainties estimated from the original population; (*iii*) analyze the generated data set using the five different models (always picking the best model based on model selection criteria and a model averaging approach); and (*iv*) evaluate the uncertainty arising from model selection and the performance of Bayesian model averaging by the probability of finding the true model and the relative estimation error (REE) (for the simulation design, see Table 2).

Steps *ii* through *iv* above were repeated 200 times to yield 200 sets of estimated population growth rates and population densities from each model. The sum of mean square errors (SSE) for population growth rate estimated in year t, $\widehat{\lambda}_t$, was calculated as

Table 3. Relative estimation error (REE) of Clinch and Powell River population growth rate estimates (λ) and the probability of being the best model out of 200 simulation runs.

		Clinch River				Powell River			
		Estimates				Estimates			
		REE				REE			
1*	2*	Median	5% CI	95% CI	*P**	Median	5% CI	95% CI	*P**
EG model	EG	0.0046	0.0001	0.0361	0.325	0.0019	0.0000	0.0154	0.375
	AR(1)	0.0413	0.0007	0.2738	0.130	0.0207	0.0001	0.2146	0.065
	RW	0.0747	0.0152	0.1790	0.135	0.0466	0.0109	0.1835	0.155
	RAR(1)	0.1072	0.0218	0.2557	0.045	0.0774	0.0157	0.2446	0.110
	HEG	0.0841	0.0174	0.2174	0.365	0.0553	0.0104	0.1797	0.295
AR(1) model	EG	0.1586	0.0615	0.3084	0.290	0.0208	0.0001	0.1337	0.215
	AR(1)	0.0329	0.0009	0.2675	0.125	0.0428	0.0008	0.3166	0.130
	RW	0.1389	0.0658	0.2657	0.260	0.0879	0.0255	0.2249	0.285
	RAR(1)	0.1574	0.0679	0.2938	0.045	0.1163	0.0317	0.2845	0.125
	HEG	0.1551	0.0661	0.2861	0.280	0.0962	0.0239	0.2035	0.245
RW model	EG	0.6439	0.2317	4.3956	0.305	0.6131	0.1482	3.2490	0.230
	AR(1)	0.6052	0.2734	4.9571	0.105	0.5133	0.1681	2.6877	0.100
	RW	0.5304	0.1370	3.1621	0.165	0.5956	0.1275	3.3099	0.285
	RAR(1)	0.5146	0.0973	3.3917	0.020	0.6022	0.1407	3.1739	0.085
	HEG	0.5906	0.2973	5.2457	0.405	0.6168	0.2029	4.2012	0.300
RAR(1) model	EG	0.6274	0.1623	3.5249	0.170	0.7436	0.1880	4.0358	0.255
	AR(1)	0.6255	0.2788	3.5387	0.215	0.9124	0.2383	4.4933	0.070
	RW	0.5450	0.0976	2.4421	0.235	0.6132	0.1292	4.0410	0.235
	RAR(1)	0.5560	0.1209	2.5779	0.025	0.5865	0.1319	3.8471	0.125
	HEG	0.6391	0.2206	4.1763	0.355	0.8640	0.1670	4.7091	0.315
HEG model	EG	0.7029	0.1958	3.9843	0.285	0.6161	0.0940	3.8571	0.145
	AR(1)	0.8542	0.1811	5.4679	0.080	0.7437	0.1472	3.7772	0.070
	RW	0.6153	0.1510	3.4636	0.215	0.4596	0.0794	2.8862	0.240
	RAR(1)	0.6106	0.1444	3.3187	0.025	0.4345	0.0899	2.5100	0.135
	HEG	0.7268	0.2489	4.6225	0.395	0.7967	0.1288	6.5752	0.410

Note: CI, credible interval; 1*, true model used to generate data; 2*, models used to estimate population growth rate from 1*; *P**, probability of being the best model.

$$\mathrm{SSE}(\widehat{\lambda}_i) = \sum_{t=1}^{k} (\widehat{\lambda}_{t,i} - \lambda'_{t,i})^2 \tag{9}$$

where i indicates the ith simulation run and k is the number of years with survey data available. The relative estimation error for estimated population growth rate in year t, $\mathrm{REE}(\widehat{\lambda}_i)$, was calculated as

$$\mathrm{REE}(\widehat{\lambda}_i) = \sum_{t=1}^{5} [(\widehat{\lambda}_{t,i} - \lambda'_{t,i})/\lambda'_{t,i}]^2 \tag{10}$$

The REE calculated in eq. 10 measures the overall estimation errors, including both estimation biases and variations in estimates among the 200 simulation runs. Box plots were used to summarize the REEs derived in the 200 simulation runs. An estimation procedure with small REE suggests that it performs well and tends to have smaller error in estimating current population growth rate.

