

# Density dependence in North American ducks

L. E. Jamieson & S. P. Brooks

Jamieson, L. E. & Brooks, S. P., 2004. Density dependence in North American ducks. *Animal Biodiversity and Conservation*, 27.1: 113–128.

## Abstract

*Density dependence in North American ducks.*— The existence or otherwise of density dependence within a population can have important implications for the management of that population. Here, we use estimates of abundance obtained from annual aerial counts on the major breeding grounds of a variety of North American duck species and use a state space model to separate the observation and ecological system processes. This state space approach allows us to impose a density dependence structure upon the true underlying population rather than on the estimates and we demonstrate the improved robustness of this procedure for detecting density dependence in the population. We adopt a Bayesian approach to model fitting, using Markov chain Monte Carlo (MCMC) methods and use a reversible jump MCMC scheme to calculate posterior model probabilities which assign probabilities to the presence of density dependence within the population, for example. We show how these probabilities can be used either to discriminate between models or to provide model-averaged predictions which fully account for both parameter and model uncertainty.

Key words: Bayesian approach, Markov chain Monte Carlo, Model choice, Autoregressive, Logistic, State space modelling.

## Resumen

*Dependencia de la densidad en los ánades norteamericanos.*— La existencia o ausencia de efectos dependientes de la densidad en una población puede acarrear importantes repercusiones para la gestión de la misma. En este artículo empleamos estimaciones de abundancia obtenidas a partir de recuentos aéreos anuales de las principales áreas de reproducción de diversas especies de ánades norteamericanos, utilizando un modelo de estados espaciales para separar los procesos de observación y los procesos del sistema ecológico. Este enfoque basado en estados espaciales nos permite imponer una estructura que depende de la densidad de la población subyacente real, más que de las estimaciones, además de demostrar la robustez mejorada de este procedimiento para detectar la dependencia de la densidad en la población. Para el ajuste de modelos adoptamos un planteamiento bayesiano, utilizando los métodos de Monte Carlo basados en cadenas de Markov (MCMC), así como un programa MCMC de salto reversible para calcular, por ejemplo, las probabilidades posteriores de los modelos que asignan probabilidades a la presencia de una dependencia de la densidad en la población. También demostramos cómo pueden emplearse estas probabilidades para discriminar entre modelos o para proporcionar predicciones promediadas entre modelos que tengan totalmente en cuenta tanto la incertidumbre referente a parámetros como a modelos.

Palabras clave: Enfoque bayesiano, Métodos de Monte Carlo basados en cadenas de Markov, Autorregresivo, Logístico, Modelación de espacio de estados.

Lara E. Jamieson & S. P. Brooks, *The Statistical Laboratory – CMS, Univ. of Cambridge, Wilberforce Road, Cambridge CB3 0WB, U.K.*

## Introduction and model

National and international legislation are increasingly putting pressure on local authorities to identify and protect key wildlife species and their habitats. This brings to the forefront the design and implementation of effective management strategies and makes them of paramount importance. In order to design an appropriate management strategy, key factors affecting the population must be understood. In particular, the identification of factors affecting survival and/or population size becomes an integral part of the management design process.

The question as to whether or not population density affects population size is one of the first that must be addressed. See Nichols et al. (1995), Bulmer (1975) and Vickery & Nudds (1984), for example. Essentially density dependence within a population acts as a stabilising mechanism which tends to move the population size towards its mean level. When the population size is small, there is a natural pressure on the population, generally in the form of an abundance of resources, to increase its numbers and vice versa. Thus, density dependence increases the population's ability to cope with "shocks" to the system (i.e., rapid increases or, more often, decreases in population size). In terms of management policy, the presence of density dependence within the population indicates an ability to resist destabilising perturbations to the system often induced by human activity (Nichols et al., 1984; Massot et al., 1992). Such activities might vary from direct effects of hunting to less obvious effects from changes in land use within the population's natural habitat, for example.

Most of the work in this area has focused upon simple hypothesis tests for the presence of density dependence. Though the data available are often rich and varied, unfortunately much of it is ignored in order to facilitate model fitting and the selection process due to analytic tractability.

Data in the form of population estimates and standard errors are common in the population ecology literature and data collection authorities are often very protective of the raw data i.e., the original counts upon which the estimates are based. Though Kalman filter-based methods are available (Sullivan, 1992; Newman, 1998), most analyses simply ignore the standard errors and treat the population estimates as if they were the true population sizes. Such simplifying assumptions often

compound the problem suggesting density dependence (Vickery & Nudds, 1984; Shenk et al., 1998). The false detection of density dependence can have a substantial effect upon the corresponding management strategy with possibly dire consequences for the species concerned.

Management of population sizes feed into key policy-making initiatives for biodiversity and environmental preservation policies. The issuance of planning permits for work in key habitat areas and recreational hunting licences must consider the impact on population status. The "ideal" population level is fixed by the "North American Waterfowl Management Plan" which is, in itself, a valuable source of information. The data and population goals are summarised in fig. 1.

The datasets under consideration comprise yearly population indices for ten duck breeding populations in North America from 1955 to 2002 (Williams, 1999). The ten time series are for Mallard (*Anas platyrhynchos*), Gadwall (*Anas strepera*), American Wigeon (*Anas americana*), Green-winged Teal (*Anas crecca*), Blue-winged Teal (*Anas discors*), Northern Shoveler (*Anas clypeata*), Northern Pintail (*Anas acuta*), Redhead (*Aythya americana*), Canvasback (*Aythya valisineria*) and Scaup (Greater-*Aythya marila* and Lesser-*Aythya affinis*). Of these ten species the latter three are diving ducks, the remainder are dabblers. Data comes in the form of annual population size estimates (based upon the raw counts and then "adjusted for biases" by comparing the aerial counts with samples on the ground), together with associated standard errors. The raw count data themselves are, unfortunately, not available.

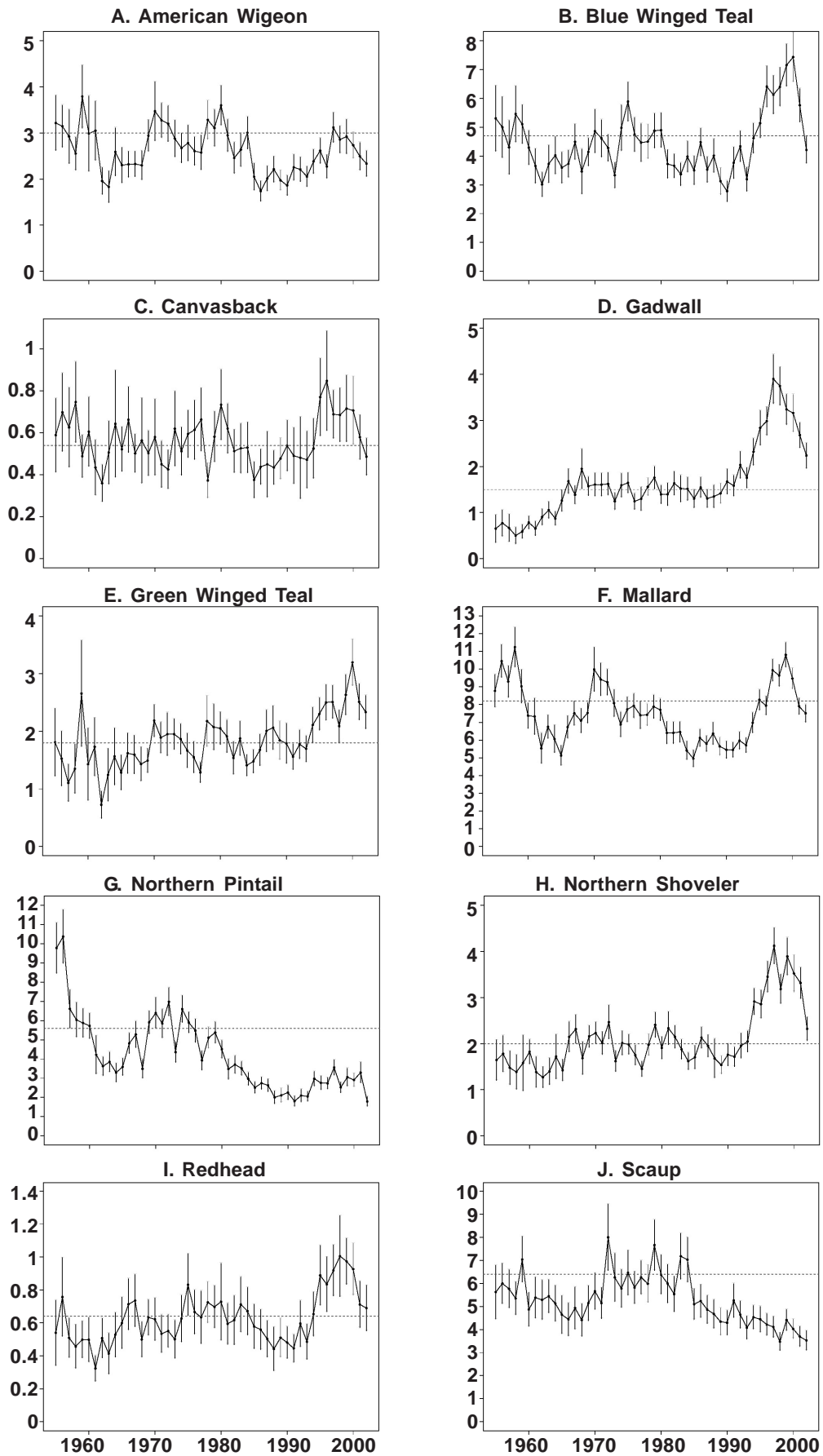
In order to make full use of the data available and to avoid this confounding effect, we seek to separate the observation error (associated with the data collection and population size estimation processes) from the system error (associated with the variability within the population itself) so that the density dependence model is fitted to the true underlying (but hidden) population size rather than the estimates provided. This can be most easily achieved through the development of a state space model separating the system and observation processes as follows.

