Weak Identifiability in Models for Mark-Recapture-Recovery Data

Olivier Gimenez, Byron J.T. Morgan, and Stephen P. Brooks

Abstract The percentage overlap between prior and posterior distributions is obtained easily from the output of MCMC samplers. A 35% guideline for overlap between univariate marginal prior and posterior distributions has been suggested as an indicator of weak identifiability of a parameter. As long as uniform prior distributions are adopted for all of the model parameters, then the suggested guideline has been found to work well for a range of models of mark-recapture-recovery data, where all the parameters are probabilities. Its use is illustrated on models for ring-recovery data on male mallards, and the Cormack-Jolly-Seber model for capture-recapture data on dippers.

Keywords Bayesian identifiability · Cormack-Jolly-Seber model · Mark-recapture-recovery models · Parameter-redundancy · Prior/posterior overlap · Sensitivity · Survival of wild animals

1 Introduction

1.1 Parameter Redundancy and Identifiability

Models may be devised for mark-recapture-recovery (mrr) data without all the parameters being estimable. A model is said to be *identifiable* if no two values of the parameters give the same maximum likelihood for the data, while *parameter-redundant* models can be re-expressed in terms of fewer than the original number of parameters (Catchpole and Morgan 1997), resulting in that case in likelihood surfaces with completely flat ridges or surfaces. The obvious way to check for parameter redundancy for a particular application is to examine the likelihood surface by computing the rank of the observed Hessian (Viallefont et al. 1998; Formann 2003). Catchpole and Morgan (1997) considered the rank of the model itself, regardless

---

B.J.T. Morgan (✉)
Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7NF, UK
e-mail: B.J.T.Morgan@kent.ac.uk

D.L. Thomson et al. (eds.), *Modeling Demographic Processes in Marked Populations,*
Environmental and Ecological Statistics 3, DOI 10.1007/978-0-387-78151-8_48,

of the data, using symbolic algebra. It is then possible to determine how many and which parameter combinations can be estimated (Catchpole et al. 1998; Catchpole et al. 2001). Applications of this approach can be found in, for example, Gimenez et al. (2003), Gimenez et al. (2005), Schaub et al. (2004), Nasution et al. (2004) and Kéry et al. (2005). In addition, Catchpole et al. (2001) considered *near redundant* models, which, while formally not parameter-redundant, can in some cases produce estimates of certain parameters with poor precision.

1.2 Weak Identifiability

The use of computational Bayesian methods for model-fitting in biology has increased in recent years (Ellison 2004; Clark 2005), including population ecology (Brooks et al. 2000a, 2002). With the increase in computing power, the temptation is to fit more and more complex models using MCMC methods. However, models that are parameter-redundant may be fitted using Bayesian methods, as a result of the information in the prior distribution, and a dramatic illustration of this is provided by Brooks et al. (2000b). Let us write a marginal posterior distribution as $\pi(\theta|Y)$, for data Y, parameter θ and prior distribution $p(\theta)$. Then the parameter θ is said to be *weakly identifiable* when $\pi(\theta|Y) \approx p(\theta)$; see Gelfand and Sahu (1999) and Garrett and Zeger (2000).

Thus in Bayesian analysis, weak identifiability arises when data supply little information about certain parameters. Weak identifiability is the counterpart of near redundancy in a classical analysis, and poses appreciable problems for a Bayesian approach. For example:

- conclusions based on the examination of weakly identified parameters can be misleading (Garrett and Zeger 2000);
- weak identifiability may result in strong correlations between parameters in the posterior distribution, which in turn implies poor mixing in the MCMC samples and very slow convergence (Carlin and Louis 1996; Rannala 2002);
- even with large sample sizes, the likelihood may be unable to overcome the prior (Neath and Samaniego 1997);
- a too-informative prior can drive posterior inference, while a prior too close to improper can yield improper posteriors (Gelfand and Sahu 1999; Bayarri and Berger 2004).

There are various ways to check for weak identifiability:

- one might conduct a classical test for parameter redundancy;
- one might undertake a detailed prior sensitivity analysis;
- one can examine the correlation matrix of the parameter estimators in the posterior distribution;
- as in Garrett and Zeger (2000), one can display the marginal prior/posterior pair plots as a visual aid;
- one can evaluate numerically, and calibrate, the overlap for each marginal prior-posterior pair.

Gimenez et al. (2006) have compared these alternative procedures, and found that the last two of these methods are simple and effective. We outline this approach in Section 2, and then illustrate its use on a data set resulting from marking male mallards, *Anas platyrhynchos*, in Section 3. In Section 4 the method is applied to a data set of capture-recapture data on dippers *Cinclus cinclus*. In Section 5 we examine the alternative approach of sensitivity analysis, while in Section 6 we present correlations between parameter estimates. The paper ends with general discussion in Section 7.