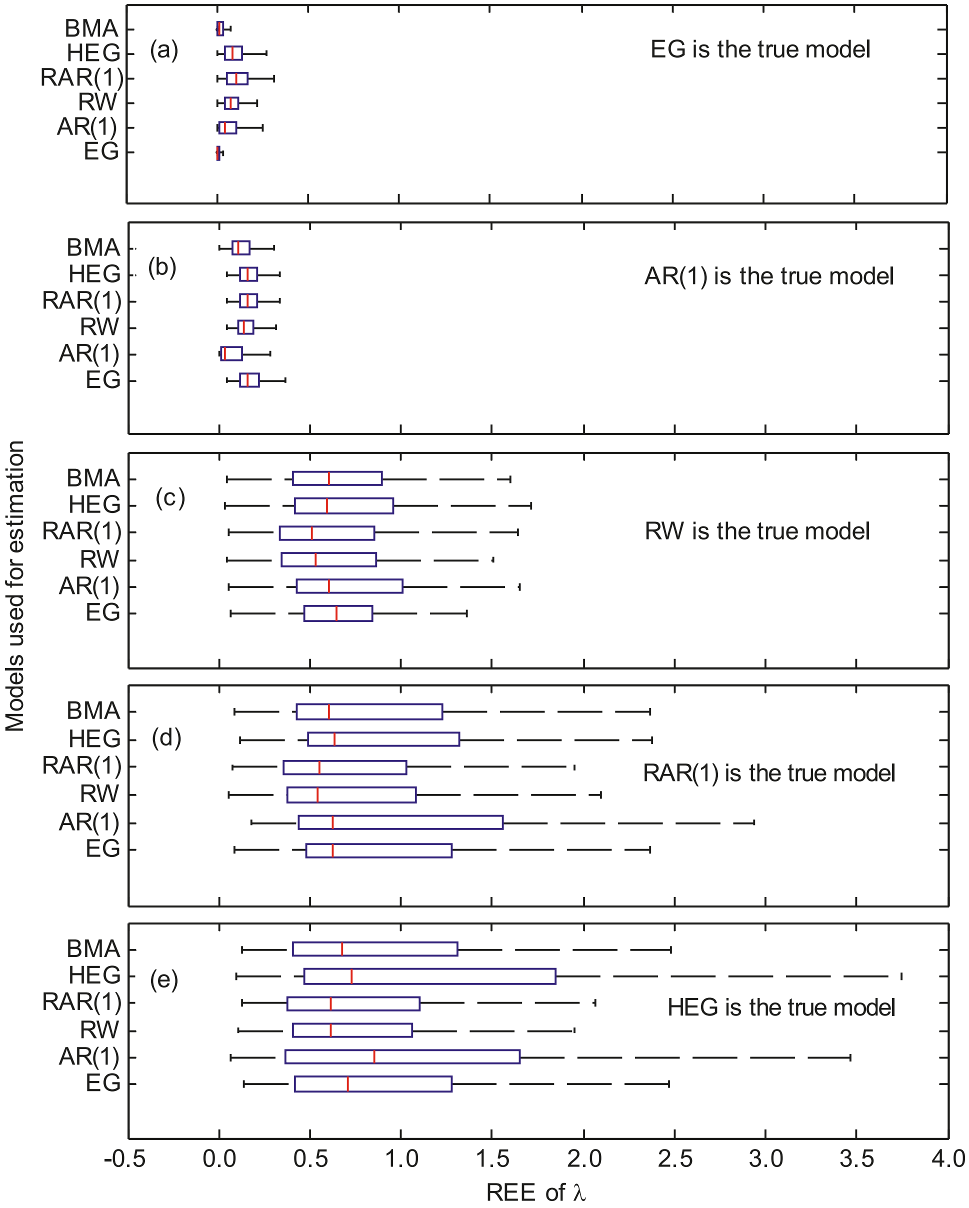
Model selection uncertainty was evaluated through a probability of choosing the "true" model as the best model, based on the lowest DIC value. For example, when the EG model was used as the true model, in each of these 200 runs, the simulation algorithm would pick the best model based on the DIC values (smallest DIC means the best model); the best model would be recorded in each of the simulation runs. After the 200 runs, the probability of each model chosen as the best model was counted. For example, if the EG model is chosen as the best model in 40 of 200 runs, then the probability is 20%. After the simulation study, a best (or robust) modeling approach was selected by comparing the REE estimates of the population growth rates.

