Bayesian analysis of the Hector's Dolphin data

R. King & S. P. Brooks

Abstract

Bayesian analysis of the Hector's Dolphin data.— In recent years there have been increasing concerns for many wildlife populations, due to decreasing population trends. This has led to the introduction of management schemes to increase the survival rates and hence the population size of many species of animals. We concentrate on a particular dolphin population situated off the coast of New Zealand, and investigate whether the introduction of a fishing gill net ban was effective in decreasing dolphin mortality. We undertake a Bayesian analysis of the data, in which we quantitatively compare the different competing biological hypotheses, determining the effect of the sanctuary upon the dolphin population.

Key words*:* Capture–recapture, Posterior model probabilities, Management schemes, Markov chain Monte Carlo.

Resumen

Análisis bayesiano de datos del delfín de Héctor.— En los últimos años, ha aumentado la preocupación por muchas poblaciones de fauna, como consecuencia del descenso observado en sus tendencias poblacionales. Ello ha llevado a la aplicación de programas de gestión orientados a aumentar las tasas de supervivencia y, por consiguiente, el tamaño de la población de numerosas especies de animales. Nos concentramos en una población de delfines concreta, situada en la costa de Nueva Zelanda, investigamos si la aplicación de una ley que prohíbe la utilización de redes de enmalle ha resultado eficaz a la hora de reducir la mortalidad de los delfines. Llevamos a cabo un análisis bayesiano de los datos, en el que comparamos cuantitativamente distintas hipótesis biológicas alternativas, y determinamos el efecto de la reserva en la población de delfines.

Palabras clave: Captura–recaptura, Probabilidades del modelo posterior, Programas de gestión, Cadena Monte Carlo de Markov.

R. King, CREEM, Univ. of St. Andrews, Buchanan Gardens, St. Andrews, Fife KY16 9LZ, U.K.– S. P. Brooks, Statistical Lab., Univ. of Cambridge, Wilberforce Road, Cambridge, CB3 0WB, U.K.

Introduction

We consider a detailed study relating to a population of Hector's dolphin (*Cephalorhynchus hectori*) located around the coast of New Zealand. These dolphins have been listed as an endangered species, and so there is particular concern and interest in identifying factors that influence their survival and the corresponding effectiveness of any management schemes to protect them. Our data concerns a colony of dolphins from the Southern Island of New Zealand around Bank's Peninsula off the coast of Akaroa. In this area, commercial gill nets are commonly used for fishing, and it is believed that these may be a contributor to dolphin mortality (Dawsson & Slooten, 1993; Slooten & Dawson, 1994). In an attempt to reduce this threat to the dolphin population, a sanctuary was placed around the peninsula in 1988. The sanctuary imposed a ban on the use of gill nets within 4 nautical miles of the shore between November to February, coinciding both with the peak inshore commercial gill netting season and the period when dolphins move closest to shore (Dawson, 1991; Cameron et al., 1999). We wish to investigate whether this sanctuary was effective in terms of decreasing dolphin mortality.

We begin in "Material and modelling" section by introducing the data that we have concerning the dolphin population and by describing the management scheme. We then describe the methodology that we use in order to answer the question of whether or not the sanctuary was effective. In "Results" section we present the results, before concluding with a Discussion.

Material and modelling

Data

Our data comprise multi–site capture–recapture records, collected annually between 1985 and 1993. The study site was divided into three locations: locations 1 and 3 lie either side of the peninsula to the South and West, respectively, and are separated by region 2, which is the harbour area around Akaroa. The data collection process involves a boat going out to observe the dolphins within the inshore waters of the different areas over a number of days each year. Within each trip, individual dolphins sighted are uniquely identified via markings on their dorsal fin and/or body (see Hammon et al., 1990 for example). Our data comprises the capture histories of each of 102 individuals, detailing the years and locations that each dolphin is observed.

A total of 668 days are spent observing the dolphins throughout the study period, with most data collected within 4 month periods during the summer season. The length of time spent by the observers at each of the sites over the different years is far from uniform. For example in 1986, 92 days are spent observing in area 2, whereas in 1989, no effort is expended in area 3. The number

of days that are spent each year observing dolphins in each of the areas is given in table 1

With the large variation in the amount of effort expended over the different years, and between the different sites, there is also a large variation in the number of dolphins that are observed over time and location. In Section 2.2, we show how we are able to explicitly account for catch–effort information in modelling the recapture rates for the dolphins.

Modelling and notation

The data are assumed to be well described by the Arnason–Schwarz model (see for example Schwarz et al., 1993; Brownie et al., 1993; Dupuis, 1995; Dupuis et al., 2002; King & Brooks, 2003b for further details). Essentially, we assume that dolphins may move freely between the different areas (independently of one another), but do not emigrate outside the study area. In addition, an individual dolphin's movement between the different areas is assumed to have a Markovian structure so that the migration of an animal depends only upon its current location and not upon its previous migration history. As usual, we make the further assumption that the observed individuals are representative of the whole population.

Under the Arnason–Schwarz model, we express the likelihood for our data as a function of the survival, recapture and migration rates (see for example, King & Brooks, 2003a). For this dataset the parameters in the model are the survival, recapture and migration rates, which we assume may depend upon time, location, neither, or both. We define: $\phi_t(r)$ – Prob (a dolphin in area $r \in R$ at time t survives until time $t + 1$); $p_{t+1}(r)$ – Prob (a dolphin in area $r \in R$ at time $t + 1$ is resighted at this time); $\psi_t(r, s)$ – Prob (a dolphin in area $r \in R$ at time t is in area $s \in R$ at time $t + 1$, given that it survives until time $t + 1$) where $R = \{1,2,3\}$ denotes the regions that the study area is divided into, and $t = 1985,...,1992$. For notational convenience, we set

$$\boldsymbol{\phi} = \{\phi_t(r),\ r \in R,\ \ t = 1985,...,1992\}$$

and similarly for $\boldsymbol{p}$ and $\boldsymbol{\psi}$.