Let us now introduce the population model we consider. Denote the population abundance at time  $t$  as  $N_t = 1, 2, \dots, T$ . We use the density dependence model of Dennis & Taper (1994) in which

$$N_t = N_{t-1} \exp\{b_0 + b_1 N_{t-1} + \sigma Z_t\} \quad (1)$$

Fig. 1. The population index values  $\hat{N}_t$  (in millions) together with approximate 95% confidence intervals derived from the standard errors and the population goal (horizontal line).

Fig. 1. Valores de los índices poblacionales  $\hat{N}_t$  (en millones), junto con los intervalos de confianza aproximados al 95%, derivados de los errores estándar y del objetivo poblacional (línea horizontal).



where  $Z_t \sim N(0,1)$ . There is considerable debate over the choice of a model for density dependence; Dennis & Taper (1994) discuss a number of models and motivate their model by considering the per-unit-abundance growth rate  $\log N_{t+1} - \log N_t$  as a linear function with a *Normal* error term. For computational ease we use a log transformation,  $p_t = \log_e N_t$ , to give the additive model

$$p_t = p_{t-1} + b_0 + b_1 \exp\{p_{t-1}\} + \sigma Z_t \quad (2)$$

We shall refer to this model as the logistic model Dennis & Taper (1994).

Clearly, setting  $b_1 = 0$  in (2) reduces the model to a simple linear trend with normal errors. The model may also be extended to include longer-range dependence by taking

$$p_t = p_{t-1} + b_0 + \sum_{\tau=1}^k b_\tau \exp\{p_{t-\tau}\} + \sigma Z_t \quad (3)$$

for some suitable value  $k$ . Turchin (1990) and Turchin et al. (1991) argue for the prevalence of second-order density dependence in ecological populations, for example. Thus, the problem of determining the nature and extent of density dependence within the population therefore reduces to a model selection problem in which the model (indexed by  $k$ ) is unknown. Berryman & Turchin (1991) discuss the inherent difficulty of determining the order of any observed density dependence and suggests the use of autocorrelation-based tests to determine the true order. Amongst other things, we will, in this paper, show how the Bayesian approach provides a very natural framework both for selecting the model order and averaging predictive inference across a range of models when the true model order is unclear.

The  $N_t$  above denote the true population sizes, which are unknown. However, the data provide us with information as to what these values might be. In particular, the data provide us with estimates  $\hat{N}_t$  of the true underlying population size together with associated standard errors  $s_t$ . The observation process, which relates the observed data to the true underlying population values, is then described by a simple Gaussian process whereby

$$\hat{N}_t | N_t, s_t^2 \sim N(\hat{N}_t | N_t, s_t^2)$$

or equivalently

$$\hat{N}_t | p_t, s_t^2 \sim N(\hat{N}_t | e^{p_t}, s_t^2) \quad (4)$$

for  $t = 1, \dots, T$ .

For notational ease we refer to the full set of abundance values  $\{p_t, t = 1, 2, \dots, T\}$  as  $\mathbf{P}$  and similarly the observed series  $\hat{\mathbf{N}}$  and associated standard errors  $\mathbf{S}$ . Let  $\mathbf{P}_{(t)}$  denote the series  $\mathbf{P}$  without the point  $p_t$ . Our time series model for  $\mathbf{P}$  requires knowledge of previous values; consider for example  $p_1$  on the left-hand side of Equation (2), then the right-hand side references  $p_0$ . Here, we use data from 1955–1960 to place priors on these values which are then imputed as part of the modelling process.

The likelihood for our problem therefore comes in two parts: one corresponding to the system and the other to the observation equations. The model in (3) can be re-written in matrix form as  $\boldsymbol{\varepsilon} = \mathbf{d} - \mathbf{Y}_k \mathbf{b}_k$  where the elements  $\varepsilon_t$  of  $\boldsymbol{\varepsilon}$  have independent normal distributions with zero mean and variance  $\sigma^2$ ,

$$\mathbf{d} = \begin{pmatrix} p_1 - p_0 \\ \vdots \\ p_t - p_{t-1} \\ \vdots \\ p_T - p_{T-1} \end{pmatrix}, \quad \mathbf{Y}_k = \begin{pmatrix} 1 & e^{p_0} & e^{p_1} & \dots & e^{p_{k+1}} \\ 1 & e^{p_1} & e^{p_0} & \dots & e^{p_{k+2}} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & e^{p_{k-1}} & e^{p_{k-2}} & \dots & e^{p_0} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & e^{p_{T-1}} & e^{p_{T-2}} & \dots & e^{p_{T-k}} \end{pmatrix}$$

$$\text{and } \mathbf{b}_k = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix} \quad (5)$$

Hence, the so-called system likelihood, relating the latent process  $\mathbf{P} = \{p_1, \dots, p_T\}$  to the parameters under model  $k$ , is given by

$$L_{\text{sys}}(\mathbf{P} | \mathbf{b}, \sigma^2, \mathbf{P}_k^-) = (2\pi\sigma^2)^{-T/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^T \varepsilon_t^2\right\} \quad (6)$$

where  $\mathbf{P}_k^- = \{p_{1-k}, \dots, p_0\}$  denote additional parameters to be estimated and

$$\varepsilon_t = p_t - p_{t-1} - b_0 - \sum_{\tau=1}^k b_\tau \exp(p_{t-\tau})$$

Similarly, from (4), the so-called observation likelihood, which relates the sequence  $\mathbf{P}$  to the observed data, is given by

$$L_{\text{obs}}(\hat{\mathbf{N}} | \mathbf{P}, \mathbf{S}) = \frac{1}{(2\pi)^{T/2} \prod_{t=1}^T s_t} \exp\left\{-\frac{1}{2} \sum_{t=1}^T \left(\frac{\hat{N}_t - e^{p_t}}{s_t}\right)^2\right\} \quad (7)$$

Of course, the interpretation of  $L_{\text{sys}}$  as a likelihood is slightly misleading in that it contains no terms corresponding to the observed data. However, it is intended to represent the likelihood of the  $\mathbf{P}$  vector had it been observed. An alternative interpretation is as a prior for the unobserved  $\mathbf{P}$ , conditioning on the remaining parameters which is consistent with a Bayesian approach to the analysis of the data. In fact, a Bayesian approach is the only viable option here since, in order to undertake the analysis, the  $N_t$  values must be removed from the joint likelihood by integration so that the model is expressed entirely without reference to the  $N_t$  values. This is impossible analytically and even classical EM-based or Kalman Filtering techniques are prohibitively complex to apply in this non-linear setting. However, the Bayesian approach integrates out the  $N_t$  numerically as part of the MCMC simulation process that we shall describe in the next section.

We discuss the Bayesian approach to statistical inference in the next section before undertaking an analysis of our data. We then extend our range of models to include an alternative density-dependent dynamic and end with some discussion of our results and of their implication for the management of these important wildlife species.

The Bayesian approach

Bayesian analyses involve the combination of the likelihood with the priors to obtain the posterior distribution as the basis for inference.

If we assume that the observed data  $\Phi$  are described by some model  $m$  with associated parameter vector  $\theta \in \Theta$ , then the Bayesian analysis is based upon the posterior distribution  $\pi(\theta | \Phi)$  where

$$\pi(\theta | \Phi) \propto f(\Phi | \theta) p(\theta),$$

$f(\Phi | \theta)$  denotes the joint probability distribution function of the data given  $\theta$  and  $p(\theta)$  denotes a prior distribution representing the analyst's beliefs about the model parameters obtained independently from the data. Markov chain Monte Carlo (MCMC; Brooks, 1998) methods can be used to explore and summarise this posterior distribution.