Results

Empirical data analysis showed that the hierarchical exponential growth model gave lowest DIC values for both the Clinch and Powell river populations (Table 2). However, the ΔDICs or DIC between the best model and the others, are between 0.35 and 1.25 for the Clinch River data and between 0.66 and 3.09 for the Powell River data.

The estimated posterior means of the Clinch River population growth rates were larger than 1 in all five models (Fig. 2*a*; Table 2). Only when the exponential growth model was used did the 95% probability credible interval of the estimated population growth rate slightly overlap with 1; all the others largely overlapped with 1. Among the five models, the HEG models resulted in the lowest DIC values, and the RAR(1) and RW models also resulted in relatively lower

Fig. 3. Box plot of the relative estimation error (REE) of population growth rate λ of the Clinch River population. From top to the bottom, the "true" models used in generating the data are as follows: EG, exponential growth model; AR(1), first-order residual autoregressive model; RW, random-walk model; RAR(1), growth rate first-order residual autoregressive model; HEG, hierarchical exponential growth model. BMA, Bayesian model averaging.



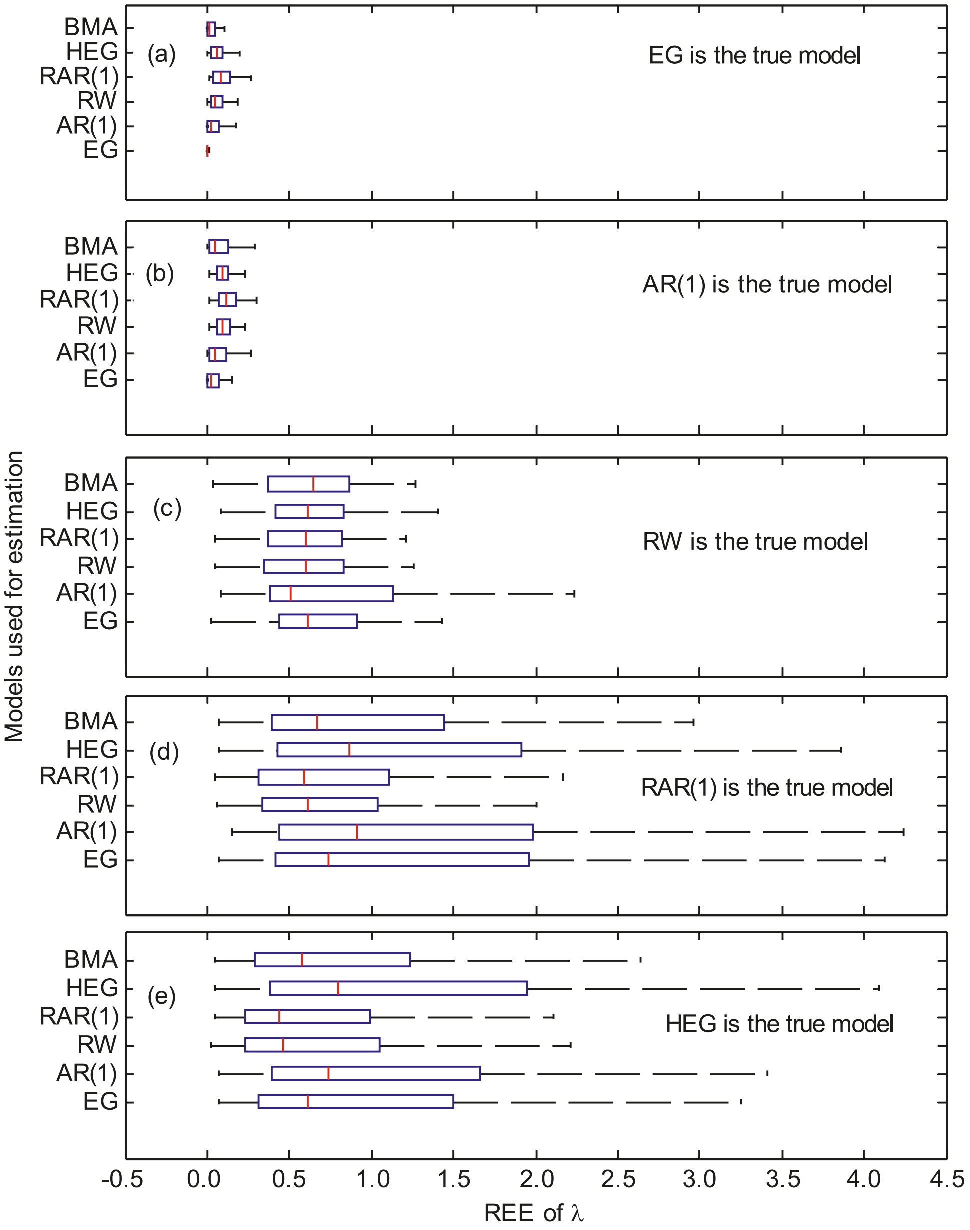
DIC values than the EG and AR(1) models. The resulting posterior means of the λ values from the better models over time were very close.