We represent different competing models, in terms of the dependence of the parameters upon the time and/or location, by placing different restrictions upon the parameters. For example, the biological hypothesis that the survival rate remains constant throughout the study, and is common to all areas would be represented by the restriction,

$$\phi_t(r) = \phi,\ \text{for all } r \in R\ \text{ and } t = 1985,...,1992$$

Clearly, this model implies that the sanctuary had no effect upon the survival rate of the dolphins. Conversely, if we believed that the survival rate did change at the time that the sanctuary was introduced, then we may wish to consider the model

Table 1. The amount of effort (in days) spent sighting dolphins in each of the study areas and years.

Tabla 1. Cuantificación del esfuerzo (en días) dedicado a avistar delfines en cada unas de las áreas de estudio y años.

Site	Year								
	1985	1986	1987	1988	1989	1990	1991	1992	1993
1	1	11	25	11	1	5	30	48	9
2	2	92	79	68	19	24	23	44	19
3	1	30	27	6	0	3	11	16	16

with a change–point, so that

$$\phi_t(r) = \phi_a(r), \text{ for } t = 1985,...,1987$$
$$\phi_t(r) = \phi_b(r), \text{ for } t = 1988,...,1992$$

for example.

In this study, we are particularly interested in whether or not the sanctuary that was introduced in 1988 had a significant impact upon the survival rates of the dolphins. Thus, we impose the constraint that there may be at most one change–point, which may or may not occur at the time that the sanctuary was introduced. Our class of models therefore comprises models for which there is either no change–point over time; or where there are the restrictions,

$$\phi_t(r) = \phi_a(r), \text{ for } t = 1985,..., T-1$$
$$\phi_t(r) = \phi_b(r), \text{ for } t = T,...,1992$$

for $T = 1986,...,1992$.

Finally, we need to define the possible location dependencies. We allow all possible restrictions upon the locations e.g., a common survival rate for all locations, distinct survival in each area, or a common survival rate for two different areas which is distinct from the third. We identify these models by defining the sets of areas with common survival rates. For example, the model with a common survival rate over areas is denoted by {1,2,3}, while the model {1}, {2,3} denotes that there is a common survival rate in areas 2 and 3, which is distinct to that in area 1, and so on. Since we consider change–point models here, we allow the location restrictions to be placed independently on the survival rates both before and after the change–point, if there is one. Thus, the dependence of the parameters upon location is conditional on the year.

For the recapture rates, we have additional information relating to the amount of effort that was expended in each location and year of the study. We wish to incorporate this into the analysis, making full use of all available information. Thus, we specify the recapture rates as a function of the corresponding effort taken in that year and location. In particular, we assume that sightings in year

$t = 1986,...,1992$ and location $r \in R$ occur as a Poisson process with general underlying recapture intensity rate $\beta_t(r)$. Thus, we set,

$$p_t(r) = 1 - \exp[-\beta_t(r) \, x_t(r)]$$

where $x_t(r)$ denotes the catch–effort in year t and location r. This can be reparameterised in the form

$$p_t(r) = 1 - [1 - a_t(r)]^{xt(r)}$$

where $a_t(r)$ is directly interpretable as the underlying recapture rate per unit time (i.e. day).

Allowing a complete spatio–temporal dependence for the recapture intensity rates (i.e. having distinct $\beta_t(r)$ parameters for each $t = 1985,...,1992$ and $r \in R$) essentially reduces the model to that with arbitrary recapture rates $p_t(r)$. However, we consider special cases for the $\beta_t(r)$ parameter, representing different possible models directly analogous to those considered above for the survival rates, (i.e. a maximum of a single change–point and all possible location dependencies). For example, the model,

$$\beta_t(r) = \beta(r), \text{ for } t = 1986,...,1993$$

represents the system where the recapture intensity rates depend only upon the location of the dolphin, suggesting that the dolphins may be inherently more observable in some areas than others; or that the observers themselves have more (or less) information relating to the areas where dolphins are most likely to be seen in each of the designated areas. Note that for all models we implicitly assume that the underlying recapture rate for any given year is homogeneous over days within that year.

We impose the same year constraints upon the migration rates, as for the survival rates for similar reasons as those discussed above. See King & Brooks (2002) for further discussion of model structures. Discriminating between these competing models tells us about the underlying dynamics of the system in terms of the possible effect of the intro-

duction of the sanctuary upon the survival and behaviour of the dolphins. We adopt a Bayesian approach here.

Bayesian analysis

Here we undertake a Bayesian analysis of the data, in which inference is based upon the posterior distribution of the parameters obtained by combining the likelihood of the data given the parameters, with the corresponding prior distribution placed upon the parameters (independently of the data), via Bayes' theorem. The prior distribution represents our beliefs concerning the parameters before observing any data. The posterior distribution is then an update of these prior beliefs having observed the data. The posterior distribution is, in general, very complex and so to obtain any inference based upon the posterior distribution we use Markov chain Monte Carlo (MCMC) in order to obtain estimates of the parameters of interest (see for example Brooks, 1998).

Within our analysis, the nature and number of parameters depends upon the model and, as well as parameter estimation, we have the additional issue of model uncertainty. This is of particular interest in our case, since the different models represent different competing hypotheses relating to the effectiveness of the sanctuary —our primary question of interest. Within the Bayesian paradigm, we are able to incorporate model uncertainty by considering the model itself to be an unknown parameter which we wish to estimate. We are then able to form the joint posterior distribution over both parameter and model space. Since the posterior distribution is defined over different dimensions (i.e. for the different models), we use Reversible jump MCMC (RJMCMC) in order to explore the distribution. See Green (1995), Richardson & Green (1997), and also King & Brooks (2003b) —in the context of multi–site capture–recapture data— for example. Using these methods, we are able to construct a single Markov chain which can explore the posterior distribution, and estimate summary statistics, such as: (1) posterior model probabilities for time and/or location dependence of the parameters; (2) the posterior probability of a change in survival rate when the sanctuary is introduced; and (3) model averaged statistics of interest, such as posterior model–averaged survival rates, taking into account model uncertainty (in addition to parameter uncertainty).