When the model itself is the subject of inference, the Bayesian posterior distribution can be extended to incorporate model as well as parameter uncertainty. By specifying a prior model probability,  $p(m)$  for models  $m \in M$ , the corresponding posterior distribution becomes

$$\pi(\theta_m, m | \Phi) \propto f_m(\Phi | \theta_m) p_m(\theta_m) p(m),$$

where  $f_m(\Phi | \theta_m)$  denotes the joint probability distribution of the data under model  $m$  given parameter vector  $\theta_m \in \Theta_m$ , and  $p_m(\theta_m)$  denotes the corresponding prior for  $\theta_m$  under model  $m$ . This more complex posterior distribution can be explored and summarised using reversible jump (RJ) MCMC methods (Green, 1995). Posterior inference is often summarised in the form of posterior moments under models of interest and marginal posterior model probabilities. These model probabilities may be used either to discriminate between competing models using Bayes factors (Kass & Raftery, 1995) or to provide model-averaged prediction for parameters that retain a coherent interpretation across models (Clyde, 1999; Madigan et al., 1996). We return to this later in the paper.

The inference problem essentially reduces to the integration of the posterior density function over what is typically a large and complex parameter space. MCMC methods overcome this problem by simulating realisations from the posterior distribution so that empirical estimates can be calculated for any statistic of interest. These realisations are obtained by simulating a Markov chain with  $\pi$  as its stationary distribution. See Gilks et al. (1996), Brooks (1998) and Brooks et al. (2000), for example.

Though posterior model probabilities provide a useful means of model comparison, formal model-checking methods are also required to ensure that the models fitted provide an adequate description of the data. There are two formal model checking procedures commonly used in the literature: the Bayesian  $p$ -value, and cross-validation.

Bayesian  $p$ -values (Gelman & Meng, 1996) can be used to check the discrepancy between the sample values and the observed. The classical discrepancy statistic is to take

$$D(x, \vartheta) = \sum_j (\sqrt{x_j} - \sqrt{e_j})^2$$

where  $x$  is the data or observed values and  $e_j$  the expected;  $j$  indexes the vectors. In our case  $\vartheta_i$  is the sample  $\mathbf{P}$  produced by the  $i^{\text{th}}$  iteration of the MCMC sampler. Using the observed likelihood of (7), we calculate

$$D(x, \vartheta_i) = L_{\text{obs}}(\hat{\mathbf{N}} | e^{\mathbf{P}}, \mathbf{S})$$

We sample  $\tilde{\mathbf{N}}$  from the observed model using  $\vartheta_i$  to obtain

$$D(x_i, \vartheta_i) = L_{\text{obs}}(\tilde{\mathbf{N}} | e^{\mathbf{P}}, \mathbf{S})$$

The Bayesian  $p$ -value is then the proportion of times that  $D(x_i, \vartheta_i) > D(x, \vartheta_i)$  which, if the model describes the data well, should be close to 1/2. See Bayarri & Berger (1998) and Brooks et al. (2000) for further details.

Cross-validation (Gelman et al., 1995; Carlin & Louis, 1996) involves treating observed values as if they were missing and investigating how well the model predicts this value. Suppose we treat observation  $N_{t^*}$  as a missing value and update the remaining parameters as before. For iteration  $i$  sample  $\tilde{N}_{t^*}^i$  using a Metropolis-Hastings step, excluding the  $\hat{N}_{t^*}$  term. We then sample  $\tilde{N}_{t^*}^i$  from the model i.e.,  $N(N_{t^*}^i, s_{t^*}^2)$ , and calculate the ratio  $\tilde{N}_{t^*}^i / \hat{N}_{t^*}$ . Repeating the same procedure for each  $t = 1, \dots, T$  we can calculate

$$CV_i = \frac{1}{T} \sum_{t=1}^T \frac{\tilde{N}_{t^*}^i}{\hat{N}_{t^*}}$$

which, if the model describes the data well, should be close to 1.

The Bayesian analysis

We begin our analysis by determining prior distributions for each of the parameters in the different models (and for the models themselves). We have full data from 1961 to 2002; additionally we also have population size estimates (and corresponding errors) for each of the duck species from 1955 to 1960. Thus we have informative priors for the latent subseries values  $\mathbf{P}^k = \{p_{1-k}, \dots, p_0\}$ . Recall that the system likelihood essentially acts as a prior for the remaining log population levels,  $\mathbf{P}$ . To



Table 1. Prior parameter values used for the data analysis for the different duck species.

Tabla 1. Valores de parámetros previos empleados para los análisis de datos correspondientes a diferentes especies de ánades.

Duck	$\mu_0$	$\sigma_0$	$\sigma_b^2$	Duck	$\mu_0$	$\sigma_0$	$\sigma_b^2$
Mallard	8.200	0.632	0.3	Northern Shoveler	2.000	0.308	2.0
Gadwall	1.500	0.274	2.0	Northern Pintail	5.600	0.707	0.5
American Wigeon	3.000	0.412	1.0	Redhead	0.640	0.141	5.0
Blue Winged Teal	4.700	0.592	3.0	Canvasback	0.540	0.130	4.0
Green Winged Teal	1.800	0.224	1.0	Scaup	6.400	0.742	0.5

obtain informative priors for  $\mathbf{b}$  we also use the additional data to obtain a suitable range from a preliminary analysis, as follows.

We first analyse the data 1955–1960 with vague priors on the unobserved  $\hat{\rho}_t$ , taking a *Normal* distribution based upon the North American Waterfowl Management Plan (Williams et al., 1999) population goal, so that

$$e^{\rho_t} \stackrel{iid}{\sim} N(\mu_0, \sigma_0^2)$$

for all  $t = 1-k, \dots, 0$  with  $\mu_0$  and  $\sigma_0^2$  given in table 1.

This initial analysis provides an estimate of the population means and variances for  $\mathbf{b}$  reflecting expert knowledge obtained independently from the data. Thus, for the main analysis we can take  $b_j \sim N(0, \sigma_b^2)$ ,  $\forall j = 0, \dots, k$ , where the  $\sigma_b^2$  are given in table 1 for the different duck species. We use the observed data as our prior for  $N_p$ ,  $t = 1-k, \dots, 0$  and, for the remaining parameters we take  $\sigma^{-2} = \tau \sim \Gamma(a, \beta)$ , with  $a = \beta = 10^{-3}$ , and adopt a uniform model prior  $k \sim U[0, k_{max}]$  where, here, we take  $k_{max} = 5$ , in order to cover a reasonable range of plausible density dependence models.

Initial runs of the code were made to investigate convergence and mixing. By mixing we mean the extent and spread with which the parameter space is explored by the chain. Using widely dispersed starting values we observed that the chains converge quickly, within 5,000 iterations. Graphical checks were made on each of the parameters updated in the chain and other standard diagnostic techniques were used (Brooks & Roberts, 1998).

Prior sensitivity was investigated, in particular the prior for  $\sigma^2$ . Although it has been common practise to use an inverse-gamma prior with both parameters set to  $10^{-3}$  there is a growing body of evidence that suggests that this could be hazardous. To check this assumption we used a variety of parameters for the inverse-gamma as well as a  $U(0, 10,000)$ . Our results did not alter substantially and we are therefore content with our choice of prior. Similar sensitivity

studies were undertaken for the remainder of the model parameters, with similar robustness observed. Finally, we tested our simulation algorithm using simulated data from a range of models and parameter values. The algorithm consistently identified the correct model and provided highly accurate parameter estimates (i.e., the posterior means were close to the values used to simulate the data). We therefore conclude that the algorithm performs well for a range of different data sets consistent with those observed.

The analyses are based upon MCMC simulations comprising 2,020,000 iterations, with the first 20,000 discarded as burn-in. Table 2 provides the posterior model probabilities for the six possible models for each of the data sets, together with cross-validation and  $p$ -values.

For many of the duck species the  $k = 0$  model attracts the highest posterior model probability. However, there is some evidence to suggest the existence of density dependence for the Redhead, Canvasback and possibly the American Wigeon. Vickery & Nudds (1984) suggest that diving duck species tend to exhibit a greater degree of density dependence than dabblers and so it is interesting to note that the Redhead and Canvasback ducks are both divers whereas the rest (apart from Scaup) are all dabbling ducks. Thus, there does indeed appear to be some evidence for a distinction in behaviour between diving and dabbling ducks, though the Scaup and American Wigeon might perhaps be anomalous species. We shall return to this point later.

Both the cross-validation and  $p$ -values suggest that most models provide an adequate fit for the data, though the  $p$ -values do appear to increase with the value of  $k$  for most species. The  $p$ -values also seem to be somewhat higher for the Green Winged Teal, Canvasback and especially the Scaup. However, all of the  $p$ -values lies within the (0.05, 0.95) range and therefore provide no evidence to suggest that any of the models perform poorly.

Table 2. Posterior model probabilities (MP), cross-validation (CV) and  $p$ -values for the logistic models with  $k = 0, \dots, 5$ .

Tabla 2. Probabilidades posteriores de los modelos (MP), validación cruzada (CV) y valores  $p$  para los modelos logísticos con  $k = 0, \dots, 5$ .