The estimated posterior means of the Powell River population growth rate were also larger than 1 in all five models (Fig. 2*b*; Table 2), but not significantly different from 1 according to the 95% credible intervals (Table 2). The HEG model resulted in the lowest DIC values. In general, the growth rate of the population in Powell River showed dynamic pattern over time, which can be seen from the population size changes and from the estimated population growth rate (Figs. 1 and 2).

Model selection uncertainties from the simulation study were high. With the Clinch River data, probabilities of determining the true model were 32.5%, 13%, 13.5%, 4.5%, and 36.5% for the EG, AR(1), RW, RAR(1), and HEG models, respectively (Table 3). With the Powell River data, probabilities of determining the true model were 37.5%, 6.5%, 15.5%, 11%, and 29.5% for the EG, AR(1), RW, RAR(1), and HEG models, respectively (Table 3). HEG tended to be selected as the best model no matter what the true models were, and RAR(1), which has the highest number of parameters, generally had the lowest probability of being selected as the best model.

The simulation study also showed that the "true" model tended to give the estimate with lowest REE, which means that the population estimates are better, but not always the lowest one (Table 3; Figs. 3 and 4). The REE values calculated from the model averaging approach were low and very close to the REE calculated when the true model was

Fig. 4. Box plot of the relative estimation error (REE) of population growth rate λ of the Powell River population. From top to the bottom, the "true" models used in generating the data are as follows: EG, exponential growth model; AR(1), first-order residual autoregressive model; RW, random-walk model; RAR(1), growth rate first-order residual autoregressive model; HEG, hierarchical exponential growth model. BMA, Bayesian model averaging.



used (Figs. 3 and 4). This implies that the model averaging approach works well in estimating population growth rate.

The sensitivity analyses to noninformative priors indicated that estimated λ (or mean of λ_t) values and their credible intervals were robust to priors of error variances in the Bayesian time series models, when uniform distribution and inverse-gamma distribution were used. Little difference was observed, so we did not present the results here. However, posterior estimates of the error variance values themselves were not robust to the priors of error variance. In general, higher lower bounds of the priors of $\sigma^2_{\varepsilon_1}$ to $\sigma^2_{\varepsilon_7}$ resulted in higher posterior means of $\sigma^2_{\varepsilon_1}$ to $\sigma^2_{\varepsilon_7}$.

Discussion

The small differences in DIC among models used to simulate population growth suggest that more than one model should be taken into consideration (Spiegelhalter et al. 2002). Spiegelhalter et al. (2002) suggested that models with ΔDIC less than 5 from the best model definitely need to be considered, whereas models with ΔDIC between 5 and 10 from the best model are substantial. Ignoring these potentially important models may lead to misinterpretation of the population trends. Using only the "best" model, without considering other possible models and the uncertainty in model selection, leads to overconfident inferences and decisions that are riskier than expected (Draper 1995). In this study, the credible intervals of population growth rate from the exponential growth (EG) models were obviously narrower than those from the other models and, therefore, overconfident compared with the BMA results. A "best" model can always be found based on model selection criteria, but the simulation study demonstrated that because of the inherent model uncertainty, the "best" model may well not be the "true" model. Bayesian model averaging provides a coherent mechanism to account for this model uncertainty. When dealing with real observation data for which the state of nature is unknown, comparison between the BMA and commonly used individual models such as the EG model protects against overconfidence in the results (Draper 1995; Katsanevakis 2006).

