ALMA TICINENSIS UNIVERSITAS · UNIVERSITÀ DI PAVIA

SCUOLA DI DOTTORATO IN SCIENZE E TECNOLOGIE 'ALESSANDRO VOLTA'

DOTTORATO DI RICERCA IN ECOLOGIA SPERIMENTALE E GEOBOTANICA
XXV CICLO

HIERARCHICAL BAYESIAN MODELLING: APPLICATIONS IN ANIMAL POPULATION ECOLOGY

SIMONE TENAN

A Thesis Submitted for the Degree of Doctor of Philosophy

TUTORS:

DR. PAOLO PEDRINI
DR. MARCO GIRARDELLO

EXTERNAL TUTOR:

DR. GIACOMO TAVECCHIA

PH.D. COORDINATOR:

PROF. GIUSEPPE BOGLIANI

2012

General abstract

Modelling and inference are fundamental to the science of ecology, and the hierarchical modelling framework is a conceptual approach that can be applied to a large and diverse set of problems. In particular, hierarchical models that explicitly describe a process model are useful for gaining insight into the form and function of an ecological system or process.

With this thesis I addressed different ecological hypotheses in the population ecology field, by exploiting the conceptual clarity and practical utility of the hierarchical modelling framework, together with the benefits of Bayesian methods as a mode of analysis and inference. I used each case study to explore the advantages of a Bayesian formulation of hierarchical models and their flexibility in addressing finer questions. Throughout empirical cases I was able to extended the current methods to multiple groups, to include external covariates, to solve problems linked with variable or model selection and to explore new applications.

More specifically I focused on the application, and performance assessment in some cases, of individual covariate models to estimate population size, of integrated population models to increase precision in vital rates and estimates of demographic quantities that were not measured in the field, followed by a close examination on Bayesian variable and model selection. Finally I explored an innovative application to the study of the occupancy dynamics of multiple species over time and space.

Robust estimates of population density, or abundance, are difficult to obtain due to the interaction between experimental design, sampling processes and behaviour responses. Setting the number of sampling occasions, to estimate closed population size from capture-recapture data, is often a question of balancing objectives and costs. Moreover, accounting for trap-response and individual heterogeneity in recapture probability can be problematic when data derive from less than four sampling occasions. Using a set of simulated capture-mark-recapture (CMR) data, I first assessed whether a Bayesian formulation of capture-recapture models based on Data Augmentation can be successfully used in three sample session study, in presence of individual recapture heterogeneity. I then applied the method to real data to estimate the population size and size-dependent population structure in the endemic Balearic Lizard, *Podarcis lilfordi*. Moreover, I extended the method to the simultaneous analysis of males and females and contrasted hypotheses on sex-ratio and sex-by-size population structure.

Estimates of population density is of primary interest for many applications of population ecology, such as conservation biology. As second application, I used individual covariate models to derive an estimate for the population density of the sessile Noble Pen shell (*Pinna nobilis*) from CMR data on this bivalve, the largest of the Mediterranean Sea, typically associated with seagrass meadows. I investigated whether habitat characteristics or physical forcing, both natural or anthropogenic, are determining the spatial differences in population density and structure. The approach is similar to the one previously outlined, but this time I extended the analysis including external covariates.

Evidence for decline or threat of wild populations typically come from multiple sources and methods that allow optimal integration of the available information represent a major advance in planning management actions. As third example of application, I used integrated population modelling and perturbation analyses to assess the demographic consequences of the illegal use of poison for an insular population of Red Kite, *Milvus milvus*. I first pooled into a single statistical framework the annual census of breeding pairs, the available individual-based data, the average productivity and the number of birds admitted annually to the local rehabilitation centre. By combining these four types of information I was able to increase estimate precision and to obtain an estimate of the proportion of breeding adults, an important parameter that was not directly measured in the field and that is often difficult to assess. Subsequently, I used perturbation analyses to measure the expected change in the population growth rate due to a change in poison-related mortality.

To close the loop of modelling and inference in ecology, the growing enthusiasm for Bayesian hierarchical model fitting has also to deal with the more technical issue of selecting competing models and variables, a procedure that has a fundamental role in the inferential process but has not yet received enough attention when applied in a Bayesian framework. To this end, I presented two possible procedures, the Gibbs variable selection and the product space method, with emphasis on the practical implementation into a general purpose software in BUGS language. To clarify the related theoretical aspects and practical guidelines, I explained the methods through applied examples on the selection of variables in a logistic regression problem and on the comparison of non-nested models for positive continuous response variable.

Finally, I explored an innovative approach to the study of the dynamics of multiple species over space and time, using a state-space formulation of dynamic occupancy models. I presented this work, that is still in progress, by using presence-absence data from a medium-term ringing scheme on migrating birds. To my knowledge this is the first attempt to model dynamical processes of local extinction and colonization in migrating birds and along the altitudinal gradient.

To my father

Acknowledgments

This experience would not have been possible without the financial means and the bureaucratic support provided by the Museo delle Scienze and the Università di Pavia, through Paolo Pedrini and Giuseppe Bogliani. Moreover, I have been privileged to stand on the shoulders of brilliant scientists, without whom this work would never have been. My thesis work owes an unpayable intellectual debt to the work of Marco Girardello (Centre for Ecology and Hydrology, Oxford, UK) and Giacomo Tavecchia (Population Ecology Group, IMEDEA (CSIC-UIB), Mallorca, Spain). They offered invaluable assistance, support and guidance in the different phases of my doctoral experience.

I am also indebted to friends and colleagues with whom I have had the good fortune of working in Spain, at the Population Ecology Group of the Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB): Daniel Oro, Meritxell Genovart, Alejandro Martínez-Abraín, Mike Fowler, Albert Fernández Chacón, Noelia Hernández, Andreu Rotger, Jose Manuel Igual Gómez, Roger Pradel (C.N.R.S., C.E.F.E., France), Fabrizio Sergio (Estacion Biologica de Doñana, CSIC, Seville), Iris Hendriks (Global Change Research, IMEDEA).

I also wish to express my gratitude to J. Andrew Royle for his suggestions and comments that led to a substantial improvement of the first chapters of the thesis.

Thanks to Antonello Provenzale and Jost von Hardenberg for their assistance. I am also grateful to Fernando Spina, Iris Hendriks, Jaume Adrover and the Grup Balear d'Ornitologia i Defensa de la Naturalesa (GOB) for providing data and photos.

Special thanks to my brother in law, Alexander Garcia Aristizabal, for sharing ideas and materials about Bayesian Statistics.

I don't have words to thank my family, Ilaria, Jole, and Giulio, for their love, support and patience.

Simone Tenan

Contents

General abstract **iii**

Table of Contents **xi**

List of Figures **xv**

List of Tables **xix**

Introduction **1**

Thesis aims and outline 3

1 Population abundance in an endemic lizard **7**

Abstract 7

1.1 Introduction 8

1.2 Materials and Methods 9

1.2.1 Model formulation 9

1.2.2 Simulated data 11

1.2.3 Capture-recapture data 12

1.2.4 Bayesian estimation 12

1.2.5 Size dependent population structure 13

1.3 Results 13

1.3.1 Simulated data 13

1.3.2 Capture-recapture data 14

1.4 Discussion 16

1.4.1 Survey design and the estimate of animal abundance by capture-recapture models 16

1.4.2 Population size, sex-ratio and size-dependent structure 17

1.5 Acknowledgments 18

2 Recreational tourism and population structure of an endangered bivalve **19**

Abstract 19

2.1 Introduction . . . 20
2.2 Methods . . . 21
2.2.1 Study area . . . 21
2.2.2 Data collection, habitat descriptors and anchoring . . . 22
2.2.3 Model formulation and parameter estimation . . . 24
2.2.4 Population structure in relation to shell width . . . 26
2.3 Results . . . 26
2.4 Discussion . . . 31
2.5 Conclusions and Recommendations . . . 32
2.6 Acknowledgements . . . 32

3 Demographic cost of illegal poisoning 33
Abstract . . . 33
3.1 Introduction . . . 34
3.2 Materials and Methods . . . 36
3.2.1 Data collection . . . 36
3.2.2 Integrated population model . . . 37
3.2.3 Likelihood for the population count data . . . 37
3.2.4 Likelihood for radio-tracking data . . . 38
3.2.5 Likelihood for reproductive success data . . . 42
3.2.6 Likelihood for unmarked birds found dead . . . 43
3.2.7 Likelihood of the integrated model . . . 43
3.2.8 Parameter estimation and model implementation . . . 45
3.2.9 Modelling the effect of poisoning on population growth rate . . . 46
3.3 Results . . . 47
3.4 Discussion . . . 49
3.4.1 An analytical framework to help a 'crisis discipline' . . . 49
3.4.2 Illegal poison and population trajectories . . . 53
3.4.3 Facultative, occasional and obligate scavengers and the illegal use of poison . . . 54
3.4.4 Conservation measures . . . 54
3.5 Acknowledgements . . . 54

4 Selecting Bayesian hierarchical models 57
Abstract . . . 57
4.1 Introduction . . . 57
4.2 Theoretical background . . . 59
4.2.1 The core of the matter . . . 59
4.2.2 The product space method . . . 60
4.2.3 Gibbs variable selection . . . 61
4.3 Examples with a real dataset . . . 63
4.3.1 Model selection with the product space method . . . 63
4.3.2 Variable selection using GVS . . . 64

4.4 Relevant points for practical implementation 65
4.4.1 Product space 65
4.4.2 GVS 67
4.5 Prior sensitivity 68
4.6 Conclusions 69
4.7 Acknowledgements 69

5 Altitudinal occupancy dynamics in migratory birds 71
Note 71
5.1 Introduction 71
5.2 Materials and Methods 72
5.2.1 Data collection 72
5.2.2 Sampling design and dynamic occupancy models 72
5.2.3 Hierarchical formulation 74
5.3 Preliminary results and discussion 75

6 General Conclusions 79

A The Bayesian paradigm in brief 85
A.0.1 Bayes' Theorem 86
A.0.2 Markov chain Monte Carlo 87
A.0.3 Gibbs sampling 89
A.0.4 Monte Carlo simulations using BUGS 90

B Population abundance in an endemic lizard 91
B.0.5 Model selection 91
B.0.6 Posterior summaries of model parameters for the simulated datasets . 92
B.0.7 Posterior distributions 93
B.0.8 R and BUGS codes 96

C Recreational tourism and population structure of Noble Pen Shell: code 99

D Demographic cost of illegal poisoning: data and codes 103

E Selecting Bayesian hierarchical models 123
E.0.9 Product space code 123
E.0.10 GVS code 126

F Occupancy dynamics in migratory birds 129
F.1 Advantages of the state-space formulation 129
F.2 Dynamic occupancy model: R and BUGS codes 129

List of Figures

1.1 Balearic Lizard, *Podarcis lilfordi* (Photograph by S. Tenan). 10

1.2 Relationship between detection probability p and body length, for first captured lizards (solid line) and for lizards captured at least once in the previous occasions (dashed line). Shaded areas represent 95%CRI. 15

1.3 Size dependent population structure, as density of individuals in relation to body length for sampled lizards (dashed line) and for population estimates (solid line). 15

2.1 Map with the locations of the sites where surveys were conducted around the island of Majorca, Balearic Islands, Spain. 22

2.2 Noble Pen Shell, *Pinna nobilis* (Photograph by I. Hendriks). 24

2.3 Site-specific densities (individuals/100 m^2) of Noble Pen Shell in the island of Majorca, Balearic Islands, Spain. 27

2.4 Relationship between Noble Pen Shell detection probability (p) and the individual covariate, shell width. The shaded area represents *95%CRI*. 29

2.5 Densities (individuals/100 m^2) of Noble Pen Shell in the island of Majorca (Balearic Islands, Spain), in relation to presence/absence of anchoring. . . . 30

2.6 Size (shell width) dependent population structure of Noble Pen Shell for each sampling site, in the island of Majorca (Balearic Islands, Spain). Estimated proportions of individuals for different dimensional classes are reported. Note the different y-axis scale for Es Cargol. 30

2.7 Kernel density estimates for Noble Pen Shell size (shell width) in relation to presence/absence of anchoring, for sampled individuals (dashed line) and population estimates (solid line). 31

3.1 Tagged Red kite, *Milvus milvus*, found dead (©GOB Mallorca). 37

3.2 Diagram of possible states of a marked red kite. Transitions between two subsequent states, from time t to $t + 1$, are denoted with arrows and correspond to parameters in the transition matrix of eq. 3.6. For the sake of clarity, the parameters and the "unobserved dead" state are not reported. Notation: DP: dead by poison; DO: dead by other causes. 40

3.3 Graphical representation of the integrated population model. Data are symbolized by small rectangles, parameters by ellipses, the relationships between them by arrows and sub-models by open rectangles. Notation: J: annual number of fledglings; R: numbers of surveyed broods whose final fledging success was known; Dp: number of unmarked birds found dead by poisoning; Do: number of unmarked birds found dead by causes other than poisoning; T: radio-tracking data; Y: population count data; b: fecundity; mdp: average probability, across age groups, of dying because of poisoning; mdo: average probability, across age groups, of dying because of other causes; N_{pop}: total number of individuals in the population; N_{dp}: expected number of unmarked birds found dead by poisoning; N_{do}: expected number of unmarked birds found dead by causes other than poisoning; s: survival probability; β_{juv}: probability of death due to poisoning given that an animal died in its first year of life; β_{1y}: probability of death due to poisoning given that an animal died in its second year of life; β_{2my}: probability of death due to poisoning given that an animal died after its second year of life; p: recapture probability of an animal with a functioning radio; c: recapture probability of a radio-tagged animal which is alive but without an active radio signal; d_1: probability of encounter of a radio-tagged animal dead by poisoning but without an active radio signal; d_2: probability of encounter of a radio-tagged animal dead by other causes and without an active radio signal; α_1, α_2, α_3: radio signal retention probability during the first three, the fourth and the fifth or more year of life, respectively; br_{ad}: proportion of breeding females relative to the total number of females older than 2 years; N_{pairs}: number of breeding females in the population; λ: population growth rate. Priors are excluded from this graph to increase visibility. 44

3.4 Observed and estimated sizes of the Red kite population of Mallorca (Spain), with a future projection of the number of breeding pairs. The solid line represents the surveyed population size, the dashed line the predicted spring population sizes along with their 95%CRI (grey shading). 49

3.5 Changes of population growth rate in relation to changes in the proportional decrease of age-specific survival probability. The black solid line represents the relationship with proportional reduction in juvenile survival (δ_{juv}), the red dashed line refers to δ_{1y}, and the green dotted line refers to δ_{2my}. Current age-specific values of δ are indicated by the arrows with the same colour of the curve to which they refer. 50

3.6 Difference between the proportional changes in age-specific survival probability due to illegal poisoning, fecundity, and population growth rate. The bold line represents population stability. The asterisks refer to the current parameter estimates, while arrows represent a theoretical increase in δ up to the level of population stability. a) juveniles. b) 1 year old. c) 2 or more year old red kites. 51
3.7 Age-independent relationship between population growth rate, fecundity, and proportional change in survival probability due to illegal poisoning. The blue horizontal plane represents population stability. 52

5.1 Distribution of the 32 sampling sites (dots) along the southern Alps. 73
5.2 Temporal variation of occupancy probability (mean and 95% credible interval) for the three migratory species. 77

B.1 Posterior distribution of population size N and its mean value (dashed line) for data of all sampled lizards. 94
B.2 Posterior distribution of population size and its mean value (dashed line) for data of sexed lizards. 94
B.3 Prior (dashed line) and posterior (solid line) distributions of parameter for trap response (tr). A N(0,1000) (panel a), U(-5,5) (panel b), and a U(-10,10) (panel c) were used to model data on sexed lizards. The probabilities that tr was positive are 0.925, 0.935, and 0.936 respectively. 95

List of Tables

1.1 Posterior summaries of model parameters for an endemic *Podarcis* lizards data under a model containing permanent behavioural response and individual heterogeneity on detection probability. The parameters p_0 and p_1 are the mean detection probabilities (on the probability scale) for the first capture event and subsequent occasions respectively, β is the coefficient on body length, μ_x and σ_x are respectively the mean and sd of population body length, ψ is the 'zero-inflation' parameter associated with data augmentation. 14

1.2 Posterior summaries of model parameters for data on sexed individuals of an endemic *Podarcis* lizards, under a model containing sex-specific permanent behavioural response and common individual heterogeneity on detection probability. For sex u, the parameters $p_{0,u}$ and $p_{1,u}$ are the sex-specific mean detection probabilities (on the probability scale) for the first capture event and subsequent occasions respectively, β is the coefficient on body length, μ_x and σ_x are respectively the mean and sd of population body length, ψ_u is the sex-specific 'zero-inflation' parameter associated with data augmentation. m denotes males, f females. 16

2.1 Site characteristics. Maximum wind speed is the directional component (taking into account exposure of the sites) of the wind velocity p95 (95th percentile) averaged over 2008, 2009 and 2010. Fragmentation was measured in 2009, while shoot density and coverage are measured between 2001 and 2008. . . . 23

2.2 Posterior summary of model parameters for data of the Noble Pen Shell aggregated at site level. Densities were derived parameters expressed as individuals/100 m^2, α is the detection probability (on the logit scale) for an average size individual, β is the slope for the relationship between detectability and shell width (cm), ψ is the "zero-ination" parameter associated with data augmentation, $\mu_{x,site}$ is the site-specific mean shell width, $\sigma_{x,site}$ is the shell width standard deviation. Note that α, β, and ψ are site-independent. Posterior mean and related 95% credible interval are reported for each parameter. 28

2.3 Posterior summary of model parameters; Noble Pen Shell in relation to anchoring. Densities were derived parameters expressed as individuals/100 m^2, α is the detection probability (on the logit scale) for an average size individual, β is the slope for the relationship between detectability and shell width, $\psi_{no\ anchoring/anchoring}$ are the "zero-ination" parameters associated with data augmentation and specific for locations without or with anchoring, $\mu_{x,no\ anchoring/anchoring}$ are the site-specific means for shell width (cm), $\sigma_{x,no\ anchoring/anchoring}$ are the shell width standard deviations. Note that α and β are anchoring-independent. Posterior mean and related 95% credible interval are reported for each parameter. 29

3.1 Estimated demographic parameters of the Red kite population of the island of Mallorca (Spain). We show the posterior mean and 95% credible interval (95%CRI, lower and upper limit) of the estimates, obtained by a full integrated model (IPM2), an integrated model without considering data of unmarked birds found dead (IPM1), and a multi-state model with only radio-tracking data. For parameter notation see Methods. 48

3.2 Sensitivity and elasticity of population growth rate of the Red kite population of the island of Mallorca (Spain). For parameter notation see Methods. 50

4.1 Approaches to variable and model selection implemented in BUGS language (following *O'Hara and Sillanpää*, 2009) with the related main reference. . . . 59

4.2 Posterior variable inclusion probabilities, $p(\gamma_j = 1|\boldsymbol{y})$, obtained under four different priors for regression coefficients α and β (assumed drawn from $N(0, \tau^{-1})$ with a varying precision τ), and for two prior distributions on random effect hyperparameter σ_ϵ. γ_1 and γ_2 are the inclusion indicators. 66

4.3 Posterior model probabilities from GVS example, under the different prior sets as in Table 4.2. 66

5.1 Sampling period (as starting and ending date), and number of 5-days periods within each year. The latter is the number of repeated surveys (or sampling replicates J) within each primary period T (see 'Sampling design' section for further explanations). 73

5.2 Posterior summary of model parameters for the best dynamic occupancy model fitted to data on three migratory species sampled while crossing the souther Alps, during the period 1997-2010. Reported are the mean and the 95% credible interval of parameters for the effect of elevation on local colonization (β_γ) and survival (β_ϕ). 76

B.1 Posterior summaries of parameters for models containing only standardized body length as a covariate on detection probability (p) or the latter in addition with a permanent behavioural response (p_0, p_1). All detection probability values are given in probability scale. 92

B.2 Posterior summaries of parameters for models with a constant detection probability (p) or a permanent behavioural response (p_0, p_1). Unlike models reported in Table B1, individual heterogeneity (β) was here never modelled. All detection probability values are given in probability scale. 93

Introduction

"In matters of scientific inquiry, simplicity is a virtue. Not necessarily procedural simplicity, but conceptual simplicity – and clearly assailable assumptions."

J. Andrew Royle and Robert M. Dorazio (Hierarchical Modeling and Inference in Ecology, 2008)

A variety of problems in ecology, conservation biology, fisheries and wildlife management require inference on species occurrence and richness, whose variation may exist over space and time (*Royle and Dorazio*, 2008). These ecological inferences are typically made at different levels, ranging from a group of individuals to spatially organized community systems, that is metacommunities. Inference begins with the data observed in a sample, and statistical theory provides the conceptual and methodological framework for making conclusions on the hypotheses of interest. Parametric inference and probability modelling form a cohesive approach to compute inference both in theoretical and applied problems at the different ecological scales of organization. Ecological systems have a nested, hierarchical organization, from genes within individuals to communities within metacommunities (e.g. *Begon et al.*, 1990), and the correspondence between ecological levels of organization and the levels of a hierarchical, statistical model is highly relevant to modelling and inference in ecological problems. Hierarchical or multi-level models are becoming increasingly common in population analysis (*Halstead et al.*, 2012). They allow an explicit and formal representation of the data into constituent models of the observations and of the underlying ecological or state process that is the focus of inference (*Royle and Dorazio*, 2008). The process model describes the dynamics of the ecological process (e.g. spatial and temporal variation) and is defined by one or several state variables, which are typically unobserved (latent). The description of the state process is based on our understanding of the study system, without regards to the sampling process and the data. The observation model, conditional on the ecological process, contains a probabilistic description of the mechanisms thought to be responsible for the observable data. Besides, an additional structure can be present to explicit model assumptions relating the parameters of the two models (*Berliner*, 1996). What is relevant for ecologists is

that hierarchical modelling is more than a mere technical approach to model formulation. By focusing on conceptually and scientifically distinct components of a system, this conceptual framework helps clarify the nature of inference problem in a mathematically and statistically precise way (*Kéry and Schaub*, 2011). The term 'state-space' model is commonly used in ecological literature to indicate a hierarchical model that contain an explicit process model (e.g. *De Valpine and Hastings*, 2002). On the contrary, in the widely used hierarchical models containing an implicit process model (e.g. generalized linear mixed models) the 'process', the random effects, lacks an explicit ecological interpretation.

A hierarchical modelling framework is then motivated by two issues related to the sampling process, regardless of the scale or level of organization. The first problem derives from the fact that individuals and species are detected imperfectly. Detectability or detection bias is a pervasive theme in animal sampling, even in the case of sessile organisms (e.g. *MacKenzie and Kendall*, 2002, see also Chapter 2 of this document). Modelling detectability is important because ecological concepts, as well as scientific hypotheses, are formulated in terms of ecological state variables (e.g. abundance, N) and not in terms of difficulty of detecting animals. Imperfect detection induces a component of variation in the observation process that can reduce our ability to measure processes of interest (e.g. *Gimenez et al.*, 2008). In the context of hierarchical models, the observation model often involves an explicit characterization of detection bias, i.e. describes the detectability of individuals or species.

The second problem may arise in the case of spatial sampling (i.e. if our sample units are spatially referenced), when we have to combine or aggregate information across space, while accounting for both 'observation variance' and 'spatial variance'. The latter is the variation in abundance across 'replicate' populations or spatial sample units, and is most relevant when we want to make predictions. Once again, hierarchical models facilitate a formal model-based accounting for this, as well other, components of variation (*Royle and Dorazio*, 2008).

Regardless of form, the model represents an abstraction from which we hope to learn about a phenomenon, a process or a system. In a model-based view of inference, probability has a central role. Probability models provide the basis for describing variation in what we can observe and cannot observe, that is data and ecological processes. Moreover, probability is used to express uncertainty about inference (*Link and Barker*, 2009). Inference is an inductive process where we attempt to make general conclusions from a collection of data. This consists in fitting models to data in order to estimate parameters, to carry out an inference (such as hypothesis test, model selection and evaluation) or to make predictions. If probability is a widely accepted tool for model-based inference, the way by which probability should be used to conduct inference is still debated. In the classical frequentist paradigm, based on the idea of hypothetical collection of repeated samples or experiments, we never make direct probability statements about model parameters. Conversely, the Bayesian approach uses probability with this intention about all unknown (unobserved) quantities (*King*, 2009). It would be impossible both to fully explain what differentiates the two schools of thought, and to describe the principles of Bayesian inference herein, but several recent books exhaustively address these topics (e.g. *Congdon*, 2006; *Clark and Gelfand*, 2006; *Gelman et al.*, 2003; *McCarthy*, 2007) and, in addition, some notes are provided in Appendix A. More importantly is

that the mode of analysis and inference really stands independent of the formulation of the model. If inference is independent from model formulation, why choosing a Bayesian framework for the analysis? A Bayesian approach to inference often provides some benefits, by means of a flexible and coherent framework to analyse even very complex hierarchical models, and by a transparent accounting for all sources of variation in estimates or predictions (*Gelman and Hill*, 2006). However, the practical and conceptual utility of hierarchical models exists independent of the choice of inference framework. As stressed by *Royle and Dorazio* (2008), Bayesian is mostly an adjective that is used to describe the method by which a model can be analysed, and there is much more to hierarchical modelling than the fact that such models can easily be fitted by adopting a Bayesian approach and using a versatile Markov chain Monte Carlo (MCMC) algorithm (*Robert and Casella*, 2004).

A parametric inference based on approximating models that encompass an explicit partition of the ecological and observation processes has a conceptual and practical usefulness for the estimation of demographic quantities that lie at the heart of the science of ecology (*Kéry and Schaub*, 2011). Especially in conservation biology, a rigorous scientific approach to conservation must be based on quantitative evidence and rely on the best estimates (together with their uncertainty) of key state variables like abundance, occurrence, species richness, along with the parameters that govern their dynamics (i.e. survival/extinction, fecundity, colonization and dispersal).

Thesis aims and outline

Hierarchical models are an area of intense statistical research, and many applications relevant for ecologists have yet to be realized (*Halstead et al.*, 2012; *Cam*, 2012). With this theses I pretend to investigate the use of hierarchical Bayesian models to address different ecological hypotheses and provide some further application in the population ecology field, by addressing questions related to different ecological problems, with particular emphasis on the Bayesian approach for inference.

The advantage of using hierarchical models and obtaining inference through Bayesian methods is both, conceptual and practical. The hierarchical approach yields conceptual clarity in the distinction between model components for the ecological processes and the observation processes. The Bayesian approach is relevant because it facilitates parameter estimation by means of MCMC algorithms. In this thesis I use each case to explore the advantages of a Bayesian formulation of hierarchical models and their flexibility in addressing finer questions. Throughout empirical cases I was able to extended the current methods to multiple groups, to include external covariates, to solve problems linked with variable or model selection and to explore new applications.

More specifically I focus on the application, and performance assessment in some cases, of individual covariate models to estimate population size (Chapters 1 and 2) and integrated population models to increase precision in vital rates and estimates of demographic quantities that were not measured in the field (Chapter 3), followed by a close examination on Bayesian variable and model selection (Chapter 4). Finally I explore an innovative application to the

study of the dynamics of multiple species over time and space (Chapter 5). A general outline of each chapter is as follows.

In Chapter 1, using a set of simulated capture-mark-recapture (CMR) data, I first assessed whether a Bayesian formulation of capture-recapture models based on Data Augmentation (DA) can be successfully used in three sample session study, in presence of individual recapture heterogeneity to obtained unbiased estimates of population abundance. Individual heterogeneity has for long time prevented the use of classical closed population models. Recent solutions imply complex models and typically several capture-mark-recapture sessions. Models based on DA provide a simple solution. After exploring their performance through simulations I applied the method to real data to estimate the population size and size-dependent population structure in the endemic Balearic Lizard, *Podarcis lilfordi*. Moreover, I extended the method to the simultaneous analysis of males and females and contrasted hypotheses on sex-ratio and sex-by-size population structure.