2 Testing for Weak Identifiability

2.1 Theory

In order to check the weak identifiability of any parameter θ, Garrett and Zeger (2000) compared its marginal prior distribution to its marginal posterior distribution by directly evaluating the overlap between the two distributions. This quantity, denoted τ_θ, can be computed as

$$\tau_\theta = \int \min(p(\theta), \pi(\theta|Y))d\theta. \qquad (1)$$

Values of τ_θ lie in the interval [0,1], and when τ_θ is above some pre-determined threshold then θ is declared weakly identifiable. The *ad-hoc* threshold of 0.35 has been suggested by Garrett and Zeger (2000), and used in applications of this method.

2.2 Bayesian Inference

The mrr models that we shall consider only involve probabilities, and in this paper we have only presented results arising from taking uniform distributions on the interval [0,1] as priors for all of the model probabilities. Based on preliminary runs, we generated four chains of length 50,000, discarding the first 25,000 as burn-in. Convergence was assessed using the Brooks/Gelman/Rubin statistic (Gelman 1996), and we found that in general the Markov chains exhibited good mixing and moderate autocorrelation.

Simulations were performed using WinBUGS (Gimenez et al. 2008; Lunn et al. 2000), and the R (R Development Core Team 2008) package R2WinBUGS (Sturtz et al. 2005) was used both to call WinBUGS and examine results in R.

2.3 Practical Computation of τ

The computation of τ follows suggestions by Schmid and Schmidt (2006): we estimate the posterior distribution $\pi(\theta|Y)$ by means of a kernel density estimator $\hat{f}$ based on the MCMC generated values x_i, $i = 1, \ldots, n$ (we took $n = 1000$, corresponding to the last 1000 MCMC values obtained), namely $\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K\left(\frac{x - x_i}{h}\right)$ where K is a kernel function centered at the data points x_i and h is the bandwidth. We used a standard Gaussian kernel $K(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ with its associated optimal bandwidth $h^* = 1.06\,\hat{\sigma}\,n^{-\frac{1}{5}}$, where $\hat{\sigma}$ = min (standard deviation, interquartile range/1.34) (Silverman 1986, page 48).

We then obtained a sample from the distribution of the *min* function in Equation (1) by calculating $y_i = \min(1, \hat{f}(x_i))$ for all i. Finally, a Monte Carlo approximation to τ is given by $\hat{\tau} = \sum_i y_i/n$.

3 Mark-Recoveries: Application to the Freeman-Morgan Model

3.1 The Freeman-Morgan Model

As an example of the approach, we consider a model developed by Freeman and Morgan (1992; FM model hereafter) involving the survival probability $\phi_{1,i}$ of birds in their first year of life, possibly varying with the year i, the probability ϕ_a of survival of adult individuals (i.e. of age $\geq$ 1), taken as constant over time, and constant probabilities of reporting of rings from dead birds in their first year λ_1, or older λ_a.

Taking the FM model as an illustration, Catchpole et al. (2001) showed that probability models that are formally not parameter redundant may behave poorly when fitted to data. The main reason for this near-singularity is that the FM model contains as a sub-model the model with constant first year survival, which is parameter redundant. As a consequence, the smallest eigenvalue of the expected information matrix may be very small rather than zero as is the case in parameter-redundant models.

When Catchpole et al. (2001) applied this model to ring-recovery data obtained from animals marked as young, very poor results were sometimes obtained, with unrealistic estimates of $\phi_{1,i}$ and λ_1 and large associated standard errors.

We fitted the FM model to data on mallards, with 9 years of recovery (Table 1). The data are the result of a ringing study of males ringed as young in the San Luis Valley, Colorado, 1963–1971; Brownie et al. (1985), p48.

Displays of the marginal prior-posterior distribution pairs for the FM model parameters are given in Fig. 1, and the corresponding estimated percentages of overlap are given in Table 2, corresponding to the shaded areas in Fig. 1.

The marginal prior-posterior distribution pairs in Fig. 1 clearly suggest that all parameters except two may exhibit weak identifiability problems, viz., ϕ_a and λ_a which have relatively sharp marginal posterior distributions. Examination of Table 2

1, . . . , n (we took
obtained), namely
on centered at the
rd Gaussian kernel
th $h^* = 1.06\,\hat{\sigma}\,n^{-\frac{1}{5}}$,
4) (Silverman 1986,

in function in Equa-
Monte Carlo approxi-

Morgan Model

ped by Freeman and
bability $\phi_{1,i}$ of birds
he probability ϕ_a of
stant over time, and
their first year λ_1, or

01) showed that prob-
behave poorly when
the FM model con-
l, which is parameter
expected information
parameter-redundant

ring-recovery data
lts were sometimes
associated standard

f recovery (Table 1).
oung in the San Luis

s for the FM model
nated percentages of
s in Fig. 1.
early suggest that all
ems, viz., ϕ_a and λ_a
xamination of Table 2