However, before we can undertake our analysis we need to place suitable priors on the models and their associated parameters.

Prior distributions

We do not have any prior information relating to the survival or migration rates, and so we use vague priors. In particular, for each model, we place a $U[0, 1]$ prior on each of the survival rates. For the migration rates we use a $Dir\ (\frac{1}{2},\frac{1}{2},\frac{1}{2},)$ prior, corre-

sponding to an uninformative Jeffrey's prior (Carlin & Thompson, 1998, section 2.2.3; Jeffreys, 1961, p 181). However, we consider the prior on the recapture intensities in more detail.

Placing a prior on the recapture intensity rates implicitly imposes a prior on the recapture rates, which may be more directly interpretable. Thus, we wish to place a prior on the recapture intensity which is consistent with our beliefs concerning the recapture rates. Since the recapture intensity is positive, an obvious prior that we may wish to use is a Gamma prior. We are then able to calculate the corresponding prior on the recapture rates, using the usual transformation of variables, as follows. Suppose that we specify a $\Gamma(a,b)$ prior on $\beta_t\ (r)$. For ease of notation, we assume that there is a common underlying β for each year and location, and note that the same prior will be used for the recapture intensity rates across all possible models (i.e. area and time dependence structures). Then, it can be easily shown that the corresponding prior on the recapture rates are of the form,

$$f_p\left(p_t\left(r\right)\right) = -\frac{b^a}{\Gamma(a)}\frac{1}{x_t(r)^{a-1}}\left(1-p_t\left(r\right)\right)^{\frac{b}{x_t(r)}-1}\left(-\ln\left(1-p_t\left(r\right)\right)\right)^{a-1}$$

$$\propto \left(1-p_t\left(r\right)\right)^{\frac{b}{x_t(r)}-1}\left(-\ln\left(1-p_t\left(r\right)\right)\right)^{a-1}$$

Placing a vague prior on β does not impose a similarly vague prior on $p_t(r)$; for example, $\Gamma(0.001, 0.001)$ a prior is often considered to be vague. However, in this case using such a prior on β produces a prior on $p_t(r)$ (assuming $x_t(r) \neq 0$), which essentially places point masses at zero and one.

Conversely, it is obvious from equation (1) that if we set $a = 1$ in the prior for β, then,

$$p_t\left(r\right) \sim Beta\left(1, \frac{b}{x_t\left(r\right)}\right)$$

Suppose that we set,

$$b = \overline{x} = \frac{1}{24}\sum_{i=1986}^{1993}\sum_{r\in R}x_t\left(r\right)$$

then, the prior for the recapture rate will be flat (i.e. $U[0,1]$), when the mean amount of time is spent observing dolphins in the given year and location, i.e. when $x_t(r) = \overline{x}$. If a larger amount of time is spent (i.e. $x_t(r) > \overline{x}$), then the prior for the recapture rate is skewed to the right, and there is more prior mass on larger recapture rates. Otherwise, if a lower amount of time is spent (i.e. $x_t(r) < \overline{x}$), then the prior for $p_t(r)$ is skewed to the left, and there is more prior mass on lower recapture rates. We can also explicitly calculate the prior mean and variance for the recapture rates, since,

$$E_p\left(p_t\left(r\right)\right) = \frac{1}{1+\frac{b}{x_t\left(r\right)}} = \frac{x_t\left(r\right)}{x_t\left(r\right)+b}$$

$$Var_p\left(p_t\left(r\right)\right) = \frac{b}{x_t\left(r\right)}\left(1+\frac{b}{x_t\left(r\right)}\right)^{-2}\left(2+\frac{b}{x_t\left(r\right)}\right)^{-1}$$

Clearly, there are many possible values of **b** that we may wish to consider. For example, we may wish to set b = median $x_t (r)$, or b = $\max_{t,r} x_t (r)$. Prior information may be able used to discern the most appropriate form of *b* to take. Since we do not have any prior information, we specify that $\beta \sim \Gamma(1, \bar{x})$ *a priori*, but note that the posterior distribution is fairly insensitive to sensible choices of prior.

Finally, we need to place a prior on the models themselves. For each of the parameters, we place an equal prior probability on each possible age dependence structure. For the recapture intensities and the survival rates we place a flat prior across each possible combination of strata independently within each set of ages. Note that for the recapture intensities and survival rates, the prior is not flat over each individual model. Placing a flat prior over the whole of the model space, so that each individual model was equally likely *a priori*, would result in a greater amount of prior mass on models with a single change–point compared to those with a constant rate over time. For example, when there are no change–points there are a total of five possible location dependence structures as illustrated by the column headings of table 3. However, when there is a single change–point there are five location–dependent models both for before and after that change–point and so 5^2 models in total. Thus, putting a flat prior across all models would make a change–point model five times more likely than having no change–point, since there are five times more change–point models than there are models without a change–point.

We can then calculate the posterior distribution of the parameters and use an MCMC procedure to obtain estimates of the posterior statistics that we are interested in. We assume that there is model uncertainty relating to the survival and migration parameters, and use the reversible jump algorithm to move between the different possible models, as above. Details of the MCMC procedure are given in the appendix.

Results

We ran the simulations for a total of 1 million iterations, with the first 100,000 discarded as burn–in, and consider each set of parameters in turn. Convergence is rapid for these simulations and this is confirmed using standard diagnostic techniques, see Brooks & Roberts (1998).

Survival rates

For the survival rates, we are particularly interested in whether there is any evidence that they were increased by the introduction of the sanctuary. Thus, we are particularly interested in whether there is a change–point in the survival rates, and if so, when this change occurred in relation to the introduction of the sanctuary. Within the Bayesian framework, we are able to quantitatively compare the different possible models, in terms of their posterior (marginal) model probabilities. These are presented in table 2.

Table 2. The posterior marginal probabilities of no change–point and a change–point in each year for the survival probabilities, recapture intensities and migration rates. Recall that the survival and migration rates are defined for years 1985–1992, and the recapture intensities for years 1986–1993.