In Chapter 2, I used individual covariate models to derive an estimate for the population density of the sessile Noble Pen shell (*Pinna nobilis*) from CMR data on this bivalve, the largest of the Mediterranean Sea, typically associated with seagrass meadows. I investigated whether habitat characteristics or physical forcing, both natural or anthropogenic, are determining the spatial differences in population density and structure. The approach is similar to the one outlined in Chapter 1, but this time I extended the analysis including external covariates.

The underlying idea in Chapter 3 is to use multiple sources of data, each with a specific observational process, to infer shared state variables. Specifically, I used integrated population modelling and perturbation analyses to assess the demographic consequences of the illegal use of poison for an insular population of Red Kite, *Milvus milvus*. I first pooled into a single statistical framework the annual census of breeding pairs, the available individual-based data, the average productivity and the number of birds admitted annually to the local rehabilitation centre. By combining these four types of information I was able to increase estimate precision and to obtain an estimate of the proportion of breeding adults, an important parameter that was not directly measured in the field and that is often difficult to assess. Subsequently, I used perturbation analyses to measure the expected change in the population growth rate due to a change in poison-related mortality.

Chapter 4 deals with the more technical issue of selecting competing models and variables, a procedure that has a fundamental role in the inferential process but has not yet received enough attention when applied in a Bayesian framework. I present two possible procedures, the Gibbs variable selection and the product space method, with emphasis on the practical implementation into a general purpose software in BUGS language. To clarify the related theoretical aspects and practical guidelines, I explain the methods through applied examples on the selection of variables in a logistic regression problem and on the comparison of non-nested models for positive continuous response variable.

In Chapter 5 I explore an innovative approach to the study of the dynamics of multiple species over space and time. Specifically, I used the presence-absence data of three migrating species collected over several alpine sites to assess the dynamical processes of extinction and

colonization along the altitudinal gradient. This analysis used a hierarchical formulation of dynamic occupancy models (*Royle and Kéry*, 2007). I presented this work as the last application because it merges some of the concepts and model structures outlines in the previous chapters. Also, I would like to warn the reader that it has to be considered as a work in progress because more models have to be explored and the results are meant to be extended to other 35 migratory species. Nevertheless the idea, the procedure and the first results are presented. To my knowledge this is the first time that dynamic occupancy models are applied to data from a medium-term ringing scheme on migrating birds.

In the final chapter I draw some general conclusions stressing the communal theme that joins the different applications outlined in chapters 1 to 5. Once again, I put emphasis on the advantage of describing biological processes disentangling their hierarchical nature and the flexibility of the Bayesian approach for statistical inference in complex models.

Supporting information for chapters from 1 to 5 is provided in the appendixes from B to F, respectively.

Chapter 1

Population abundance, size-dependent structure and sex-ratio in an insular lizard

Simone Tenan, Andreu Rotger, José Manuel Igual, Óscar Moya, J. Andrew Royle, Giacomo Tavecchia (in review). *Ecology*.

Abstract

The distribution and abundance of individuals over space and time is a central issue in ecological and evolutionary studies. However, robust estimates of population density, or abundance, are difficult to obtain due to the interaction between experimental design, sampling processes and behaviour responses. Setting the number of sampling occasions, to estimate closed population size from capture-recapture (CR) data, is often a question of balancing objectives and costs. Moreover, accounting for trap-response and individual heterogeneity in recapture probability can be problematic when data derive from less than four sampling occasions. We first assessed, using a set of simulated data with a medium-to-low probability of recapture and individual recapture heterogeneity, whether a Bayesian formulation of CR-models based on Data Augmentation can be successfully used in three sample session study. Results showed that under the correct model, estimates of population size, population structure and recapture probabilities were close to those used in the simulations. We then applied the method to real data to estimate the population size and size-dependent population structure in the endemic Balearic Lizard, *Podarcis lilfordi*. Moreover, we extended the method to the simultaneous analysis of males and females and contrasted hypotheses on sex-ratio and sex-by-size population structure. Results from real data indicated a negative permanent trap response and a positive effect of size on detection probability, so that the estimated size-dependent population structure was more skewed toward smaller size than the observed sample. We estimated mean density at about 800 lizards/ha and a nearly even sex-ratio. We found that

data augmentation provides a flexible and robust framework to include complex recapture processes, when analysing three session capture-mark-recapture studies, with possible extension to between-groups comparisons.

Key-words: population estimates; data augmentation; insular lizard; simulations; sex-ratio; capture-mark-recapture; photoidentification.

1.1 Introduction

The distribution of individuals over time and space is a central theme in evolutionary and ecological theories (*Begon et al.*, 1990), but robust estimates of population size are notoriously difficult to obtain (*Kendall*, 1999, see review in *Seber*, 1982, 1992; *Schwarz and Seber*, 1999). Exhaustive counts are not possible and the number of individuals should be inferred from a partial sampling of the population using appropriate statistical models (*Seber*, 1982; *Skalski et al.*, 2005b; *Williams et al.*, 2002). Capture-mark-recapture (CMR) models, based on multiple observations of marked individuals, are nowadays common methods in ecology and evolutionary studies to estimate survival and recruitment parameters in open populations, but have been originally developed to estimate animal abundance (*Seber*, 1982; *Schwarz and Anderson*, 2001; *Williams et al.*, 2002). When capture sessions are conducted within a short interval of time, the population can be considered closed, i.e. birth or death do not occur during sampling, and simple CMR models can be used to estimate animal abundance. The simplest of these models is the Lincoln-Petersen index based on two capture-recapture sessions (*Seber* 1982). The method has three important advantages that make it particularly appealing for ecological studies. It relies on the minimum number of capture-mark-recapture occasions, i.e. two, it provides a simple formulation for the estimate of population size and its variance, and it does not require the use of individually distinguishable marks. The Lincoln-Petersen method can be extended to K capture sessions, with $K > 2$. In this case population estimates can be obtained using the approximation proposed by Schnabel and Schumacher-Eschymeyer (in *Seber* 1982). Whether to sample the population two or more times is often a question of balancing objectives and costs. However choices underlying the survey design are not without consequences. For example, in two CMR occasions the sampling effort is minimized, but sample size might not be enough to reach the desired precision. Using a set of simulated data, *Rees et al.* (2011) showed that a population should be sampled 5 to 10 times to obtain unbiased estimates of its size. If multiple samples, i.e. > 2 occasions, may guarantee a large enough sample size (*Rees et al.* 2011), extending the sampling over a long period has the risk to violate the closure hypothesis, that is to say that no births, deaths or permanent movements occur during sampling.

Survey design should also consider that CMR models rely on the hypotheses that marks are not lost and that all individuals are independent and equally likely to be captured. While the first hypothesis is generally met, the assumption of independence and equal capture probability across individuals is not always verified. In a study of elk *Cervus elaphus* population

abundance *Skalski et al.* (2005a) found that the probability of being marked covaried with herd size, so that animals from the same herd tend to have the same encounter history. Heterogeneity in detection probability might be due to individual characteristics, such as size or sex. The recapture probability of the sessile Pen Shell *Pinna nobilis*, for example, depends on shell width and assuming equal detection would lead to underestimate recruitment rate (*Kéry and Schaub*, 2011; *Hendriks et al.*, 2012 in press). Unequal catchability rises also when recapture probability at a given time depends on whether an animal was captured before or not. This trap-response can be positive, when captured animals are more likely to be captured again, or negative, when captured animals are less likely to be captured in the future (*Pollock et al.* 1990). Trap-response might be common in studies with baited traps or in which animals have to be physically recaptured. Unequal catchability across individuals leads, more typically, to underestimate the true animal abundance (*Pollock et al.* 1990).
There are multiple ways to correct for a trap-dependence in closed population models (*Pledger*, 2000; *Chao et al.*, 2008) however, none of them seems flexible enough to allow trap-response and recapture heterogeneity in studies with less than four occasions. Recently, *Royle* (2009) has showed how abundance estimates can be derived by K capture-recapture occasions through individual covariate models analysed in a Bayesian framework by data augmentation (henceforth DA). The method permits incorporation of heterogeneity of recapture due to individual characteristics, as well as trap-response, and has recently been extended to include between occasion survival parameters (*Gardner et al.* 2010a). The estimation of population size through data augmentation has been applied to data with multiple capture-recapture occasions and a moderate-to-high recapture probability (*Royle and Dorazio* 2008). However in many fieldwork studies these 'ideal' conditions are not met and surveys may not be repeated more than two or three times due to cost or logistic problems (see for example *Hendriks et al.* 2012 in press).
Here we wanted to assess whether the individual covariate model approach can be successfully used in small sample sizes that result from three sample sessions. We first assessed the performance of this method using a set of simulated data with a medium-to-low probability of recapture and individual recapture heterogeneity, i.e. size-dependent recapture and trap-response. We then apply the method to real data to estimate the population size and size-dependent population structure in the Balearic Lizard *Podarcis lilfordi* (Fig. 1.1), a lizard endemic to the Balearic archipelago (Spain). Moreover, we extended the method to the simultaneous analysis of males and females and contrast hypotheses on sex-ratio and sex-by-size population structure.

1.2 Materials and Methods

1.2.1 Model formulation

We refer to the classical closed population situation in which a population of N individuals is sampled J times, yielding encounter histories on $n \leq N$ individuals. If we assume that detection probability does not vary over J occasions we can consider the capture frequencies

Figure 1.1: Balearic Lizard, *Podarcis lilfordi* (Photograph by S. Tenan).

of the sample of n unique individuals, where individuals $i = 1, 2, \ldots, n$ were captured $\{y_i\}_{i=1}^{n}$ times. An auxiliary individual variable (x_i, in this case body length) is thought to influence the detectability of individuals. We assumed that captures are independent and identically distributed (i.i.d.) Bernoulli trials with parameter $p(x_i; \theta_1) \equiv p_i$, with $logit(p_i) = \alpha + \beta x_i$, where the parameter θ_1 is the vector $\theta_1 = (\alpha, \beta)$. We adopted a Bayesian formulation of the individual covariate model based on parameter-expanded data augmentation (e.g. *Royle and Dorazio* 2011). The general concept is to physically augment the observed data set with a fixed, known number, say $M - n$, of "all zero" capture-recapture histories, and to analyse the augmented dataset (of size M) with a new model. Given the augmented dataset, we introduced a set of latent variables z_i for $i = 1, 2, \ldots, M$ which are Bernoulli trials with the parameter ψ. This parameter is the inclusion probability, that is the probability that an individual from the augmented data list is an element of the exposed population. Thus, $1 - \psi$ is the zero-inflation parameter, quantifying the number of excess zeros in the augmented data list. Parameter ψ is related to N in the sense that $N \sim Binomial(M, \psi)$ under the model for the augmented data. Conceptually, M represents the size of some hypothetical super-population of individuals, from which the population of N individuals exposed to sampling represent a subset of. If $z_i = 0$, then individual i (from the super-population of size M) does not correspond to an individual in the population exposed to sampling, whereas if $z_i = 1$ individual i is a member of the population of size N (*Royle and Dorazio* 2008). We assert that M is sufficiently large so that the posterior of N was not truncated (achieved by trial and error with no philosophical or practical consequence; *Royle and Young* 2008). Under this formulation, population size is a derived parameter $N = \sum_{i=1}^{M} z_i$, and the estimation problem is converted from one of estimating N to one of estimating the parameter ψ and summaries of the latent variables z.
The model for the augmented data is composed of three components:

1. $z_i \sim Bernoulli(\psi)$;

2. $[y_i|p(x_i)] \sim Binomial(J, z_i\, p(x_i))$, with $logit(p_i) = \alpha + \beta x_i$;

3. $[x_i] \sim Normal(\mu_x, \sigma_x^2)$.

We further extended the model to account for a behavioural effect due to a possible permanent trap response (*Royle and Dorazio* 2008), so that the detection model is

$$logit(p_{ij}) = \alpha_0(1 - x_{1,ij}) + \alpha_1 x_{1,ij} + \beta x_{2,i}.$$

The covariate $x_{1,ij}$ indicates if the individual i was captured at some previous time ($x_{1,ij} = 1$ if the individual was captured previous to sample j), and $x_{2,i}$ is the body length covariate. In this model parametrisation, α_0 is the mean for individuals that have not previously been captured, and α_1 is the intercept for previously captured individuals. A further step was to express α_0 as the product of a trap response parameter (tr) and α_1 as follows:

$$logit(p_{ij}) = tr\, \alpha_1\, (1 - x_{1,ij}) + \alpha_1\, x_{1,ij} + \beta\, x_{2,i}.$$

Expressing α_0 as a function of α_1 allowed us to save one parameter when modelling real data using sex-specific detection probabilities (see below), assuming that the degree of trap response was equal for males and females.

1.2.2 Simulated data

We simulated a population of $N = 200$ individuals along with their standardized body length (drawn from a Normal distribution with mean 0 and variance 1) and subjected them to sampling considering detection probability positively affected by body length (with a covariate coefficient $\beta = 0.5$ on the logit scale) and two different levels for the mean detection probability ($p = 0.3$ and $p = 0.6$). In addition, we sampled the simulated population in the presence of a behavioural response after initial capture, a negative permanent trap response (trap shyness) again for two distinct levels of detection probability. More specifically, we fixed that the mean probability of being detected for the first time ($p_0 = 0.3$ and $p_0 = 0.6$) was twice the probability at next occasions (hence $p_1 = 0.15$ and $p_1 = 0.30$, respectively). Three sampling occasions were considered, and the observed data consist of capture histories and covariate values for a variable number of individuals for each of the four different datasets obtained. Simulated data were modelled considering or not the behavioural response on detection probability and the individual covariate. Therefore, concerning the behavioural effect, we obtained the following combinations: (i) data simulated and modelled without trap response, (ii) data simulated without trap response but modelled with a permanent trap response, (iii) data simulated and modelled with a permanent trap response, and (iv) data simulated with a permanent trap response but modelled without it. All these combinations were done for a lower and a higher level of detection probability (obtaining eight different cases), keeping the value of the parameter for individual heterogeneity, β, constant.

1.2.3 Capture-recapture data

Individuals were captured using pit traps positioned along the edges of bushes over an area of c. 0.25 ha. Captured lizards were measured and sexed by the inspection of phemoral pores. Recent work has shown that individuals of *Podarcis muralis* can be recognized by the highly variable and individually unique pattern of pectoral scales (*Sacchi et al.* 2010). In a pilot study conducted during the period 2007-2008 we found the same high variability in the Balearic lizard (results not shown). Individuals were identified from their pectoral scale patterns using the computer aided APHIS procedure (Moya et al. in prep.). Each captured lizard was photographed using a digital camera and released on the same trap where it was captured. Here we used the capture-photo-recapture data of 130 lizards (64 males, 66 females) collected in June 2010 during three consecutive days.
We initially considered four potential models. Model 1: p constant across individuals and time; Model 2: p affected only by a possible permanent behavioural effect; Model 3: p affected only by lizards' body length; Model 4: p affected by both trap response and individual heterogeneity. For the analyses, the body length covariate was centred, by subtracting the mean. We then selected the best supported model (see Appendix B) and extended it considering sex-specific inclusion probabilities as follows, for an individual i of sex u:

1. $z_{iu} \sim Bernoulli(\psi_u)$;
2. $[y_i|p(x_i)] \sim Binomial(J, z_{iu}\, p(x_i))$,
   with two competing linear predictors:
   $logit(p_{iju}) = tr\, \alpha_{1,u}\, (1 - x_{1,ij}) + \alpha_{1,u}\, x_{1,ij} + \beta\, x_{2,i}$, and
   $logit(p_{iju}) = tr\, \alpha_{1,u}\, (1 - x_{1,ij}) + \alpha_{1,u}\, x_{1,ij} + \beta_{1,u}\, x_{2,i}$;
3. $[x_i] \sim Normal(\mu_x, \sigma_x^2)$.

The two candidate models had in common a sex-specific detection probability and differed for the coefficient for the body length effect (sex-dependent or not). For further details on model selection see Appendix B.

1.2.4 Bayesian estimation

We augmented both simulated and real data with $M - n = 500$ observations of $y = 0$, and corresponding missing covariates. Models with sex-specific parameters were run on a single dataset previously augmented of 200 zeros for each sex. Posterior masses for the estimates of population size N were located well away from the upper bounds, indicating that sufficient data augmentation was used. As in *Royle* (2009) we adopted conventional default priors which, ostensibly, express little prior information about the model parameters. For the body length mean parameter (μ_x), a normal prior with mean 0 and variance 1000 was used (replicating the analyses with a uniform prior between -10 and 10), whereas for precision ($\tau_x = 1/\sigma_x^2$) a gamma prior with shape and scale both equal to 0.01 was used. For the trap response parameter (tr) we replicated the analyses using a normal prior with mean 0 and variance 1000, and two uniform priors between -5, 5 and -10, 10 respectively. For

the α_1 and β parameters we repeated the analyses using both a normal prior with mean 0 and variance 1000, and a uniform prior between -10 and 10. For the inclusion parameters, ψ, a uniform prior between 0 and 1 was used. Summaries of the posterior distribution were calculated from three independent Markov chains initialized with random starting values, run 50,000 times after a 20,000 burn-in and re-sampling every 20 draws for simulated data. For modelling real data 100,000 iterations, a 50,000 burn-in and a thinning rate of 30 were used. We computed the Brooks-Gelman-Rubin convergence diagnostic ($\hat{R}$; *Brooks and Gelman* 1998) for which values near 1.0 indicate convergence. For our data, the $\hat{R}$ for each parameter was less than 1.009. The models were implemented in program WinBUGS (*Lunn et al.* 2000), that we executed from R (*R Development Core Team* 2011) with the package R2WinBUGS (*Sturtz et al.* 2005). An R script with the WinBUGS model specification for the sex-specific model is provided as supplement in Appendix B.

1.2.5 Size dependent population structure

Once a sample of the joint posterior distribution was obtained, the estimation of lizards' encounter frequency in relation to body length is straightforward. From the superpopulation of latent variables z_i we can extract and tabulate data for individuals that are members of the population of N individuals exposed to sampling (those with $z = 1$). From this sample we summarized size dependent population structure for the whole sample of individuals in the study area.

1.3 Results

1.3.1 Simulated data

Models that contain the same effects on detection probability considered when data were simulated, in this case individual heterogeneity and presence/absence of a behavioural response, achieved good estimates of population size N and other parameters of interest (Appendix B, Table B.1). However, as expected estimate precision is lower in the presence of a trap response and a low detection probability ($p_0 = 0.3$, $p_1 = 0.15$). When data generated without trap response were modelled considering a permanent trap response effect, precise estimates were obtained only for a moderate-to-high mean detection probability ($p = 0.6$). Conversely, when data simulated with a negative trap response were analysed without any behavioural effect, parameters estimates were imprecise, regardless of the mean value of detection probability. The effect of individual heterogeneity on detection probability (β) can be correctly estimated with or without a behavioural response. Furthermore, parameter estimates seemed closer to the reference values when the model does not take individual heterogeneity into account (Appendix B, Table B.2).

1.3.2 Capture-recapture data

The best supported model included both behavioural response and effect of body length (model 4), with a posterior probability of 0.953. Model 2, with only trap response received a weak support (0.047), while the other two models (constant and with only individual heterogeneity) received 0 posterior probability. The sensitivity analysis for testing the influence of parameter priors on model selection revealed only minimal changes on posterior model probabilities, ranged from 0.898 to 0.965 for model 4, and from 0.035 to 0.102 for model 2, leaving unaffected the relative importance of covariates.
The posterior distribution for population size N in the most supported model is shown in Appendix B (Fig. B.1). Posterior summaries of parameters are given in Table 1.1. Posterior distributions of p_0, p_1, and β were concentrated above zero. Thus, the results indicate a decrease in detection probability once an individual is captured, and a positive effect of size on the detection probability (Fig. 1.2). The estimated population mean and standard deviation of the body length covariate indicate that the sample of measured covariate values (mean=6.42, sd=0.66) was slightly biased towards greater values. Back-transforming the posterior mean of the estimate for μ gave a population mean $E[x] = 6.29$, with $SD[x] = 0.70$ (Fig. 1.3). The proportion of lizards with body length lower than the sample mean was estimated to be 13% higher than that derived by the sampled data.

Table 1.1: Posterior summaries of model parameters for an endemic *Podarcis* lizards data under a model containing permanent behavioural response and individual heterogeneity on detection probability. The parameters p_0 and p_1 are the mean detection probabilities (on the probability scale) for the first capture event and subsequent occasions respectively, β is the coefficient on body length, μ_x and σ_x are respectively the mean and sd of population body length, ψ is the 'zero-inflation' parameter associated with data augmentation.

Parameter	Mean	SD	2.5%	Median	97.5%
N	179.141	32.155	145.000	171.000	259.000
p_0	0.415	0.072	0.265	0.420	0.548
p_1	0.264	0.036	0.197	0.263	0.335
β	0.774	0.263	0.291	0.760	1.314
μ_x	-0.132	0.105	-0.370	-0.122	0.037
σ_x	0.696	0.052	0.606	0.693	0.808
ψ	0.285	0.054	0.216	0.275	0.416

We extended the best supported model to males and females, considering p_0, p_1 and ψ as sex-specific and looking at the posterior model probabilities to test a possible difference in the individual heterogeneity between sexes. The model with sex-independent β received a posterior probability of 0.822, with some variation in relation to different prior on parameters (0.800–0.995). Considering this latter model, posterior distributions for sex-specific population sizes are reported in Appendix B (Fig. B.2). The estimates indicate an even sex-ratio, with widely overlapping 95%CRI for the sex-specific population sizes (Table 1.2).
Sensitivity analysis gave similar posterior mean and 95%CRI for the trap response parameter

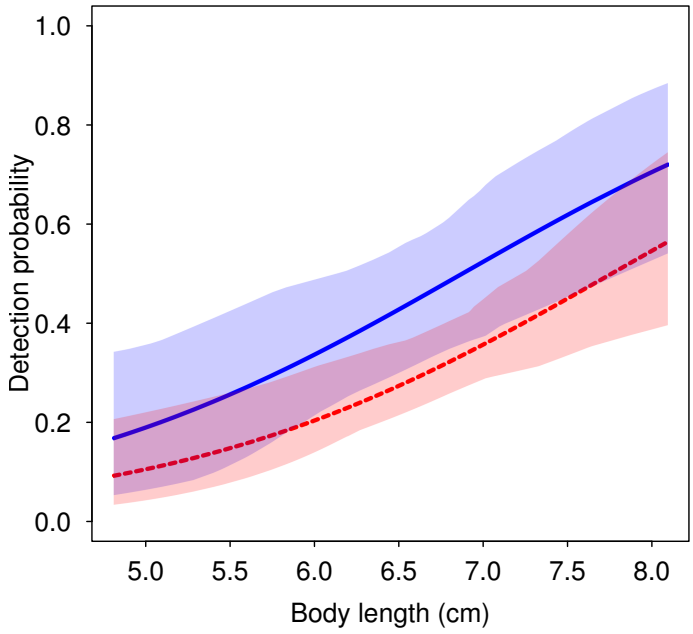

Figure 1.2: Relationship between detection probability p and body length, for first captured lizards (solid line) and for lizards captured at least once in the previous occasions (dashed line). Shaded areas represent 95%CRI.

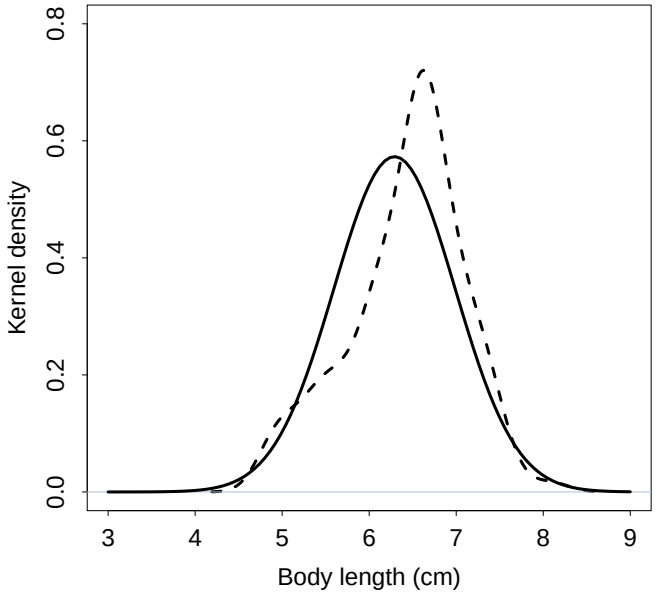

Figure 1.3: Size dependent population structure, as density of individuals in relation to body length for sampled lizards (dashed line) and for population estimates (solid line).

(tr) under different priors [posterior mean (95%CRI), under N(0,1000): 0.357 (-0.271–1.237); under U(-5,5): 0.350 (-0.291–1.243); under U(-10,10): 0.351 (-0.270–1.241); Appendix B, Fig. B.3]. As before, posterior distributions of mean detection probabilities (in this case sex-specific) and β were concentrated above zero. Results indicates a slightly stronger decrease in detection probability for females once an individual is captured and, in common to both sexes, a positive effect of individual size on detection probability.

Table 1.2: Posterior summaries of model parameters for data on sexed individuals of an endemic *Podarcis* lizards, under a model containing sex-specific permanent behavioural response and common individual heterogeneity on detection probability. For sex u, the parameters $p_{0,u}$ and $p_{1,u}$ are the sex-specific mean detection probabilities (on the probability scale) for the first capture event and subsequent occasions respectively, β is the coefficient on body length, μ_x and σ_x are respectively the mean and sd of population body length, ψ_u is the sex-specific 'zero-inflation' parameter associated with data augmentation. m denotes males, f females.

Parameter	Mean	SD	2.5%	Median	97.5%
N_m	92.521	22.104	71.000	86.000	153.000
N_f	101.338	27.774	73.000	93.000	182.000
$p_{0,m}$	0.401	0.073	0.236	0.410	0.526
$p_{0,f}$	0.375	0.087	0.192	0.381	0.532
$p_{1,m}$	0.287	0.050	0.196	0.286	0.388
$p_{1,f}$	0.239	0.046	0.157	0.236	0.337
β	0.866	0.332	0.274	0.838	1.592
μ_x	-0.147	0.119	-0.448	-0.128	0.026
σ_x	0.652	0.051	0.565	0.647	0.764
ψ_m	0.352	0.088	0.247	0.332	0.589
ψ_f	0.382	0.108	0.255	0.355	0.689

1.4 Discussion

1.4.1 Survey design and the estimate of animal abundance by capture-recapture models

The use of CMR models to estimate animal abundance in closed populations relies on the validity of model assumptions, the main ones being that population is closed, marks are not lost and individuals are independent and equally capturable. In some cases, results are robust to deviation from these assumptions and assessing the performance of CMR model in different situations would provide important guidelines for survey design (e.g. *Skalski et al.* 2005a). *Rees et al.* (2011) simulated a survey of an hypothetical closed population of small mammals with traps positioned at random, on a grid or along the habitat preferred by animals. They concluded that, when traps are randomly positioned, it would be necessary to

use about ten capture-recapture sessions to obtain an accurate estimate of population size, but only half if traps were positioned along the preferred habitat. We found that individual covariate models, in a Bayesian formulation, provide a flexible and robust framework to include complex recapture processes when analysing three session capture-mark-recapture studies. Results from simulated data have showed that under the correct model, estimates of population size, population structure and recapture probabilities were close to those used in the simulations. In real data we can test whether a model with trap-response is appropriate through model selection. Besides, to our knowledge, there are not informative goodness of fit test available to detect trap-response or heterogeneity. (Note that temporary trap-response can be test using available software such as U_CARE (*Choquet et al.* 2002) or MARK (*White and Burnham* 1999)). Interestingly a general model including an 'unnecessary' trap-response effect performed well when the recapture probability was moderately high ($p = 0.6$), but not when it was moderately low ($p = 0.30$). Applying this model to data would be 'playing safe' only when capture probability is high. The opposite was never true; simpler models, i.e. with no trap-response, were never adequate when trap-response was present and population size tended to be overestimated . A possible way to avoid the model selection would be to compute the likelihood of the data given a particular model. A likelihood ratio test (LRT) can thus be used to detect the presence of a trap-response (results not shown; *Pollock et al.* 1990). This approach can be done in three-sessions study as the likelihood of the data given a particular model can be easily computed, but it becomes cumbersome if the number of occasions increases.