Table 1 Recovery data for male mallards Anas platyrhynchos

Year of ringing	Year of recovery (−1962) 1	2	3	4	5	6	7	8	9	Number never seen again
1963	83	35	18	16	6	8	5	3	1	787
1964		103	21	13	11	8	6	6	0	534
1965			82	36	26	24	15	18	4	927
1966				153	39	22	21	16	8	942
1967					109	38	31	15	1	1005
1968						113	64	29	22	927
1969							124	45	22	940
1970								95	25	786
1971									38	315

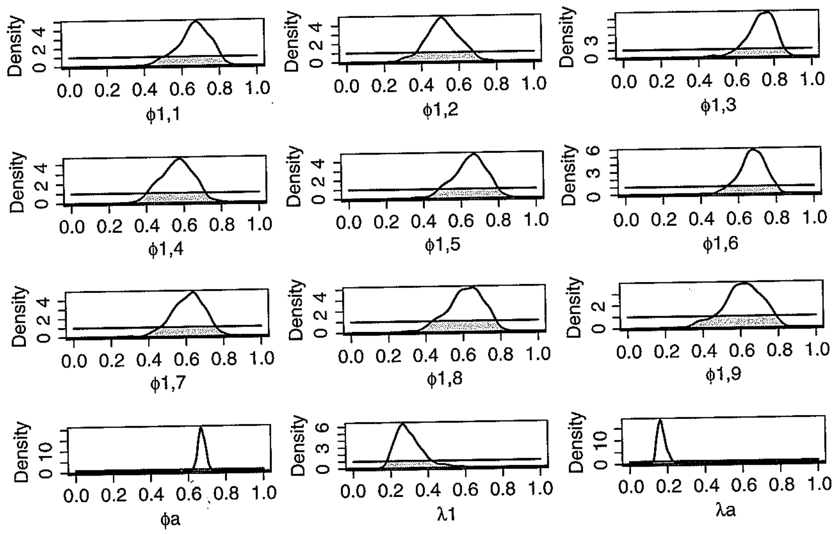


Fig. 1 Display of the prior-posterior distribution pairs for $\phi_{1,i}$, ϕ_a, λ_1 and λ_a in the FM model applied to the mallard data. In order to compute τ, the shaded area of overlap has to be calculated (Equation 1). The priors for all parameters were chosen as $U(0, 1)$ distributions. Note that scales in panels may differ

leads to the same conclusion with all τ values greater than or close to 0.35, except for ϕ_a and λ_a and the young reporting rate λ_1 which was also found to be identifiable using the 35% threshold. This is in agreement with the results obtained by Catchpole et al. (2001), Table 5(a).

3.2 *Fitting the Parameter-Redundant Sub-model*

It is of interest here to fit the parameter-redundant sub-model, which arises when there is just a single survival probability, ϕ_1, for birds in their first year of life.

Table 2 τ values expressed as percentages for the FM model fitted to the mallard data

Parameter	τ
$\phi_{1,1}$	40.2
$\phi_{1,2}$	41.5
$\phi_{1,3}$	34.6
$\phi_{1,4}$	40.1
$\phi_{1,5}$	41.7
$\phi_{1,6}$	34.7
$\phi_{1,7}$	41.5
$\phi_{1,8}$	43.4
$\phi_{1,9}$	45.3
ϕ_a	10.7
λ_1	30.8
λ_a	13.4

The resulting graph showing the overlaps between priors and posteriors is shown in Fig. 2. It appears from this graph that the only parameter that is not weakly identifiable is ϕ_a. This result is in agreement with the classical methodology of Catchpole et al. (2001), which formally identifies the three estimable parameters in this case as ϕ_a, and the two parameter combinations $\phi_1\lambda_a$, and $\phi_1(1-\lambda_1)$. The values for the

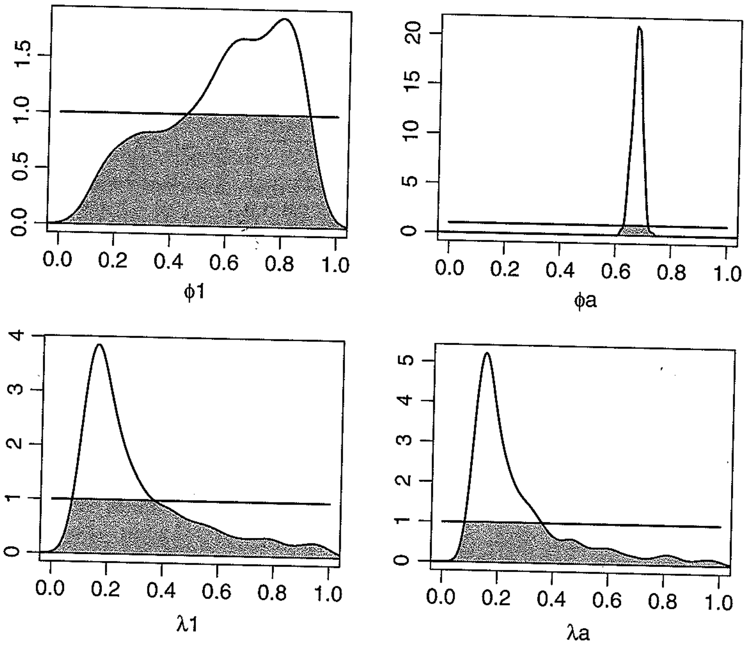
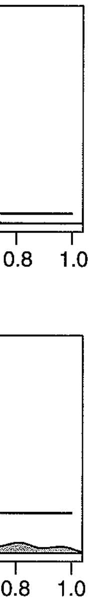