Tabla 2. Probabilidades marginales posteriores de ausencia de punto de cambio y presencia de punto de cambio en cada año para las probabilidades de supervivencia, intensidades de recaptura y tasas de migración. Recuérdese que las tasas de supervivencia y migración se han definido para los años 1985–1992, y las intensidades de recaptura para los años 1986–1993.

Change–point	Posterior probability		
	Survival rates	Recapture intensities	Migration rates
None	0.314	0.175	0.180
1986	0.088	–	0.036
1987	0.036	0.241	0.155
1988	0.160	0.035	0.549
1989	0.136	0.421	0.041
1990	0.112	0.087	0.024
1991	0.069	0.069	0.010
1992	0.084	0.012	0.005
1993	–	0.012	–

The marginal model with most posterior support has a constant survival rate over all time, suggesting that the introduction of the sanctuary had no impact upon the survival rates of the dolphins. However, the second most probable model *a posteriori* places a change–point at the time that the sanctuary is introduced.

The model–averaged estimates of the survival rates (together with 95% highest posterior density intervals – HPDI's) are given in figure 1, and suggest a slow decreasing trend post–1987, in each area. Thus, there appears to be little evidence that the sanctuary was effective in decreasing the mortality rate of the dolphins. The similarity between survival rates within the different areas across all years are indicative of the large posterior probabilities that the survival rates are constant across areas. These are presented in table 3). Clearly, the most posterior support has a common survival rate over the different areas: approximately 50% for all years. Conversely, there is very little posterior support for distinct survival rates within each area.

Recapture intensities

There is some evidence that the recapture intensities change over time within the study, as shown in table 2. In particular the strongest support is for a change–point in 1989 (42%) or 1987 (24%). The evidence for a change–point is clearly seen in the posterior means for the recapture intensities over time in figure 2A, where there appears to be a large increase in the recapture intensity from 1989 in area 3. This not only coincides with the sanctuary being introduced the previous year but also to a relatively small amount of effort made in area 3 from 1989 onwards, compared with previous years (see table 1), resulting in a much greater number of dolphins observed in this area. Before 1989 only 3 dolphins are observed in area 3; from 1989, 24 dolphins are observed in area 3. There also appears to be a change in the posterior probabilities for the location dependence from 1989 onwards for the recapture rates, as shown in table 3.

The corresponding mean recapture rates are given in figure 2B. We can see that the recapture rates in area 2 are generally higher than those in other areas. This is the harbour area of the peninsula, where most effort was actually made in observing the dolphins. The dip in the recapture rates for all areas within the middle of the study corresponds to generally decreased effort spent in observing the dolphins (see table 1).

Migration rates

The marginal posterior model probabilities for the change–point models for the migration rates are also given in table 2. Most posterior support (55%) is for the model where there is a change–point in the migration rates the year that the sanctuary is introduced. This suggests that the sanctuary may have influenced the behaviour of
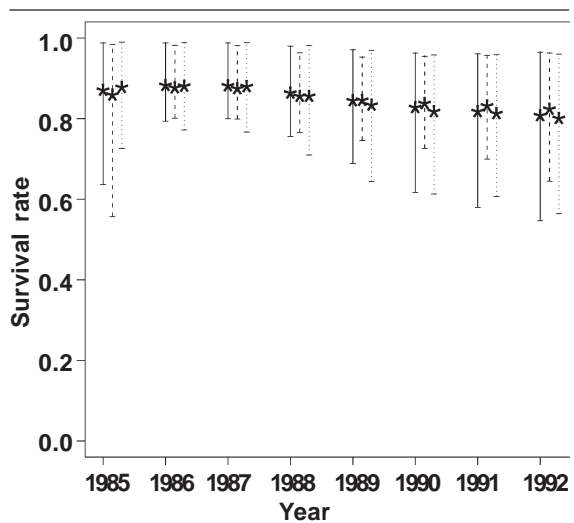

Fig. 1. The posterior model–averaged mean (*) and 95% HPDI (vertical lines) for the survival rates over time: — Area 1; - - - Area 2; ····· Area 3.

Fig. 1. Media posterior del modelo promediado (*) y 95% de HPDI (intervalos máximos de densidad posterior) (líneas verticales) para las tasas de supervivencia a lo largo del tiempo: — Área 1; - - - Área 2; ···· Área 3.

the dolphins, in terms of their movement around the peninsula. The change in the movement of the dolphins is clearly seen in the model–averaged estimates of the migration rates, presented in figure 3. One possible reason for this may be an indirect link with possible changes in fish stocks in the area, due to the gill net fishing ban, which may also have an impact upon the abundance of different breeds of fish and marine animals. However, further research would be necessary in order to investigate this possibility.

The dolphins appear to predominantly stay in the same areas, although there are some movements between the different regions. This demonstrated by the general dominance of the top line in each of figures 3A–C.

Discussion

We consider multi–site capture–recapture data of the Hector's dolphins around the Bank's Peninsula in New Zealand. This species is endangered, and we wish to assess the impact (if any) of the management scheme placing a fishing gill net ban around the shore–line of the peninsula during the summer months. We have applied a Bayesian analysis to the data collected. However, within our analysis, there is little evidence to suggest that

Table 3. Posterior marginal probabilities or the arrangement of strata over time for the survival probabilities and recapture intensities.

Tabla 3. Probabilidades marginales posteriores o disposición de los estratos a lo largo del tiempo para las probabilidades de supervivencia y las intensidades de recaptura.