1.4.2 Population size, sex-ratio and size-dependent structure

Pérez-Mellado et al. (2008) used line transect methods (*Buckland* 2001) to estimate the density of the Balearic lizard in 41 islands of the Balearic archipelago. They found that density varies across the islets of the archipelago from a minimum of 32 to a maximum of about 8000 lizards ha^1 with most islands having a density lower than 1000 ha^1 (median = 717; Table 1 in *Pérez-Mellado et al.* 2008). The reason of this variability is not fully known and it is likely to be the result of multiple factors such as island size, density compensation and predation relaxation in islands (*Pérez-Mellado et al.*, 2008; *Salvador*, 2009). We estimated mean density at about 800 lizards ha^1, close to the median density in the archipelago but a third of the one found by *Pérez-Mellado et al.* (2008) for the same island. Part of the difference of our estimates with those reported may be due to the natural fluctuations of the population from one year to the next. However part might arise from the systematic biases of the two methodologies. In habitat with dense vegetation, where animals are more difficult to detect, accurate density estimates by line transect are hardly achievable. Capture-mark recapture methods, are thought to be more accurate that visual methods in monitoring elusive species (*Wanger et al.* 2009). However, in our case the capture method select for individuals large enough to reach and fall into the traps (body length ≥ 4.8). This might lead to underestimating total population size, depending on the presence and proportion of very small individuals. We found that recapture probability covaried positively with lizard size so that

the estimated population is more skewed toward smaller sizes than the observed sample. A limitation in our case was that we assumed lizards's body size to be normally distributed, and hence symmetric. Finally, we have shown how capture-recapture models of the ordinary sense, but analysed using a Bayesian formulation with data augmentation, can be extended to the simultaneous analysis of multiple (two, in our case) groups. This allowed us to estimate lizard sex-ratio taking advantage of the parameters shared between groups (see also *Gardner et al.* 2010b for another example with spatial capture-recapture models). Adult sex-ratio in lizards is often reported to be females biased, as expected in polygynous vertebrates (*Massot et al.*, 1992; *Le Galliard et al.*, 2005; *Buckley and Jetz*, 2007). However this varies substantially in time and space (*Massot et al.*, 1992; *Galán*, 2004). *Schoener and Schoener* (1980) proposed a mechanistic model in which the number of females changed with per capita resource availability while the number of males depended on habitat quality. The interaction between resource availability and habitat quality would generate spatio-temporal changes in the sex-ratio. Hence populations are expected to be female-skewed in good habitats and/or when resources are abundant and male-skewed when habitat are bad and/or resources are scarce. Indeed, biased adult sex-ratio can arise temporally due to the interaction between ephemeral resources and despotic behaviour (Pérz-Mellado pers. comm.) or might result from a higher permanent emigration or recapture probability of females or an higher mortality of males (*Schoener and Schoener*, 1980; *M'Closkey et al.*, 1998; *Galán*, 2004). We found that recapture probability was slightly lower in females, but once corrected for this difference, the estimates of females number was only slightly higher than the estimated number of males. In *Schoener and Schoener* (1980)'s model this would correspond to a high level of per capita resources. The area surveyed is limited by a small beach regularly visited by tourist during the all summer period. It is possible that the extra food provided by tourists increases the per capita resources. If this was true, we expect the 1:1 ratio to change throughout the year.

1.5 Acknowledgments

The research was founded by the project BFU-2009-09359 from the Minister of Economy and Innovation of the Spanish Government. We thank Lucia Bonnet for facilitating the access to the study area. Thanks to IB port and "ExcursionBoat" Colonia St. Jordi, for their help with the logistics. The Conselleria de Medi Ambient of the Balearic Governement for the permission to carry the study.

Chapter 2

Recreational tourism structures coastal populations of the largest Mediterranean bivalve

Iris E. Hendriks, **Simone Tenan**, Giacomo Tavecchia, Núria Marbà, Gabriel Jordà, Salud Deudero, Elvira Álvarezg, Carlos M. Duarte (in review). *Biological Conservation.*

Abstract

The decline of important coastal habitats, like seagrass meadows, is likely to influence populations of associated species, like the Noble Pen Shell, *Pinna nobilis*. Here we used a Bayesian formulation of individual covariate models to derive a reliable estimate of populations of *Pinna nobilis* in shallow, and thus usually most impacted, areas around the island of Majorca, Balearic Islands, Spain. At six evaluated sites we find quite distinct densities ranging from 1.4 to 10.0 individuals/100 m^2. These differences in density could not be explained by habitat factors like shoot density and meadow cover, nor did dislodgement by storms (evaluated by maximum wind speeds at the sites) seem to play an important role. However, Noble Pen Shell density was related to anchoring as at sites where anchoring was not permitted the average density was 7.9 individuals/100 m^2 while in sites where ships anchored the density was on average 1.7 individuals/100 m^2. As for the conservation of *Posidonia oceanica* meadows, for the associated population of *Pinna nobilis* it would be of utmost importance to reduce anchoring pressure as a conservation measure for these endangered and protected bivalves.

Key-words: Pinna nobilis, habitat, population structure, hierarchical models, Bayesian analysis, data augmentation, capture-recapture, population size, individual covariate.

2.1 Introduction

Coastal marine biodiversity is expected to decrease as a consequence of the biotic and abiotic changes resulting from anthropogenic activities (*Hendriks et al.*, 2006; *Jordà et al.*, 2012a; *Vaquer-Sunyer and Duarte*, 2008; *Waycott et al.*, 2009). Yet, despite a general consensus on this scenario, the current state of many benthonic populations and the factors threatening their persistence are still poorly understood (*Irish and Norse*, 1996; *Kochin and Levin*, 2003; *Lawler et al.*, 2006). A consequence of this knowledge gap are major uncertainties concerning adequate managerial strategies to address the emerging conservation problems (*Norse and Crowder*, 2005) for benthonic populations in littoral areas. Marine communities in coastal areas are characterized by the presence of 'ecosystem engineers', species able to modify the physical and geochemical conditions in their environment, facilitating the life of other organisms in the community (*Bouma et al.*, 2009; *Jones et al.*, 1994). The reduction of ecosystem engineers is likely to create an extinction cascade difficult to evaluate (*Coleman and Williams*, 2002; *Gutiérrez and Jones*, 2008; *Gutiérrez et al.*, 2012; *Ormerod*, 2003). The most important engineer species in the Mediterranean Sea are corals, bivalves and seagrasses. Seagrass, particularly *Posidonia oceanica*, meadows directly modify the nature and complexity of sediment composition and contribute to increase water clarity (*Duarte*, 2000; *Gutiérrez et al.*, 2012; *Hendriks et al.*, 2010). However, *Posidonia* meadows are declining (*Marbà and Duarte*, 2010; *Marbà et al.*, 2005) in parallel to mounting impacts of human activities in Mediterranean coastal ecosystems (e.g. *Vaquer-Sunyer and Duarte*, 2008). The decline in *Posidonia* meadows, resulting from compounded local and global effects, is so severe as to possibly drive these ecosystems to functional extinction before the end of this century (*Jordà et al.*, 2012a). This will impact the populations of species associated with *Posidonia*, which harbours species of particular conservation importance.

The species associated with *Posidonia oceanica* which status is most compromised is arguably the Noble Pen Shell, *Pinna nobilis*, the largest bivalve of the Mediterranean Sea. *Pinna nobilis* threatened by ocean acidification, habitat loss and/or direct human disturbance. Throughout the Mediterranean basin typical Noble Pen Shell densities are in the range of a few (1 to 10) individuals per 100 m^2 (*Moreteau and Vicente*, 1982; *Vicente et al.*, 1980; *Zavodnik et al.*, 1991). Its populations are in decline, and the species is listed as endangered and protected under the European Council Directive 92/43/EEC (*EEC*, 1992). *Pinna nobilis* is particularly vulnerable to anchoring impacts, associated with the increasing use of Mediterranean coastal areas (*Milazzo et al.*, 2004). The Noble Pen Shell is typically associated with meadows of the seagrass *Posidonia oceanica*. Even though there are populations of Noble Pen Shell that are not associated to seagrass meadows (*Katsanevakis*, 2005, 2007), this is an exception and normally populations are closely linked to seagrass habitats. Seagrass provides shelter for small animals from storms that can dislodge them (*García-March et al.*, 2007; *Hendriks et al.*, 2011), increases food supply for filter feeders by reducing current flow and trapping particles (*Hendriks et al.*, 2008; *Peterson et al.*, 1984) and provides shelter from predators. In littoral areas used for recreational tourism, many meadows have been impacted by anchoring and pollution from recreational boating, which is insufficiently regulated around the islands (*Procaccini et al.*, 2003; *Sánchez-Camacho*, 2003). Seagrass habitats are sensitive to damage from

dragging anchors (*Backhurst and Cole*, 2000; *Ceccherelli et al.*, 2007; *Duarte*, 2002; *Walker et al.*, 1989). *Pinna nobilis*, with relatively fragile shells, stand upright in the seagrass meadows, protruding up to 70 cm above the sediments, and can be damaged directly by the anchor track. The persistence of *Pinna nobilis* populations is dependent on anthropogenic impacts, but also on habitats properties. However, the latter are poorly understood.

Our first objective was to derive an estimate for the population density of *Pinna nobilis* around the Balearic Islands and investigated whether habitat characteristics or physical forcing are determining the spatial differences in population density and structure (*Hedley et al.*, 2004). Capture-mark-recapture (CMR) models, based on multiple observations of marked individuals, can be used to estimate animal abundance (*Seber*, 1982; *Williams et al.*, 2002). CMR models include a set of parameters to account for the observational process, such as detection failures (*Schwarz and Anderson*, 2001; *Williams et al.*, 2002), expected to be a problem for organisms living in seagrass meadows. These models rely on the hypotheses that all individuals are equally likely to be captured. If not corrected, unequal catchability leads to biased estimates of the animal abundance (*Pollock et al.*, 1990). *Hendriks et al.* (2012) found that the probability of detection of Noble Pen Shells in Posidonia meadows is positively associated with shell size, but that this association is similar across sites. In contrast, we expect population density and structure to vary spatially. *Royle* (2009) have showed how size-dependent recapture can be incorporated into models for population abundance using data augmentation techniques (for other examples see e.g. *Kéry and Schaub*, 2011; *Royle and Dorazio*, 2008). Here we extended the model to stratified data and simultaneously analysed the CMR data from five different sites. We then test whether site dependent differences were influenced by site-specific anchoring pressures or by the physical characteristic of the habitat. This information can be used to focus conservation efforts for the population of endangered bivalves in coastal areas.

2.2 Methods

2.2.1 Study area

We conducted a survey along the coastline of Majorca, Baleares, Spain at six sites, Magalluf ($39^\circ\,30.100'\,\mathrm{N}$, $2^\circ\,32.360'\,\mathrm{E}$), Cala d'Or ($39^\circ\,22.164'\,\mathrm{N}$, $3^\circ\,13.887'\,\mathrm{E}$), Pollença ($39^\circ\,53.792'\,\mathrm{N}$, $3^\circ\,05.523'\,\mathrm{E}$), Es Cargol ($39^\circ\,16.394'\,\mathrm{N}$, $3^\circ\,02.476'\,\mathrm{E}$), St. Maria ($39^\circ\,09.000'\,\mathrm{N}$, $2^\circ\,56.960'\,\mathrm{E}$) and Es Castell ($39^\circ\,09.120'\,\mathrm{N}$, $2^\circ\,55.480'\,\mathrm{E}$); Fig. 2.1). The six sites had a uniform depth between 5-6 m but contrasting anchoring pressure and physical characteristics. Magalluf is an area with important tourist development with associated pollution (*Medina*, 2004) and physical disturbance of the habitat by anchoring of boats. Cala d'Or is also an important tourist destination, but at the time of the surveys, the site was closed for anchoring in summer to prevent spreading of the invasive alga *Caulerpa taxifolia*. Pollença is an extensive bay in the North of the island, which gently slopes so the surveyed area at 5m depth is far from the typical recreational areas. Es Cargoll is located at an exposed area on a south-eastern tip of the island. St. Maria is a protected area in the archipelago of Cabrera with no anthropogenic

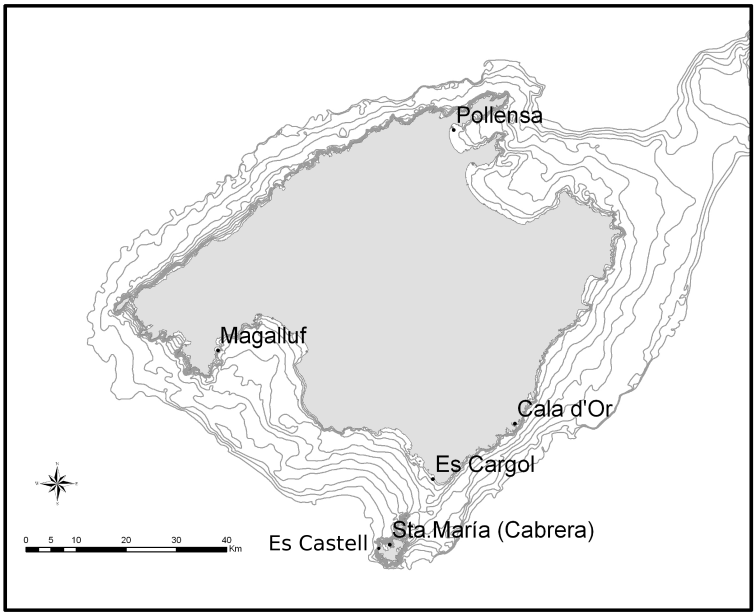

Figure 2.1: Map with the locations of the sites where surveys were conducted around the island of Majorca, Balearic Islands, Spain.

pressures, while Es Castell is located in the same archipelago at the entrance of the enclosed area (Es Port) where in summer many boats enter, but mooring is only allowed at supplied permanent mooring buoys so no physical damage is expected in this area (see Table 2.1).

2.2.2 Data collection, habitat descriptors and anchoring

Noble Pen Shell surveys were carried out in two years (2007 and 2010). In each site, we randomly positioned underwater tape of 30 m length as transect line. Each transect was randomly assigned a team of two divers, with at least one experienced diver on each team. Capture-recapture data were collected along a strip with 1.5 m width at each side of this line. Each diver marked all Noble Pen Shells found within this strip using a metal peg inserted until level with the sediment with a discrete tag displaying a unique alphanumeric code. Once at the end of the transect line, divers switched side and searched for already marked Noble Pen Shells marked by the previous diver ('re-capture'). The width of each marked individual was noted on a PVC bar and measured in the laboratory. On subsequent surveys, diver teams changed randomly to minimize a possible 'diver' effect on recapture probability. A total of 13 different SCUBA divers participated to the project for a total of 15 different two-diver teams. Team composition did not affect the counts (more information on the methodology in *Hendriks et al.*, 2012). Capture-mark sessions were organized on different days throughout the year but mostly concentrated either in summer (water temperature allows for a longer or more comfortable immersion time) or winter (shorter leaf length of the seagrass facilitates

Table 2.1: Site characteristics. Maximum wind speed is the directional component (taking into account exposure of the sites) of the wind velocity p95 (95th percentile) averaged over 2008, 2009 and 2010. Fragmentation was measured in 2009, while shoot density and coverage are measured between 2001 and 2008.

	Es Castell	St. Maria	Cala d'Or	Pollença	Magalluf	Es Caragol
	Physical forcing					
Anchoring	absent	absent	absent	medium	high	high
Wind (m/s)	10.5 ± 0.6	10.5 ± 0.6	10.9 ± 0.2	11.3 ± 0.3	10.5 ± 0.4	10.5 ± 0.6
	Habitat					
Shoot density	569.8	865.9	384	612.9	494.7	1012.5
Meadow coverage	93.9	93.5	95.7	97.7	87.7	52.1

searching effort).

At each site we evaluated shoot density and seagrass coverage (as % area) of *Posidonia oceanica* meadows as possible predictors for population structure of *Pinna nobilis* (Fig. 2.2). Coverage was estimated from scoring habitat along a 30 m transect. Shoot density was obtained from a permanent monitoring program established on the same locations (*Marbà et al.*, 2005). Strong wind events may lead to dislodgement of individuals through the enhancement of waves or wind-induced currents. We used wind speed as a proxy to characterize the physical stress present on each site. Hourly wind data were obtained from the outputs of the HIRLAM model run by the Spanish meteorological agency (AEMET) at 0.05° (about 5 km) resolution. In order to identify areas subject to stronger physical stresses we compute the 95th percentile of wind intensity over the years 2008, 2009, 2010. In particular we use the wind velocity from the direction to which each particular site was exposed to get an idea of the maximal stresses.

Each site has a different anchoring regime and regulation and as consequence, data on the anchoring pressure from recreational tourism were difficult to obtain. In 2008, the number of registered recreational boaters on the only island of Mallorca was 324,522 (*CITTIB*, 2009). There is no legislation on the total number of boats that can access the anchoring areas around the islands of the archipelago (*Balaguer et al.*, 2011), but in most sites there is a delimitation of bathing areas and a no navigation area is established 200 m from the coast when a beach is present, or 50 m for other types of coast. However, many boats disregard or are unaware of this legislation and anchor in shallower restricted areas close to shore (*Balaguer et al.*, 2011). *Balaguer et al.* (2011) divided the coast of Mallorca in three areas, Eastern, Central and Northern, and estimated the number of boats that navigate these waters. We used their estimates as a index of anchoring pressures for the four site that fall into the areas considered (Cala d'Or, Es Escargol fall into the East area, Magalluf in the Midlle and Pollença in the North). The estimated number of boats would come down to 258, 200 and 584 boats per km^2 of seabed area (seagrass beds included) available and commonly used for anchoring in these areas for the Eastern, Central and Northern area, respectively (Table 1 in *Balaguer et al.*, 2011). However, Cala d'Or was closed off during our survey period so now anchoring was

Figure 2.2: Noble Pen Shell, *Pinna nobilis* (Photograph by I. Hendriks).

allowed in this area. The remaining two sites are inside the National Park of Cabrera Islands where anchoring is off-limits (St. Maria) or strictly regulated (Es Castell, Table 2.1).

2.2.3 Model formulation and parameter estimation

Each site was considered as a closed population in which N individuals are sampled on J occasions with $J = 2$, leading to a sample of n unique individuals. We considered the detection constant over the occasions, and converted the encounter histories to capture frequencies of the sample of n unique individuals (*Royle and Dorazio*, 2008), where each individual i were captured y times (with $y = 1$ or 2). Data were analysed using a Bayesian formulation of individual covariate models based on parameter-expanded data augmentation technique (DA, hereafter; *Liu and Wu*, 1999; *Royle et al.*, 2007). The general concept is to physically augment the observed data set with a fixed, known number, say $M - n$, of "all zero" encounter histories, and to analyse the augmented dataset (of size M) with a new model. This new model is a zero-inflated version of the conventional known-N model, and could be easily fitted using Markov chain Monte Carlo (MCMC) sampling (see e.g. *Royle and Dorazio*, 2011 for further details on DA). Given the augmented dataset, we introduced a set of latent variables z_i for $i = 1, 2, \ldots, M$ which are Bernoulli trials with the parameter ψ. This parameter is the probability that an individual from the augmented data list is an element of the population. Conceptually, the population of N individuals represents a subset of some hypothetical super-population of individuals M. By means of DA technique, the problem of estimating population size (N) is converted into that of estimating inclusion probability (ψ), since the expecting value for N is equal to $M\psi$ (*Kéry and Schaub*, 2011). Population size N could

potentially be any integer between 0 and M, and DA just induces for N a discrete uniform prior on the interval $(0, M)$. By estimating the value for each z_i we can estimate Noble Pen Shell abundance in a way that naturally excludes the structural zeros in the augmented data. If $z_i = 0$, then individual i from the super-population of size M does not correspond to an individual in the population exposed to sampling, whereas if $z_i = 1$ the individual is a member of the population of size N. An estimator of the total population size is then simply derived as $N = \sum_{i=1}^{M} z_i$ (*Royle*, 2009). We can estimate the total number of individuals in each specific site or for groups of sites sharing similar characteristics, the only difference being the indexes we use in the summation. Because the total sampled area was different from site to site we derived the density (individuals/100 m^2) of Noble Pen Shell by dividing N for the specific sampled area at each model iteration to obtain credible interval of N. To model the stratified population size we augmented each group-specific dataset, and then fitted a model with group-specific inclusion probability (ψ_{group}) to the ensemble dataset (see below). In this way we estimated the latent variable for the unobserved individuals in each specific group of interest, as well as the related shell width.

Individual covariate for unobserved individuals was estimated by assuming shell width as normally distributed, with a mean and a variance to be estimated. Then, by introducing in the observation process the relationship between shell width and detection probability, the population mean and variance of the individual covariate was corrected for the size-biased sampling. We therefore specified this dependence as $logit(p_i) = \alpha + \beta x_i$ with $x_i \sim Normal(\mu_x, \sigma_x^2)$, where α parameterizes the detection probability (on the logit scale) for an average size individual while β parameterizes the variation in detectability in relation to shell width (vector x, previously centred by subtracting the mean). Given this basic model formulation, we aggregated data in different ways to address specific questions by building the corresponding models. First, we tested for a seasonal effect on detection probability, aggregating data from all sites in relation to sampling season (winter or summer) considering $logit(p_{i,season}) = \alpha_{season} + \beta x_i$ with $x_i \sim Normal(\mu_x, \sigma_x^2)$. A second model assumed a site-specific mean and variance for shell width and the basic constraint for detection probability, $logit(p_{i,site}) = \alpha + \beta x_{i,site}$ with $x_{i,site} \sim Normal(\mu_{x,site}, \sigma_{x,site}^2)$. To test the effect of environmental covariates (shoot density, fragmentation, meadow coverage and maximum wind speed) and an anthropogenic factor (presence/absence of anchoring) on Noble Pen Shell density we assumed the probability that the ith individual is a member of the population exposed to sampling to depend on meadow parameters, wind speed, or anchoring. These effects were modelled separately, with the same model formulation. In this case, given the site-specific standardized values (by subtracting the mean and divided for the standard deviation) of each predictor, we let ψ_{cov} denote the probability that an individual from a site with a specific covariate value (cov) is a member of the population of pen shells exposed to sampling (i.e. group-specific inclusion probability). Thus, the model of z_i, for the ith individual detected in a site with a covariate value cov_{site}, can be written as $z_i|cov \sim Bernoulli(\psi_{cov})$, with $logit(\psi_{cov}) = \alpha_\psi + \beta_\psi \, cov_{site}$. Parameter β_ψ represents the slope for the relationship between the covariate value of a specific site and the number (and/or density) of Noble Pen Shells present in that site. Similarly, we let ψ_{anch} denote the probability that an individual, from a

site with or without anchoring, is a member of the population of Noble Pen Shells exposed to sampling.

As before the model for the latent state of the ith individual in a certain site with anchoring present or not was $z_i|anch \sim Bernoulli(\psi_{anch})$, which provides an explicit connection between the presence of anchoring in a site and the number (and thus density) of Noble Pen Shells present in that site. In the model we then derived the difference in the estimated densities of individuals between sites with and without anchoring, together with its related uncertainty, as a direct measure of the effect of anchoring on Noble Pen Shell density. In the models with the continuous predictors we assumed a site-specific population mean and standard deviation for shell width, $x_{i,site} \sim Normal(\mu_{x,site}, \sigma^2_{x,site})$, while in relation to anchoring we were interested in modelling and evaluating the shell width population structure in relation to this human-related factor, as $x_{i,anch} \sim Normal(\mu_{x,anch}, \sigma^2_{x,anch})$.

Posterior masses for the estimates of population size N were located well away from the upper bounds, indicating that sufficient data augmentation was used. For shell width mean parameter (μ_x), a normal priors with mean 0 and variance 1000 was used (replicating the analyses with a uniform prior between -10 and 10), whereas for precision ($\tau = 1/\sigma^2$) a gamma prior with shape and scale both equal to 0.001 was used. For the α and β parameters we repeated the analyses using both a normal prior with mean 0 and variance 1000, and a uniform prior between -10 and 10. For the inclusion parameters ψ a uniform prior between 0 and 1 was used.

Summaries of the posterior distribution were calculated from three independent Markov chains initialized with random starting values, run 100,000 times after a 50,000 burn-in and re-sampling every 30 draws. For our analyses the Brooks-Gelman-Rubin convergence diagnostic (*Brooks and Gelman*, 1998) was less than 1.003 for all parameters, which indicates convergence. Model formulations were implemented in program WinBUGS (*Lunn et al.*, 2000), executed from R (*R Core Team*, 2012) with the package R2WinBUGS (*Sturtz et al.*, 2005). An R script with the WinBUGS model specification (for anchoring effect) is provided as supporting information in Appendix C.

2.2.4 Population structure in relation to shell width

From the marginal posterior distribution of parameter z_i we summarized Noble Pen Shells' frequency in relation to shell width. Thus, from the super-population of latent variables z_i we extracted and tabulated data for individuals that are members of the population of N individuals exposed to sampling (those with $z = 1$). We then summarized size dependent population structure for the different sampling sites and in relation to anchoring.

2.3 Results

We marked a total of 356 individuals, with an average shell width of 14.57 cm $\pm$ 0.27 SE. Average detection probability did not differ between the two sampling seasons, with widely overlapped 95% credible intervals (hereafter 95%CRI) for the two estimates (p_{winter}

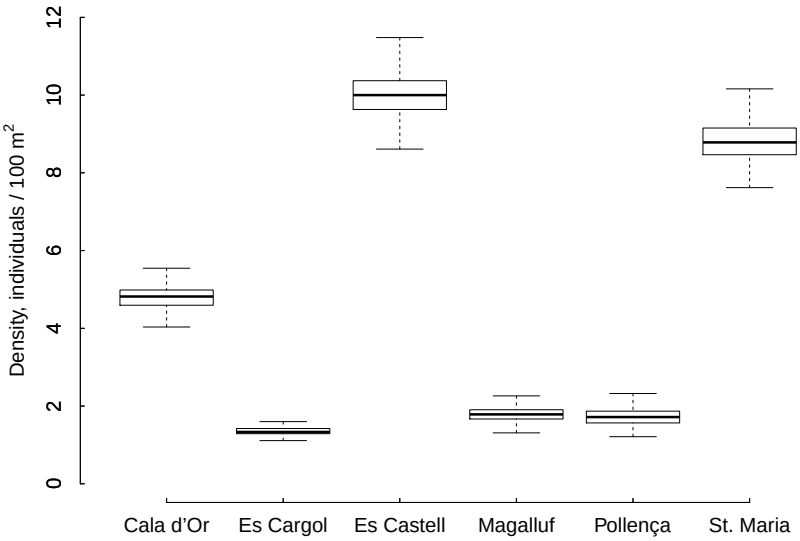

Figure 2.3: Site-specific densities (individuals/100 m^2) of Noble Pen Shell in the island of Majorca, Balearic Islands, Spain.