Exponential growth models have been widely used in analyzing population abundance and (or) density data (Hilborn and Walters 1992). There was very limited population trend and growth rate analyses for all the endangered mussel species in the Clinch, Powell, and other rivers (J. Jones, unpublished data). The hierarchical exponential growth model and other colored-noise models have been suggested for investigating endangered freshwater mussels, as well as other aquatic invertebrates and fish species, whose recruitment and population status vary over time (Heino et al. 2000; Peterman et al. 2003; Wichmann et al. 2003).

For the current populations of dromedary pearlymussel in the Clinch and Powell rivers, the estimated 95% credible intervals of population growth rates from the model-averaging approach overlap with value 1, although the posterior mean values of all the models were higher than 1. This indicates that there is not enough evidence to support the hypothesis of populations increasing. The status of the population in the Powell River is more pessimistic than that in the Clinch River, which showed a lower population density, a more dynamic pattern of population growth rate, and a very limited overall population increase with high uncertainty.

Investigated populations of the dromedary pearlymussel were distributed at Swan Island, Brooks Island, and Kyles Ford in the Clinch River and Buchanan Ford, McDowell Shoal, and Bales Ford in the Powell River, where local demes have fluctuated at relatively low levels of abundance over the last 25 years (Wolcott and Neves 1994; Ahlstedt and Tuberville 1997; Ahlstedt et al. 2005). To understand and improve the population status at other locations in the Clinch and Powell rivers, further ecological studies on local habitats and environmental conditions are suggested.

Our knowledge of these endangered freshwater mussels is still limited, especially regarding the influence of biotic and abiotic factors on their life history, such as the effects of water discharge and fish host usage on reproductive success. We still lack knowledge to proceed to more mechanistic approaches, such as individual-based models (IBM), for understanding mussel population dynamics. Field observations and monitoring experiments are currently underway to provide information that can be used to simulate the dynamics of these species using mechanistic models (J. Jones, unpublished data). For example, a coupled biological–physical model or a model with habitat and spatial heterogeneity would help to further understand and protect these mussel species (Werner et al. 2001; Grimm et al. 2005*a*, 2005*b*).

Hierarchical Bayesian methods are especially applicable to small-area estimation problems, such as this one (He and Sun 1998). However, we also realized that HEG resulted in high variance of relative estimation error, which means that in cases when the goodness-of-fit is low, it can result in a highly biased result. Comparison with the other models is still suggested. The exponential growth model is widely used because of its simplicity, and in this study, both field data analysis and simulation study showed that it seems robust to noise. Random noise caused by environmental variation can be more important than the "true" population dynamic pattern. This is especially true and crucial for endangered species (Petchey 2000; Schwager et al. 2006). The corresponding risks of extinction are different when different models are used (Halley and Kunin 1999; Morales 1999; Schwager et al. 2006). Colored noise can decrease or increase the extinction risk (Heino et al. 2000; Wichmann et al. 2003), and the model structures used in generating the colored noise are relevant (Morales 1999; Wichmann et al. 2005). For these endangered freshwater mussels, because insufficient data were available to estimate population size, risk of population extinction can be evaluated by using the predicted trajectory of the density from different models to obtain an estimate of local extinction that could be couched in the probability of finding no mussels in a sample of size *n*. For increased precaution against extinction, models such as AR(1), RAR, and RW are highly recommended to be incorporated into the model-averaging framework because they are better at capturing environmental noise than the EG model. We recommend a simulation study for estimating population growth rate of any endangered species, as this would lower the risk of misinterpreting population status and thus allow better management of the rare biological resource.

High model selection uncertainty implies that estimating population growth rate and evaluating the population status based on one single model, or even one "best" model, is of high risk. Because model selection uncertainty is high, the model averaging approach is recommended for dealing with growth rate estimation and population trend analysis of endangered mussel populations. The model averaging approach balances the needs for best models and for model selection uncertainty and minimizes the risk of underestimating or overestimating population extinction risk (Morales 1999; Johnson and Omland 2004; Schwager et al. 2006).