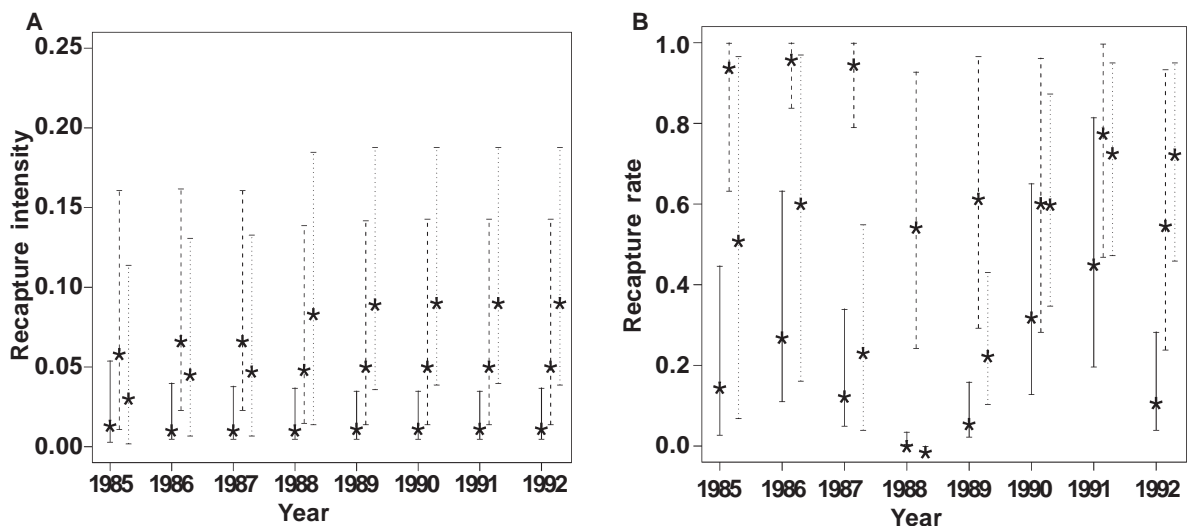|	Strata groupings				
Survival probabilities					
Year	{1}, {2}, {3}	{1,2}, {3}	{1,3}, {2}	{1}, {2,3}	{1,2,3}
1985	0.046	0.138	0.144	0.135	0.537
1986	0.035	0.138	0.132	0.133	0.562
1987	0.035	0.140	0.133	0.132	0.561
1988	0.041	0.152	0.134	0.138	0.534
1989	0.051	0.166	0.135	0.145	0.503
1990	0.059	0.170	0.136	0.154	0.482
1991	0.063	0.172	0.137	0.159	0.468
1992	0.070	0.178	0.142	0.164	0.446
Recapture intensities					
1986	0.294	0.084	0.308	0.251	0.062
1987	0.322	0.055	0.243	0.356	0.025
1988	0.323	0.054	0.232	0.370	0.022
1989	0.322	0.252	0.047	0.368	0.011
1990	0.324	0.260	0.014	0.396	0.001
1991	0.324	0.262	0.012	0.398	0.005
1992	0.325	0.264	0.010	0.397	0.003
1993	0.324	0.263	0.011	0.396	0.005

Fig. 2. The posterior model–averaged mean (*) and 95% HPDI (vertical lines) for the recapture intensities (A) and recapture rates (B) over time: — Area 1; - - - Area 2; ···· Area 3.

Fig. 2. Media posterior del modelo promediado (*) y 95% de HPDI (líneas verticales) para las intensidades de recaptura (A) y las tasas de recaptura (B) a lo largo del tiempo: —. Área 1; - - - Área 2; ···· Área 3.
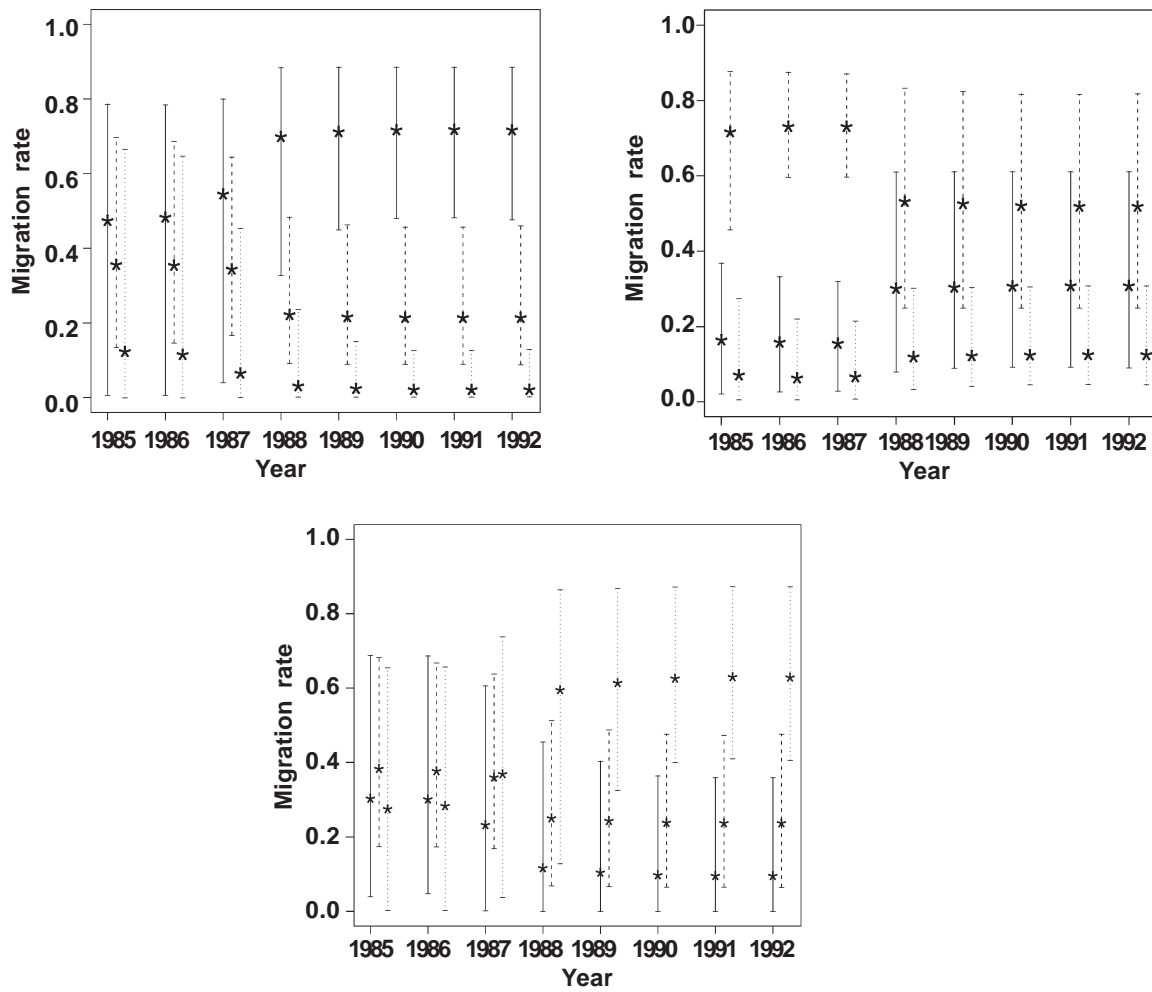
Fig. 3. The model-averaged (*) and 95% HPDI for the migration rates from area 1 (A), area 2 (B) and area 3 (C). Lines denote movement to : — Area 1; - - - Area 2; ···· Area 3.