Fig. 2 Display of the prior-posterior distribution pairs for ϕ_1, ϕ_a, λ_1 and λ_a in the parameter redundant sub model of the FM model, applied to the mallard data. The priors for all parameters were chosen as $U(0, 1)$ distributions. Note that scales in panels differ

rd data

iors is shown in
weakly identifi-
gy of Catchpole
ters in this case
e values for the



in the parameter
for all parameters

---

overlaps are, $\tau_{\phi_1} = 75.93$, $\tau_{\phi_a} = 10.78$, $\tau_{\lambda_1} = 60.66$, $\tau_{\lambda_a} = 51.33$. The difference between Figs. 1 and 2 is one of degree, with the weakly-identified parameters of Fig. 2 exhibiting more overlap with the prior than in the case of Fig. 1.

4 Mark-Recapture: Application to the Cormack-Jolly-Seber Model

We consider now the fully time-dependent capture-mark-recapture Cormack-Jolly-Seber (CJS) model originally developed by Cormack (1964), Jolly (1965) and Seber (1965), for which all parameters are time dependent. We define ϕ_i as the probability that an animal survives to time t_{i+1} given that it is alive at time t_i, and p_j the probability of being recaptured at time t_j. Even though it is parameter-redundant, the model can be useful for analyzing capture-mark-recapture data. We fitted the CJS model to the dipper data set which consists of 7 capture occasions (Table 3). The data are the result of a recapture study of both male and female birds ringed as adults in eastern France, in 1981–1986. In this application, the last survival (ϕ_6) and detection (p_7) probabilities are known to be confounded while all the other parameters are estimable.

Displays of the prior-posterior distribution pairs for survival and detection probabilities are given in Fig. 3.

The overlap percentages are given in Table 4 and correspond to the shaded areas in Fig. 3.

From examining Fig. 3, it appears that almost all CJS model parameters are well identified given the small overlap between the prior and posterior distributions. However, as expected, the last survival and detection probabilities are the exception, their posterior distributions being relatively flat and therefore providing more coverage of the uniform priors. In addition, it should be noted that the first survival and the first detection probabilities are also clearly weakly identifiable, due to the fact that very few individuals were marked at the first sampling occasion (approximately 7% of the full data set). The visual diagnostic of Fig. 3 is confirmed by looking at the numerical values of τ in Table 4. In addition, parameter p_3 seems to be marginally weakly identifiable ($\tau_{p_3} = 35.9$).

Table 3 Recapture data for European dippers *Cinclus cinclus* (data taken from Lebreton et al. 1992)

Year of release	Year of recapture (1981+)						Number never recaptured
	1	2	3	4	5	6	
1981	11	2	0	0	0	0	9
1982		24	1	0	0	0	35
1983			34	2	0	0	42
1984				45	1	2	32
1985					51	0	37
1986						52	46