= 0.578, 0.502–0.653, *95%CRI*; p_{summer} = 0.587, 0.522–0.652, *95%CRI*). Site-specific estimates of Noble Pen Shell density varied, on average, from 1.4 (1.2–1.6, *95%CRI*) to 10.0 (9.0–11.2, *95%CRI*) individuals/100 m^2 (Table 2.2, Fig. 2.3). As expected, detectability was positively affected by shell width, with a *95%CRI* for the slope parameter that did not encompass zero (β = 0.126, 0.078–0.174, *95%CRI*; Table 2.2, Fig. 2.4). Environmental parameters were not related to Noble Pen Shell density as the effect of shoot density was not relevant and the *95%CRI* for the related parameter did encompass zero (β_ψ = -0.028, -0.183–0.126, *95%CRI*). Moreover, Noble Pen Shell density was not affected by meadow coverage (β_ψ = 0.044, -0.128–0.217, *95%CRI*). Dislodgement by storms did not seem to be an issue in our populations as Noble Pen Shell density was not affected by wind speed (β_ψ = 0.061, -0.150–0.292, *95%CRI*). In contrast, average Noble Pen Shell density was different in relation to anchoring, with 7.9 (7.1–8.9, *95%CRI*) individuals/100 m^2 in sites without anchoring pressure and 1.7 individuals/100 m^2 (1.4–2.1, *95%CRI*; Table 2.3, Fig. 2.5) in sites where anchoring was permitted. Site-specific population structure in relation to shell width showed certain variability in both mean and standard deviation (Table 2.2, Fig. 2.6). Average shell width varies from 8.116 cm (6.691–9.394, *95%CRI*) for Es Cargol to 17.090 cm (16.033–18.058, *95%CRI*) for Es Castell. Average shell width standard deviation was smaller in Es Cargol (3.139, 2.343–4.388, *95%CRI*) and wider in Pollença (9.770, 6.965–13.880, *95%CRI*; Table 2.2). Size-dependent population structure showed differences also in relation to anchoring (Fig. 2.7). Mean estimated shell width was higher in sites without anchoring (14.440, 13.608–15.187, *95%CRI*) with no overlapping credible intervals between the two estimates as mean size in presence of anchoring was 9.362, 6.667–11.468, *95%CRI* (Table 2.3). The estimate for shell width standard deviation was distinctly lower in locations without anchoring (4.761, 4.327–5.259, *95%CRI*) than in those with this physical stressor present (6.905, 5.587–8.779, *95%CRI*; Fig. 2.7).

Table 2.2: Posterior summary of model parameters for data of the Noble Pen Shell aggregated at site level. Densities were derived parameters expressed as individuals/100 m^2, α is the detection probability (on the logit scale) for an average size individual, β is the slope for the relationship between detectability and shell width (cm), ψ is the "zero-ination" parameter associated with data augmentation, $\mu_{x,site}$ is the site-specific mean shell width, $\sigma_{x,site}$ is the shell width standard deviation. Note that α, β, and ψ are site-independent. Posterior mean and related 95% credible interval are reported for each parameter.

Parameter	mean	2.5%	97.5%
Density Cala d'Or	4.8	4.3	5.5
Density Es Cargol	1.4	1.2	1.6
Density Es Castell	10.0	9.0	11.2
Density Magalluf	1.8	1.4	2.2
Density Pollença	1.7	1.4	2.2
Density St. Maria	8.8	8.0	9.9
α	0.350	0.123	0.574
β	0.126	0.078	0.174
ψ	0.315	0.279	0.356
$\mu_{x,\textit{Cala d'Or}}$	12.653	11.618	13.605
$\mu_{x,\textit{Es Cargol}}$	8.116	6.691	9.394
$\mu_{x,\textit{Es Castell}}$	17.090	16.033	18.058
$\mu_{x,\textit{Es Magalluf}}$	11.688	8.894	14.129
$\mu_{x,\textit{Pollença}}$	9.891	4.595	14.172
$\mu_{x,\textit{St. Maria}}$	14.237	13.239	15.126
$\sigma_{x,\textit{Cala d'Or}}$	3.866	3.247	4.629
$\sigma_{x,\textit{Es Cargol}}$	3.139	2.343	4.388
$\sigma_{x,\textit{Es Castell}}$	4.307	3.686	5.083
$\sigma_{x,\textit{Es Magalluf}}$	5.441	3.878	7.984
$\sigma_{x,\textit{Pollença}}$	9.770	6.965	13.880
$\sigma_{x,\textit{St. Maria}}$	4.668	4.090	5.376

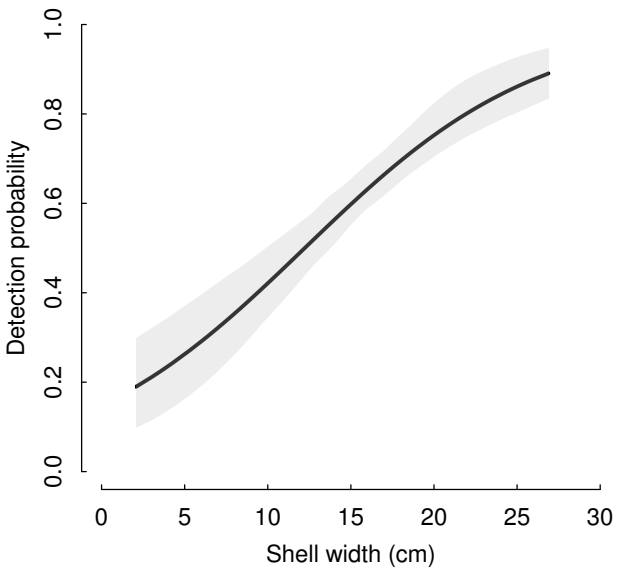

Figure 2.4: Relationship between Noble Pen Shell detection probability (p) and the individual covariate, shell width. The shaded area represents *95%CRI*.

Table 2.3: Posterior summary of model parameters; Noble Pen Shell in relation to anchoring. Densities were derived parameters expressed as individuals/100 m^2, α is the detection probability (on the logit scale) for an average size individual, β is the slope for the relationship between detectability and shell width, $\psi_{no\ anchoring/anchoring}$ are the "zero-ination" parameters associated with data augmentation and specific for locations without or with anchoring, $\mu_{x,no\ anchoring/anchoring}$ are the site-specific means for shell width (cm), $\sigma_{x,no\ anchoring/anchoring}$ are the shell width standard deviations. Note that α and β are anchoring-independent. Posterior mean and related 95% credible interval are reported for each parameter.

Parameter	mean	2.5%	97.5%
Density with no anchoring	7.9	7.1	8.9
Density with anchoring	1.7	1.4	2.1
Density difference	6.2	5.5	7.1
α	0.335	0.103	0.563
β	0.143	0.096	0.195
$\psi_{no\ anchoring}$	0.329	0.287	0.382
$\psi_{anchoring}$	0.356	0.271	0.469
$\mu_{x,no\ anchoring}$	14.440	13.608	15.187
$\mu_{x,anchoring}$	9.362	6.667	11.468
$\sigma_{x,no\ anchoring}$	4.761	4.327	5.259
$\sigma_{x,anchoring}$	6.905	5.587	8.779

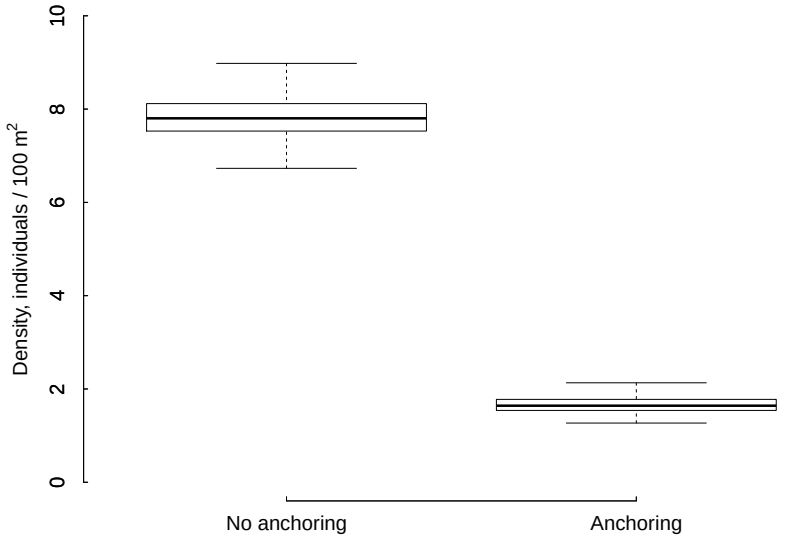

Figure 2.5: Densities (individuals/100 m^2) of Noble Pen Shell in the island of Majorca (Balearic Islands, Spain), in relation to presence/absence of anchoring.

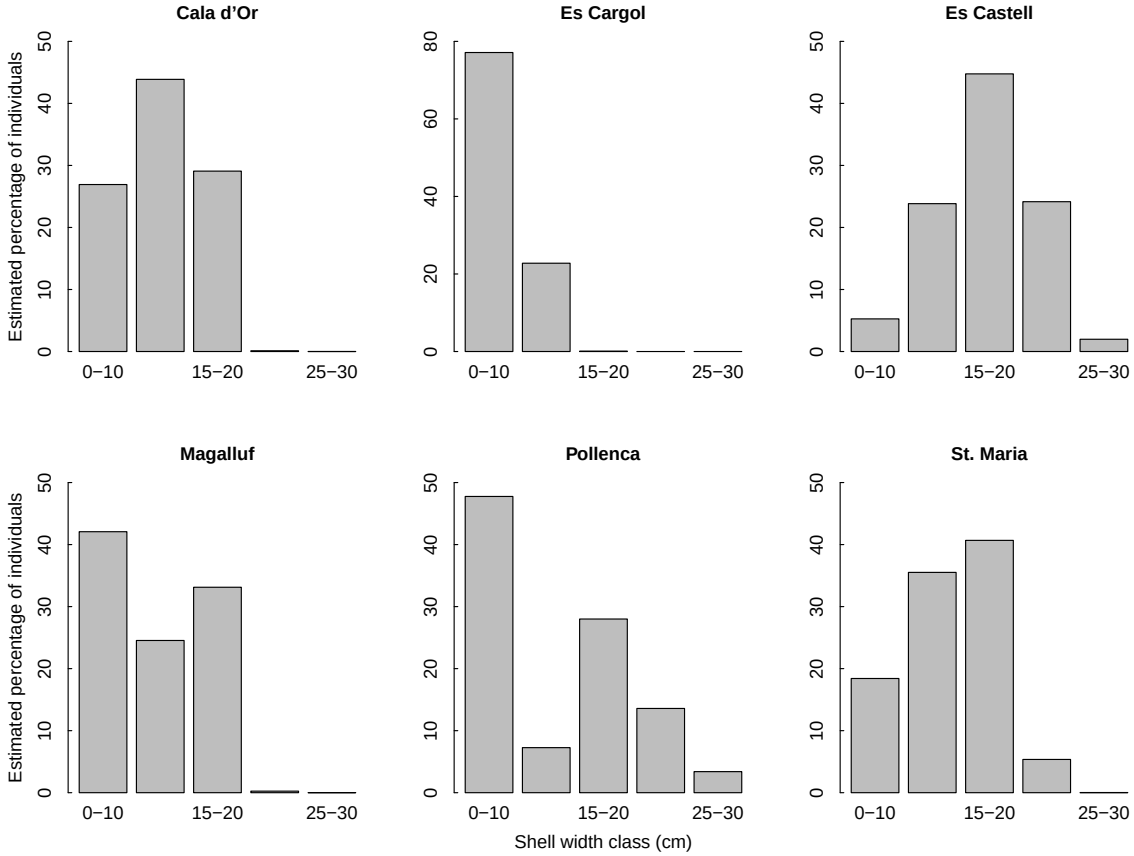

Figure 2.6: Size (shell width) dependent population structure of Noble Pen Shell for each sampling site, in the island of Majorca (Balearic Islands, Spain). Estimated proportions of individuals for different dimensional classes are reported. Note the different y-axis scale for Es Cargol.

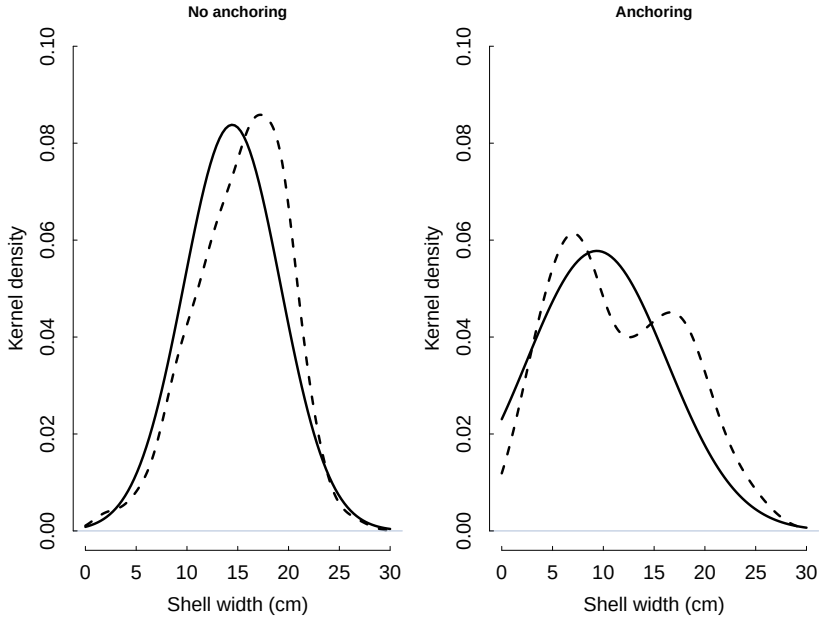

Figure 2.7: Kernel density estimates for Noble Pen Shell size (shell width) in relation to presence/absence of anchoring, for sampled individuals (dashed line) and population estimates (solid line).

2.4 Discussion

We used a Bayesian formulation of individual covariate models to investigate the difference in structure and abundance of the Noble Pen Shell populations in six coastal sites of the archipelago of Balearic islands (Spain). The technique of data augmentation allowed us to derive reliable estimates for number of individuals as well as for the population structure, even if there was an effect of shell width on the detectability of the individuals. Processing the data according to normal capture-mark protocols, i.e. without individual heterogeneity in detection probability, would have resulted in an under estimation of small individuals while over estimating the percentage of large animals in the population (*Hendriks et al.*, 2012). We found that population density and the number of large Noble Pen Shells were less in sites with anchoring of recreational boats. Site-specific differences in population size and structure however can be attributed to many factors, including food availability.

The close association of the Noble Pen Shell with its seagrass habitat and the facilitation seagrass provides in terms of mechanical shelter, food increase and shelter against predation leads to believe that meadow structure would affect the (sustainable) number of bivalves living within its boundaries. However, we did not find a significant effect of simple variables like shoot density or spatial cover on population density or size structure of *Pinna nobilis*. It is possible that the processes that structure of the meadows of *Posidonia oceanica* act on different time scales from those that structure the populations of the Noble Pen Shell. *Posidonia* meadows can be very old and grow very slowly while recruitment and development of the population of *Pinna nobilis* would take place over time scales like 10 to maximum 30 years.

On the other hand, shoot mortality has increased very rapidly and most probably the current regression of seagrass meadows, will not have a direct effect on Noble Pen Shell populations already established until a critical thresholds of shoot density is met.

We used wind speed as a proxy for storminess and wave action since we did not encounter suitable data on waves for all our sites. Climate models project less storms over the Mediterranean basin for the end of the century (*Giorgi and Lionello*, 2008; *Marcos et al.*, 2011). This would reflect on lower wave heights (*Jordà et al.*, 2012b; *Medina*, 2004), which will even decrease the pressure of dislodgement caused by wave action. We thus believe that natural physical forcing dislodging individuals during storms with high wave action is not a likely factor in structuring the populations of the Noble Pen Shell around the Balearic archipelago, not now, nor in the near future.

The major determinant of shallow populations of *Pinna nobilis* appears to be anchoring. Average Noble Pen Shell density was different in relation to anchoring, with a difference of 6.2 (5.5–7.1, *95%CRI*) individuals/100 m^2 between sites with anchoring compared to sites without this pressure. In 2008, the number of recreational boaters on the island of Mallorca was 324522 (*CITTIB*, 2009). This cause of structural damage has far more effect than dislodgement by storms or habitat quality.

2.5 Conclusions and Recommendations

Physical dislodgement by anchoring causes fast and unpredictable mortality on larger Noble Pen Shells. Selective mortality of large individuals might have important consequences for the future of the population (*Coltman et al.*, 2003). Conservation efforts for the Noble Pen Shell, *Pinna nobilis*, should prioritize the installation of permanent mooring buoys to decrease physical damage to seagrass meadows and the associated Noble Pen Shell population.

2.6 Acknowledgements

This is a contribution to the MEDEICG project, funded by the Spanish Ministry of Economy and Competitiveness (contract no. CTM2009-07013). The authors thank the many divers helping out with the field surveys; R. Martinez, L. Basso, M. Cabanellas-Reboredo, M. Noguera, E. Diaz, L. Royo, S. Sardu, A. Canepa. I.E.H. was supported by a grant from the Juan de la Cierva program, Spanish government (JCI-2007-123-844). G.J. acknowledges a JAE-DOC contract funded by the Spanish Research Council (CSIC). ST was funded by a PhD grant from the Science Museum (Trento) in collaboration with the University of Pavia.

Chapter 3

Demographic consequences of poison-related mortality in a threatened bird of prey

Simone Tenan, Jaume Adrover, Antoni Muñoz Navarro, Fabrizio Sergio, Giacomo Tavecchia (2012). PLoS ONE 7(11): e49187. doi:10.1371/journal.pone.0049187

Abstract

Evidence for decline or threat of wild populations typically come from multiple sources and methods that allow optimal integration of the available information represent a major advance in planning management actions. We used integrated population modelling and perturbation analyses to assess the demographic consequences of the illegal use of poison for an insular population of Red Kites, *Milvus milvus*. We first pooled into a single statistical framework the annual census of breeding pairs, the available individual-based data, the average productivity and the number of birds admitted annually to the local rehabilitation centre. By combining these four types of information we were able to increase estimate precision and to obtain an estimate of the proportion of breeding adults, an important parameter that was not directly measured in the field and that is often difficult to assess. Subsequently, we used perturbation analyses to measure the expected change in the population growth rate due to a change in poison-related mortality. We found that poison accounted for 0.43 to 0.76 of the total mortality, for yearlings and older birds, respectively. Results from the deterministic population model indicated that this mortality suppressed the population growth rate by 20%. Despite this, the population was estimated to increase, albeit slowly. This positive trend was likely maintained by a very high productivity (1.83 fledglings per breeding pair) possibly promoted by supplementary feeding, a situation which is likely to be common to many large obligate or facultative European scavengers. Under this hypothetical scenario of double societal costs (poisoning of a threatened species and feeding programs), increasing poison

control would help to lower the public cost of maintaining supplementary feeding stations.

Key-words: Bayesian; capture-recapture; demography; integrated population model; matrix population model; Milvus milvus; poisoning; reproductive success; radiotracking; state-space model.

3.1 Introduction

The current rate of biodiversity loss has generated concerns on the future of many wild populations and increased the need for population monitoring and risk assessment. The expected long-term trend of a population and its probability of extinction are typically obtained through population viability analyses (*Beissinger and Westphal*, 1998; *Brook et al.*, 2000; *Beissinger and McCullough*, 2002). The core of this analysis is a mathematical model that projects the current state of the population into the future and estimates population extinction, or quasi-extinction, probability. The population model, which includes survival and fertility parameters and projects the population state, is usually parameterised using estimates of demographic parameters derived from individual-based information, i.e. capture-recapture data (*Caswell*, 2001). Despite much effort to increase the precision and realism of capture-recapture analyses (e.g. *Pradel*, 2009; *Tavecchia et al.*, 2012), the use of a single dataset to parameterise the population model poses the problem of model validation (*Coulson et al.*, 2001b). The latter can be implemented by retrospective analyses (*Caswell*, 2001), for example by comparing model predictions with population surveys (*Coulson et al.*, 2001a), but this ignores sampling errors associated with counts (*Clark and Bjørnstad*, 2004; *Tavecchia et al.*, 2009). Also, the many sources of variation in large capture-recapture datasets generally violate the assumptions of population dynamics models (see *Tavecchia et al.*, 2008; *Sanz-Aguilar et al.*, 2010), propagating errors into the estimate of extinction risk (*Maunder*, 2004). Recently various computational approaches have been proposed to integrate data from different sources of information, such as the P-system based models (e.g. *Cardona et al.*, 2009; *Margalida et al.*, 2011) and integrated population models (*Besbeas et al.*, 2002; *Maunder*, 2004; *Tavecchia et al.*, 2009; *Schaub and Abadi*, 2010). In particular, the latter allows the incorporation of counts and individual-based data into a single analysis through a joint likelihood. In this integrated analysis population counts are linked to population state by an observation equation, while a state equation describes the link between population state and demographic processes through a population model (hereafter 'transition model') based on per-capita survival and fecundity taken from individual-based data (*Tavecchia et al.*, 2009). The transition model is structurally similar to the one used in population projections but is constructed through parameters that integrate information simultaneously acquired from different sources. Such integrated models yield multiple advantages: (1) their integrated structure reduces parameter uncertainty (*Besbeas et al.*, 2002); (2) their consensual estimates increase the realism of population state forecasting and incorporate into model predictions the variance and covariance between different demographic parameters; and (3) they allow estimation of latent parameters, i.e. parameters that appear in the biological process, i.e. the

population model, but not measured empirically (*Tavecchia et al.*, 2009). At present, integrated modelling represents a useful extension of the classical analyses based on a single type of data. This is particularly evident in those cases where information on population threats is available for different data sources or at different spatial scales. For example, individual life-history data of long-lived seabirds come mainly from observations at the breeding colony, whereas the main threat to population persistence is the mortality at sea due to fishery bycatch (*Igual et al.*, 2009). Integrated modelling allows the integration of these two types of data that are typically analysed separately (see e.g. *Belda and Sanchez*, 2001; *Laneri et al.*, 2010; *Igual et al.*, 2009). Similarly, *Schaub and Abadi* (2010) integrate capture-recapture data of eagle owls *Bubo bubo* with the number of owls found dead on the roads and reported by the public.

This sort of analysis is ideally suited to the assessment of conservation threats to endangered organisms, such as many top predatory taxa. The charismatic nature of these species makes them the focus of attention and monitoring by multiple figures, including professional researchers, public administrations and amateurs, leading to simultaneous but heterogeneous sources of information. A good example in this context is offered by birds of prey, a group of species typically monitored by different entities and frequently subject to direct or indirect human related mortality such as illegal hunting (*Smart et al.*, 2010), primary and secondary poisoning (*Whitfield et al.*, 2003), habitat destruction (*Tilman et al.*, 1994), prey depletion, collision with windmills and electrocution on power lines (*Sergio et al.*, 2004; *Lehman et al.*, 2007; *Schaub and Abadi*, 2010). Despite their legal protection in several countries, many raptor populations continue to be at risk, as showed by a worldwide deterioration of their conservation IUCN index (*Butchart et al.*, 2004). Therefore, methods allowing optimal integration of available information to estimate the relative impact of human-related mortality on population growth could represent major advances in our capability to plan management and halt declines. To this aim, we offer an example of implementation of integrated modelling to estimate the impact of illegal poisoning on a threatened raptor. Illegal poisoning is a form of persecution usually generated by conflicts with human interests associated with livestock rearing or hunting, and indiscriminately affects birds or mammals that occasionally or regularly feed on carcasses, or other poison-soaked baits (e.g. pesticides as carbofuran or alpha-chlorolose; *Whitfield et al.*, 2003; *González et al.*, 2007). Toxicoses related to this illegal activity have been identified as the main threat for the conservation of different species of raptors in Europe and Asia (e.g. *Margalida et al.*, 2008). Poison-related mortality often affects breeding adults, and several studies have documented the detrimental effects of this factor on population dynamics, especially for long-lived species with low reproductive rates and delayed maturity (*Hernández and Margalida*, 2009). Our model species, the Red Kite *Milvus milvus* is a medium-sized raptor distributed exclusively in the western Palearctic (*Del Hoyo et al.*, 1996). Since the 19th century, the species has declined throughout Europe, and many of its populations are nowadays considered endangered due to the illegal use of poisoning baits to control predators of game species (*Whitfield et al.*, 2003; *IUCN*, 2010; *Smart et al.*, 2010). In Spain, which holds one of the largest breeding and wintering populations of Europe, Red Kites have been added to the red list of species at risk of extinction in 2011. On

the 3,640 km^2 island of Mallorca of the archipelago of Balearics (Spain), the population was reduced to only 7-8 pairs in the year 2000 (*Adrover et al.*, 2002). The population has recently increased to 19 breeding pairs, but its small size makes it still vulnerable to stochastic peaks of adult mortality, such as those caused by poisoning. A previous analysis of intensive radio-tracking data from this population showed that illegal poisoning accounted on average for 53% of the mortality (*Tavecchia et al.*, 2012). However, this estimate was based on marked birds only and its effect on the population growth rate is unknown.

Our first aim was to obtain a more precise and 'consensual' estimate of mortality due to illegal poisoning. We did so by combining four different types of information: detailed monitoring on radio-marked birds, the number of breeding pairs from annual surveys, the number of fledglings, and the number of birds found poisoned and brought to the local rehabilitation centre by the public. These four data sources were mathematically combined into a population model incorporating the age-dependent demographic parameters. By combining separate datasets we were able to increase estimate precision and estimate the proportion of breeding birds, a parameter that was not directly measured in the field. We then used the consensual estimates derived from the integrated model to parameterised the age-structured population model and assess the demographic consequences of poison related mortality using perturbation analyses.

3.2 Materials and Methods

3.2.1 Data collection

The field data were collected from the Red Kite population of the island of Mallorca, in the archipelago of Balearic Islands (Spain). From 1999 to 2010, the whole island was intensively surveyed annually to count the number of breeding pairs. Territorial pairs were censused throughout the whole island and by visiting formerly occupied breeding sites during the spring courtship period. For our analysis we retained only the number of active nests where at least one egg was laid (hereafter breeding pairs). All broods were intensively monitored and each year the number of fledglings was recorded. Moreover, 142 red kites were equipped with VHF radio-tags (TW-3 model by Biotrack; lifespan: c. 3-4 years) just before fledgling during the period 2000-2010. Individuals were handled following rules and permissions by Conselleria d'Agricultura Medi Ambient i Territori of the Government of the Balearic Islands. In addition to radio-tags, all chicks were marked using PVC wing-tags with a unique alpha-numeric code, one on each wing. The wing-tags were used to assess the loss of radio signal ceased by mechanical or electrical failures. Simultaneous loss of both types of tags was never observed, and all dead birds, found with or without transmitters, had retained at least one wing-tag (*Tavecchia et al.*, 2012). All tagged birds were searched monthly throughout the whole island by car or, occasionally, helicopter. Here, we retained the information on live resightings made from April to June only, whilst we gathered information on birds recovered dead throughout the year. All recovered carcasses (Fig. 3.1) were examined post-mortem to establish the cause of death. Exposure to a toxic substance was confirmed by toxicology

Figure 3.1: Tagged Red kite, *Milvus milvus*, found dead (©GOB Mallorca).

analyses. Additional observations were obtained at feeding stations, territories and roost sites to record the presence of birds whose radio signal had been lost. Finally, we compiled the number of unmarked kites brought to the local wildlife rehabilitation centre between 1999-2010 and killed by poisoning or other causes. These two time-series were formed by a total of four poisoned birds and eleven individuals killed by causes other than poisoning (electrocution, aircraft collision, drowning in artificial water reservoir and other unknown causes). Recovered birds whose radio failed before death were discarded to avoid dependence with radio-tracking data.

3.2.2 Integrated population model

The different sources of demographic information (population surveys, number of fledglings, radio-tracked birds and recoveries of dead individuals) were combined into a single model. The major advantage of analysing all data sets within a single model simultaneously is that the precision of parameter estimates is increased and parameters for which no explicit data are sampled can be estimated (*Brooks et al.*, 2004; *Besbeas et al.*, 2002; *Abadi et al.*, 2010a; *Cave et al.*, 2010; *Kéry and Schaub*, 2011; *Schaub and Abadi*, 2010; *Schaub et al.*, 2010, 2007, 2011). The integrated model was fitted in the Bayesian framework because this provides more flexibility than the frequentist framework and exact measures of parameter uncertainty (*Besbeas et al.*, 2002; *Kéry and Schaub*, 2011; *Schaub et al.*, 2010).