$k$	MP	CV	$p$ -value	$k$	MP	CV	$p$ -value
<b>Mallard</b>				<b>Northern Shoveler</b>			
0	0.848	0.999	0.583	0	0.807	0.998	0.604
1	0.136	0.999	0.556	1	0.177	0.999	0.564
2	0.015	1.003	0.660	2	0.015	1.004	0.548
3	0.001	0.998	0.692	3	0.001	0.999	0.567
4	0.000	0.999	0.645	4	0.000	0.998	0.616
5	0.000	0.999	0.679	5	0.000	0.998	0.610
<b>Gadwall</b>				<b>Northern Pintail</b>			
0	0.835	0.995	0.649	0	0.850	0.998	0.621
1	0.108	0.996	0.629	1	0.104	0.997	0.583
2	0.037	0.997	0.855	2	0.043	0.998	0.545
3	0.008	0.994	0.897	3	0.003	0.996	0.543
4	0.010	0.993	0.820	4	0.000	0.998	0.558
5	0.002	0.993	0.850	5	0.000	0.996	0.566
<b>American Wigeon</b>				<b>Redhead</b>			
0	0.463	0.993	0.673	0	0.020	0.996	0.518
1	0.397	0.994	0.585	1	0.016	0.997	0.468
2	0.125	0.997	0.687	2	0.658	0.998	0.518
3	0.013	0.992	0.768	3	0.179	0.993	0.689
4	0.002	0.993	0.804	4	0.100	0.995	0.613
5	0.000	0.993	0.814	5	0.027	0.995	0.615
<b>Blue Winged Teal</b>				<b>Canvasback</b>			
0	0.729	0.997	0.599	0	0.117	0.98	0.780
1	0.254	0.998	0.557	1	0.324	0.983	0.652
2	0.016	1.002	0.626	2	0.353	0.980	0.700
3	0.001	0.997	0.655	3	0.126	0.975	0.860
4	0.000	0.997	0.788	4	0.056	0.979	0.862
5	0.000	0.997	0.769	5	0.025	0.982	0.897
<b>Green Winged</b>				<b>Scaup</b>			
0	0.702	0.988	0.807	0	0.980	0.989	0.885
1	0.252	0.989	0.767	1	0.018	0.989	0.848
2	0.034	0.991	0.796	2	0.002	0.996	0.821
3	0.007	0.986	0.834	3	0.000	0.989	0.860
4	0.004	0.986	0.846	4	0.000	0.989	0.917
5	0.001	0.989	0.798	5	0.000	0.990	0.902

Table 3 provides the posterior means and 95% HPDI's for four typical duck species. In each case, the posterior mean for  $b_0$  is fairly close to zero for the  $k = 0$  model reflecting the fairly constant nature

of the population level of these four species, though we note that the value for the Northern Pintail is negative reflecting the population decline in figure 1. For the  $k = 1$  models, the value of  $b_1$  is

Table 3. Parameter estimates and 95% HPDI's for four duck species under the logistic models with  $k = 0, \dots, 5$ .

Tabla 3. Estimaciones de parámetros e intervalos de predicción de alta densidad (HPDI) al 95%, para cuatro especies de ánades con arreglo a los modelos logísticos con  $k = 0, \dots, 5$ .

$k$	$b_0$	$\sigma^2$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$
<b>Blue Winged Teal</b>							
0	0.001 <sub>(-0.045,0.047)</sub>	0.022 <sub>(0.011,0.035)</sub>					
1	0.226 <sub>(0.023,0.437)</sub>	0.022 <sub>(0.011,0.034)</sub>	-0.051 <sub>(-0.098,-0.006)</sub>				
2	0.247 <sub>(0.045,0.455)</sub>	0.020 <sub>(0.008,0.033)</sub>	-0.012 <sub>(-0.101,0.077)</sub>	-0.044 <sub>(-0.132,0.044)</sub>			
3	0.230 <sub>(0.003,0.459)</sub>	0.019 <sub>(0.008,0.032)</sub>	-0.004 <sub>(-0.102,0.097)</sub>	-0.059 <sub>(-0.217,0.089)</sub>	0.011 <sub>(-0.100,0.124)</sub>		
4	0.247 <sub>(0.010,0.458)</sub>	0.014 <sub>(0.002,0.029)</sub>	0.067 <sub>(-0.085,0.070)</sub>	-0.192 <sub>(-0.495,0.070)</sub>	0.138 <sub>(-0.106,0.425)</sub>	-0.065 <sub>(-0.215,0.074)</sub>	
5	0.344 <sub>(0.109,0.579)</sub>	0.013 <sub>(0.003,0.025)</sub>	0.022 <sub>(-0.083,0.137)</sub>	-0.084 <sub>(-0.282,0.080)</sub>	-0.001 <sub>(-0.161,0.202)</sub>	0.101 <sub>(-0.074,0.254)</sub>	-0.118 <sub>(-0.226,-0.006)</sub>
<b>Northern Pintail</b>							
0	-0.026 <sub>(-0.085,0.034)</sub>	0.038 <sub>(0.020,0.059)</sub>					
1	0.126 <sub>(-0.052,0.309)</sub>	0.038 <sub>(0.021,0.058)</sub>	-0.040 <sub>(-0.084,0.005)</sub>				
2	0.085 <sub>(-0.100,0.269)</sub>	0.038 <sub>(0.021,0.057)</sub>	-0.101 <sub>(-0.180,-0.021)</sub>	0.071 <sub>(-0.007,0.147)</sub>			
3	0.073 <sub>(-0.114,0.263)</sub>	0.038 <sub>(0.021,0.058)</sub>	-0.108 <sub>(-0.195,-0.023)</sub>	0.057 <sub>(-0.042,0.153)</sub>	0.024 <sub>(-0.062,0.107)</sub>		
4	0.075 <sub>(-0.117,0.265)</sub>	0.039 <sub>(0.021,0.059)</sub>	-0.106 <sub>(-0.194,-0.020)</sub>	0.060 <sub>(-0.044,0.162)</sub>	0.031 <sub>(-0.067,??)</sub>	-0.012 <sub>(-0.98,0.072)</sub>	
5	0.086 <sub>(-0.096,0.268)</sub>	0.034 <sub>(0.018,0.053)</sub>	-0.115 <sub>(-0.200,-0.030)</sub>	0.069 <sub>(-0.028,0.165)</sub>	0.055 <sub>(-0.040,0.150)</sub>	0.037 <sub>(-0.056,0.131)</sub>	-0.073 <sub>(-0.139,-0.006)</sub>
<b>Redhead</b>							
0	0.013 <sub>(-0.018,0.045)</sub>	0.010 <sub>(0.004,0.018)</sub>					
1	0.179 <sub>(0.020,0.347)</sub>	0.010 <sub>(0.004,0.018)</sub>	-0.265 <sub>(-0.528,-0.013)</sub>				
2	0.194 <sub>(0.070,0.324)</sub>	0.006 <sub>(0.002,0.011)</sub>	0.358 <sub>(-0.101,0.734)</sub>	-0.652 <sub>(-1.031,-0.196)</sub>			
3	0.171 <sub>(0.070,0.280)</sub>	0.002 <sub>(0.000,0.005)</sub>	0.834 <sub>(-0.161,1.803)</sub>	-0.832 <sub>(-2.678,1.080)</sub>	-0.278 <sub>(-1.410,0.746)</sub>		
4	0.227 <sub>(0.089,0.374)</sub>	0.002 <sub>(0.000,0.006)</sub>	0.662 <sub>(-0.382,1.742)</sub>	-1.034 <sub>(-3.034,1.037)</sub>	0.642 <sub>(-1.592,2.431)</sub>	-0.640 <sub>(-1.553,0.454)</sub>	
5	0.227 <sub>(0.041,0.446)</sub>	0.003 <sub>(0.000,0.006)</sub>	0.594 <sub>(-0.557,1.698)</sub>	-0.820 <sub>(-2.858,0.965)</sub>	0.435 <sub>(-1.418,2.406)</sub>	-0.600 <sub>(-2.255,1.194)</sub>	0.021 <sub>(-0.008,1.085)</sub>
<b>Canvasback</b>							
0	0.002 <sub>(-0.027,0.030)</sub>	0.008 <sub>(0.001,0.017)</sub>					
1	0.334 <sub>(0.039,0.669)</sub>	0.011 <sub>(0.002,0.022)</sub>	-0.625 <sub>(-1.253,-0.074)</sub>				
2	0.345 <sub>(0.063,0.664)</sub>	0.009 <sub>(0.002,0.019)</sub>	-0.239 <sub>(-1.254,0.571)</sub>	-0.407 <sub>(-1.121,0.486)</sub>			
3	0.277 <sub>(0.052,0.567)</sub>	0.004 <sub>(0.000,0.013)</sub>	0.305 <sub>(-1.027,1.932)</sub>	-0.532 <sub>(-2.726,1.137)</sub>	-0.295 <sub>(-1.249,0.820)</sub>		
4	0.312 <sub>(0.026,0.653)</sub>	0.004 <sub>(0.000,0.012)</sub>	0.136 <sub>(-1.287,1.569)</sub>	-0.374 <sub>(-2.205,1.348)</sub>	0.031 <sub>(-1.645,0.089)</sub>	-0.382 <sub>(-1.754,0.876)</sub>	
5	0.340 <sub>(-0.007,0.669)</sub>	0.003 <sub>(0.000,0.010)</sub>	0.013 <sub>(-1.222,1.524)</sub>	-0.294 <sub>(-2.188,1.366)</sub>	0.429 <sub>(-1.499,2.260)</sub>	-0.538 <sub>(-2.173,1.310)</sub>	0.254 <sub>(-1.623,1.039)</sub>

always negative, as we would expect. Perhaps most interesting are the parameter estimates for the *a posteriori* most probable model for the Redhead duck. The second-order model is suggested here with a positive coefficient for the population size in the previous year and a negative (and larger) coefficient for the year before that. This suggests that the population size two years earlier has the greatest effect on the current population size. We shall return to discuss the underlying dynamics of the Redhead population later.