Fig. 3. Modelo promediado (*) y 95% de HPDI para las tasas de migración de (a) el área 1 (A), el área 2 (B) y el área 3 (C). Las líneas indican movimiento hacia: — Área 1; - - -. Área 2; ···· Área 3.

this sanctuary achieved its aim. On the contrary, if there has been a trend in the survival rate this appears to have been negative overall.

This, perhaps surprising, result may be a consequence of other factors that may not have been accounted for. For example, it is unclear as to the effect that such a sanctuary would have upon the whole of the local ecosystem. The increased fish stocks that would be present within the inshore waters once the ban was imposed may impact upon both predators higher in the food chain (possibly attracting them to the area), and also other marine animals competing for the same resources. In addition, Slooten & Dawson (1994) identify pollution as influencing the survival rates of the dolphins. Thus, the sanctuary alone appears to be ineffective in its attempt to improve the survival

rates of the dolphins, although it is possible that had the sanctuary not have been introduced, there may have been an even greater decline in the survival rate of the dolphins. Some or all of these factors could have been investigated by adopting an alternative design incorporating controls for each of these potential effects. Such designs (often referred to as BACI —before, after, control and impact— designs) have significant advantages for assessing management impacts and would have improved the ability of the study to determine the true effect of the sanctuary.

The movement of the dolphins between the different areas does appear to change at the same time that the sanctuary is placed around the peninsula. Within our Bayesian analysis, this marginal model has a posterior probability of 55%. This may be a

result of the effect that the sanctuary had upon the fish stocks in the study area, which may be more evenly spread around the peninsula with the imposed fishing bans. Alternatively, the dolphins may move more freely to areas which were previously heavily populated with gill nets. Clearly, any such hypothesis would need further investigation.

Finally, we note that there is some evidence that the recapture intensities changed from 1989. It is clear from figure 2A that this appears to be largely due to the change in the recapture intensity in area 3. This may be as a result of the observers learning where the dolphins are more abundant in area 3 from previous boat surveys in the study; or possibly as a result of the sanctuary being introduced and affecting the dolphins behaviour (possibly moving further inshore or to slightly different waters), and making them easier to observe.

Clearly, the sanctuary placed around the peninsula does not appear to be enough to improve the survival rates of the dolphins. It would appear unlikely that it would increase the mortality rate of the dolphins, so this analysis suggests that there are other (more predominant) factors that are affecting the dolphins survival rates which need to be addressed, in order to conserve this dolphin population.

Acknowledgements

References

Brooks, S. P., 1998. Markov Chain Monte Carlo Method and its Application. *The Statistician,* 47: 69–100.

Brooks, S. P. & Roberts, G. O., 1998. Diagnosing Convergence of Markov Chain Monte Carlo Algorithms. *Statistics and Computing,* 8: 319–335.

Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H. & Hestbeck, J. B., 1993. Capture–recapture Studies for Multiple Strata including Non–Markovian Transition Probabilities. *Biometrics,* 49: 1173–1187.

Cameron, C., Barker, R., Fletcher, D., Slooten, E. & Dawson, S., 1999. Modelling Survival of Hector's Dolphins around Banks Peninsula, New Zealand. *Journal of Agricultural, Biological and Environmental Statistics,* 4: 126–135.

Carlin, B. P. & Louis, T. A., 1998. *Bayes and Empirical Bayes Methods for Data Analysis.* Chapman and Hall/CRC

Dawson, S., 1991. Incidental catch of Hector's dolphin's in inshore gillnets. *Marine Mammal Science,* 7: 283–295.

Dawson, S. & Slooten, E., 1993. Conservation of Hector's Dolphins: The Case and Process which led to Establishment of the Bank's Peninsula Marine Mammal Sanctuary. *Aquatic Conservation: Marine and Freshwater Ecosystems,* 3: 207–221.

Dupuis, J. A., 1995. Bayesian Estimation of Movement and Survival Probabilities from Capture–recapture Data. *Biometrika,* 82: 761–772.

Dupuis, J. A., Badia, J., Maublanc, M. & Bon, R., 2002. Survival and Spatial Fidelity of Mouflons (*Ovis gmelini*): a Bayesian Analysis of an Age–dependent Capture–Cecapture Model. *Journal of Agricultural, Biological and Environmental Statistics,* 7: 277–298.

Green, P. J., 1995. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika,* 82: 711–732.

Hammon, P. S., Mizroch, S. A. & Donovan, G. P., 1990. *Individual Recognition of Cetaceans: Use of Photo–Identification and other techniques to estimate Population Size.* Technical report, International Whaling Commission.

Jeffreys, H., 1961. *Theory of Probability, third edition.* Oxford, Univ. Press.

King, R. & Brooks, S. P., 2002. Bayesian Model Discrimination for Multiple Strata Capture–Recapture Data. *Biometrika,* 89: 785–806.