Acknowledgements

This study was supported by Virginia Tech new faculty startup funding to Y. Jiao and by the US Department of Agriculture Cooperative State Research, Education and Extension Service, Hatch project No. 0210510 to Y. Jiao. The authors thank mussel biologist Steve Ahlstedt of the US Geological Survey (retired) for sharing survey data collected over his career.

References

Ahlstedt, S.A., and Tuberville, J.D. 1997. Quantitative reassessment of the freshwater mussel fauna in the Clinch and Powell rivers, Tennessee and Virginia. *In* Conservation and management of freshwater mussels. II. *Edited by* K.S. Cummings, A.C.

Buchanan, C.A. Mayer, and T.J. Naimo. Upper Mississippi River Conservation Committee, Rock Island, Illinois. pp. 72–97.

Ahlstedt, S.A., Fagg, M.T., Butler, R.S., and Connell, J.F. 2005. Long-term trend information for freshwater mussel populations at twelve fixed-station monitoring sites in the Clinch and Powell rivers of Eastern Tennessee and Southwestern Virginia 1979–2004. Final Report, US Fish and Wildlife Service, Cookeville, TN 38501, USA.

Beamish, R.J., Noakes, D.J., McFarlane, G.A., Klyashtorin, L., Ivanov, V.V., and Kurashov, V. 1999. The regime concept and natural trends in the production of Pacific salmon. Can. J. Fish. Aquat. Sci. **56**: 516–526. doi:10.1139/cjfas-56-3-516.

Burnham, K.P., and Anderson, D. 2002. Model selection and multimodel inference. Springer-Verlag, New York.

Caswell, H., and Cohen, J.E. 1995. Red, white and blue: environmental variance spectra and coexistence in metapopulations. J. Theor. Biol. **176**: 301–316. doi:10.1006/jtbi.1995.0200.

Clark, J.S. 2003. Uncertainty and variability in demography and population growth: a hierarchical approach. Ecology, **84**: 1370–1381. doi:10.1890/0012-9658(2003)084[1370:UAVIDA]2.0.CO;2.

Draper, D. 1995. Assessment and propagation of model uncertainty (with discussion). J. R. Stat. Soc. B, **57**: 45–97.

Gelman, A. 2005. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. **2**: 1–19.

Gilks, W.R. 1996. Full conditional distributions. *In* Markov chain Monte Carlo in practice. *Edited by* W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. Chapman and Hall, London, UK. pp. 75–78.

Glantz, M.H. 1992. Climate variability, climate change and fisheries. Cambridge University Press, Cambridge and New York.

Grimm, V., Revilla, E., Groeneveld, J., Kramer-Schadt, S., Schwager, M., Tews, J., Wichmann, M.C., and Jeltsch, F. 2005*a*. Importance of buffer mechanisms for population viability analysis. Conserv. Biol. **19**: 578–580. doi:10.1111/j.1523-1739.2005.000163.x.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H., Weiner, J., Wiegand, T., and DeAngelis, D.L. 2005*b*. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science (Washington, D.C.), **310**: 987–991. doi:10.1126/science.1116681. PMID:16284171.

Halley, J.M. 1996. Ecology, evolution and 1/f-noise. Trends Ecol. Evol. **11**: 33–37. doi:10.1016/0169-5347(96)81067-6.

Halley, J.M., and Kunin, W.E. 1999. Extinction risk and the 1/f family of noise models. Theor. Popul. Biol. **56**: 215–230. doi:10.1006/tpbi.1999.1424. PMID:10607517.

He, Z., and Sun, D. 1998. Hierarchical Bayes estimation of hunting success rates. Environ. Ecol. Stat. **5**: 223–236. doi:10.1023/A:1009669302755.

Heino, M., Ripa, J., and Kaitala, V. 2000. Extinction risk under coloured environmental noise. Ecography, **23**: 177–184. doi:10.1034/j.1600-0587.2000.230203.x.

Hilborn, R., and Walters, C. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York.

Hoeting, J.A., Madigan, D., Raftery, A.E., and Volinsky, C.T. 1999. Bayesian model averaging: a tutorial. Stat. Sci. **14**: 382–417. doi:10.1214/ss/1009212519.

Houseman, E.A. 2005. A robust regression model for a first-order autoregressive time series with unequal spacing: application to water monitoring. Appl. Stat. **54**: 769–780.