3.2.3 Likelihood for the population count data

To describe the model, we began by describing the likelihoods components for the different demographic parameters and subsequently defined how they would be linked and estimated in a single overall model. Survey data were modelled by a state-space model (*Brooks et al.*, 2004), which consisted of a set of states that described the true but unknown development of

the population and an observation process linking the observed population counts to the true population size assuming an observation error (*Kéry and Schaub*, 2011; *Schaub et al.*, 2010). The state process was described deterministically by a female-based, pre-breeding matrix projection model (*Caswell*, 2001) with three age classes (1, 2 and ≥ 3 years old respectively) as

$$\begin{bmatrix} N_1 \\ N_2 \\ N_{ad} \end{bmatrix}_{t+1} = \begin{bmatrix} 0 & s\,\frac{1}{2}\,b_t\,0.1 & s\,\frac{1}{2}\,b_t\,br_{ad} \\ s & 0 & 0 \\ 0 & s & s \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_{ad} \end{bmatrix}_t \tag{3.1}$$

where $N_{1,t}$ is the number of 1 year old females at time t, $N_{2,t}$ is the number of females of 2 years old at time t, and $N_{ad,t}$ is relative to females older than 2 years at time t. Survival probabilities of a female between time t and $t+1$ is denoted s, and b_t is the fecundity at time t. Although the model was female-based, fecundity refers to the complete reproductive output and it was halved to account for the number of females raised per breeding female. This was justified by the even sex ratio observed for a sub-sample of genetically sexed fledglings (J. Adrover *unpublished data*).

Based on intensive monitoring of radio-tagged birds, Red Kites usually begin breeding at 3 years old or later, although in Mallorca c. 10% of females bred in their second year of life (*Tavecchia et al.*, 2012). However, the proportion of breeding females older than 2 years, br_{ad}, was estimated as a latent parameter, because it was impossible to obtain a figure from our own data due to the limited lifespan of the radio-tags (see below).

Since 1999 we are aware of only one case of emigration from Mallorca (from a sample of 230 marked birds). As a consequence we assumed that no immigration or permanent emigration from the island were present (J. Adrover *unpublished data*).

To account for demographic stochasticity, we used Poisson and binomial distributions to describe the dynamics of the true population size over time, already described by the population model in eq. 3.1. Specifically, the age-specific numbers of females in year $t+1$ were modelled as

$$N_{1,t+1} \sim Po\left((0.1\,N_{2,t} + br_{ad}\,N_{ad,t})\,s\,\frac{1}{2}\,b_t\right) \tag{3.2}$$

$$N_{2,t+1} \sim Bin\,(N_{1,t},\,s) \tag{3.3}$$

$$N_{ad,t+1} \sim Bin\,((N_{2,t} + N_{ad,t}),\,s)\,. \tag{3.4}$$

The observation process is conditional on the state process. We assumed the counts of breeding females in year t (y_t) to follow a Poisson distribution (*Kéry and Schaub*, 2011; *Schaub et al.*, 2010; *Abadi et al.*, 2010b) as

$$y_t \sim Po(0.1\,N_{2,t} + br_{ad}\,N_{ad,t}) \tag{3.5}$$

The likelihood of the population count data is $L_{counts}(\boldsymbol{y}|\boldsymbol{b}, s, \boldsymbol{N}, br_{ad})$.

3.2.4 Likelihood for radio-tracking data

We implemented the model outlined in *Tavecchia et al.* (2012) as a multi-state capture-recapture model in a Bayesian framework (*Kéry and Schaub*, 2011), to estimate an age-

independent survival probability, the age-dependent mortality of marked birds, the incidence of tag loss and the relative magnitude of different sources of mortality.

The observation of live and dead birds, together with the information on tag loss, formed the set of observable events from which we estimated the proportion of birds that died by poisoning or by other (natural) causes. We considered that individuals can move across three main states: alive, dead by poison, and dead because of other causes. Given that individuals can lose their radio transmitter, we considered the above states for birds with and without a functioning radio. Moreover, we included an additional dead state that corresponds to an unobserved dead state (*Kéry and Schaub*, 2011; *Lebreton et al.*, 1999). Therefore, observable "recently dead" individuals move to state "unobserved dead" at the next occasion. This latter state is absorbing, meaning that once individuals are in this state they will remain there (*Kéry and Schaub*, 2011). This differentiation assumes that corpses are found soon after death and allows us to estimate the reporting rate associated with the observable dead states and the probability of dying from different causes (*Schaub and Pradel*, 2004). For alive birds we distinguished six age classes: juveniles (noted 'juv') spanning the time from tagging as nestlings up to the end of the first year of life, one-year olds ('1y') to the time between 1 and 2 yr old, two-year olds ('2y') between 2 and 3 years old, three-year olds ('3y') between 3 and 4 years old, four-year olds ('4y') between 4 and 5 years old, and five or more year olds ('5my') from 5 yr old to all following years. We considered six age-specific states in relation to lifespan of tag batteries, that did not exceed 3-4 years. A marked bird may survive from year t to year $t+1$ with probability s, or it may die with probability $1-s$ some time during the year. If it dies, this is either because of poisoning with probability β_z (subscript z refers to the following age-classes, juv, juvenile; $1y$, one year old; $2my$, 2 years or older) or because of any other cause with probability $1-\beta_z$. The fate of a marked individual also accounts for the radio signal retention α_k. The subscript k define three age-classes on the basis of radio signal decay probability, as estimated in *Tavecchia et al.* (2012) (1 to the time between tagging as nestlings up to the end of the third year of life, 2 between 3 and 4 years old, 3 for all the following years).

The above states, in relation with tag retention and different age classes lead to a 16×16 transition matrix (eq. 3.6). Between any interval, individuals might change state according to the transitions in Fig. 3.2.

At any given time, we could observe 16 types of mutually exclusive events, arbitrary coded with numbers from 1 to 16 (eq. 3.7). The events coded from '1' to '6' refer to encounters of individuals alive with a functioning radio and belonging to one of the six age-classes mentioned above. Similarly, codes from '9' to '13' refer to birds alive and without a functioning radio. Note that the latter codes are referred only to the five age-classes of non-juveniles birds. Codes '7' and '14' refer to individuals found poisoned with and without a functioning radio respectively. Similarly, '8' and '15' code for birds found dead for causes other than poisoning, with and without a functioning radio respectively. These codes do not distinguish whether the radio was physically lost or not functioning. The last possible event (coded '16') refers to cases when the radio signal cannot be heard and the animal cannot be seen. This may correspond to any underlying state: for example, the animal may have lost the radio or

be carrying one that ceased to function, or it may be dead having lost the radio and remaining undetected. Each of the other events can happen only with one state. Conditional on

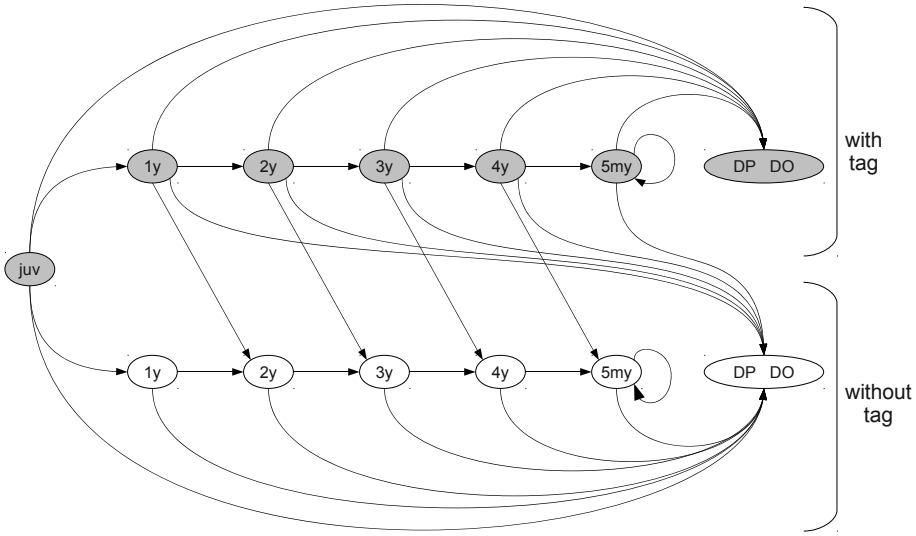

Figure 3.2: Diagram of possible states of a marked red kite. Transitions between two subsequent states, from time t to $t+1$, are denoted with arrows and correspond to parameters in the transition matrix of eq. 3.6. For the sake of clarity, the parameters and the "unobserved dead" state are not reported. Notation: DP: dead by poison; DO: dead by other causes.

the different fates, an animal with a functioning radio may be encountered with probability p, while an animal without radio signal may be encountered with probability c if alive, d_1 if dead by poisoning, and d_2 if dead by other causes.

The likelihood of this sub-model is categorical and we used a state-space parameterization to implement the model (*Kéry and Schaub*, 2011; *Gimenez et al.*, 2007). The likelihood of the radio-tracking data is $L_{rt}(\boldsymbol{T}|s, \alpha_1, \alpha_2, \alpha_3, \beta_{juv}, \beta_{1y}, \beta_{2my}, p, c, d_1, d_2)$.

	$juv.t$	$1y.t$	$2y.t$	$3y.t$	$4y.t$	$5my.t$	$DP.t$	$DO.t$	$1y.nt$	$2y.nt$	$3y.nt$	$4y.nt$	$5my.nt$	$DP.nt$	$DO.nt$	UD
$juv.t$	0	$s\alpha_1$	0	0	0	0	$(1-s)\beta_{juv}\alpha_1$	$(1-s)(1-\beta_{juv})\alpha_1$	$s(1-\alpha_1)$	0	0	0	0	$(1-s)\beta_{juv}(1-\alpha_1)$	$(1-s)(1-\beta_{juv})(1-\alpha_1)$	0
$1y.t$	0	0	$s\alpha_1$	0	0	0	$(1-s)\beta_{1y}\alpha_1$	$(1-s)(1-\beta_{1y})\alpha_1$	0	$s(1-\alpha_1)$	0	0	0	$(1-s)\beta_{1y}(1-\alpha_1)$	$(1-s)(1-\beta_{1y})(1-\alpha_1)$	0
$2y.t$	0	0	0	$s\alpha_1$	0	0	$(1-s)\beta_{2my}\alpha_1$	$(1-s)(1-\beta_{2my})\alpha_1$	0	0	$s(1-\alpha_1)$	0	0	$(1-s)\beta_{2my}(1-\alpha_1)$	$(1-s)(1-\beta_{2my})(1-\alpha_1)$	0
$3y.t$	0	0	0	0	$s\alpha_2$	0	$(1-s)\beta_{2my}\alpha_2$	$(1-s)(1-\beta_{2my})\alpha_2$	0	0	0	$s(1-\alpha_2)$	0	$(1-s)\beta_{2my}(1-\alpha_2)$	$(1-s)(1-\beta_{2my})(1-\alpha_2)$	0
$4y.t$	0	0	0	0	0	$s\alpha_3$	$(1-s)\beta_{2my}\alpha_3$	$(1-s)(1-\beta_{2my})\alpha_3$	0	0	0	0	$s(1-\alpha_3)$	$(1-s)\beta_{2my}(1-\alpha_3)$	$(1-s)(1-\beta_{2my})(1-\alpha_3)$	0
$5my.t$	0	0	0	0	0	$s\alpha_3$	$(1-s)\beta_{2my}\alpha_3$	$(1-s)(1-\beta_{2my})\alpha_3$	0	0	0	0	$s(1-\alpha_3)$	$(1-s)\beta_{2my}(1-\alpha_3)$	$(1-s)(1-\beta_{2my})(1-\alpha_3)$	0
$DP.t$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
$DO.t$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
$1y.nt$	0	0	0	0	0	0	0	0	0	s	0	0	0	$(1-s)\beta_{1y}$	$(1-s)(1-\beta_{1y})$	0
$2y.nt$	0	0	0	0	0	0	0	0	0	0	s	0	0	$(1-s)\beta_{2my}$	$(1-s)(1-\beta_{2my})$	0
$3y.nt$	0	0	0	0	0	0	0	0	0	0	0	s	0	$(1-s)\beta_{2my}$	$(1-s)(1-\beta_{2my})$	0
$4y.nt$	0	0	0	0	0	0	0	0	0	0	0	0	s	$(1-s)\beta_{2my}$	$(1-s)(1-\beta_{2my})$	0
$5my.nt$	0	0	0	0	0	0	0	0	0	0	0	0	s	$(1-s)\beta_{2my}$	$(1-s)(1-\beta_{2my})$	0
$DP.nt$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
$DO.nt$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
UD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

(3.6)

Transition matrix for the multi-state sub-model. From the state at t (rows) to state at $t+1$ (columns) different transition probabilities could encompass the following probabilities: annual survival (s), radio signal retention during the first three, the fourth and the fifth or more year of life (α_1, α_2, and α_3, respectively), and mortality due to poisoning given that an animal has died during its first, second or more than second year of life (β_{juv}, β_{1y}, and β_{2my} respectively). State abbreviations are a combination of a prefix referred either to the six age-classes (from 'juv', for juveniles, to '5my' for 5 or more year olds) or to the cause of death ('DP' for dead by poison, 'DO' for dead by other reasons), and a suffix that specifies the presence of a functioning radio ('.t' and '.nt' for with and without radio signal, respectively).

$$
\left[\begin{array}{c|cccccccccccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\
\hline
juv.t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1y.t & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
2y.t & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
3y.t & 0 & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
4y.t & 0 & 0 & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
5my.t & 0 & 0 & 0 & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
DP.t & 0 & 0 & 0 & 0 & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
DO.t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\
1y.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c & 0 & 0 & 0 & 0 & 0 & 0 & 1-c \\
2y.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c & 0 & 0 & 0 & 0 & 0 & 1-c \\
3y.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c & 0 & 0 & 0 & 0 & 1-c \\
4y.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c & 0 & 0 & 0 & 1-c \\
5my.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c & 0 & 0 & 1-c \\
DP.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & d_1 & 0 & 1-d_1 \\
DO.nt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & d_2 & 1-d_2 \\
UD & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right] \tag{3.7}
$$

Observation matrix for the multi-state sub-model. The matrix specifies the probability of each event (in column, coded with numbers from 1 to 16) conditional on each state (rows). Codes from '1' to '6' refer to encounters of individuals alive with a functioning radio and belonging to one of the six age-classes, from juvenile to 5 or more years old birds. Codes from '9' to '13' refer to birds alive and without a functioning radio. Codes '7' and '14' refer to individuals found poisoned with and without a functioning radio respectively, while '8' and '15' code for birds found dead for causes other than poisoning, with and without a functioning radio respectively. Code '16' refers to cases when the radio signal cannot be heard and the animal cannot be seen. p is the probability of encounter of an animal with a functioning radio, c is the probability of encounter of an animal alive without an active radio signal, d_1 is the probability of encounter of an animal dead by poisoning and without an active radio signal, d_2 is the probability of encounter of an animal dead by other causes and without an active radio signal. For state abbreviations see transition matrix in eq. 3.6.

3.2.5 Likelihood for reproductive success data

We derived fecundity from the yearly counts of fledglings. The fecundity rate (b_t) was defined as the number of offspring (J) produced per mature female in year t. We assumed that J_t followed a Poisson distribution with parameter written as a product of the number of recorded reproducing females (R_t) and fecundity rate (b_t), hence, $J_t \sim Po(R_t\, b_t)$. The likelihood of this sub-model is denoted as $L_{rp}(\boldsymbol{J}, \boldsymbol{R}|\boldsymbol{b})$.

3.2.6 Likelihood for unmarked birds found dead

Within the state-space sub-model for population count data we included yearly counts of unmarked birds recovered dead by the local wildlife rescue centre. The estimated number of unmarked birds found dead by poisoning (N_{dp}) and by other causes (N_{do}) at time $t+1$ were modelled as drawn from a Multinomial distribution with sample size $N_{pop,t} = (N_{1,t} + N_{2,t} + N_{ad,t})$ equal to the total number of individuals in the population (see *Spiegelhalter et al.*, 2007 for the practical implementation of a Multinomial distribution with an unknown order N), and a probability vector made up of the following probabilities

$$\begin{aligned} mdp &= (1-s)\,\hat{\beta} && (3.8)\\ mdo &= (1-s)\,(1-\hat{\beta}) && (3.9)\\ s &= 1 - mdp - mdo && (3.10) \end{aligned}$$

where mdp and mdo are the average probabilities (across age groups) of dying because of poisoning or other causes respectively, while the complementary probability with respect to one is the survival probability. $\hat{\beta}$ is the arithmetic mean of the age-specific proportion of deaths due to poisoning (β_{juv}, β_{1y}, and β_{2my}).

We then assumed the total number of dead birds recovered in year t (dp_t for poisoned and do_t for other causes) to follow a Binomial distribution as

$$\begin{aligned} dp_t &\sim Bin(N_{dp,t},\, d_1) && (3.11)\\ do_t &\sim Bin(N_{do,t},\, d_2). && (3.12) \end{aligned}$$

The likelihood of unmarked birds found dead by poisoning was $L_{dp}(\boldsymbol{dp}|s, \boldsymbol{N}, \beta_{juv}, \beta_{1y}, \beta_{2my}, d_1)$, while the one for birds found dead by other causes was $L_{do}(\boldsymbol{do}|s, \boldsymbol{N}, \beta_{juv}, \beta_{1y}, \beta_{2my}, d_2)$.

3.2.7 Likelihood of the integrated model

The likelihoods of the four types of data have parameters in common, as displayed graphically in Fig. 3.3. By combining these data sources into a single analysis, and by using an integrated population model, more information can be used to estimate demographic parameters (*Abadi et al.*, 2010b). Assuming that the different data types are independent, the joint likelihood of the complete integrated model is the product of the different parts (*Brooks et al.*, 2004; *Besbeas et al.*, 2002, 2003), thus

$$\begin{aligned} L_{IPM} = {} & L_{counts}(\boldsymbol{y}|\boldsymbol{b}, s, \boldsymbol{N}, br_{ad}) \times L_{rt}(\boldsymbol{T}|s, \alpha_1, \alpha_2, \alpha_3, \beta_{juv}, \beta_{1y}, \beta_{2my}, p, c, d_1, d_2) \\ \times & L_{rp}(\boldsymbol{J}, \boldsymbol{R}|\boldsymbol{b}) \times L_{dp}(\boldsymbol{dp}|s, \boldsymbol{N}, \beta_{juv}, \beta_{1y}, \beta_{2my}, d_1) \times L_{do}(\boldsymbol{do}|s, \boldsymbol{N}, \beta_{juv}, \beta_{1y}, \beta_{2my}, d_2). \end{aligned} \quad (3.13)$$

Because the population was small, some individuals were likely to occur in different data sets, violating the assumption of independence between different likelihoods. However, a simulation study, which combined capture-recapture, population count and reproductive success data, showed that the violation of this assumption has only minimal impact on accuracy of parameter estimates (*Abadi et al.*, 2010a). Therefore, although the structure of our data slightly differ from such simulation study, we assumed a similarly minimal impact. The same assumption was employed in another recent study (*Schaub et al.*, 2010).

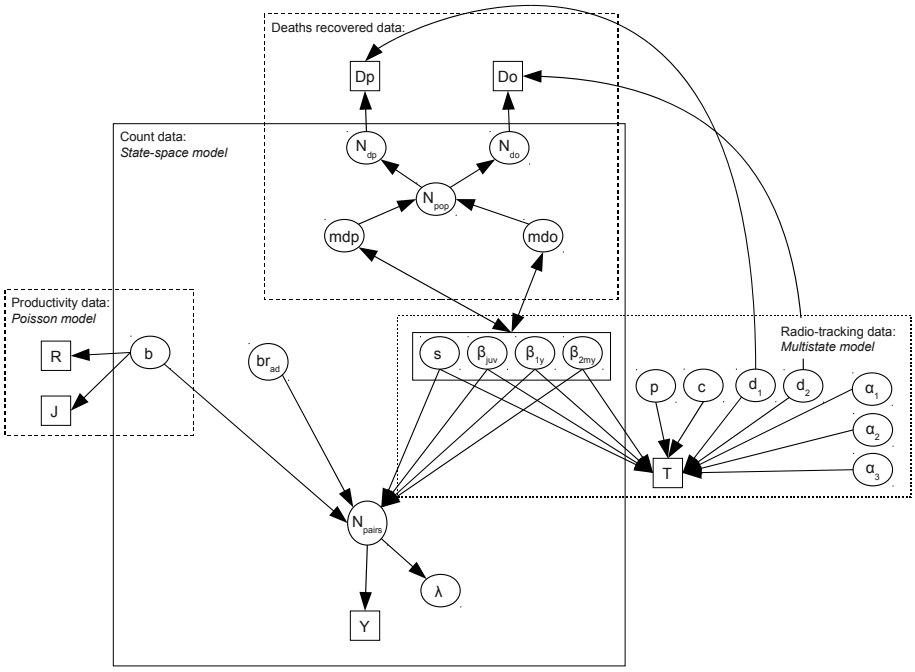

Figure 3.3: Graphical representation of the integrated population model. Data are symbolized by small rectangles, parameters by ellipses, the relationships between them by arrows and sub-models by open rectangles. Notation: J: annual number of fledglings; R: numbers of surveyed broods whose final fledging success was known; Dp: number of unmarked birds found dead by poisoning; Do: number of unmarked birds found dead by causes other than poisoning; T: radio-tracking data; Y: population count data; b: fecundity; mdp: average probability, across age groups, of dying because of poisoning; mdo: average probability, across age groups, of dying because of other causes; N_{pop}: total number of individuals in the population; N_{dp}: expected number of unmarked birds found dead by poisoning; N_{do}: expected number of unmarked birds found dead by causes other than poisoning; s: survival probability; β_{juv}: probability of death due to poisoning given that an animal died in its first year of life; β_{1y}: probability of death due to poisoning given that an animal died in its second year of life; β_{2my}: probability of death due to poisoning given that an animal died after its second year of life; p: recapture probability of an animal with a functioning radio; c: recapture probability of a radio-tagged animal which is alive but without an active radio signal; d_1: probability of encounter of a radio-tagged animal dead by poisoning but without an active radio signal; d_2: probability of encounter of a radio-tagged animal dead by other causes and without an active radio signal; α_1, α_2, α_3: radio signal retention probability during the first three, the fourth and the fifth or more year of life, respectively; br_{ad}: proportion of breeding females relative to the total number of females older than 2 years; N_{pairs}: number of breeding females in the population; λ: population growth rate. Priors are excluded from this graph to increase visibility.

3.2.8 Parameter estimation and model implementation

We used a hierarchical formulation of the integrated model to estimate temporal variability of fecundity (b), while keeping the other demographic rates constant over time. Thus, the annual estimates of fecundity were thought to originate from a random process with a common mean and a constant temporal variance. For the $\log$ of this parameter we assumed

$$\log(b_t) = \gamma_0 + \epsilon_t, \text{ with } \epsilon_t \sim N(0, \sigma_b^2) \tag{3.14}$$

where γ_0 is the mean fecundity on the log scale and σ_b^2 is the temporal variance of fecundity on the $\log$ scale.

The joint likelihood of the model (eq. 3.13) is based on data of females only. However, we had also tracking data of males. These data were also included and modelled, but they contributed to the joint likelihood only improving the precision of parameter estimates that are all common in both sexes (*Abadi et al.*, 2010b). We used the Bayesian approach and Markov chain Monte Carlo (MCMC) simulation to estimate the parameters. We therefore based inference on the posterior distribution, which is proportional to the likelihood and the prior distribution. For the initial population sizes we used weakly informative priors (*Kéry and Schaub*, 2011; *Schaub et al.*, 2010). See model code in Appendix D for the exact specification of the priors for all parameters. Some experimentation with different prior choices suggested they had low impact on the parameter estimates, indicating that the inferences were mainly determined by the information contained in the data. For the latent parameter br_{ad} (proportion of breeding females older than 2 years), for which no explicit data were available, we specified three prior distributions to assess whether the integrated model provides an identifiable estimate of the parameter. For the latter, the posterior distribution was almost the same under the different sets of priors (a uniform distribution between 0.5 and 1 [$U(0.5, 1)$], a normal distribution with mean 0.75 and variance 1000 truncated to the values between 0.5 and 1 [$N(0.75, 1000)\mathrm{I}(0.5, 1)$], and a normal distribution with mean 0.5 and variance 0.25 truncated to the values between 0.5 and 1 [$N(0.5, 0.25)\mathrm{I}(0.5, 1)$]). Furthermore a trial integrated analysis with br_{ad} fixed to 0.8 (*Mougeot and Bretagnolle*, 2006) did not emphasize substantial changes in the other parameter estimates.

MCMC simulations were implemented in program WinBUGS (*Lunn et al.*, 2000), that we executed from R (*R Development Core Team*, 2011) with the package R2WinBUGS (*Sturtz et al.*, 2005). We ran three chains for 1,000,000 iterations of which we discarded the first 500,000 iterations as burn-in, and thinned the remaining every 20th sample for parameter estimation. We assessed the convergence of the MCMC simulations to the posterior distribution using the Brooks-Gelman-Rubin criterion, $\hat{R}$, (*Brooks and Gelman*, 1998). The $\hat{R}$ values were ≤ 1.01 for all parameters by the end of the burn-in period. A $\hat{R} < 1.05$ suggested that convergence may be assumed, and our burn-in period and run lengths were adequate (*Spiegelhalter et al.*, 2007). Furthermore, the annual population growth rate (λ_t) was estimated as a derived parameter, calculated as the ratio of the number of females in year t to the number of females in year $t + 1$. The growth rate averaged over the study period was calculated as the geometric mean of all year-specific values (*Schaub et al.*, 2007). Then for each age class z, we derived the survival rate in the absence of illegal poisoning from the

age-independent survival probability and the age-specific proportion of birds which died by poisoning (β_z) as

$$S_{np,z} = 1 - (1 - s)(1 - \beta_z). \tag{3.15}$$

One particularly useful feature of integrated models in the Bayesian framework is that predictions of the population sizes in the future can be made within the MCMC samples, thus fully accounting for all uncertainty in the parameter estimates (*Kéry and Schaub*, 2011). We thus estimated population sizes for three further years (2011-2013) after the last one for which real data were available. A 3-year time interval reflects the typical duration of the decision-making processes related to management actions, and allows estimation of future population sizes while avoiding excessive increases of uncertainty.

We assessed the magnitude of the improvements in the estimates of demographic parameters by comparing the precision (standard error and 95% credible interval) of these estimates obtained from (i) a stand-alone multi-state model (MS) including only radio-tracking data, (ii) an Integrated Population Model including all data sets but those referred to unmarked birds found dead on the local rehabilitation centre (IPM1), and (iii) a full model based on all data sets available (IPM2).

3.2.9 Modelling the effect of poisoning on population growth rate

To assess the demographic consequences of poison related mortality we used a deterministic Leslie matrix population model (as defined in eq. 3.1) with parameter estimates obtained from the full integrated model (IPM2, Table 3.1), to simulate the demography for different levels of poisoning, all other things being equal. We thus used perturbation analysis to compute the sensitivity and elasticity of the population growth rate to different demographic parameters and the proportional decrease in survival due to illegal poisoning. The latter was calculated, for the z-th age class, as $\delta_z = 1 - (s/S_{np,z})$. In its standard form, the sensitivity measures the impact of changes in matrix elements (a_{ij}) on population growth rate (λ):

$$Sens_{ij} = \frac{\partial \lambda}{\partial a_{ij}}. \tag{3.16}$$

Sensitivities can also be applied to the vital rates (low-level parameters; *Caswell*, 2001). This is done by tracking the changes in λ resulting from changes in the vital rates implicit in the matrix elements a_{ij}.

Similarly, elasticity values can also be calculated for vital rates. Standard elasticity considers the proportional change in λ due to a proportional change in a parameter:

$$Elas_{ij} = \frac{\partial \log \lambda}{\partial \log a_{ij}} = \frac{\partial \lambda / \lambda}{\partial a_{ij} / a_{ij}} = \frac{a_{ij}}{\lambda} S_{ij} \tag{3.17}$$

where E_{ij} is the elasticity of the matrix element a_{ij}, and S_{ij} is its sensitivity. In analogy, elasticity values of vital rates can be obtained by multiplying vital rate sensitivity by x/λ, where x is the value of the vital rate under consideration (*Caswell*, 2001). Unlike the values for matrix elements, vital rates sensitivity and elasticity may be negative, but as it is the magnitude of the change that is of interest, rather than its sign, absolute values are quoted

throughout the paper. Note that the elasticities of matrix cells sum to 1, but those for all matrix elements do not (*Caswell*, 2001).