– 2003a. A Note on Closed Form Likelihoods for Arnason–Schwarz models. *Biometrika,* 90: 435–444.

– 2003b. Survival and Spatial Fidelity of Mouflon: the Effect of Location, Age and Sex. *Journal of Agricultural, Biological and Environmental Statistics,* 8: 486–513.

Richardson, S. & Green, P. J., 1997. On Bayesian Analysis of Mixtures with an Unknown Number of Components. *Journal of the Royal Statistical Society, Series B,* 59: 731–792.

Schwarz, C. G., Schweigert, J. F. & Arnason, A. N., 1993. Estimating Migration Rates using Tag–Recovery Data. *Biometrics,* 49: 177–193.

Slooten, E. & Dawson, S., 1994. *Hector's Dolphin, in Delphinidae and Phocoenidae (Vol V. of Handbook of Marine Mammals).* Academic Press, N.Y.

Appendix. In order to explore and summarise the posterior distribution we use a reversible jump Markov chain Monte Carlo algorithm. Essentially this involves constructing a Markov chain with stationary distribution equal to the posterior distribution. Then, following an initial burn–in period, so that the stationary distribution is reached, realisations of the chain can be regarded as a sample from the posterior distribution and used to estimate summary statistics of interest. The reversible jump MCMC algorithm consists of two different move types: one for updating the parameters; and the other for updating the model itself.

Apéndice. Con la finalidad de explorar y resumir la distribución posterior, hemos utilizado un algoritmo de Monte Carlo de cadena de Markov de salto reversible. En esencia, este algoritmo implica construir una cadena de Markov con una distribución estacionaria equivalente a la distribución posterior. Posteriormente, y después de un periodo inicial que permita llegar a una distribución fija, las partes finales de la cadena pueden ser utilizadas como una muestra de la distribución posterior, siendo entonces utilizada para estimar los estadísticos resumen en los que estamos interesados. El algoritmo reversible de la cadena de Monte Carlo es de dos tipos: uno para actualizar los parámetros y el otro para actualizar al propio modelo.

Within–model MCMC updates

We update each of the parameters in the model using the standard Metropolis–Hastings algorithm, within each iteration of the Markov chain. In particular, we use a random walk Metropolis–Hastings algorithm. Suppose that we are proposing to update the parameter $\phi_t(r)$. Then we propose parameter $\phi'_t(r)$, such that $\phi'_t(r) = \phi'_t(r) + \xi$ where, $\xi \sim U[-\delta, \delta]$, with δ chosen via pilot tuning. We accept the new proposed value with the standard acceptance probability (see for example Brooks, 1998). In practice $\delta = 0.1$ appears to work well for both the survival rates and recapture intensities. This is simply generalised when we have sets of times and/or strata grouped together.

We use an alternative Metropolis–Hastings update for the migration rates, since we need to retain the sum to unity constraints. Suppose that for a given time t, we wish to update the parameter $\psi_t(r, s)$. Then we randomly choose $u \in R \setminus s$, and set,

$$\psi'_t(r, s) = \psi_t(r, s) + \xi$$
$$\psi'_t(r, u) = \psi_t(r, u) - \xi$$

where $\xi \sim U[-v, v]$

In practice we set $v = 0.05$. This move is accepted with the standard probability. Within each iteration of the Markov chain we cycle through each age group and propose to update all $r, s \in R$ using the above procedure.

Between–model (RJ) MCMC updates

To update the dependence structure of the recapture parameters, we need to use the reversible jump algorithm to move between the different models, since the models are of different dimensions. The reversible jump algorithm can be seen to be an extension of the Metropolis–Hastings algorithm, allowing for movements between different states. Within each step of the Markov chain, we propose to update: (i) the number of change–points on the survival rates, recapture intensities and migration rates; (ii) the location of a change–point (if any) for each of the parameters; and (iii) the area dependence on the survival rates and recapture intensities for each age group.

We consider the different types of reversible jump updates in turn.

Adding/removing change–point

We initially consider the survival rates. If there is a constant survival rate, we propose to add a change–point. Otherwise if there is already a change–point, we propose to remove it, since we only consider models with a maximum of a single change–point. Initially, suppose that we propose to add a change–point, so that the current model, denoted by m has a common survival rate over time, and for simplicity we assume that the survival rate is also common over all areas, i.e.,

Appendix. (Cont.)

$$\phi_t(r) = \phi \text{ for all } r \in R \text{ and } t = 1985,...,1992$$

Then we propose to move to the new model, m', with parameters,

$$\phi'_t(r) = \phi + \varepsilon \text{ for all } r \in R \text{ and } t = 1985,...,T{-}1$$
$$\phi'_t(r) = \phi - \varepsilon \text{ for all } r \in R \text{ and } t = T,...,1992$$

where T is randomly chosen in the interval $\{1986,...,1992\}$, and $\varepsilon \sim N(0,\sigma^2)$, for σ^2 chosen via pilot tuning. Then, if any $\phi'_t(r) \notin [0,1]$ the move is automatically rejected, else it is accepted with probability,

$$\min\left(1, \frac{\pi(\boldsymbol{\phi'}, \boldsymbol{\psi}, \boldsymbol{\beta}, m'|data)|J|}{\pi(\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\beta}, m|data) P(m'|m) q(\varepsilon)}\right)$$

where $\pi(\cdot|\cdot)$ is the posterior distribution over parameter and model space evaluated at the given parameter values; $|J|$ a Jacobian term, which is equal to 2 in this case; $P(m'|m) = 1/7$ is the probability that given that we are in model m, we propose to move to model m' with a change–point at time T (for which there are seven possible choices, each chosen with equal probability), and $q(\varepsilon)$ is the corresponding Normal proposal density. Clearly this approach can be generalised for any given location dependence, with the restriction that the new survival rates before and after the change–point have this same location dependence, with the corresponding changes to the Jacobian term. Note that in general $|J| = 2^k$, where k is simply the number of distinct survival rates over the locations.