Johnson, J.B., and Omland, K.S. 2004. Model selection in ecology and evolution. Trends Ecol. Evol. **19**: 101–108. doi:10.1016/j.tree.2003.10.013. PMID:16701236.

Jones, J., Neves, R.J., Ahlstedt, S.A., and Mair, R.A. 2004. Life history and propagation of the endangered dromedary pearlymussel (*Dromus dromas*) (Bivalvia: Unionidae). J. North Am. Benthol. Soc. **23**: 515–525. doi:10.1899/0887-3593(2004)023<0515:LHAPOT>2.0.CO;2.

Katsanevakis, S. 2006. Modelling fish growth: model selection, multi-model inference and model selection uncertainty. Fish. Res. **81**: 229–235. doi:10.1016/j.fishres.2006.07.002.

Morales, J.M. 1999. Viability in a pink environment: why "white noise" models can be dangerous. Ecol. Lett. **2**: 228–232. doi:10.1046/j.1461-0248.1999.00074.x.

Morrison, J.P.E. 1942. Preliminary report on mollusks found in the shell mounds of the Pickwick Landing Basin in the Tennessee River valley. Bur. Am. Ethnol. Bull. **129**: 339–392.

Ortmann, A.E. 1918. The nayades (freshwater mussels) of the upper Tennessee drainage with notes on synonymy and distribution. Proc. Am. Philos. Soc. **57**: 521–626.

Parmalee, P.W., and Bogan, A.E. 1998. The freshwater mussels of Tennessee. University of Tennessee Press, Knoxville, Tennessee.

Parmalee, P.W., Klippel, W.E., and Bogan, A.E. 1980. Notes on the prehistoric and present status of the naiad fauna of the middle Cumberland River, Smith County, Tennessee. Nautilus, **94**: 93–105.

Parmalee, P.W., Klippel, W.E., and Bogan, A.E. 1982. Aboriginal and modern freshwater mussel assemblages (Pelecypoda: Unionidae) from the Chickamauga Reservoir, Tennessee. Brimleyana, **8**: 75–90.

Petchey, O.L. 2000. Environmental colour affects aspects of single-species population dynamics. Proc. R. Soc. Lond. B, Biol. Sci. **267**: 747–754. doi:10.1098/rspb.2000.1066.

Peterman, R.M., Pyper, B.J., and MacGregor, B.W. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci. **60**: 809–824. doi:10.1139/f03-069.

Schwager, M., Johst, K., and Jeltsch, F. 2006. Does red noise increase or decrease extinction risk? Single extreme events versus series of unfavorable conditions. Am. Nat. **167**: 879–888. doi:10.1086/503609.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A. 2002. Bayesian measures of model complexity and fit (with discussion). J. R. Stat. Soc. Ser. B, Methodol. **64**: 583–640. doi:10.1111/1467-9868.00353.

Spiegelhalter, D.J., Thomas, A., Best, N., and Lunn, D. 2004. WinBUGS user manual (version 1.4.1). MRC Biostatistics Unit, Cambridge, UK.

Terwilliger, K. 1991. Virginia's endangered species. McDonald and Woodward Publishing Company, Blacksburg, Virginia.

US Fish and Wildlife Service. 1983. Dromedary pearlymussel recovery plan. US Fish and Wildlife Service, Atlanta, Georgia. Available from Fish and Wildlife Reference Service, 5430 Grosvenor Lane, Suite 110, Bethesda, MD 20814, USA.

Vasseur, D.A., and Yodzis, P. 2004. The color of environmental noise. Ecology, **85**: 1146–1152. doi:10.1890/02-3122.

Werner, F.E., Quinlan, J.A., Lough, R.G., and Lynch, D.R. 2001. Spatially explicit individual-based modeling of marine populations: a review of the advances in the 1990s. Sarsia, **86**: 411–421.

Wichmann, M.C., Johst, K., Moloney, K.A., Wissel, C., and Jeltsch, F. 2003. Extinction risk in periodically fluctuating environments. Ecol. Model. **1**: 29–40.

Wichmann, M.C., Johst, K., Schwager, M., Blasius, B., and Jeltsch, F. 2005. Extinction risk, coloured noise and the scaling of variance. Theor. Popul. Biol. **68**: 29–40. doi:10.1016/j.tpb.2005.03.001.

Wolcott, L.T., and Neves, R.J. 1994. Survey of the freshwater mussel fauna of the Powell River, Virginia. Banisteria, **3**: 3–14.