Figure 2 provides a model-averaged plot of the estimated population sizes across time and compares them with the observed data for four duck species. The Canvasback and Redhead plots are clearly much smoother than the other two since the posterior is dominated by density dependent models which have a natural smoothing effect. Note for example the occasional quite substantial differences between the population estimates  $\hat{N}_t$  and the posterior means for the  $N_t$  especially for the Canvasback data set. This is a result of the smooth-



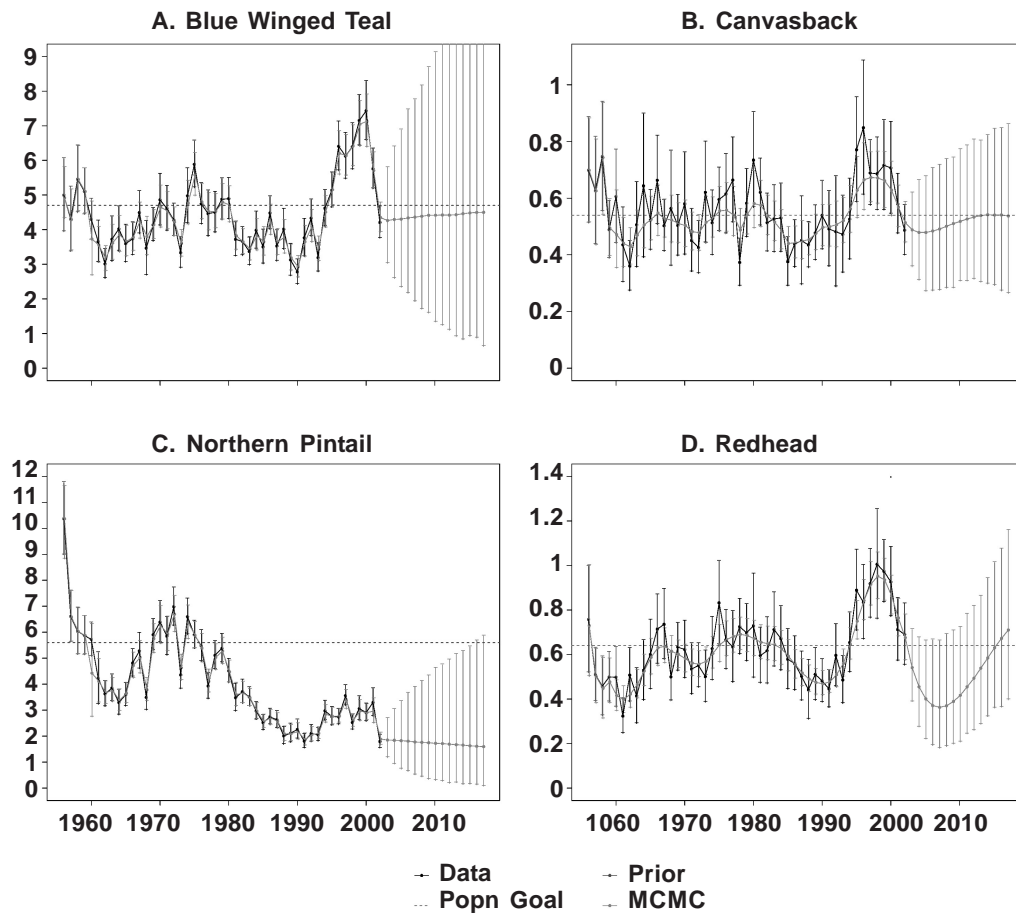


Fig. 2. Plot of observed population estimates  $\hat{N}_t$  (in millions) together with model-averaged posterior means and 95% HPDI's for the  $N_t$  under the logistic models with  $k = 0, \dots, 5$  (i.e., six models per species). Posterior predictive values for the next fifteen years together with corresponding 95% HPDI's are also included.

Fig. 2. Representación gráfica de las estimaciones poblacionales observadas  $\hat{N}_t$  (en millones), junto con las medias posteriores de los promedios de predicciones según los modelos y los HPDI al 95% para  $N_t$  con arreglo a los modelos logísticos con  $k = 0, \dots, 5$  (es decir, seis modelos por especie). También se incluyen los valores posteriores de predicción para los próximos quince años, junto con los correspondientes HPDI al 95%.

ing effect of the density dependent system process which moderates the observed estimates removing any unlikely sharp changes in the population level.

Figure 2 also provides the posterior means for the elements of  $\mathbf{P}_k^-$  together with 95% HPDI's as well as a model-averaged posterior predictive plot for the true population size for the 15 years following the completion of the study. Though the 95% HPDI's for the predicted population levels are fairly wide, the plots for both the Blue Winged Teal and the Canvasback suggest that the current management of these species is very much in keeping with the aim of the population goals set. However, the prediction for the Northern Pintails suggests

that the decline observed over recent years is likely to continue with very little chance of achieving the agreed population goal under the current management regime. On the other hand, the Red-head predictions suggest that the recent decline may soon end and that the population goal may well be achieved within the medium term. Thus, whilst the management of the Blue Winged Teal and Canvasbacks appears to be performing well, that of the Northern Pintail and Redhead might well be improved. In fact, similar predictive plots for the remaining duck species, suggests that the population levels for the Scaup and American Wigeon are also set to fall well below their popu-

lation goals, whilst those for the Gadwall, Green Winged Teal and Northern Shoveler will continue to exceed them.

Of course, these analyses are based upon the assumption that growth depends linearly upon population size. In the next section we challenge this assumption by considering an alternative model that assumes that the growth rate depends only logarithmically upon population size.

#### An alternative density-dependent model

There is considerable debate over the best model for describing density dependence in models of this sort. The model of Equation (1) is a recent addition to the literature and is based upon the assumption that the per-unit-abundance growth rate can be defined in discrete time as  $\log N_{t+1} - \log N_t$  plus noise, see Dennis & Taper (1994). An alternative first-order population model was suggested by Reddingius (1971) and sets

$$N_t = N_{t-1} \exp \{b_0 + b_1 \log N_{t-1} + \sigma Z_t\} \quad (8)$$

Using the log transformation introduced above, the model becomes

$$\begin{aligned} p_t &= p_{t-1} + b_0 + b_1 p_{t-1} + \sigma Z_t = \\ &= b_0 + (1 + b_1) p_{t-1} + \sigma Z_t \end{aligned}$$

Trivial manipulations can be used to recast this model as a simple first-order autoregressive model which can be easily analysed using most standard statistical packages. We therefore refer to this model as the autoregressive model. Royama (1981) extends this model to consider higher order density dependence providing models equivalent to the general  $k$ th-order autoregressive process. Whilst this model has convenient statistical properties (e.g., with  $k = 1$ , taking  $|1 + b_1| < 1$  leads to a stationary model) and is easy to fit to observed data, it relies on the assumption that growth depends only logarithmically on population density and is therefore somewhat weaker than the logistic model.

To consider the suitability of the autoregressive model for describing our data, we can repeat the analysis described in previous sections to fit this model to our data and to determine the order of the density dependence. We also extend our analysis to discriminate between the logistic and autoregressive models by calculating posterior model probabilities for suitable values of  $k$  under both models.

As with the logistic model, we analysed the 1955–1960 data separately first to obtain priors for  $b$  for the subsequent analysis of the 1961–2002 data sets of primary interest. The remaining priors are unchanged. We ran the RJMCMC simulations for a total of 2,020,000 iterations, discarding the first 20,000 iterations as before.

Table 4 provides the posterior model probabilities for the logistic and autoregressive models of different orders, together with CV and  $p$ -values for the autoregressive models (the corresponding values for the logistic models are provided in table 2). (Note that when  $k = 0$ , the logistic and autoregressive models are identical and so they have identical CV and  $p$ -values.) The broad interpretation of the results of table 4 are similar to those of table 2 in that the majority of species show little evidence of being density dependent apart from the Redhead, Canvasback, American Wigeon and now also the Blue Winged Teal. The posterior model probabilities for the  $k = 0$  model have generally decreased with the introduction of the autoregressive model since, for most species, the autoregressive models attract higher posterior model probabilities than the corresponding logistic models. This suggests that the majority of species are best modelled by the autoregressive model, though the Redhead and Canvasback species appear to be best described by the logistic model.