Alternatively, in the reverse move, to retain the reversibility conditions, we only propose to remove the change–point if the location dependence is the same over all times. Then the proposed survival rate is simply taken to be the mean of the survival rates either side of the change–point for each area. The corresponding acceptance probability is simply the reciprocal of equation (2).

The analogous update is used for the recapture intensity, with the restriction that the proposed recapture rates are positive. We set $\sigma^2 = 0.1$ and 0.01 for the survival rates and recapture intensities respectively.

However, we need to consider a different updating procedure for the migration rates, since we need to retain the restriction that the migration rates sum to unity. Again, suppose that we propose to add in a change–point to the current model m, with migration rates $\psi(r, s)$, $r, s \in R$ (recall that $R = \{1,2,3\}$). Then, for $r \in R$ and $s = 1,2$, we propose the parameters in the new model to be,

$$\psi'_t(r, s) = \psi_t(r, s) + \eta(r, s) \text{ for all } r \in R \text{ and } t = 1985,...,T{-}1$$
$$\psi'_t(r, s) = \psi_t(r, s) \text{ for all } r \in R \text{ and } t = T,...,1992$$

with probability 1/2; else we set,

$$\psi'_t(r, s) = \psi_t(r, s) \text{ for all } r \in R \text{ and } t = 1985,...,T{-}1$$
$$\psi'_t(r, s) = \psi_t(r, s) + \eta(r, s) \text{ for all } r \in R \text{ and } t = T,...,1992$$

Here, T chosen uniformly in $[1986,...,1992]$, and, $\eta(r, s) \sim N(0, \sigma^2)$, where σ^2 is chosen via pilot tuning. In this case we set $\sigma^2 = 0.1$. Essentially, we are simulating a new set of migration rates for either before or after the change–point, which are similar to their current values, while the others remain the same. We also set,

$$\psi'_t(r,3) = 1 - \sum_{r=1}^{2} \psi'_t(r,s)$$

to ensure that the migration rates sum to unity. If any $\phi'_t(r, s) \notin [0,1]$, then we automatically reject the move, else we accept the move with the acceptance probability,

$$\min\left(1, \frac{\pi(\boldsymbol{\phi}, \boldsymbol{\psi'}, \boldsymbol{\beta}, m'|data) P(m|m') |J|}{\pi(\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\beta}, m|data) P(m'|m) q(\eta)}\right)$$

Here $|J| = 1$, $P(m'|m) = 1/2 \times 1/7$; $P(m'|m) = 1/2$ and $q(\boldsymbol{\eta})$ denotes the Normal proposal density for the set of parameters $\boldsymbol{\eta} = \{\eta(r, s): r \in R, s = 1,2\}$.

Appendix. (Cont.)

Location of change–point

Initially, consider the survival rates. Then, within this updating procedure, we propose to update the location of the change–point if there is one present. Suppose that we are in the model with the change–point at time T, so that,

$$\phi_t(r) = \phi_a(r) \text{ for all } r \in R \text{ and } t = 1985,...,T-1$$
$$\phi'_t(r) = \phi_b(r) \text{ for all } r \in R \text{ and } t = T,...,1992.$$

Then we propose to move to model m' by updating the change–point to time $T' = T \pm 1$. If $T' \notin [1985,...,1992]$, we reject the proposal; else we set,

$$\phi'_t(r) = \phi_a(r) \text{ for all } r \in R \text{ and } t = 1985,...,T'-1$$
$$\phi'_t(r) = \phi_b(r) \text{ for all } r \in R \text{ and } t = T',...,1992$$

We then accept the proposed move with the standard Metropolis–Hastings acceptance probability, i.e.,

$$\min\left(1, \frac{\pi(\boldsymbol{\phi'}, \boldsymbol{\psi}, \boldsymbol{\beta}, m'|data)}{\pi(\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\beta}, m|data)}\right)$$

We use the analogous updating procedure for the recapture intensities and migration rates.

Updating area dependence

Initially consider the survival rates and assume that there is no change–point; else, if there is a change–point, then we update the location dependence before and after the change–point independently of each other. We assume that there is a common survival rate over time and location, i.e. the area dependence is denoted by {1,2,3}. Then, we propose to update this area dependence, by splitting the group into two: there are three possibilities {1}, {2,3}; {1,3}, {2}; and {1,2}, {3}. We choose each one with equal probability, without loss of generality suppose that we propose to move to model {1}, {2,3}. Then, for the new model m', we propose the parameters,

$$\phi'_t(r) = \phi_t(r) + \omega \text{ for } r = 1 \text{ and } t = 1985,...,1992$$
$$\phi'_t(r) = \phi_t(r) - \omega \text{ for } r = 2, 3 \text{ and } t = 1985,...,1992$$

where, $\omega \sim N(0, \tau)$, for τ chosen via pilot tuning. We reject the proposed move if any $\phi'_t(r) \notin [0,1]$, else, we accept the move with probability,

$$\min\left(1, \frac{\pi(\boldsymbol{\phi'}, \boldsymbol{\psi}, \boldsymbol{\beta}, m'|data) P(m|m')|J|}{\pi(\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\beta}, m|data) P(m'|m) q(\omega)}\right)$$

where the Jacobian $|J| = 2$; $P(m'|m)$ denotes the probability of moving to model m' from model m and $q(\omega)$ the Normal proposal density for the simulated parameter ω. Here, $P(m'|m) = 1/3$, since we can only increase the dimension of the model, and, $P(m|m') = 1/2$, if in model $m' = \{1\}, \{2,3\}$ we propose to decrease, or increase, the dimension of the model to {1,2,3}, or {1}, {2}, {3}, with equal probability. Once more, to retain the reversibility condition, the reverse move is deterministically defined by this move. In particular, we set the new survival rate to be the mean of the current values, and accept the move with probability equal to the reciprocal of the above. The analogous move holds when proposing to move to the model with distinct survival rates over all areas.

We apply the analogous move to the recapture intensities, restricting the parameter values to be simply > 0. Pilot tuning suggests setting $\tau = 0.1$ for both the survival rates and recapture intensities.