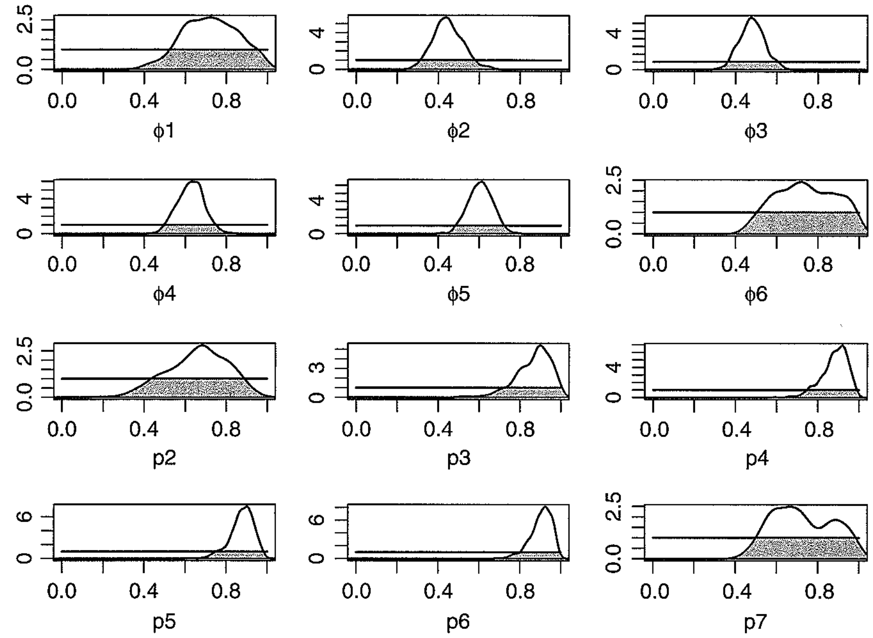


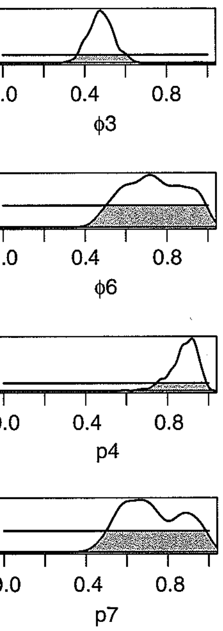
Fig. 3 Display of the prior-posterior distribution pairs for ϕ and p in the CJS model applied to the dipper data. The priors for all parameters were chosen as $U(0, 1)$ distributions. Note that scales in panels differ

Table 4 τ values expressed as percentages for the CJS model applied to the dipper data. For all parameters we use $U(0, 1)$ priors throughout

Parameter	τ
ϕ_1	53.6
ϕ_2	33.4
ϕ_3	29.2
ϕ_4	27.9
ϕ_5	27.4
ϕ_6	57.2
p_2	54.7
p_3	35.9
p_4	28.4
p_5	26.4
p_6	23.4
p_7	56.5

5 Sensitivity Analysis

If parameters are weakly identifiable, then we would expect this to be revealed by a sensitivity analysis, in which we repeat the analysis several times, each time for a different configuration of prior distributions. To illustrate this approach, we consider the CJS model applied to the dipper data set. We take a $U(0, 1)$ prior for all of the



CJS model applied to the
utions. Note that scales in

to the dipper data. For all

this to be revealed by
times, each time for a
approach, we consider
, 1) prior for all of the

Table 5 Sensitivity analysis for the CJS model applied to the dipper data. In rows we show the parameter for which the prior is changed in turn (from a $U(0, 1)$ to a Beta(1,9)). In columns we give the posterior means for all the parameters. The last column gives the Euclidean distance between each row and the case where all priors are $U(0, 1)$

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	p_2	p_3	p_4	p_5	p_6	p_7	eucl. dist.
ϕ_1	0.47	0.46	0.48	0.63	0.60	0.72	0.75	0.87	0.88	0.87	0.90	0.74	0.27
ϕ_2	0.74	0.39	0.49	0.62	0.61	0.73	0.65	0.89	0.88	0.88	0.90	0.73	0.07
ϕ_3	0.72	0.46	0.43	0.63	0.60	0.73	0.67	0.86	0.89	0.88	0.91	0.74	0.05
ϕ_4	0.72	0.45	0.48	0.56	0.61	0.73	0.67	0.87	0.87	0.89	0.90	0.74	0.07
ϕ_5	0.72	0.45	0.48	0.63	0.55	0.74	0.66	0.87	0.88	0.87	0.92	0.74	0.06
ϕ_6	0.72	0.45	0.48	0.63	0.61	0.53	0.67	0.87	0.88	0.88	0.90	0.91	0.28
p_2	0.83	0.43	0.48	0.63	0.60	0.73	0.42	0.87	0.88	0.88	0.91	0.74	0.26
p_3	0.72	0.55	0.45	0.63	0.60	0.73	0.67	0.60	0.88	0.87	0.90	0.74	0.30
p_4	0.72	0.45	0.53	0.60	0.61	0.74	0.67	0.87	0.67	0.87	0.90	0.73	0.21
p_5	0.73	0.45	0.48	0.69	0.57	0.74	0.66	0.87	0.88	0.71	0.91	0.73	0.18
p_6	0.72	0.45	0.48	0.63	0.67	0.71	0.67	0.87	0.88	0.87	0.73	0.71	0.19
p_7	0.72	0.45	0.48	0.63	0.61	0.91	0.67	0.87	0.88	0.87	0.89	0.53	0.27

parameters except for one, which is given a $Beta(1, 9)$ prior, and we then change, in turn, which parameter has the different prior. As we can see from the results displayed in Table 5, the conclusions are not straightforward.

However it appears that there is little sensitivity of parameters to the prior, except for parameters ϕ_1, p_2, ϕ_6, p_3 and p_7. This is in line with the findings from using the overlap measure τ, but there is a difference in order. As with the overlap measure, there is an issue of calibration here. Carrying out the sensitivity analysis by changing the prior for one parameter at a time is time consuming since we need to run the MCMC chains as many times as the number of parameters (note also that to check for convergence of the MCMC, we always run two chains in parallel). Thus this approach could soon become intractable with more complex models, particularly as one would typically use several alternative beta distributions.

Because of the relative complexity of a detailed sensitivity analysis, we now consider a simpler alternative approach, of examining correlations between parameters, obtained from the MCMC output in the usual way.

6 Correlations Between Estimates

6.1 The Dipper Data

The correlation matrix of the parameters for the CJS model applied to the dipper data is given in Table 6.