In terms of density dependence, the Canvasback and Redhead species both have high posterior probability on density-dependent models, as do the American Wigeon and Blue Winged Teal. We note that for the American Wigeon and Blue Winged Teal, the posterior support for the first-order autoregressive model is considerably stronger than the support for the first-order logistic model which now dominates the zeroth order model. According to the CV and  $p$ -values, all models appear to fit reasonably well though, as before, the  $p$ -values seem to increase with  $k$  and are particularly high for the higher-order Canvasback and Scaup models.

Table 5 provides the posterior means and 95% HPDI's for the autoregressive models for our four key species. We first note the strong agreement between this and table 3 in terms of the system error,  $\sigma^2$ . We also see that the first-order models are all stationary with  $b_1$  negative. Though somewhat more complex to check it can also be shown that the *a posteriori* most probable model for the Blue Winged Teal (the  $k = 1$  model) is also stationary (roots are 0.941 and 0.047; see Diggle, 1990) so that the population should neither explode nor die away but settle to a value around 4.4 million.

Figure 3 provides the model-averaged predictive plots for the four key duck species. In fact there is very little difference between these plots and those provided in figure 2. This is because, in both figures, essentially the same models dominate the overall plot. The one exception is for the Blue Winged Teal, where the first-order autoregressive model now dominates the  $k = 0$  model. Whereas the predictive plot for this species appears to be slowly rising in figure 2A (since  $b_0$  in the  $k = 0$  model is positive), it appears to quickly settle to the value predicted above in figure 3A. Very similar predictive plots are obtained for the remaining duck species and so the management conclusions drawn from the original analysis remain largely unchanged with the introduction of the autoregressive models.

Table 4. Posterior model probabilities (MP) for the logistic (LM) and autoregressive (AM) models for  $k = 0, \dots, 5$  and both with and without the state space element to the models i.e., the observation process model is removed and the  $N_t$  replaced by  $\hat{N}$  in the system process model. The cross-validation and  $p$ -values for the autoregressive state space models are also included.

Tabla 4. Probabilidades posteriores de los modelos (MP) para los modelos logísticos (LM) y autorregresivos (AM) para  $k = 0, \dots, 5$  ambos con y sin el elemento de espacio de estados correspondiente a los modelos; es decir, el modelo de procesos de observación ha sido eliminado, mientras que  $N_t$  ha sido substituido por  $\hat{N}$  en el modelo de procesos del sistema. También se incluyen la validación cruzada y los valores  $p$  para los modelos de estado espaciales autorregresivos.

	$k$	With state space				Without state space	
		MP		Autoregressive		MP	
		LM	AM	CV	$p$ -value	LM	AM
Mallard	0	0.439				0.582	
	1	0.070	0.421	0.998	0.554	0.088	0.220
	2	0.008	0.058	0.999	0.644	0.015	0.049
	3	0.001	0.002	0.998	0.688	0.003	0.012
	4	0.000	0.000	0.999	0.641	0.004	0.021
	5	0.000	0.000	0.999	0.659	0.001	0.006
Gadwall	0	0.629				0.807	
	1	0.081	0.245	0.996	0.619	0.075	0.070
	2	0.028	0.002	0.995	0.748	0.027	0.014
	3	0.006	0.000	0.995	0.805	0.003	0.001
	4	0.008	0.000	0.992	0.643	0.002	0.001
	5	0.001	0.000	0.990	0.852	0.001	0.000
American Wigeon	0	0.224				0.165	
	1	0.192	0.481	0.994	0.584	0.370	0.371
	2	0.061	0.032	0.993	0.705	0.045	0.041
	3	0.006	0.004	0.993	0.734	0.005	0.004
	4	0.001	0.000	0.993	0.745	0.001	0.000
	5	0.000	0.000	0.993	0.770	0.000	0.000
Blue Winged Teal	0	0.273				0.438	
	1	0.095	0.623	0.998	0.543	0.156	0.334
	2	0.006	0.003	0.998	0.588	0.019	0.042
	3	0.000	0.000	0.998	0.609	0.003	0.006
	4	0.000	0.000	0.999	0.650	0.000	0.001
	5	0.000	0.000	0.998	0.741	0.000	0.000
Green Winged Teal	0	0.470				0.276	
	1	0.169	0.328	0.990	0.753	0.341	0.259
	2	0.023	0.001	0.989	0.786	0.076	0.031
	3	0.005	0.002	0.987	0.844	0.011	0.003
	4	0.003	0.000	0.989	0.850	0.002	0.000
	5	0.000	0.000	0.992	0.761	0.001	0.000
Northern Shoveler	0	0.541				0.655	
	1	0.119	0.318	0.999	0.559	0.135	0.145
	2	0.010	0.012	0.999	0.545	0.034	0.024
	3	0.001	0.000	0.999	0.548	0.004	0.002
	4	0.000	0.000	0.999	0.612	0.001	0.000
	5	0.000	0.000	0.998	0.621	0.000	0.000

Table 4. (Cont.)

	<i>k</i>	With state space				Without state space	
		MP		Autoregressive		MP	
		LM	AM	CV	<i>p</i> -value	LM	AM
Northern Pintail	0	0.530				0.497	
	1	0.065	0.225	0.998	0.585	0.078	0.134
	2	0.027	0.150	0.998	0.5437	0.076	0.146
	3	0.002	0.001	0.998	0.541	0.017	0.036
	4	0.000	0.000	0.999	0.612	0.003	0.007
	5	0.000	0.000	0.997	0.558	0.002	0.006
Redhead	0	0.019				0.360	
	1	0.015	0.013	0.997	0.464	0.366	0.186
	2	0.629	0.031	0.995	0.63	0.055	0.016
	3	0.171	0.000	0.995	0.641	0.007	0.001
	4	0.096	0.000	0.995	0.602	0.006	0.001
	5	0.026	0.000	0.996	0.589	0.002	0.000
Canvasback	0	0.095				0.017	
	1	0.262	0.186	0.983	0.631	0.562	0.295
	2	0.285	0.004	0.977	0.809	0.079	0.030
	3	0.102	0.000	0.977	0.835	0.011	0.003
	4	0.045	0.000	0.979	0.861	0.002	0.000
	5	0.020	0.000	0.982	0.906	0.001	0.000
Scaup	0	0.896				0.420	
	1	0.017	0.083	0.989	0.850	0.085	0.173
	2	0.002	0.003	0.989	0.826	0.051	0.151
	3	0.000	0.000	0.989	0.855	0.019	0.078
	4	0.000	0.000	0.991	0.881	0.003	0.017
	5	0.000	0.000	0.989	0.894	0.001	0.004

## Discussion

In this paper we provide a detailed analysis of the North American duck census data collected over the past forty years. We analyse the data for ten separate species and investigate the possibility that each species exhibits density dependent fluctuations in population level. We discuss the two distinct density dependence models proposed in the literature and conclude that both provide adequate fits to our data. However, some species appear to be better described by one model rather than the other.

The imputation of the true underlying population levels from the observed estimates has a significant impact on the results of the analysis. The rightmost columns of table 4 provide the corresponding posterior model probabilities when we treat the  $\hat{N}_t$  as if they were the true population levels i.e., removing the state space element of the

model and ignoring the estimated standard errors. These can be compared with the corresponding results under the state space models provided in the leftmost columns. Table 4 suggests that without the state space element to the model, the American Wigeon are equally well described by the logistic and autoregressive models and the Redhead population have significant posterior support for the model without density dependence with very little support for the second-order model. Also, without the state space element, the Canvasback population have high posterior support for the first-order logistic model with moderate support for the first-order autoregressive model as well. Though not provided, the corresponding CV and *p*-values have also been calculated for the non state-space models. Whilst most of the CV values remain high (at around 0.85), all of the models have extremely low *p*-values (all less than

Table 5. Parameter estimates and 95% HPDI's for four duck species under the autoregressive models with  $k = 0, \dots, 5$ .

Tabla 5. Estimaciones de parámetros y HDPI al 95% para cuatro especies de ánades con arreglo a los modelos autorregresivos con  $k = 0, \dots, 5$ .