3.3 Results

To verify the precision enhancement yielded by the integration of multiple datasets, we estimated three sets of parameters by integrating the information sequentially. The first set incorporates the individual data only (MS). The second integrates them with the survey of breeding pairs and the number of fledglings (IPM1), while the third adds to the previous sets the information on birds found dead and reported to the local rehabilitation centre by the general public (IPM2). As expected, the additional information resulted in increased parameter precision (Table 3.1). More specifically, precision was most improved for the estimate of the proportion of adult birds dying because of poison (β_{2my}, -19% in the $95\%CRI$) and for the reporting rates of birds found dead by poisoning without a functioning radio-tag (d_1, -40%). The difference between the mean estimates obtained from the full integrated model (IPM2) and those from IPM1 and MS was larger for the probability of encounter of dead animal without a functioning radio (d_1, $+42\%$; d_2, $+103\%$). The temporal variability of fecundity was slightly different from zero ($\hat{\sigma}_b^2 = 0.110$, $95\%CRI$: 0.004, 0.328), but the pattern of average fecundity, b, showed no obvious temporal trend.

In the transition matrix, we specified the proportion of breeding females older than 2 years, br_{ad}, as a latent parameter, i.e. a parameter for which information was not available. The adult breeding proportion was difficult to estimate because birds loose the radio-transmitters when about four years old. This hidden, or latent, parameter was estimated on the basis of the remaining integrated information to be 0.63 and, despite a relatively large level of uncertainty, sensitivity analyses indicated that it was estimable (see Parameter estimation and model implementation subsection in Methods). The age-independent survival rate (s) was 0.808 ($95\%CRI$: 0.772, 0.841). The probability of dying because of poisoning was age-dependent, being higher for birds older than one year ($\beta_{1y} = 0.764$, $95\%CRI$: 0.508, 0.942; $\beta_{2my} = 0.764$, $95\%CRI$: 0.464, 0.943) than for juveniles ($\beta_{juv} = 0.428$, $95\%CRI$: 0.274, 0.590). Assuming that human-related mortality was additive, survival probabilities in the absence of illegal poisoning can be derived from s and each age-specific β (eq. 3.15). Simulated, poison-free survival was higher in adult birds (≥ 1 year old; $S_{np,1y} = 0.955$, $95\%CRI$: 0.904, 0.989; $S_{np,2my} = 0.955$, $95\%CRI$: 0.895, 0.989) than in juveniles ($S_{np,juv} = 0.890$, $95\%CRI$: 0.852, 0.925). From these estimates we calculated that illegal poisoning reduced survival probability by 15% in adults (δ_{1y} and δ_{2my}) and 9% in juveniles (δ_{juv}).

The smoothed estimates of the total annual number of breeding pairs, as well as the observed population sizes, showed a positive trend with a population growth higher than one ($\lambda = 1.136$, $95\%CRI$: 1.063, 1.222) and roughly constant throughout the study (Fig. 3.4). Predictions of population size over the next three years (2011-2013) suggest a slow increase in the number of breeding pairs, although the $95\%CRI$ expands over time reflecting increasing uncertainty (Fig. 3.4). The combination of mean estimates of demographic parameters (IPM2, Table 3.1) in the deterministic model suggested a slightly positive population growth

Table 3.1: Estimated demographic parameters of the Red kite population of the island of Mallorca (Spain). We show the posterior mean and 95% credible interval (95%CRI, lower and upper limit) of the estimates, obtained by a full integrated model (IPM2), an integrated model without considering data of unmarked birds found dead (IPM1), and a multi-state model with only radio-tracking data. For parameter notation see Methods.

	IPM2			IPM1			MS		
Parameter	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper
s	0.808	0.772	0.841	0.814	0.776	0.849	0.821	0.782	0.857
β_{juv}	0.428	0.274	0.590	0.429	0.273	0.593	0.428	0.270	0.593
β_{1y}	0.764	0.508	0.942	0.755	0.491	0.942	0.758	0.493	0.942
β_{2my}	0.764	0.464	0.943	0.688	0.356	0.923	0.690	0.365	0.924
α_1	0.862	0.820	0.900	0.862	0.820	0.900	0.862	0.820	0.900
α_2	0.333	0.191	0.493	0.333	0.191	0.493	0.334	0.191	0.494
α_3	0.045	0.002	0.097	0.045	0.002	0.097	0.045	0.002	0.097
p	0.990	0.972	0.999	0.990	0.972	0.999	0.990	0.973	0.999
c	0.328	0.255	0.405	0.331	0.253	0.413	0.315	0.245	0.391
d_1	0.075	0.036	0.099	0.053	0.009	0.097	0.054	0.009	0.097
d_2	0.461	0.168	0.899	0.227	0.022	0.745	0.250	0.025	0.793
br_{ad}	0.631	0.504	0.898	0.676	0.507	0.953	–	–	–
$S_{np,juv}$	0.890	0.852	0.925	0.894	0.856	0.928	–	–	–
$S_{np,1y}$	0.955	0.904	0.989	0.954	0.904	0.989	–	–	–
$S_{np,2my}$	0.955	0.895	0.989	0.942	0.877	0.986	–	–	–
b	1.825	1.529	2.153	1.778	1.484	2.103	–	–	–
$\hat{\sigma}_b^2$	0.110	0.004	0.328	0.111	0.004	0.330	–	–	–
λ	1.136	1.063	1.222	1.142	1.067	1.230	–	–	–

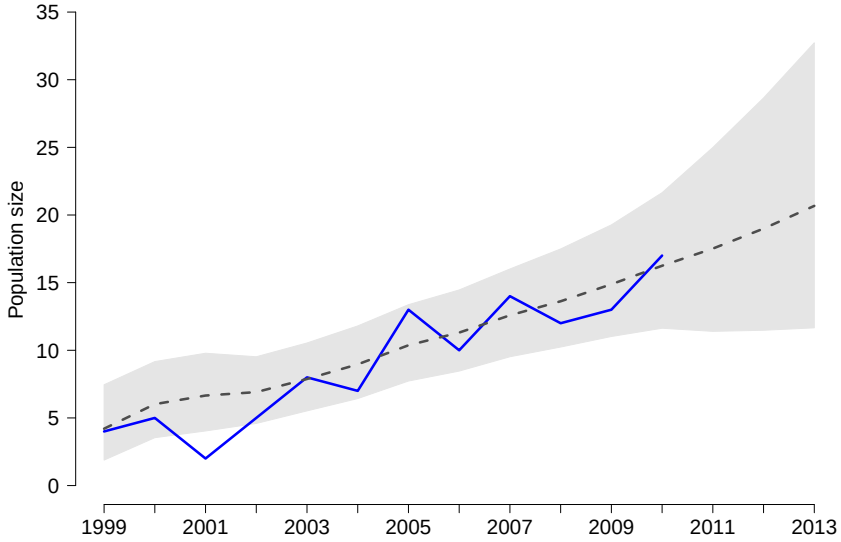

Figure 3.4: Observed and estimated sizes of the Red kite population of Mallorca (Spain), with a future projection of the number of breeding pairs. The solid line represents the surveyed population size, the dashed line the predicted spring population sizes along with their 95%CRI (grey shading).

($\lambda_1 = 1.082$) which matches well the slow increase in counts of breeding pairs. The population is thus projected to increase by 8.2% per year. When we explored how this rate was affected by an increase of survival probability due to lowered poisoning (δ), assuming that such mortality was additive, we found that the population would decline ($\lambda < 1$) if survival will be reduced by 45%, 49%, and 25% for juveniles, 1 year old, and 2 or more years old individuals, respectively (that would lead to survival probabilities of 0.49, 0.49, 0.72; Fig. 3.5). Sensitivity of population growth to the different demographic rates was higher for adult survival in the absence of illegal poisoning ($S_{np,2my}$) and the related proportional decrease due to this mortality cause (δ_{2my}; Table 3.2). The deterministic model indicated that a further reduction in survival probability, with $\delta_{juv} > 0.54$, $\delta_{1y} > 0.57$, $\delta_{2my} > 0.29$ would not be compensated even by the maximum fecundity recorded for the species (2.2 young fledged per breeding pair *Mougeot and Bretagnolle*, 2006; Fig. 3.6).

Finally, we explored an average age- and time-independent relationship between population growth rate, fecundity, and the proportional reduction in survival probability due to illegal poisoning. Results indicated that a decline in survival is proportionally more difficult to be compensated by an increase in per capita fecundity (Fig. 3.7).

3.4 Discussion

3.4.1 An analytical framework to help a 'crisis discipline'

Conservation biology is often referred to as a crisis discipline because anthropogenic alterations and the rate of population extinctions 'do not give the luxury of time' (*Soulé*, 1985;

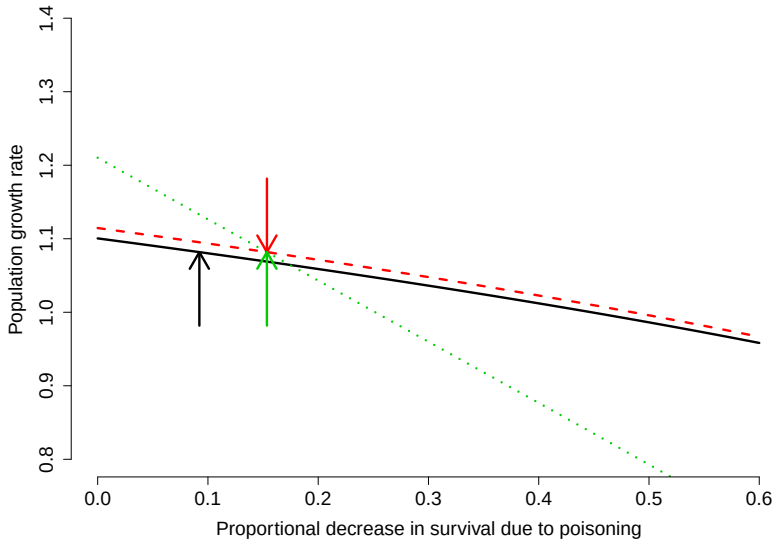

Figure 3.5: Changes of population growth rate in relation to changes in the proportional decrease of age-specific survival probability. The black solid line represents the relationship with proportional reduction in juvenile survival (δ_{juv}), the red dashed line refers to δ_{1y}, and the green dotted line refers to δ_{2my}. Current age-specific values of δ are indicated by the arrows with the same colour of the curve to which they refer.

Table 3.2: Sensitivity and elasticity of population growth rate of the Red kite population of the island of Mallorca (Spain). For parameter notation see Methods.

Parameter	Sensitivity	Elasticity
δ_{juv}	0.207	0.018
δ_{1y}	0.222	0.036
δ_{2my}	0.834	0.118
br_{ad}	0.283	0.165
$S_{np,juv}$	0.211	0.174
$S_{np,1y}$	0.197	0.174
$S_{np,2my}$	0.739	0.652
b	0.103	0.174

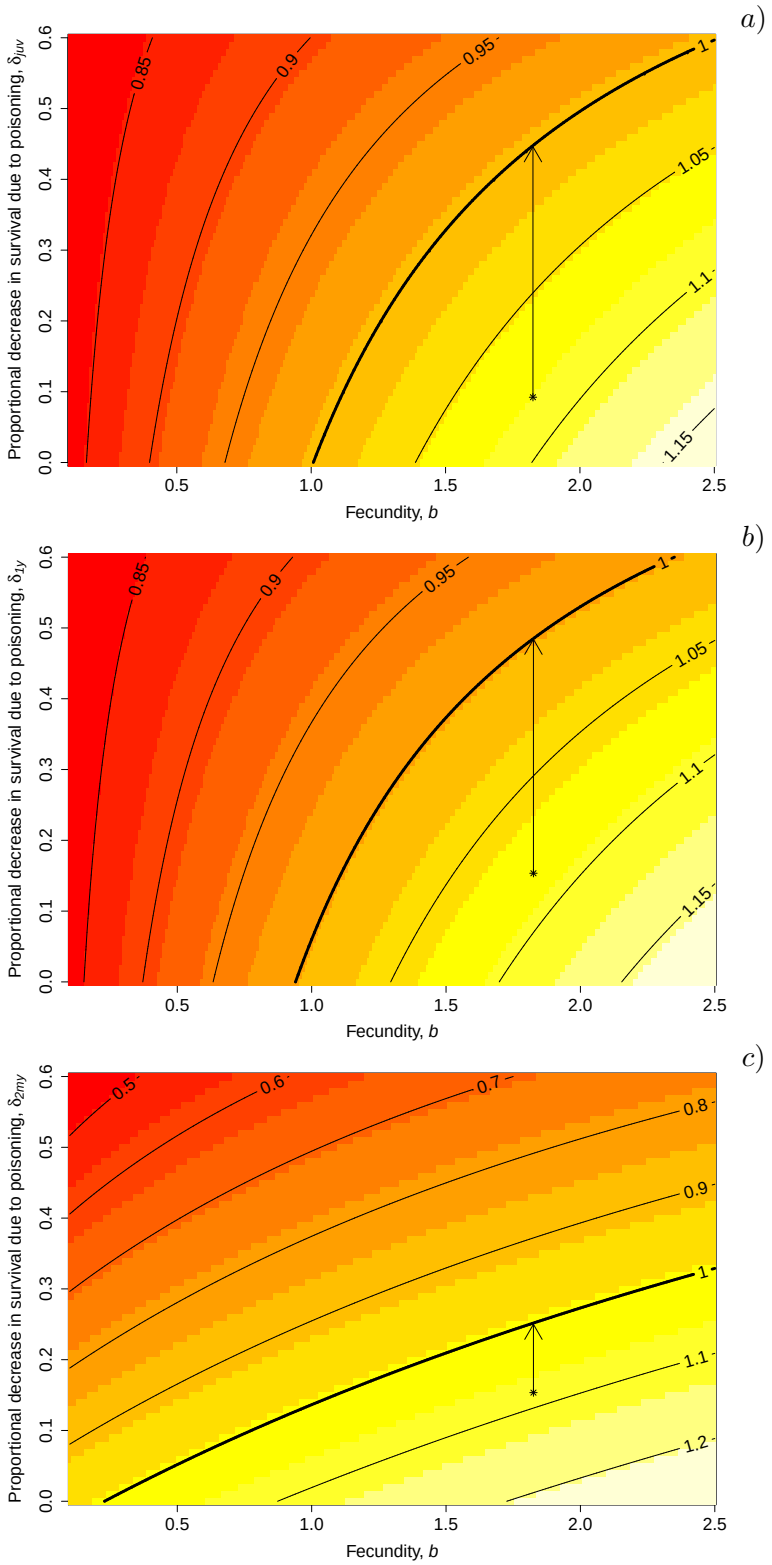

Figure 3.6: Difference between the proportional changes in age-specific survival probability due to illegal poisoning, fecundity, and population growth rate. The bold line represents population stability. The asterisks refer to the current parameter estimates, while arrows represent a theoretical increase in δ up to the level of population stability. a) juveniles. b) 1 year old. c) 2 or more year old red kites.

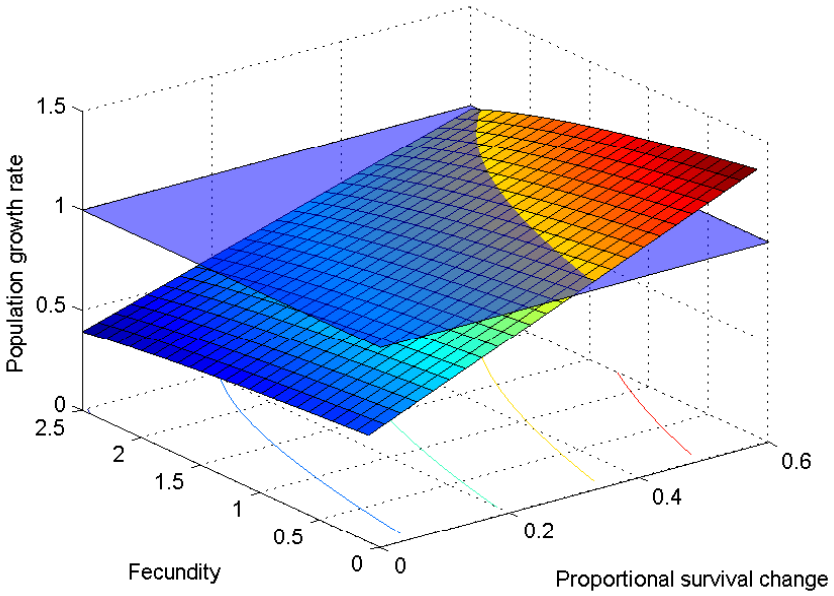

Figure 3.7: Age-independent relationship between population growth rate, fecundity, and proportional change in survival probability due to illegal poisoning. The blue horizontal plane represents population stability.

Pullin et al., 2004). For this reason, diagnosis of population threats is often based on small sample size and evidence may come from scattered sources. As a consequence, marked uncertainty accompanies inferences on population trajectories or on the relative importance of different mortality causes. Here, we showed a statistical framework that increases estimate precision by analysing multiple data simultaneously. In particular, we integrated radio-tracking data, nest and fledglings counts and time series of birds found dead by the general public into a single analysis to obtain consensus estimates and explore the demographic impact of poisoning on an endangered raptor. The joint analysis delivered more precise estimates, which incorporated all available information and led to a more accurate assessment of the importance of specific causes of mortality and short-term population forecasting. Moreover, the integrated analysis allowed the estimation of the breeding and non-breeding sector of the population, an elusive parameter that can be measured only under special circumstances and usually through a pronounced survey (e.g. *Kenward et al.*, 2000). Finally, it is interesting to note that the raw frequency of animals reported to have died from a specific cause is an information frequently available from local authorities but generally considered as too coarse to yield any meaningful estimate of mortality impact. In our case, even if this information was scarce, it contributed to improve parameter estimates.

Information from multiples sources can be joined into a single analysis as long as data are independent, at least partially (*Besbeas et al.*, 2002; *McCrea et al.*, 2010). The assumption of independence causes a trade-off between statistical needs and ecological realism. Indeed, joining independent data may help to meet model assumptions, but it increases the variance

components, for example because separate information collected at different spatial scales may not be fully comparable. In our case, data came from an insular population, which minimises the noise related to immigration, emigration, or other sources of variance potentially included in the analyses when data are drawn from different populations. On the other hand, radio-tracked birds were only a portion of the total population and we considered only unmarked birds in the time series of kites reported dead by the public, in order to employ as independent datasets as possible.

3.4.2 Illegal poison and population trajectories

The aim of this work was to measure the demographic consequences of illegal poisoning for a threatened raptor. The proportion of dead kites which were killed by poisoning in our study was considerable as it varied from 0.43 to 0.76 for yearlings and older birds, respectively. *Smart et al.* (2010) reported a similar figure for yearlings in the UK, but a smaller value for older birds. Despite an increase of 15% in mortality probability due to poisoning, our population seems to be slowly growing and is expected to continue to increase in the immediate future. We assumed a constant sampling variance during the study period and, although unlikely, we think it is a reasonable assumption given the small number of breeding pairs each year. The positive trend is likely to be maintained by a very high productivity (one of the highest registered for the species, *Mougeot and Bretagnolle*, 2006; *Smart et al.*, 2010; *Sergio et al.*, 2005), which seemed to compensate for small declines in adult survival (Fig. 3.6). The high productivity was probably promoted by the local presence of a rubbish dump (closed in 2009) and by supplementary feeding stations established by local authorities in recent years. Unfortunately, no specific data are currently available to measure the exact role of supplementary feeding in our population. The effectiveness of artificial feeding sites in improving productivity and sustaining population viability is still debated (*González et al.*, 2006; *Carrete et al.*, 2006; *Robb et al.*, 2008; *Oro et al.*, 2008), for example *Oro et al.* (2008) reported that feeding stations suppressed the dispersal of immature birds, which increased the competition among individuals and resulted in a negative impact on population growth. In our case, this mechanism seems unlikely because the population is not habitat-limited and we only know few cases of pre-breeding dispersal outside the study area. If supplementary food promoted population growth this could generate a paradoxical scenario where anthropogenic mortality is compensated by food supplemented by humans. Under this scenario of double societal costs, increasing poison control might help to lower the public cost of maintaining supplementary feeding stations. Finally, even though the population is slowly increasing, the number of breeding pairs is still small and we estimated that the population growth would be 20% higher without the additional mortality due to poisoning. Clearly, our analysis confirmed the extreme sensitivity of this extinction-threatened species to poisoning, as suggested by previous authors (e.g. *Smart et al.*, 2010; *Whitfield et al.*, 2003; *Sergio et al.*, 2005).

3.4.3 Facultative, occasional and obligate scavengers and the illegal use of poison

As a facultative scavenger, the Red Kite is frequently attracted to small poisoned baits that are illegally used to control mammalian predators of game species. The demographic impacts outlined for our population are likely to apply to other occasionally scavenging raptors, such as Spanish imperial eagles *Aquila adalberti*, or to obligate scavengers such as bearded vultures *Gypaetus barbatus* or Egyptian vultures *Neophron percnopterus* (e.g. *González et al.*, 2007; *Ortega et al.*, 2009; *Oro et al.*, 2008; *Hernández and Margalida*, 2009)). In our case, perturbation analyses indicated that population growth was more sensitive to adult survival in the absence of poisoning ($S_{np,2my}$) and to the related proportional decline due to this mortality cause (δ_{2my}). A similar outcome is expected for obligate scavengers, such as vultures, whose survival in the absence of poisoning is typically high (*Sarrazin and Legendre*, 2000; *Oro et al.*, 2008; *Martínez-Abraín et al.*, 2012). In these species, characterized by extended longevity and low natality, the proportional decline in survival caused by poisoning is likely to have a stronger impact on population growth because their typically low fecundity cannot function as a buffering trait. Therefore, the scenario outlined for our population could be considered as a simulation of minimum impact, when compared to many larger, longer-lived and more obligate scavengers.

3.4.4 Conservation measures

Our results suggest that illegal poisoning was the most important cause of mortality of Red Kites in the Island of Mallorca and was able to suppress population growth by 20%. Nevertheless, a high fecundity rate seemed helpful to counter-sustain population viability. Unfortunately, it is not possible to evaluate to what degree this could be ascribed to the beneficial effect of artificial feeding sites, and further research is needed to understand their role as a population management tool. As already stressed by other authors, future management activities should concentrate on the eradication of illegal poisoning (*Hernández and Margalida*, 2009), and on devising techniques of predator control compatible with the conservation of the Red Kite and other vulnerable species (*González et al.*, 2007). Improving the monitoring, surveillance and the post-mortem lab-analyses of illegal poisoning episodes could further help to build the necessary pressure to reduce the cases of malicious or negligent actions (*Guitart et al.*, 2010). In this way, an accurate forensic diagnosis has a key role in the investigation of poisoning (*García-Fernández et al.*, 2006) and could be supported by techniques that facilitate rapid mortality detection, such as radio tagging and other remote sensing devices.

3.5 Acknowledgements

We thank A. Margalida for helpful review of the manuscript. We also thank the many people that voluntarily helped in collecting the data. We are grateful to the Natura Park foundation, the Grup d'Ornitologia Balear i defensa de la naturalesa, the Conselleria d'Agricultura Medi Ambient i Territori of the Government of the Balearic Islands and the Cuerpo Forestal

del Estado for the logistic support. This project was partially funded by the Spanish Government (project BFU2009-09359). We also thank the Conselleria d'Agricultura Medi Ambient i Territori of the Government of the Balearic Islands, the Fundació 'La Caixa', the Fundación Biodiversidad, the Fundación Caja Madrid and Gesa-Endesa for financing data collection over the years.

Chapter 4

Bayesian Hierarchical models : is all that glitters gold?

Simone Tenan, Iris Hendriks, Giacomo Tavecchia (*to be submitted*).

Abstract

Following the advent of MCMC engines Bayesian hierarchical models are becoming increasingly common in modelling ecological data. However, the great enthusiasm for model fitting has not yet encompassed the selection of competing models, despite its fundamental role in the inferential process. We present two possible procedures for model and variable selection in a Bayesian framework, the product space method and the Gibbs variable selection, with emphasis on the practical implementation into a general purpose software in BUGS language. To clarify the related theoretical aspects and practical guidelines, we explain the methods through applied examples on the comparison of non-nested models for positive continuous response variable and on the selection of variables in a logistic regression problem. Despite the relatively wide range of available techniques and the difficulties related to the maximization of sampling efficiency, for their conceptual simplicity and ease of implementation the proposed methods can represent potentially useful tools for ecologists and conservation biologists that want to close the loop of a Bayesian analysis.

Key-words: BUGS, Hierarchical modelling, Markov Chain Monte Carlo, Model selection, Variable selection.

4.1 Introduction

With the recent advance in computational statistic and the refinement of ecological questions there is an increasing interest in a hierarchical approach to ecological modelling (e.g. *Clark*,

2005; *Royle and Dorazio*, 2008; *Halstead et al.*, 2012; *Kéry and Schaub*, 2011). This approach puts emphasis on the distinct components, i.e. processes of ecological systems, leading to the hierarchical models that explicitly incorporate variances from the multiple levels of the information (*Royle and Dorazio*, 2008; *Gelman et al.*, 2003). The development of Markov Chain Monte Carlo (MCMC) framework (*Robert and Casella*, 2004) and the advent of MCMC engines in the BUGS language (*Lunn et al.*, 2000; *Plummer*, 2003) have greatly helped to overcome the computational problems due to the complex structure of the model and contributed to the fast-growing use of Bayesian hierarchical models. The Bayesian approach offers several advantages in model implementation and fitting (*Schaub and Kéry*, 2012) and have generated a high enthusiasm around Bayesian hierarchical models in ecology and conservation biology (e.g. *Halstead et al.*, 2012; *Kéry and Schaub*, 2011). Surprisingly this enthusiasm has been mainly confined to model fitting (*Schaub and Kéry*, 2012; *Halstead et al.*, 2012) and only marginally considering their validation and selection. In wildlife and ecological applications model validation and model selection is an important, if not fundamental, part of the inferential process (*Link and Barker*, 2006; *Burnham and Anderson*, 2002). However, model selection procedures for Bayesian hierarchical models have not generated the same enthusiasm as model building or fitting. Rather the opposite. Widely used and easy to implement information criteria like DIC (*Spiegelhalter et al.*, 2002) cannot be safely used to compare hierarchical models (*Millar*, 2009; *Kéry and Schaub*, 2011). At present, and possibly unknown to many enthusiastic ecologists, model selection in hierarchical models is complex, computationally challenging and no consensus has emerged in the literature on a single approach (*Link and Barker*, 2006). The problem is not new to statisticians and there are examples of possible alternatives for Bayesian model choice (e.g. *Sisson*, 2005; *Congdon*, 2006; *Lunn et al.*, 2009a; *Ando*, 2010; *Ntzoufras*, 2002), but the ecological applications are still few and computationally challenging to implement (*King*, 2009; *Link and Barker*, 2006; *O'Hara and Sillanpää*, 2009; *Royle and Dorazio*, 2008; *Spiegelhalter et al.*, 1996). An exhaustive comparison of the performance of different methods implementable in BUGS is provided by *O'Hara and Sillanpää* (2009).

Here, conscious of the practical needs of many ecologists, we provide an example to address the problems of both Bayesian variable and model selection using two Gibbs sampler based strategies, taken from the wider survey of methods already implemented using a BUGS software (*O'Hara and Sillanpää*, 2009; Table 4.1). The first method we shall discuss is the so called product space method (*Carlin and Chib*, 1995), which is a model determination strategy. From the extension of the general idea of *Carlin and Chib* (1995) derived the second method we will explain, called Gibbs variable selection (GVS), introduced by *Dellaportas et al.* (2000). The two methods are based on the so-called trans-dimensional Markov chains sampling frameworks (see *Sisson*, 2005 for a general review) that permit the construction of Markov chains which simultaneously traverse both parameter and model space. After a theoretical description we illustrated the product space method with the comparison of two non-nested hierarchical models of different distributional form, and the GVS approach with an example of logistic regression with random effects.