We have high negative correlation between ϕ_6 and p_7 (-0.89) as expected, moderate negative correlation between ϕ_1 and p_2 (-0.50), and all remaining pairs of parameters give correlations in the range (-0.32, 0.16). We note that the value of -0.32 relates to parameter p_3, as well as to parameter ϕ_2. Thus in this example

Table 6 Correlation matrix for the parameters of the CJS model applied to the dipper data

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	p_2	p_3	p_4	p_5	p_6	p_7
ϕ_1	1.00	−0.15	0.02	−0.03	−0.03	0.00	−0.50	0.02	0.01	−0.02	−0.05	0.00
ϕ_2	−0.15	1.00	−0.08	0.00	0.00	−0.03	0.09	−0.32	0.05	−0.03	0.00	0.05
ϕ_3	0.02	−0.08	1.00	−0.06	0.02	0.03	−0.03	0.16	−0.18	0.03	0.01	−0.04
ϕ_4	−0.03	0.00	−0.06	1.00	−0.09	−0.02	0.03	−0.01	0.15	−0.32	0.00	−0.02
ϕ_5	−0.03	0.00	0.02	−0.09	1.00	0.01	0.02	−0.02	0.01	0.15	−0.30	−0.06
ϕ_6	0.00	−0.03	0.03	−0.02	0.01	1.00	0.04	0.05	−0.02	−0.07	0.02	−0.89
p_2	−0.50	0.09	−0.03	0.03	0.02	0.04	1.00	0.01	−0.02	0.04	−0.02	−0.03
p_3	0.02	−0.32	0.16	−0.01	−0.02	0.05	0.01	1.00	−0.04	−0.01	0.02	−0.06
p_4	0.01	0.05	−0.18	0.15	0.01	−0.02	−0.02	−0.04	1.00	−0.02	−0.04	0.01
p_5	−0.02	−0.03	0.03	−0.32	0.15	−0.07	0.04	−0.01	−0.02	1.00	−0.04	0.07
p_6	−0.05	0.00	0.01	0.00	−0.30	0.02	−0.02	0.02	−0.04	−0.04	1.00	0.03
p_7	0.00	0.05	−0.04	−0.02	−0.06	−0.89	−0.03	−0.06	0.01	0.07	0.03	1.00

considering the correlation structure between estimates has proved to be useful, and is relatively easy to implement.

6.2 The Mallard Data and the FM Model

The correlation matrix of the parameters for the FM model applied to the Mallard data is given in Table 7.

In this case the correlation matrix is not so easy to interpret. As expected, there are generally low correlations between ϕ_a and all of the other parameters, but that is not true of parameters λ_1 and λ_a, which is therefore out of line with the findings of Section 3.1.

6.3 The Mallard Data and the Parameter-Redundant Sub-model

The correlation matrix of the parameters for the FM submodel applied to the mallard data is given in Table 8.

Table 7 Correlation matrix for the parameters of the FM model applied to the mallard data

	ϕ_a	$\phi_{1,1}$	$\phi_{1,2}$	$\phi_{1,3}$	$\phi_{1,4}$	$\phi_{1,5}$	$\phi_{1,6}$	$\phi_{1,7}$	$\phi_{1,8}$	$\phi_{1,9}$	λ_1	λ_a
ϕ_a	1.00	−0.06	−0.07	−0.05	−0.04	−0.03	−0.02	0.00	−0.01	−0.03	−0.02	0.22
$\phi_{1,1}$	−0.06	1.00	0.85	0.87	0.88	0.88	0.87	0.88	0.86	0.77	0.86	−0.89
$\phi_{1,2}$	−0.07	0.85	1.00	0.82	0.87	0.86	0.84	0.86	0.84	0.73	0.89	−0.83
$\phi_{1,3}$	−0.05	0.87	0.82	1.00	0.88	0.88	0.87	0.88	0.86	0.77	0.85	−0.90
$\phi_{1,4}$	−0.04	0.88	0.87	0.88	1.00	0.90	0.88	0.89	0.88	0.78	0.91	−0.87
$\phi_{1,5}$	−0.03	0.88	0.86	0.88	0.90	1.00	0.88	0.90	0.88	0.78	0.88	−0.89
$\phi_{1,6}$	−0.02	0.87	0.84	0.87	0.88	0.88	1.00	0.88	0.86	0.77	0.88	−0.87
$\phi_{1,7}$	0.00	0.88	0.86	0.88	0.89	0.90	0.88	1.00	0.88	0.77	0.90	−0.87
$\phi_{1,8}$	−0.01	0.86	0.84	0.86	0.88	0.88	0.86	0.88	1.00	0.76	0.87	−0.85
$\phi_{1,9}$	−0.03	0.77	0.73	0.77	0.78	0.78	0.77	0.77	0.76	1.00	0.77	−0.78
λ_1	−0.02	0.86	0.89	0.85	0.91	0.88	0.88	0.90	0.87	0.77	1.00	−0.81
λ_a	0.22	−0.89	−0.83	−0.90	−0.87	−0.89	−0.87	−0.87	−0.85	−0.78	−0.81	1.00