$k$	$b_0$	$\sigma^2$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$
<b>Blue Winged Teal</b>							
0	0.001 (-0.045, 0.047)	0.022 (0.011, 0.035)					
1	0.383 (0.023, 0.437)	0.022 (0.011, 0.034)	-0.263 (-0.098, -0.006)				
2	0.247 (0.045, 0.455)	0.020 (0.008, 0.033)	-0.012 (-0.101, 0.077)	-0.044 (-0.132, 0.044)			
3	0.383 (0.037, 0.738)	0.020 (0.009, 0.033)	-0.128 (-0.534, 0.284)	-0.175 (-0.773, 0.404)	0.040 (-0.404, 0.465)		
4	0.401 (0.027, 0.781)	0.019 (0.005, 0.033)	-0.073 (-0.589, 0.546)	-0.268 (-1.293, 0.494)	0.151 (-0.570, 0.974)	-0.087 (-0.567, 0.373)	
5	0.546 (0.198, 0.904)	0.014 (0.004, 0.026)	0.007 (-0.414, 0.427)	-0.277 (-0.907, 0.281)	-0.035 (-0.558, 0.562)	0.408 (-0.191, 0.940)	-4.824 (-0.894, -0.044)
<b>Northern Pintail</b>							
0	-0.026 (-0.085, 0.034)	0.038 (0.020, 0.059)					
1	0.157 (-0.064, 0.380)	0.038 (0.021, 0.059)	-0.144 (-0.311, 0.024)				
2	0.110 (-0.114, 0.335)	0.038 (0.021, 0.057)	-0.403 (-0.727, -0.068)	0.292 (-0.035, 0.610)			
3	0.097 (-0.132, 0.326)	0.038 (0.021, 0.058)	-0.424 (-0.775, -0.075)	0.222 (-0.192, 0.627)	0.100 (-0.251, 0.449)		
4	0.103 (-0.129, 0.337)	0.038 (0.021, 0.059)	-0.417 (-0.760, -0.061)	0.232 (-0.197, 0.665)	0.135 (-0.278, 0.548)	-0.055 (-0.402, 0.291)	
5	0.127 (-0.095, 0.350)	0.035 (0.019, 0.053)	-0.441 (-0.779, -0.096)	0.269 (-0.140, 0.668)	0.210 (-0.182, 0.604)	0.169 (-0.236, 0.563)	-0.320 (-0.632, -0.002)
<b>Redhead</b>							
0	0.013 (-0.018, 0.045)	0.010 (0.004, 0.018)					
1	-0.072 (-0.160, 0.013)	0.010 (0.004, 0.018)	-0.173 (-0.336, -0.014)				
2	-0.072 (-0.114, -0.032)	0.002 (0.000, 0.004)	0.646 (0.311, 0.936)	-0.795 (-1.063, -0.477)			
3	-0.076 (-0.138, -0.016)	0.002 (0.000, 0.005)	0.579 (-0.128, 1.259)	-0.640 (-1.817, 0.698)	-0.951 (-0.851, 0.545)		
4	-0.099 (-0.184, -0.014)	0.002 (0.000, 0.006)	0.520 (-0.480, 1.660)	-0.769 (-2.995, 0.878)	0.344 (-1.335, 1.760)	-0.292 (-0.863, 0.446)	
5	-0.091 (-0.212, 0.003)	0.002 (0.000, 0.006)	0.501 (-0.387, 1.316)	-0.696 (-2.195, 1.102)	0.381 (-1.032, 2.037)	-0.492 (-1.508, 0.753)	0.126 (-0.478, 0.662)
<b>Canvasback</b>							
0	0.002 (-0.027, 0.030)	0.008 (0.001, 0.017)					
1	-0.226 (-0.449, -0.026)	0.011 (0.002, 0.023)	-0.355 (-0.699, -0.048)				
2	-0.178 (-0.371, -0.033)	0.005 (0.000, 0.016)	0.197 (-0.611, 0.821)	-0.474 (-0.989, 0.181)			
3	0.192 (-0.406, -0.029)	0.005 (0.000, 0.014)	0.111 (-0.651, 0.943)	-0.263 (-1.414, 0.682)	-0.148 (-0.685, 0.476)		
4	-0.214 (-0.434, -0.016)	0.004 (0.000, 0.012)	0.014 (-0.739, 0.723)	-0.151 (-1.167, 0.832)	0.115 (-0.817, 1.202)	-0.309 (-1.037, 0.454)	
5	-0.234 (-0.439, -0.022)	0.002 (0.000, 0.010)	0.049 (-0.658, 1.414)	-0.240 (-1.949, 0.734)	0.280 (-0.883, 1.207)	-0.162 (-1.089, 1.453)	0.285 (-0.939, 0.472)

0.15 and most less than 0.05) suggesting that these models provide a particularly poor fit.

A more practical comparison is obtained by comparing the predictive plots between our models and those without the state space element. Figure 4 provides the corresponding plot for the Redhead ducks and can be compared with figure 3D. Treating the population estimates as if they were the true values suggests a far more stable dynamic than the corresponding state space model and far greater uncertainty in future population levels. Of course,

the state space models are very difficult to fit within the classical paradigm which is, perhaps why they have often been ignored in the analysis of data of this sort. However, the Bayesian approach we adopt here easily deals with the imputation of the true underlying population level, providing a far better framework for population management which properly accounts for uncertainties in the data and provides a better model of the true underlying dynamic, essentially removing the observation process from the system model.

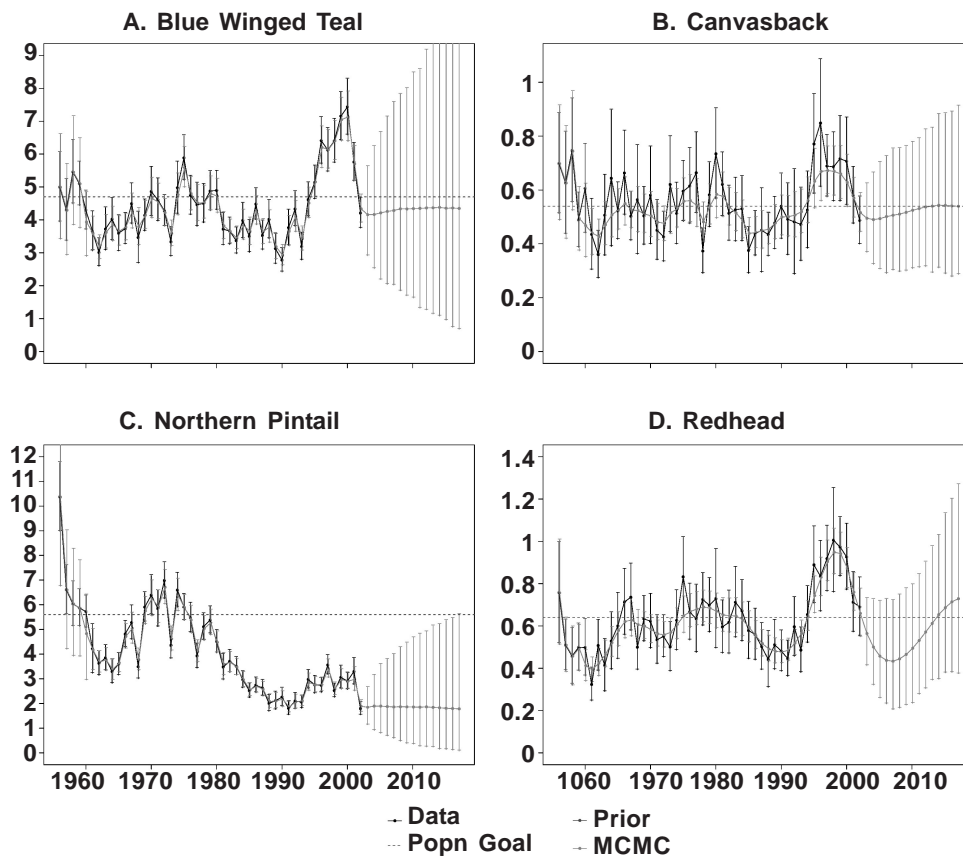


Fig. 3. Plot of observed population estimates  $\hat{N}_t$  (in millions) together with model-averaged posterior means and 95% HPDI's for the  $N_t$  under the logistic and autoregressive models with  $k = 0, \dots, 5$  (i.e., 11 models per species). Posterior predictive values for the next fifteen years together with corresponding 95% HPDI's are also included.

*Fig. 3. Representación gráfica de las estimaciones poblacionales observadas  $\hat{N}_t$  (en millones), junto con las medias posteriores de los promedios de predicciones según los modelos y los HPDI al 95% para  $N_t$  con arreglo a los modelos logísticos y autorregresivos con  $k = 0, \dots, 5$  (es decir, 11 modelos por especie). También se incluyen los valores posteriores de predicción para los próximos quince años, junto con los correspondientes HPDI al 95%.*

An additional advantage of the Bayesian approach is that it can be directly extended to consider alternative observation and system processes. For example, over-dispersed Poisson or log-normal distributions might provide a better model of the observation process. Extensions of this sort are easy to implement by simply adjusting the relevant posterior parameter updates within the MCMC simulation. Within the classical paradigm, such extensions prevent the use of Kalman filtering equations that might otherwise be used for some of the models described within this paper.

Our analyses suggest that the Northern Pintail, Redhead and Canvasback ducks exhibit some form of density dependence. Dynamics of this sort can be explained in any number of ways. For example the process of home-range establish-

ment during the breeding season; restricted resources such as food or prime habitat; or environmental factors such as weather or predator levels affecting mortality in a density dependent manner. The Redhead and Canvasback are both diving ducks and it is popularly believed that divers exhibit a greater tendency for density dependent behaviour than dabbling ducks (Bailey, 1981; Johnson & Grier, 1988; Viljugrein et al., 2004). However, these effects are often masked by the use of overly-simplistic models and, in particular, by treating the observed estimates of population levels as if they were the true underlying level. By separating the observation and system process, we have been able to detect a fairly strong density dependent signal in three duck species and little evidence for density dependence in the rest.