Table 4.1: Approaches to variable and model selection implemented in BUGS language (following *O'Hara and Sillanpää*, 2009) with the related main reference.

Indicator model selection	
a) Gibbs variable selection (GVS)	*Dellaportas et al.* (2000)
b) Unconditional Priors for variable Selection	*Kuo and Mallick* (1998)
Stochastic search variable selection (SSVS)	*George and McCulloch* (1993)
Adaptative shrikage	
a) Jeffreys' prior	*Hobert and Casella* (1996)
b) Laplacian shrinkage	*Park and Casella* (2008)
Model space approach	
a) Reversible jump MCMC	*Green* (1995)
b) Product space	*Carlin and Chib* (1995)
c) Composite model space (CMS)	*Godsill* (2001)

4.2 Theoretical background

For the sake of clarity, for each method we adopted the general notation of the related main reference, in order to ease an in-depth examination into the statistical literature. We just made probability notation uniform throughout the paper, using $p(\cdot|\cdot)$ for conditional probability density and $p(\cdot)$ for marginal distribution, using the terms 'distribution' and 'density' interchangeably, as in *Gelman et al.* (2003).

4.2.1 The core of the matter

Bayesian multimodel inference can be viewed as an extension of the Bayesian inference we already know: in this case models are the unknown quantities and we just want to make inference about them on the basis of their posterior distribution, given the data (*Link and Barker*, 2009). In a set of models $\mathcal{M} = \{M_1, M_2, \ldots, M_K\}$, with $j = 1, \ldots, K$, inference on model uncertainty can be done by obtaining observations from the joint posterior distribution of a model and the related parameters $(M_j, \boldsymbol{\beta}_{M_j})$ and consequently estimate $p(M_j|\boldsymbol{y})$ and $p(\boldsymbol{\beta}_{M_j}|M_j, \boldsymbol{y})$. In principle the process of calculating posterior model probabilities is quite simple: for M_j we firstly choose prior model probability $p(M_j)$ independently of the data, and priors for model parameters $p(\boldsymbol{\beta}_{M_j}|M_j)$. Then, the observed data contribute to the posterior model probabilities through $p(\boldsymbol{y}|M_j)$ as follows

$$p(M_j|\boldsymbol{y}) = \frac{p(M_j)\, p(\boldsymbol{y}|M_j)}{\sum\limits_{M_j \in \mathcal{M}} p(M_j)\, p(\boldsymbol{y}|M_j)}, \quad M_j \in \mathcal{M}.$$

The marginal likelihood, $p(\boldsymbol{y}|M_j)$, is calculated using the likelihood of the data under the model, and the conditional prior distribution of the model parameters as

$$p(\boldsymbol{y}|M_j) = \int p(\boldsymbol{y}|M_j, \boldsymbol{\beta}_{M_j})\, p(\boldsymbol{\beta}_{M_j}|M_j)\, d\boldsymbol{\beta}_{M_j}.$$

In the simple case of two competing models (M_1, M_2) their relative probability is

$$\frac{p(M_1|\boldsymbol{y})}{p(M_2|\boldsymbol{y})} = \frac{p(M_1)}{p(M_2)} \frac{\int p(\boldsymbol{y}|M_1, \boldsymbol{\beta}_{M_1})\, p(\boldsymbol{\beta}_{M_1}|M_1)\, d\boldsymbol{\beta}_{M_1}}{\int p(\boldsymbol{y}|M_2, \boldsymbol{\beta}_{M_2})\, p(\boldsymbol{\beta}_{M_2}|M_2)\, d\boldsymbol{\beta}_{M_2}} \tag{4.1}$$

where the second ratio on the right hand side is the Bayes factor (B_{12}), that is the ratio of the two marginal likelihoods (*Kass and Raftery*, 1995). We can therefore read eq. 4.1 as "posterior model odds = B_{12} × prior model odds". Thus, the Bayes factor can be interpreted as the change in the model odds resulting from observing the data (*Lodewyckx et al.*, 2011), providing a mechanism for converting prior model probabilities to posterior model probabilities (*Link and Barker*, 2006). We can naturally base the choice of the best supported model or producing a model-averaged prediction just on posterior probabilities, bearing in mind that the probability of a model is always conditional on the model set, and can be interpreted as a relative degree of support within that set.

4.2.2 The product space method

Currently one of the most supported and effective procedure for model selection of Bayesian hierarchical models is the reversible jump MCMC, a flexible technique introduced by *Green* (1995). Since few years a WinBUGS application is available for the reversible jump MCMC (*Lunn et al.*, 2009a) but traditionally this technique cannot be implemented in BUGS language and requires complex analytical derivation of a mapping function and a Jacobian matrix to maximize sampling efficiency. As the reversible jump MCMC, the product space method that we illustrate here relies on combining the models to be compared within an hierarchical "supermodel". To describe how this latter method works, we mainly followed *Lodewyckx et al.* (2011) and the original paper by *Carlin and Chib* (1995). Model combination is obtained through the use of a single model index variable M that controls which model generated the observed data vector $\boldsymbol{y}$. In regard to the Bayes factor the prior of the model index M corresponds to the prior model odds, while the posterior of M corresponds to the posterior model odds, in this case estimated through a Gibbs sampler. In the problem of choosing between K models for the observed data vector $\boldsymbol{y}$, where at each model corresponds a distinct parameter vector $\boldsymbol{\theta}_j$, with $j = 1, \ldots, K$, our main interest lies in the posterior probabilities of each of the K models, that is $p(M = j|\boldsymbol{y})$. For the product space method, we define a (mixture) model with a composite parameter vector $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_K\}$ which takes any value from the Cartesian product of the models' parameter spaces, that is $\mathcal{M} \times \prod_{M \in \mathcal{M}} \Theta_M$, where Θ_M is the set of all possible values for the coefficient of model M. For model j, part of the mixture model, the joint distribution is

$$p(\boldsymbol{y}, \boldsymbol{\theta}|M = j) = p(\boldsymbol{y}|\boldsymbol{\theta}, M = j)p(\boldsymbol{\theta}|M = j). \tag{4.2}$$

Eq. 4.2 can be rewritten as

$$p(\boldsymbol{y}, \boldsymbol{\theta}|M = j) = p(\boldsymbol{y}|\boldsymbol{\theta}_j, M = j)p(\boldsymbol{\theta}_j|M = j)p(\boldsymbol{\theta}_{i \neq j}|M = j) \tag{4.3}$$

since $\boldsymbol{\theta}_{i \neq j}$ is not relevant under $M = j$ and independent of $\boldsymbol{\theta}_j$ (the $M \neq j$ part is specified similarly). Note that in eq. 4.3 the terms $p(\boldsymbol{y}|\boldsymbol{\theta}_j, M = j)$ and $p(\boldsymbol{\theta}_j|M = j)$ are respectively

the likelihood and the prior corresponding to model j. Given $M = j$, the model defined in eq. 4.3, even with added parameters $\boldsymbol{\theta}_{i\neq j}$, becomes essentially model j with respect to its marginal likelihood, and the same hold for $M \neq j$. Independent 'pseudoprior' or linking densities $p(\boldsymbol{\theta}_{i\neq j}|M = j)$ (or similarly $p(\boldsymbol{\theta}_j|M \neq j)$) have to be specified in order to define the mixture model with the parameter vector $\boldsymbol{\theta}$. Pseudopriors do not influence the marginal posterior distribution, but their accurate definition is important for the mobility of the Gibbs sampler.

Carlin and Chib (1995) proposed using a Gibbs sampler to generate from the posterior distribution $p(\boldsymbol{\theta}_j, M = j|\boldsymbol{y})$. To implement the Gibbs sampler we derive the full posterior conditional distribution, of both each $\boldsymbol{\theta}_j$ and M. For sampling model parameters the distribution is given by

$$p(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{i\neq j}, M, \boldsymbol{y}) \propto \begin{cases} p(\boldsymbol{y}|\boldsymbol{\theta}_j, M = j)p(\boldsymbol{\theta}_j|M = j) & \text{if } M = j \\ p(\boldsymbol{\theta}_j|M \neq j) & \text{if } M \neq j, \end{cases} \tag{4.4}$$

which implies that a sample of $\boldsymbol{\theta}_j$ is generated from the model j full conditional distribution when the model index $M = j$, while when $M \neq j$ the sample is generated from the related pseudoprior. We then refer to another conditional distribution for sampling the model index M:

$$p(M = j|\boldsymbol{\theta}, \boldsymbol{y}) = \frac{A_j}{\sum_{k=1}^{K} A_k} \tag{4.5}$$

where

$$A_j = p(\boldsymbol{y}|\boldsymbol{\theta}_j, M = j) \prod_{i=1}^{K} p(\boldsymbol{\theta}_i|M = j)\, p(M = j) \tag{4.6}$$

and $p(M = j)$ is the prior model probability. Finally, the posterior model probability for model j is estimated as

$$\hat{p}(M = j|\boldsymbol{y}) = \frac{\text{number of occurence of } M = j}{\text{total number of iterations.}} \tag{4.7}$$

4.2.3 Gibbs variable selection

Assume that we want to explain an outcome y_i for individual i $(i = 1, \ldots, N)$ using p covariates. Clearly, these variables can be continuous or discrete, and are candidate for the inclusion in the linear predictor. If we assume that model uncertainty is restricted to variable selection, each model $M \in \mathcal{M}$ can be represented by a vector of binary indicators $\boldsymbol{\gamma} \in \{0, 1\}^p$. This vector indicates which of the possible sets of covariates are present in the model. For a generalized linear model, for which we assume distribution, link function and variance function known, the linear predictor can be written as

$$\boldsymbol{\eta} = \sum_{j=0}^{p} \gamma_j \boldsymbol{X}_j \boldsymbol{\beta}_j \tag{4.8}$$

where $\boldsymbol{X}_j$ and $\boldsymbol{\beta}_j$ are the design matrix and the parameter vector for the jth term. We now concentrate on the estimation of the posterior distribution of $\boldsymbol{\gamma}$.

The likelihood of each model is denoted by $p(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\gamma})$ and the prior by the hierarchical structure $p(\boldsymbol{\beta},\boldsymbol{\gamma}) = p(\boldsymbol{\beta}|\boldsymbol{\gamma})p(\boldsymbol{\gamma})$. The latter is thus composed by the prior for the parameter vector $\boldsymbol{\beta}$ conditional on the model structure $\boldsymbol{\gamma}$ and the prior of the corresponding model $p(\boldsymbol{\gamma})$. We can consider the partition of $\boldsymbol{\beta}$ into two vectors $(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}})$ corresponding to the components of $\boldsymbol{\beta}$ included ($\gamma_i = 1$) or excluded ($\gamma_i = 0$) from the model. Hence the vector $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ corresponds to the active parameters of the model (i.e. $\boldsymbol{\beta}_M$), while $\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}}$ corresponds to the remaining parameters, which are not included in the model defined by $\boldsymbol{\gamma}$ (*Ntzoufras*, 2009). Under this partition we can rewrite the prior $p(\boldsymbol{\beta},\boldsymbol{\gamma})$ as the product of the model parameter prior $p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\boldsymbol{\gamma})$, the model parameter pseudoprior $p(\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}}|\boldsymbol{\beta}_{\boldsymbol{\gamma}},\boldsymbol{\gamma})$, and the prior of the corresponding model $p(\boldsymbol{\gamma})$. The parameter vector $\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}}$ does not gain any information from the data and does not influence the actual posterior of the parameters of each model $p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\boldsymbol{\gamma},\boldsymbol{y})$. However, we want to stress again that pseudopriors influences the performance of the MCMC algorithm and can be chosen as linking densities to increase the efficiency of the Gibbs sampler.

The full conditional posterior distributions for the (included and excluded) model parameters are given by

$$\begin{aligned} p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}},\boldsymbol{\gamma},\boldsymbol{y}) &\propto p(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\gamma})\, p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\boldsymbol{\gamma})\, p(\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}}|\boldsymbol{\beta}_{\boldsymbol{\gamma}},\boldsymbol{\gamma}) \\ p(\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}}|\boldsymbol{\beta}_{\boldsymbol{\gamma}},\boldsymbol{\gamma},\boldsymbol{y}) &\propto p(\boldsymbol{\beta}_{\setminus\boldsymbol{\gamma}}|\boldsymbol{\beta}_{\boldsymbol{\gamma}},\boldsymbol{\gamma}) \end{aligned}$$

and for the variable indicator γ_j by

$$\gamma_j|\boldsymbol{\beta},\boldsymbol{\gamma}_{\setminus j},\boldsymbol{y} \propto Bernoulli\left(\frac{O_j}{1+O_j}\right)$$

with

$$O_j = \frac{p(\boldsymbol{y}|\boldsymbol{\beta},\gamma_j=1,\boldsymbol{\gamma}_{\setminus j})}{p(\boldsymbol{y}|\boldsymbol{\beta},\gamma_j=0,\boldsymbol{\gamma}_{\setminus j})}\,\frac{p(\boldsymbol{\beta}|\gamma_j=1,\boldsymbol{\gamma}_{\setminus j})}{p(\boldsymbol{\beta}|\gamma_j=0,\boldsymbol{\gamma}_{\setminus j})}\,\frac{p(\gamma_j=1,\boldsymbol{\gamma}_{\setminus j})}{p(\gamma_j=0,\boldsymbol{\gamma}_{\setminus j})}.$$

We may avoid the dependence of the full conditional posterior distribution of the active parameters $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ on the pseudoprior by assuming that the prior for $\boldsymbol{\beta}_j$ depends only on $\boldsymbol{\gamma}_j$. This restrictive assumption may be realistic when priors are intended to be non-informative and most appropriate when the columns of different $\boldsymbol{X}_j$ in eqn. 4.8 are orthogonal. This approach is given by

$$p(\boldsymbol{\beta}_j|\boldsymbol{\gamma}_j) = (1-\gamma_j)\, p(\boldsymbol{\beta}_j|\gamma_j=0) + \gamma_j\, p(\boldsymbol{\beta}_j|\gamma_j=1) \tag{4.9}$$

and results in the following full conditional posterior

$$p(\boldsymbol{\beta}_j|\boldsymbol{\gamma},\boldsymbol{\beta}_{\setminus j},\boldsymbol{y}) \propto \begin{cases} p(\boldsymbol{y}|\boldsymbol{\gamma},\boldsymbol{\beta})\, p(\boldsymbol{\beta}_j|\gamma_j=1) & \text{if } \gamma_j = 1 \\ p(\boldsymbol{\beta}_j|\gamma_j=0) & \text{if } \gamma_j = 0. \end{cases}$$

For prior and pseudoprior model parameters we can use a mixture of Normal distribution as follows:

$$p(\boldsymbol{\beta}_j|\gamma_j=1) \equiv N(0,\Sigma_j) \qquad \text{and} \qquad p(\boldsymbol{\beta}_j|\gamma_j=0) \equiv N(\bar{\mu}_j, S_j).$$

The choice of pseudopriors can be done in the same way as explained for the product space method. Hyperparameters $\bar{\mu}_j$ and S_j may be obtained from a pilot run of the full model.

For the variable indicator γ_j we can set a common prior as $p(\gamma_j) \sim Bernoulli(0.5)$ when there are no restrictions on the model space, whereas in other cases like hierarchical or graphical log-linear models $p(\gamma_j|\boldsymbol{\gamma}_{\setminus j})$ must depend on $\gamma_{\setminus j}$ (for further details see *Chipman*, 1996 and *Ntzoufras*, 2002).

4.3 Examples with a real dataset

We present two applications by using the data from *Hendriks et al.* (2012), where capture-mark-recapture and individual body size data on a large bivalve (Noble pen shell, *Pinna nobilis*) were used. We considered a subset of the original data, gathered on five sites along the coast of the islands Majorca and Cabrera (Balearic Islands, Spain). In each site, transects of 30 m length were randomly positioned underwater and each transect was randomly assigned to a team of two divers. Capture-recapture data were collected along this line. Each diver marked all noble pen shells found along a side of the transect line using a metal peg with a unique alphanumeric code. Once at the end of the transect line, divers switched side and searched for already marked noble pen shells marked by the previous diver ('re-capture'). The shell width of each marked individual was measured. On subsequent surveys, diver teams changed randomly to minimize a possible 'diver' effect on recapture probability. We considered data for 234 marked individuals, with an average shell width of 15.28 $\pm$ 4.84 cm (mean $\pm$ SD), ranging from 2.3 to 26.6 cm. The number of individuals changed across sites as follows: 48, 63, 14, 12, 97 marked individuals for each site respectively. For further details see *Hendriks et al.* (2012).

4.3.1 Model selection with the product space method

Using shell width data collected by *Hendriks et al.* (2012) in five different sites, we could be interested in whether populations differ. The problem can be treated as a random-effects ANOVA, assuming that the expected shell width in the five populations are not independent. In the case of positive continuous response variables, like shell width, the normal distribution can be used when the sampling distribution $p(\boldsymbol{y}|\boldsymbol{\theta})$ is relatively symmetric and far away from zero. However, other distributions such as the Gamma can be adopted (*Ntzoufras*, 2009). Thus, we focus our attention on the comparison of two non-nested models, a Normal and a Gamma random-effects ANOVA, through the product space method. For each pen shell i of a given width, measured at site s (with $s = 1, \ldots, 5$) the models under comparison are:

$$\begin{array}{lcccc} M_1: & width_i \sim N(\mu_i, \sigma^2), & \mu_i = \alpha_{s(i)} & \alpha_{s(i)} \sim N(\mu_\alpha, \sigma^2_\alpha) \\ M_2: & width_i \sim gamma(\mu_i\,\tau, \tau), & log(\mu_i) = \gamma_{s(i)} & \gamma_{s(i)} \sim N(\mu_\gamma, \sigma^2_\gamma). \end{array}$$

The following four sets of prior distributions were chosen to evaluate sensitivity of the Bayes factors to prior assumptions:

		set 1	set 2	set 3	set 4
M_1	σ	$U(0,10)$	$U(0,10)$	$gamma(0.001,0.001)$	$U(0,10)$
	μ_α	$N(0,0.001)$	$U(-20,20)$	$U(-20,20)$	$N(0,1e-6)$
	σ^2_α	$U(0,10)$	$U(0,10)$	$gamma(0.001,0.001)$	$U(0,10)$
M_2	$\tau^{-0.5}=\sigma$	$U(0,10)$	$U(0,10)$	$gamma(0.001,0.001)$	$U(0,10)$
	μ_γ	$N(0,0.001)$	$U(-10,10)$	$U(-10,10)$	$N(0,1e-6)$
	σ^2_γ	$U(0,10)$	$U(0,10)$	$gamma(0.001,0.001)$	$U(0,10)$.

We estimated posterior model probabilities and the corresponding log Bayes factor to quantify the relative evidence between the competing models. Prior model probabilities were calibrated manually. This operation can be time-consuming when one or more models are strongly supported compared to the others. This means that less favoured models can be almost never selected by the Gibbs sampler. Therefore, we need to find prior model probabilities that boost the achievement of approximately equal number of posterior model activations. As in our example, these priors can be strongly asymmetric, and hence difficult to spot. In consequence, it is difficult (or even impossible) achieving equal posterior model activation strictly speaking. The ratio of posterior to prior model odds (i.e. the Bayes factor) does not depend on prior model probabilities. Therefore, chosen priors ($p(M=j)$) and observed posterior model probabilities ($\hat{p}(M=j|\boldsymbol{y})^{obs}$) are easily transformed into corrected posterior model probabilities as follows

$$\hat{p}(M=j|\boldsymbol{y})^{corr} = \left(\frac{\hat{p}(M=j|\boldsymbol{y})^{obs}}{p(M=j)}\right) \Big/ \sum_{j=1}^{K}\left(\frac{\hat{p}(M=j|\boldsymbol{y})^{obs}}{p(M=j)}\right)$$

and see supplementary material for further details. Corrected posterior model probabilities represent the values we would obtain if we setted a uniform prior for model index (*Lodewyckx et al.*, 2011). In our specific case the best prior model probabilities had very extreme values, $p(M=1)=10^{-10}$, 10^{-8}, 10^{-8}, and 10^{-10} for the four sets of priors, leading to the following log Bayes factors $log(B_{12})=24.4$, 20.9, 22.3, and 29.5. These figures indicate very strong support in favour of M_1 (see e.g. *Kass and Raftery*, 1995 for an interpretation scheme for values of the Bayes factor). In practice, if we would keep equal prior model probabilities we needed about $e^{log(B_{12})} \approx 39$, 1.3, 5.2, and 6.8 billion iterations (depending on the priors set) to have at least one M_2 activation. Details for model implementation are given in the supplemented R and BUGS code (Appendix E). The models were implemented in JAGS (*Plummer*, 2003) through the R2jags package (*Su and Yajima*, 2012).

4.3.2 Variable selection using GVS

Hendriks et al. (2012) used logistic regression models where the recaptured outcome (1=recaptured, 0=missed) was the response variable. As potential predictors of the variability in

the probability of recapture they considered the difference across sites and the shell width as a continuous individual covariate (centred for the analysis). To illustrate the GVS method we considered the following full model

$$y_{is} \sim Bernoulli(p_{is}), \quad log\left(\frac{p_{is}}{1-p_{is}}\right) = \alpha + \beta\, width_i + \epsilon_s \tag{4.10}$$

for a pen shell i, of a given width, at site s (with $s = 1, \ldots, 5$), where p_{is} is the site-specific recapture probability and ϵ_s is the random effect for the sth site, assuming $\epsilon_s \sim N(0, \sigma^2_\epsilon)$.

We used eight sets of priors for model parameters, in the light of the related sensitivity of posterior model probabilities (see specific section below). More specifically we chose four different priors for regression coefficients α and β, assumed drawn from a $N(0, \tau^{-1})$ for $\tau \in \{0.1, 0.01, 0.001, 1e-6\}$, and two different prior distributions on the random effect hyperparameter σ_ϵ, a uniform (0,5) and a half-Cauchy with scale 1 (see *Gelman* (2006) for a justification of this prior, also in relation to the small number of levels for the site effect).

The pseudoprior parameters were taken from a pilot run of the saturated model. The priors for the inclusion indicators were defined as $\gamma_j \sim Bernoulli(0.5)$ with $j = 1, 2$, with γ_1 related to the β parameter and γ_2 to the random effects.

In order to better understand how this approach works, we can briefly start by considering a simplified version of the model in eq. 4.10 with one single covariate, $y_i \sim Bernoulli(p_i)$, $logit\,(p_i) = \alpha + \beta\, width_i$. In order to estimate the posterior probability that the shell width covariate is "in" the model (i.e. the posterior inclusion probability) we can adopt one single inclusion indicator for the β parameter, which means that there are only two possible models, one with only intercept α and another with α and β. In this case, the data give 'very strong' evidence in favour of the model with the shell width effect, as the posterior inclusion probability is estimated to 0.982 (by using a Normal prior with mean 0 and variance 1000 for the parameters of the linear predictor). Note that in this particular case, with one single inclusion indicator γ, the posterior model probability corresponds to the posterior inclusion probability.

Going back to the original model (eq. 4.10) where site random effects were considered in addition to the β parameter for shell width, we estimated posterior variable inclusion probabilities and posterior model probabilities under the eight above mentioned prior regimes (Table 4.2 and 4.3). The R and BUGS code is reported as supplementary material (Appendix E).

4.4 Relevant points for practical implementation

In order to facilitate the implementation of the two methods, we can pinpoint some practical aspects.

4.4.1 Product space

Models under comparison are of different distributional form. Consequently the likelihood is defined using the "zeros trick" (*Spiegelhalter et al.*, 2007). This allows the use of any form likelihood and does not restricts in the number of distributions available (*Katsis and*

Table 4.2: Posterior variable inclusion probabilities, $p(\gamma_j = 1|\boldsymbol{y})$, obtained under four different priors for regression coefficients α and β (assumed drawn from $N(0, \tau^{-1})$ with a varying precision τ), and for two prior distributions on random effect hyperparameter σ_ϵ. γ_1 and γ_2 are the inclusion indicators.

	Priors			
	$\tau = 0.1$	$\tau = 0.01$	$\tau = 0.001$	$\tau = 1e-6$
	$\sigma_\epsilon \sim U(0,5)$			
γ_1	0.998	0.995	0.981	0.621
γ_2	0.055	0.054	0.057	0.066
	$\sigma_\epsilon \sim half-Chaucy(1)$			
γ_1	0.998	0.994	0.980	0.614
γ_2	0.150	0.146	0.139	0.173

Table 4.3: Posterior model probabilities from GVS example, under the different prior sets as in Table 4.2.

				Priors			
				$\tau = 0.1$	$\tau = 0.01$	$\tau = 0.001$	$\tau = 1e-6$
	γ_1	γ_2	Model	$\sigma_\epsilon \sim U(0,5)$			
M_1	0	0	α	0.001	0.004	0.017	0.345
M_2	1	0	$\alpha + \beta\, width_i$	0.944	0.941	0.926	0.589
M_3	0	1	$\alpha + \epsilon_s$	0.000	0.001	0.002	0.034
M_4	1	1	$\alpha + \beta\, width_i + \epsilon_s$	0.055	0.054	0.055	0.032
				$\sigma_\epsilon \sim half-Chaucy(1)$			
M_1	0	0	α	0.002	0.005	0.016	0.301
M_2	1	0	$\alpha + \beta\, width_i$	0.849	0.849	0.845	0.526
M_3	0	1	$\alpha + \epsilon_s$	0.000	0.001	0.004	0.085
M_4	1	1	$\alpha + \beta\, width_i + \epsilon_s$	0.149	0.145	0.135	0.088

Ntzoufras, 2005). The likelihood structure of the two competing models is reported and the parameter space of the trans-dimensional model includes the model index and the parameter vectors, $\{M, \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K\}$ with $K = 2$. At each MCMC iteration, the model index activates one of the two models, and consequently determines the parameter vector that has to be connected to the likelihood. In other words activation of a particular model determines how data, parameters and pseudopriors are connected to each other. Parameter vector $\boldsymbol{\theta}_j$, of the activated model $M = j$, is assigned a prior distribution and updated using prior information and observed data. If model $M \neq j$ is activated, $M = j$ is not connected to the likelihood and its parameters are assigned a pseudoprior distribution such that sampling continues. Therefore, as seen in eq. 4.6 the aim of using pseudopriors is to allow computation of conditional probabilities of model indexes. As we have already stressed even with GVS, pseudopriors do not have influence on posterior model probabilities because they are integrated out, but their choice is important for sampling efficiency. To improve convergence we can estimate pseudopriors' parameters by running the models separately and then using the posterior samples, to match the pseudopriors as nearly as possible to the true model-specific posteriors (*Carlin and Chib*, 1995).

To achieve good quality of posterior model probability estimates we have to combine (approximately) equal posterior model activation with frequent model switching. For the categorical parameter M, the lack of model switches in its Markov chains is comparable to a high level of autocorrelation for the Markov chain of a continuous parameter. Some possible solutions to improve switching of model index are i) adjusting prior model probabilities if one or more candidate models are much more supported by the data than the others, ii) reparameterizing models for increase the number of shared parameters, iii) improving pseudoprior estimation, iv) thinning chains (but see *Link and Eaton*, 2012). If we set parameters in common to the competing models we have to be sure that their interpretation is the same in the different models (see example in *Carlin and Chib*, 1995), and we need to check whether their posterior distributions have enough overlap. For example, in our distributions the parameter μ has exactly the same interpretation since it is the mean of $y_i = width_i$ in both the two models.