ed to the dipper data

p_5	p_6	p_7
−0.02	−0.05	0.00
−0.03	0.00	0.05
0.03	0.01	−0.04
−0.32	0.00	−0.02
0.15	−0.30	−0.06
−0.07	0.02	−0.89
0.04	−0.02	−0.03
−0.01	0.02	−0.06
−0.02	−0.04	0.01
1.00	−0.04	0.07
−0.04	1.00	0.03
0.07	0.03	1.00

ved to be useful, and

plied to the Mallard

. As expected, there
parameters, but that
ne with the findings

nt Sub-model

pplied to the mallard

l to the mallard data

$\phi_{1,9}$	λ_1	λ_a
−0.03	−0.02	0.22
0.77	0.86	−0.89
0.73	0.89	−0.83
0.77	0.85	−0.90
0.78	0.91	−0.87
0.78	0.88	−0.89
0.77	0.88	−0.87
0.77	0.90	−0.87
0.76	0.87	−0.85
1.00	0.77	−0.78
0.77	1.00	−0.81
−0.78	−0.81	1.00

Table 8 Correlation matrix for the parameters of the FM submodel applied to the mallard data

	ϕ_a	ϕ_1	λ_1	λ_a
ϕ_a	1.00	0.00	0.01	0.04
ϕ_1	0.00	1.00	0.86	−0.86
λ_1	0.01	0.86	1.00	−0.59
λ_a	0.04	−0.86	−0.59	1.00

We can see that the only parameter from the original parameter set that is estimable, ϕ_a, is essentially uncorrelated with the other estimates, however other conclusions are elusive, and this is true also if one considers the results of a principal component analysis of the correlation matrix.

7 Discussion

Experience with a range of different models and data sets in the general mrr area consistently suggests that, as proposed by Garrett and Zeger (2000), the prior/posterior overlap threshold of $\tau = 35\%$ works well as a guideline for diagnosing weak identifiability in models for the annual survival of wild animals, when the approach is confined to the case of uniform prior distributions. It has been shown by Gimenez et al. (2006) that if other priors are used then it is difficult to calibrate τ.

A difference between using a classical approach to parameter redundancy based on symbolic algebra and estimating the prior/posterior overlap, as in this paper, is that the latter approach takes account of the effect of both the data and the model. In practice, if possible it is important to understand both the redundancy structure of individual models as well as the influence of data. This has been seen in the analysis of Section 3; the estimates of reporting probabilities for the Mallard data are quite different, but in contrast the estimates of first-year survival are not. What this means is that the fitted FM model for the mallard data is essentially similar to the parameter-redundant sub-model with parameters $(\lambda_1, \lambda_a, \phi_1, \phi_a)$.

While in practice it is important to know the parameter-redundancy status of a model, that is not always possible (Jiang et al. 2007; Pradel et al. 2008; Schaub et al. 2004). As we have seen, weak identifiability may be due to the model and/or the data, and the cause of weak identifiability can be investigated once a model is fitted, by fitting the model again to a larger data set simulated from the fitted model. If that were done for the three examples of this paper, then nothing would change for the parameter-redundant sub-model of the FM model. No amount of additional data can change the fact that of the original four parameters, it is only ϕ_a that can be estimated precisely. For the FM model, increased precision for the estimates of first-year survival should improve overall performance, as in this case the model is not parameter redundant. For the CJS model, analysis of the dipper data has produced estimates of low precision either because of parameter redundancy or because of lack of data. Here increasing cohort sizes will improve the precision of parameters such as ϕ_1, but the basic parameter redundancy due to the confounding of parameters ϕ_6 and p_7 will remain.

We believe that the use of a 35% overlap threshold for τ, combined with uniform priors, is an important guide for interpreting the results of Bayesian analyses of mrr data, and we recommend its use as a simple, general guideline in the area. Of course, it is only a guideline, and needs to be interpreted sensibly.

Because the correlation matrix between parameters is easily obtained from the MCMC output, then we recommend that it is also examined. As any Bayesian analysis will involve some sensitivity analysis, then the results of such an analysis might also be of value. We note finally that when uniform priors are used then overlap with the posterior may be related to the variance of the corresponding parameter, though this will also depend on features such as skew (see for example the results of Fig. 2). A crude alternative to the measure of overlap considered in this paper is simply to use (and calibrate) the posterior inter-quartile range, which provides a measure of spread that is less affected by skew.

The R and WinBUGS programs used in this paper are available on request from the first author.

Acknowledgments The authors would like to thank R. Pradel for very stimulating and helpful discussions. The paper was improved as a result of comments from two referees. O. Gimenez's research was supported by a Marie-Curie Intra-European Fellowship within the Sixth European Community Framework Programme.

References

Bayarri MJ, Berger JO (2004) The interplay of Bayesian and frequentist analysis. Statistical Science 19:58–80.