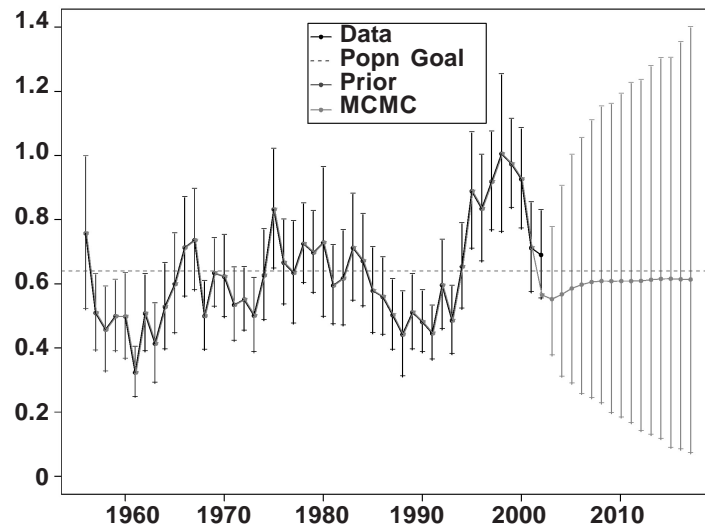


Fig. 4. Plot of observed population estimates  $\hat{N}_t$  (in millions) for the Redhead duck, together with model-averaged posterior means and 95% HPDI's for the  $N_t$  under the logistic and autoregressive models, with  $k = 0, \dots, 5$  (i.e., 11 models), but without the state space element to the models i.e., observation process model removed and  $N_t$  replaced by  $\hat{N}_t$  in the system process model. Posterior predictive values for the next fifteen years together with corresponding 95% HPDI's are also included.

Fig. 4. Representación gráfica de las estimaciones poblacionales observadas  $\hat{N}$  (en millones) para el porrón americano, junto con las medias posteriores de los promedios de predicciones según los modelos y los HPDI al 95% para  $N_t$  con arreglo a los modelos logísticos y autorregresivos, con  $k = 0, \dots, 5$  (es decir, 11 modelos), pero sin el elemento del estado espacial correspondiente a los modelos; es decir, el modelo de procesos de observación ha sido eliminado, mientras que  $N_t$  ha sido substituido por  $\hat{N}_t$  en el modelo de procesos del sistema. También se incluyen los valores posteriores de predicción para los próximos quince años, junto con los correspondientes HDPI al 95%.

By providing predictive plots of population levels, we are also able to assess the success or otherwise of the U.S. Fish and Wildlife Service's management policy in maintaining their target population levels. For many species, predictions suggest that the management policy works extremely well. However, for others, it is clear that the current policy is leading to either a steady growth or decline in the population level. Perhaps the most alarming is the predicted rapid decline of the Northern Pintail and Scaup populations, with the lower limits of the Northern Pintail predicted intervals moving towards dangerously low levels within a very short period.

Ideally, the management process would be integrated into the analysis presented here. Additional data providing information on reproductive and mortality rates together with, for example, hunting license records would enable us to extend this analysis to provide a more detailed picture of the population level dynamics and in particular, their dependence upon human activity. For example, different land-use or licensing policies could be investigated to determine their effect on the population level over

time. We hope that the analyses we present here highlight the importance of the integration of the models and methodology we describe here within the management policy process.

#### References

- Bailey, R. O., 1981. A Theoretical Approach to Problems in Waterfowl Management. *Transactions North American Wildlife and Natural Resources Conference*, 46: 58–71.
- Bayarri, M. & Berger, J., 1998. Quantifying Surprise in the Data and Model Verification (with discussion). In: *Bayesian Statistics VI*: 53–82 (J. Bernardo, J. Berger, A. Dawid & A. Smith, Eds.). Oxford Univ. Press, Oxford.
- Berryman, A. & Turchin, P., 2001. Identifying the Density-Dependent Structure Underlying Ecological Time Series. *Oikos*, 92: 265–270.
- Brooks, S. P., 1998. Markov Chain Monte Carlo Method and its Application. *The Statistician*, 47: 69–100.
- Brooks, S. P., Catchpole, E. A. & Morgan, B. J. T.,

2000. Bayesian Animal Survival Estimation. *Statistical Science*, 15: 357–376.
- Brooks, S. P. & Roberts, G. O., 1998. Diagnosing Convergence of Markov Chain Monte Carlo Algorithms. *Statistics and Computing*, 8: 319–335.
- Bulmer, M. G., 1975. The Statistical Analysis of Density Dependence. *Biometrics*, 31: 901–911.
- Carlin, B. P. & Louis, T. A., 1996. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall, London.
- Clyde, M. A., 1999. Bayesian Model Averaging and Model Search Strategies. In: *Bayesian Statistics 6*: 157–186 (J. M. Bernardo, A. F. M. Smith, A. P. Dawid and J. O. Berger, Eds.). Oxford Univ. Press, Oxford.
- Dennis, B. & Taper, M. L., 1994. Density Dependence in Time Series Observations of Natural Populations: Estimation and Testing. *Ecological Monographs*, 64: 205–224.
- Diggle, P. J., 1990. *Time Series: A Biostatistical Introduction*. Oxford Univ. Press, Oxford.
- Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B., 1995. *Bayesian Data Analysis*. Chapman and Hall, London.
- Gelman, A. & Meng, X., 1996. Model Checking and Model Improvement. In: *Markov Chain Monte Carlo in Practice*: 189–201 (W. R. Gilks, S. Richardson & D. J. Spiegelhalter, Eds.). Chapman and Hall, London.
- Gilks, W. R., Richardson, S. & Spiegelhalter, D. J., 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.
- Green, P. J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82: 711–732.
- Johnson, D. H. & Grier, J. W., 1988. Determinants of Breeding Distribution of Ducks. *Wildlife Monographs*, 100: 5–37.
- Kass, R. E. & Raftery, A. E., 1995. Bayes Factors. *Journal of the American Statistical Association*, 90: 773–795.
- Madigan, D. M., Raftery, A. E., Volinsky, C. & Hoeting, J., 1996. Bayesian Model Averaging. In: *Integrating Multiple Learned Models (IMLM-96)*: 77–83 (P. Chan, S. Stolfo & D. Wolpert, Eds.).
- Massot, M., Clobert, J., Pilorge, T., Lecomte J. & Barbault, R., 1992. Density Dependence in the Common Lizard: Demographic Consequences of a Density Manipulation. *Ecology*, 73: 1742–1756.
- Newman, K. B., 1998. State–Space Modelling of Animal Movement and Mortality with Application to Salmon. *Biometrics*, 54: 1290–1314.
- Nichols, J. D., Conroy, M. J., Anderson, D. R. & Burnham, K. P., 1984. Compensatory Mortality in Waterfowl Populations: A Review of the Evidence and Implications for Research and Management. *Transactions of the North American Wildlife Nature Reserve Conference*, 49: 535–554.
- Nichols, J. D., Johnson, F. A. & Williams, B. K., 1995. Managing North–American Waterfowl in the Face of Uncertainty. *Annual Review of Ecological Systems*, 26: 177–199.
- Reddingius, J., 1971. Gambling for Existence: A Discussion of some Theoretical Problems in Animal Population Ecology. *Acta Biotheoretica*, 20: 1–208.
- Royama, T., 1981. Fundamental Concepts and Methodology for the Analysis of Population Dynamics, with particular reference to univoltine species. *Ecological Monographs*, 51: 473–493.
- Shenk, T. M., White, G. C. & Burnham, K. P., 1998. Sampling–Variance Effects on Detecting Density Dependence from Temporal Trends in Natural Populations. *Ecological Monographs*, 68: 445–463.
- Sullivan, P. J., 1992. A Kalman Filter Approach to Catch–at–Length Analysis. *Biometrics*, 48: 237–257.
- Turchin, P., 1990. Rarity of density Dependence or Population Regulation with Lags? *Nature*, 344: 660–663.
- Turchin, P., Lorio Jr, P. L., Taylor, A. D. & Billings, R. F., 1991. Why do Populations of Southern Pine Beetles (*Coleoptera: Scolytidae*) Fluctuate? *Environmental Entomology*, 20: 401–409.
- Vickery, W. & Nudds, T., 1984. Detection of Density Dependent Effects in Annual Duck Censuses. *Ecology*, 65: 96–104.
- Viljugrein, H., Stenseth, N. C. & Steinbakk, G. H. (in press). Density Dependence in North American Ducks: A State–Space Modelling Approach. *Ecology Letters*.
- Williams, B. K., Koneff, M. D. & Smith, D. A., 1999. Evaluation of Waterfowl Management under the North American Waterfowl Management Plan. *Journal of Wildlife Management*, 63: 417–440.