Brooks SP, Catchpole EA, Morgan BJT (2000a) Bayesian animal survival estimation. Statistical Science 15:357–376.

Brooks SP, Catchpole EA, Morgan BJT, Barry SC (2000b) On the Bayesian analysis of ring-recovery data. Biometrics 56:951–956.

Brooks SP, Catchpole EA, Morgan BJT, Harris MP (2002) Bayesian methods for analysing ringing data. Journal of Applied Statistics 29:187–206.

Brownie C, Anderson D, Burnham KP, Robson DS (1985) Statistical Inference from Band Recovery Data a Handbook. Second Edition. U. S. Fish and Wildlife Service, Resource Publication 156. Washington, D. C., USA. 305pp.

Carlin BP, Louis TA (1996) Bayes and Empirical Bayes Methods for Data Analysis. Chapman & Hall/CRC, London.

Catchpole EA, Kgosi PM, Morgan BJT (2001) On the near-singularity of models for animal recovery data. Biometrics 57:720–726.

Catchpole EA, Morgan BJT (1997) Detecting parameter redundancy. Biometrika 84:187–196.

Catchpole EA, Morgan BJT, Freeman SN (1998) Estimation in parameter-redundant models. Biometrika 85:462–468.

Clark JS (2005) Why environmental scientists are becoming Bayesians. Ecology Letters 8:2–14.

Ellison AM (2004) Bayesian inference in ecology. Ecology Letters 7:509–520.

Formann AK (2003) Latent class model diagnosis from a frequentist point of view. Biometrics 59:189–196.

Garrett ES, Zeger SL (2000) Latent class model diagnosis. Biometrics 56:1055–1067.

Gelfand AE, Sahu K (1999) Identifiability, improper priors, and Gibbs sampling for generalized linear models. Journal of the American Statistical Association 94:247–253.

Gelman A (1996) Inference and monitoring convergence. In Gilks WR, Richardson S, Spiegelhalter DJ, editors, Markov Chain Monte Carlo in Practice., pages 131–143. Chapman & Hall, London.

Gimenez O, Bonner S, King R, Brooks SP, Jamieson LE, Grosbois V, Morgan BJT (2008) WinBUGS for population ecologists: Bayesian modeling using Markov chain Monte Carlo methods. In: Thomson DL, Cooch EG, Conroy MJ (eds.) Modeling Demographic Processes in Marked Populations. Environmental and Ecological Statistics, Springer, New York, 3:883–916.

Gimenez O, Brooks SP, Morgan BJT (2006) Assessing weak identifiability in mark-recapture and mark-recovery models. Technical Report UKC/IMS/06/030, University of Kent.

Gimenez O, Choquet R, Lebreton JD (2003) Parameter redundancy in multistate capture-recapture models. Biometrical Journal 45:704–722.

Gimenez O, Viallefont A, Choquet R, Catchpole EA, Morgan BJT (2005) Methods for investigating parameter redundancy. Animal Biodiversity and Conservation 27:561–572.

Jiang H, Pollock KH, Brownie C, Hightower JE, Hoeing JM, Hearn WS, (2007) Age-dependent tag return models for estimating fishing mortality, natural mortality, and selectivity. Journal of Agricultural, Biological, and Environmental Statistics 12:177–194.

Kéry M, Gregg KB, Schaub M (2005) Demographic estimation methods for plants with unobservable life-states. Oikos 108:307–320.

Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS – a Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10:325–337.

Nasution MD, Brownie C, Pollock KH, Powell RA (2004) The effect on model identifiability of allowing different relocation rates for live and dead animals in the combined analysis of telemetry and recapture data. Journal of Agricultural Biological and Environmental Statistics 9:27–41.

Neath AA, Samaniego FJ (1997) On the efficacy of Bayesian inference for nonidentifiable models. American Statistician 51:225–232.

Pradel R, Maurin-Bernier L, Gimenez O, Genovart M, Choquet R, Oro D (2008) Estimation of sex-specific survival with imperfect clues about sex. Canadian Journal of Statistics 36:29–42.

R Development Core Team (2008) R: A Language and Environment for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0.

Rannala B (2002) Identifiability of parameters in MCMC Bayesian inference of phylogeny. Systematic Biology 51:754–760.

Schaub M, Gimenez O, Schmidt BR, Pradel R (2004) Estimating survival and temporary emigration in the multistate capture-recapture framework. Ecology 85:2107–2113.

Schmid F, Schmidt A (2006) Nonparametric estimation of the coefficient of overlapping – theory and empirical application. Computational Statistics and Data Analysis 50:1583–1596.

Silverman BW (1986) Density Estimation for Statistics and Data Analysis. Chapman & Hall, London.

Sturtz S, Ligges U, Gelman A (2005) R2WinBUGS: A package for running WinBUGS from R. Journal of Statistical Software 12:1–16.

Viallefont A, Lebreton JD, Reboulet AM, Gory G (1998) Parameter identifiability and model selection in capture-recapture models: a numerical approach. Biometrical Journal 40:313–325.