4.4.2 GVS

The likelihood for the full model is specified as usual in BUGS, with the only difference being the incorporation of the binary inclusion indicators γ_1 and γ_2 in the linear predictor (defined as `g[1]` and `g[2]` in the code). When we are dealing with models using explanatory variables that do not involve interactions, the latent variable γ_j can be treated as *a-priori* independent, as we did. We then adopted independent priors for model parameters (as in eq. 4.9) using a mixture of independent normal distributions

$$\beta_j \sim \gamma_j \, N(0, \Sigma_j) + (1 - \gamma_j) \, N(\bar{\mu}_j, S_j), \qquad j = 1, 2$$

where $\bar{\mu}_j$ and S_j are the mean and the variance respectively used in the corresponding pseudoprior distributions and Σ_j is the prior variance, when the jth term is included in the model.

When categorical predictors are considered in the model the definition of the prior distribution for model parameters can be complicated. To this end, *Ntzoufras* (2002) and *Katsis and Ntzoufras* (2005) provide detailed guidelines and examples.
We directly calculated posterior model probabilities in BUGS by adding few lines of code. This is feasible when the number of models under consideration is small (and equals to 2^p if we consider all possible models). In the case of large model spaces this approach is not recommended since it involves a large amount of values stored, slowing down the BUGS software (*Ntzoufras*, 2009). Therefore we can save only the model indicator and export it into R to obtain its frequency tabulation. The model indicator (noted as `mdl` in the code) was calculated by transforming $\boldsymbol{\gamma}$ with the following formula (*Ntzoufras*, 2002) that converts numbers defined in the binary numerical system to the corresponding numbers in decimal numerical system

$$m(\boldsymbol{\gamma}) = 1 + \sum_{j=1}^{p} \gamma_j 2^{j-1}.$$

Posterior model probabilities were then calculated with the BUGS command `equals`, by tracing whether a specific model is visited at each MCMC iteration.

4.5 Prior sensitivity

Posterior model probabilities are often more sensitive to the prior specification than the posterior distribution of parameters themselves. In presence of model uncertainty the priors on the parameters $p(\boldsymbol{\theta}|M)$ need to be specified with care, and a sensitivity analysis should always be performed and discuss for a number of sensible priors (*King*, 2009). This sensitivity of posterior model probabilities and Bayes factors on priors, known as "Lindley-Bartlett paradox" is probably the main disadvantage of Bayesian variable and model selection and we must pay particular attention to that. In particular the specification of prior variance σ^2_θ is hard since, in non-informative cases, must be large to avoid prior bias within each model but not large enough to activate the above mentioned paradox and fully support the simplest model. The same issue appears in any model selection problem and it is more evident in nested model comparisons (*Ntzoufras*, 2009). Bayes factors are hence unstable in presence of improper, non-informative priors for model parameters, but the problem extends also to the use of vague proper priors (*Berger and Pericchi*, 1996). Furthermore models having more parameters allow greater prior uncertainty in the range of the data to be produced, which means more sensitivity of the Bayes factor. As *Link and Barker* (2006) pointed out, assessing the reasonableness of selecting priors is inevitably a subjective process. Therefore, there does not seem to be a general solution to this issue, but it should be admitted in any analysis and priors should be clearly stated (*Royle and Dorazio*, 2008). We illustrated this sensitivity to prior by repeating both the product space and the GVS analysis using different prior distribution sets that were sufficiently uninformative to yield very similar posterior distributions for model parameters. Note that the use of an almost fully uninformative prior (with variance of 1 000 000) tends to increase the evidence in favour of Model 1 in both product space and GVS cases, as in more complex models the vague prior have a greater prior uncertainty. This

effect would have been even more evident if we used a larger prior variance. To get a sense of this effect see the example of Table 1 in *Link and Barker* (2006).

4.6 Conclusions

We wanted to present two possible procedure for hierarchical variable and model selection in a Bayesian framework, with emphasis on the practical aspect of the approaches. Unlike reversible jump MCMC, GVS algorithm is simpler and can be easily implemented by the freely available MCMC pieces of software in BUGS language (*Katsis and Ntzoufras*, 2005). Furthermore, this approach can be used not only in relation to variable selection problems (where the models concerned differ only in the form of the linear predictor) but also to compare models of different distributional form (e.g. Poisson and negative binomial or Generalized Poisson; *Katsis and Ntzoufras*, 2005), as we did with the Carlin and Chib's method. GVS can also be easily modified into the simpler (indicator model selection) method proposed by *Kuo and Mallick* (1998), frequently used by *Royle and Dorazio* (2008) in their book which represents a benchmark for ecologists. The key differences between the GVS and the product space method, but also more broadly between the other Gibbs sampler based variable selection strategies, are in their requirements in terms of priors and/or pseudopriors. Product space method and GVS both require pseudopriors just to improve the efficiency of the sampler. GVS is less expensive in requirements of pseudopriors, but correspondingly less flexible (*Dellaportas et al.*, 2000). The main drawback of the product space method is the unavoidable specification of, and generation from, many pseudoprior distributions. Even if the pseudopriors do not enter the marginal likelihood distributions, generation from $K-1$ pseudopriors at each iteration is required and is computationally demanding (*Dellaportas et al.*, 2002).

The conceptual simplicity of the two approaches can represent an appealing feature for non-statisticians, however we have to be aware of the difficulties related to the maximization of sampling efficiency, an issue that is in any case common to all trans-dimensional MCMC methods. Furthermore, bearing in mind the sensitivity of the Bayes factors to the choice of priors on parameters, we need to articulate the reasons for prior assumptions and evaluate the related sensitivity of inferences (*Link and Barker*, 2006).

We believe that the product space method and the GVS can represent potentially useful tools for ecologists and conservation biologists that are used to fit their models with the widely used software packages in BUGS language, especially in relation to the trade-off between ease of implementation and generality of application.

4.7 Acknowledgements

We greatly thank Bob O'Hara and Mike Fowler for suggestions that truly improved a previous version of this manuscript. Funds were partially provided by the Regional Government of Balearic Islands and FEDER funding. ST was funded by a PhD grant from the Science Museum (Trento) in collaboration with the University of Pavia.

Chapter 5

Modelling altitudinal occupancy dynamics in migratory birds

Simone Tenan, Albert Fernández Chacón, Giuseppe Bogliani, Paolo Pedrini, Giacomo Tavecchia (*in prep.*)

Note

This contribution is still a work in progress. However, being part of the work conducted during the PhD, we report a synthesis with a description of the analytical framework together with some preliminary results.

Key-words: Altitudinal gradient; bird migration; colonization; extinction, Alps.

5.1 Introduction

Analysis of species occurrence and site occupancy along the elevational gradient has been poorly described for migratory birds crossing mountainous regions, with species-specific information often merely limited to observed frequencies, furthermore referred to discrete elevational points such as high and low-elevation, or passes and valleys (e.g. *DeLong et al.*, 2005; *Williams et al.*, 2001; *Wilson and Martin*, 2005). Studies examining habitat use during post-breeding and migratory periods have generally found evidence for habitat selection that in some cases differed from breeding habitats (e.g. *Moore et al.*, 1990). Furthermore, species often vary in their degree of habitat specialization (e.g. *Hutto*, 1985). In mountainous regions we expect a complex response of migrants' abundance and occurrence to topography (e.g. *Bruderer*, 1978) and altitudinal distribution of available resources. In particular, altitudinal variation in abundance can influence occupancy dynamics, i.e. survival and colonization probabilities, of sites along the elevational gradient. However, to date we are not aware of

studies that have attempted to analyse (site) occupancy dynamics in migrating birds crossing mountain chains. This is what this work aims to achieve. In the light of metapopulation theory (*Hanski*, 1999) we explicitly modelled presence-absence data to estimate survival/extinction and colonization probabilities, which represent the demographic components of occupancy dynamics. More specifically we assessed changes in site extinction and colonization probabilities along the elevational gradient, for different bird species crossing the southern portion of the Alpine chain during the post-nuptial migration. To account for imperfect detectability, as well as for other sources of heterogeneity of parameters among sites, we exploited the Bayesian dynamic occupancy modelling framework (*Royle and Dorazio*, 2008) and based our inference on ringing data spanning a 14-year period and gathered at 32 sites.

As an anticipation we provide some preliminary result of this work in progress for three migratory species: the Eurasian woodcock (*Scolopax rusticola*), the Ring ouzel (*Turdus torquatus*),and the Garden warbler (*Sylvia borin*).

5.2 Materials and Methods

5.2.1 Data collection

We analysed detection/non-detection data referred to 3 migratory species and collected during the post-nuptial migration, in the period 1997-2010, at 32 pre- and southern Alpine sites (Fig. 5.1). Birds were trapped daily by means of mist-nets that were operated according to a standardized protocol over the years. However, no trapping occurred in some days due to logistic problems or adverse weather conditions. The general monitored period ranged from the end of August to the end of October, with a variable length of the capture season between sites and years, mainly in relation to the onset of winter snowfall (for sites at higher elevation) or logistic reasons. Sampling sites are spread along an altitudinal gradient, from 190 to 1774 m a.s.l. (mean = 885 m; SD = 612.2 m), and are located at alpine passes (12 sites), mountain slopes (9), or at the bottom of alpine valleys (11) (*Pedrini et al.*, 2012). To minimize the number of local individuals not yet on migration, we defined a threshold date after which the birds were considered to be migrants, and only those were included in the analysis. Following *Schaub and Jenni* (2000) and *Jenni and Kéry* (2003), this threshold date was determined based on literature (*Cramp et al.*, 2000) and on the onset of migration at trapping sites where the species does not breed. Data were considered at a 5-day period resolution, and for each temporal unit a measure of sampling effort was derived (as the number of day of trapping activity, ranging from zero to five).

5.2.2 Sampling design and dynamic occupancy models

We consider our data as obtained from repeated detection/non-detection surveys of $i = 1, 2, \ldots, M$ sites, with $M = 32$. Each site was surveyed $j = 1, 2, \ldots, J$ times within each of $t = 1, 2, \ldots, T$ primary periods (with $T = 14$). The number of replicate samples J per primary period varies between species from 4 to 17, mainly in relation to the length of the migratory season.

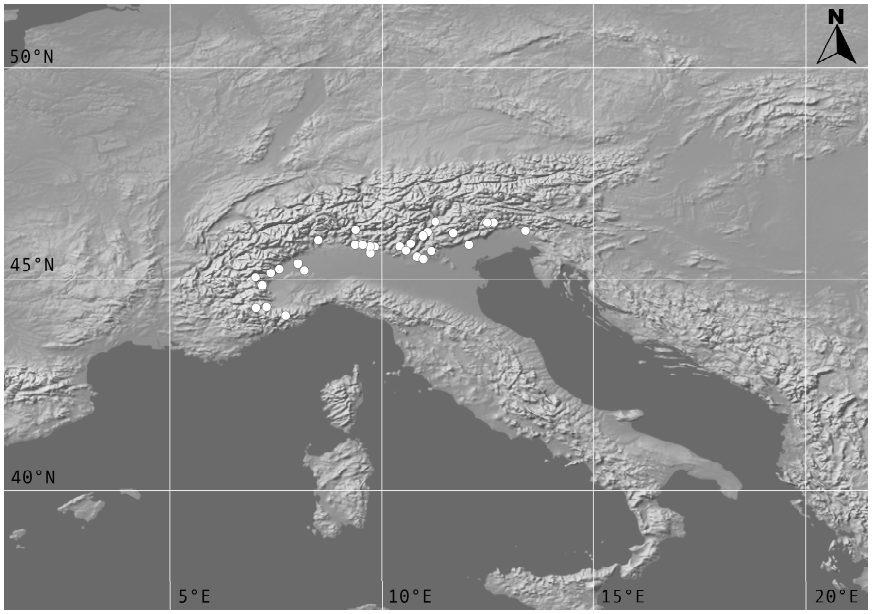

Figure 5.1: Distribution of the 32 sampling sites (dots) along the southern Alps.

Table 5.1: Sampling period (as starting and ending date), and number of 5-days periods within each year. The latter is the number of repeated surveys (or sampling replicates J) within each primary period T (see 'Sampling design' section for further explanations).

Species	Period	n replicates
Scolopax rusticola	4 Ago – 1 Nov	18
Turdus torquatus	14 Ago – 1 Nov	16
Sylavia borin	4 Ago – 17 Oct	15

Let $y_j(i,t)$ denote the observed occupancy status of site i for survey j within primary period t. We suppose that $y_j(i,t)$, $j = 1,2,\ldots,J$ are independent and identical distributed (*iid*) for each site i and primary period t. Then, let $z(i,t)$ denote the true occupancy status of site i during primary period t, having possible states 'occupied' ($z = 1$) or 'not occupied' ($z = 0$). Probability of site occupancy (or probability of occurrence) for period t is denoted by $\psi_t = \Pr(z(i,t) = 1)$. As above mentioned, we parameterised occupancy dynamics in terms of local extinction and colonization processes. The canonical parameters of the dynamic occupancy models are the initial occupancy ψ_1, the survival probabilities $\{\phi_t\}_{t=1}^{T-1}$ and the colonization probabilities $\{\gamma_t\}_{t=1}^{T-1}$. The probability that an occupied site 'survive' (i.e. remain occupied) from period t to $t+1$ is $\phi_t = \Pr(z(i,t+1) = 1|z(i,t) = 1)$. Local extinction probability is the complement of survival probability, i.e. $\epsilon_t = 1-\phi_t$ (*MacKenzie et al.*, 2003). In metapopulation systems, local colonization is the analogue of the recruitment process in a population. Local colonization probability, from period t to $t+1$, is $\gamma_t = \Pr(z(i,t+1) = 1|z(i,t) = 0)$.

Classical likelihood-based inference under dynamic occupancy models (*MacKenzie et al.*, 2003) is based on removing the latent state process from the model by marginalization (or integration). On the contrary, under the hierarchical modelling approach the state variable is retained in the model without any difficulties (*Royle and Kéry*, 2007). This is attractive since the hierarchical representation naturally leads to several important extensions (see below). In these cases a Bayesian analysis provides some important benefits to fit complex models (see also Appendix F), where even likelihood analysis is not practically feasible (*Royle and Dorazio*, 2008). In this first phase of our study, we fitted a unique model with an altitudinal effect on both local colonization and survival parameters, while detection probability was expressed as a function of an additive combination of survey date and sampling effort (number of fieldwork days within each five-day period, see above).

5.2.3 Hierarchical formulation

Following (*Royle and Kéry*, 2007), in the state model the initial occupancy states, at $t = 1$, are assumed to be *iid* Bernoulli random variables,

$$z(i,t) \sim Bern(\psi_1) \qquad \text{for} \quad i = 1,2,\ldots,M \tag{5.1}$$

whereas, in subsequent periods,

$$z(i,t)|z(i,t-1) \sim Bern(\pi(i,t)) \qquad \text{for} \quad t = 2,3,\ldots,T \tag{5.2}$$

where

$$\pi(i,t) = z(i,t-1)\phi_{t-1} + [1 - z(i,t-1)]\gamma_{t-1}. \tag{5.3}$$

For a site that is occupied at $t-1$, the survival component in eq. 5.2 determines the subsequent state and $z(i,t)$ is a Bernoulli outcome with probability ϕ_{t-1}. In the case a site is not occupied at time $t-1$, then the recruitment component in eq. 5.2 determine the subsequent state, and $z(i,t)$ is a Bernoulli outcome with probability γ_{t-1}.

We used the auto-logistic parameterization (*Royle and Dorazio*, 2008) of the model to parameterise the effect of elevation for both site survival and colonization probabilities. We

thus extended the model as follows:

$$logit(\pi_{i,t}) = a_t + b_t\, z(i,t-1) + \beta_1\, elev_i + \beta_2\, elev_i\, z(i,t-1) \tag{5.4}$$

where

$$\gamma_{t-1} = logit^{-1}(a_t)$$

and

$$\phi_{t-1} = logit^{-1}(a_t + b_t).$$

In this case, the effect of elevation on the logit of colonization probability is simply β_1 and the effect on the logit of survival probability is $\beta_1 + \beta_2$.

The observation model is specified conditional on the latent state process $z(i,t)$, as

$$y(i,j,t)|z(i,t) \sim Bern(z(i,t)p_{ijt}). \tag{5.5}$$

We accounted for variation in detection probability in relation to both sampling effort and survey dates (previously standardized) as

$$logit(p_{ijt}) = \alpha_0 + \alpha_1\, eff_{ijt} + \alpha_2\, surv_j.$$

This model can easily be implemented in BUGS; an example of model code is reported in Appendix F. It is worth mentioning that under the auto-logistic parameterization Markov chains typically exhibit very low autocorrelation (*Royle and Dorazio*, 2008).

5.3 Preliminary results and discussion

In all the three species, at least one dynamic component (i.e. local colonization or survival) is affected by elevation, with posterior estimates that do not encompass zero in the 95% credible interval (Fig. 5.2). This can be related to the altitudinal succession of more or less desirable habitats for each migratory species. More specifically, there is evidence for a positive relationships between elevation and both site colonization and survival, for the Eurasian woodcock and the Ring ouzel, leading to larger occupancy probability at higher elevation. On the contrary, in the Garden warbler colonization was negatively related to elevation, while survival probability should be constant along the elevational gradient, as its 95% credible interval does encompass zero (Fig. 5.2). For the latter, occupancy probability would hence be higher at lower elevation. Estimates for the elevational trend of occupancy dynamic components seem consistent with current ecological knowledge for the three species (*Pedrini et al.*, 2012). The Eurasian woodcock is more related to medium elevation broad-leaved forests, the Ring ouzel is present at medium-high elevation sites between coniferous forest with clearings and alpine grasslands, while the Garden warbler should be more dependent, during migration, to berry production of bushes at lower elevation (*Cramp et al.*, 2000). Occupancy probability in some years can be difficult to estimate, probably due to the low number of available data (Fig. 5.2). We then note a very small average occupancy probability for the Ring ouzel, in contrast to the very high probability for the Garden warbler, while a great uncertainty characterizes the estimates for the Eurasian woodcock (Fig. 5.2).

Table 5.2: Posterior summary of model parameters for the best dynamic occupancy model fitted to data on three migratory species sampled while crossing the souther Alps, during the period 1997-2010. Reported are the mean and the 95% credible interval of parameters for the effect of elevation on local colonization (β_γ) and survival (β_ϕ).

Species	Parameter	mean	2.5%	97.5%
Scolopax rusticola	β_γ	3.512	1.561	4.937
	β_ϕ	6.552	2.828	9.344
Turdus torquatus	β_γ	3.942	1.782	4.968
	β_ϕ	7.865	4.980	9.685
Sylavia borin	β_γ	-3.432	-4.933	-0.820
	β_ϕ	-0.519	-1.552	0.358

Model selection results provide very strong evidence, at least in the three study species, that detection probability should be modelled as dependent to sampling effort in addition to survey date (results not shown).

Further investigations are needed to better understand the relationships between species traits and altitudinal occupancy dynamics, being aware that our study rely on the interpretation of the occupancy probability as the probability of using a site sometime during the investigated migratory period (*Royle and Dorazio*, 2008).

Finally, we note that migratory species crossing the Alps usually do not display a classical metapopulation where patches are separated by unsuitable or inferior habitat. Instead, spatially separated samples are drawn from a more or less continuous distribution (*Royle and Kéry*, 2007). Therefore, the scope of inference should be the population of all possible sites in the study area, and the population estimates for dynamic parameters are consequently more appropriate that the finite sample estimates of these quantities.

To out knowledge, this work represents the first attempt to model site occupancy dynamics of birds during the migratory period, and along the elevational gradient. By analysing the changes in site use probability in relation to altitude, this contribution has some features in common to the previous works, such as the use of detection/non-detection data (as in Chapter 1 and 2), the effect of external covariates as in Chapter 2, and the need for choosing between competing models (as addressed in Chapter 4). Besides, there is a conceptual equivalence of models for different systems so that, under a hierarchical framework, for many problems the observation models and the process models translate across systems. This yields to the formal equivalence between occupancy models and closed population models, or between dynamic occupancy models and models of population dynamics. In addition, a second unifying theme is the technical device of data augmentation that formalizes the equivalence between occupancy models and models of repeated sampling of populations of individuals, both in closed and open population contexts (*Royle and Dorazio*, 2008).

With regard to this work, other objectives will be addressed in the near future: (i) relating dynamic parameters to species traits and environmental covariates, (ii) exploring time-related changes in altitudinal occupancy dynamics, (iii) analysing a quadratic elevation effect on

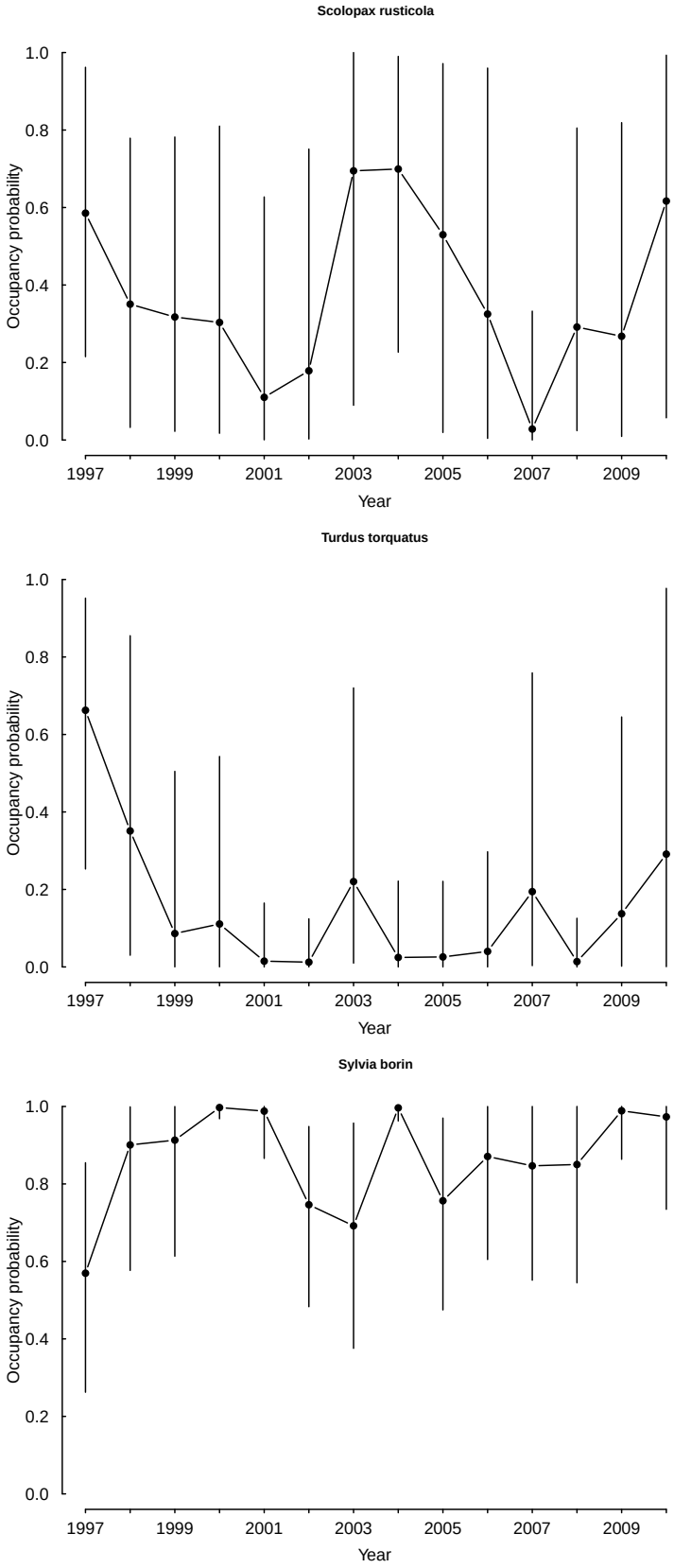

Figure 5.2: Temporal variation of occupancy probability (mean and 95% credible interval) for the three migratory species.

dynamic components of occupancy, and (iv) extending the analyses to other 35 migratory species.

Chapter 6

General Conclusions

"No matter how much planning has gone into our investigation prior to collecting data, no matter how carefully we have designed our study, no matter how familiar we are with the system we are studying, there comes a point where we must describe our observations using a mathematical model."

William A. Link and Richard J. Barker
(Bayesian Inference with Ecological Applications, 2010)

With this thesis I addressed different ecological hypotheses in the population ecology field, by exploiting the conceptual clarity and practical utility of the hierarchical modelling framework, together with the benefits of Bayesian methods as a mode of analysis and inference. The general goal was to explore their potential and identify their limitations, aiming to overcome them using new model formulations. My work is based on theories, model formulations and practical aspects developed by other researchers before me. When possible I have extended this body of work hoping that this further step will be of use in addressing finer questions for the understanding of population functioning, species distribution and ultimately conservation.

In chapters 1 and 2 I adopted a hierarchical formulation of traditional individual effects models for closed populations, with the addition of individual covariates that influence detection probability, and the extension to external covariate to model spatial variability. Models with individual covariates are very similar to models containing unstructured variation or individual heterogeneity in the form of an individual random effect. The main distinction is that in individual covariate models we have information about the covariate distribution by virtue of some covariate values being observed. The analysis was then performed by exploiting the technical device of data augmentation. Under a hierarchical formulation, data augmentation formalises the conceptual equivalence between different classes of models, such as occupancy

models (Appendix F) and models of repeated sampling of populations of individuals, both in closed and open populations (*Royle and Dorazio*, 2008).

Setting the number of sampling occasions, to estimate closed population size from capture-recapture (CR) data, is often a question of balancing objectives and costs. Moreover, accounting for trap-response and individual heterogeneity in recapture probability can be problematic when data derive from less than four sampling occasions. I found that individual covariate models, in a Bayesian formulation, provide a flexible and robust framework to include complex recapture processes when analysing three session capture-mark-recapture studies (Chapter 1). Results from simulated data have showed that under the correct model, estimates of population size, population structure and recapture probabilities were close to those used in the simulations. Interestingly a general model including an 'unnecessary' trap-response effect performed well when the recapture probability was moderately high ($p = 0.6$), but not when it was moderately low ($p = 0.30$). Applying this model to data would be 'playing safe' only when capture probability is high. The opposite was never true; simpler models, i.e. with no trap-response, were never adequate when trap-response was present and population size tended to be overestimated. In real data we can test whether a model with trap-response is appropriate through model selection. Besides, to my knowledge, there are not informative goodness of fit test available to detect permanent trap-response or heterogeneity in closed models (D. Borchers and R.M. Dorazio *pers. comm.*), a point worth investigating in the future.

Results from real CR data on the endemic Balearic Lizard indicated substantial differences in the estimates of mean density between our study and a line transect based study. Part of this difference is probably due to natural fluctuations of the population or to spatial heterogeneity in population density. However, part might arise from the systematic biases of the two methodologies. In habitat with dense vegetation, where animals are more difficult to detect, accurate density estimates by line transect are hardly achievable. Capture-mark recapture methods, are thought to be more accurate that visual methods in monitoring elusive species. However, in my case the capture method selects for individuals large enough to reach and fall into the traps. This might lead to underestimating total population size, depending on the presence and proportion of very small individuals. Additional investigations should be carried out to compare the two methods. Besides, I have shown how capture-recapture models of the ordinary sense, but analysed using a Bayesian formulation with data augmentation, can be extended to the simultaneous analysis of multiple (two, in my case) groups. This allowed us to estimate lizard sex-ratio, which was nearly even. As adult sex-ratio in lizards is often reported to be females biased, as expected in polygynous vertebrates (*Massot et al.*, 1992; *Le Galliard et al.*, 2005; *Buckley and Jetz*, 2007), further research might focus on possible spatio-temporal changes in the sex-ratio that might arise from the interaction between resource availability and habitat quality (*Massot et al.*, 1992; *Galán*, 2004).

Individual covariate models were then used to investigate the difference in structure and abundance of six populations of Noble Pen shell from CR data (Chapter 2). I tested whether site dependent differences were influenced by site-specific anthropogenic pressure (anchoring) or by the physical characteristics of the habitat. This information can be used to fo-

cus conservation efforts for the population of this endangered bivalve in coastal areas. As a technical aspect, I identified a covariate effect (related to shell width) on the individual detectability. Moreover the major determinant of shallow populations of Noble Pen shell appeared to be anchoring, both in terms of population density and size-dependent population structure. Actually, physical dislodgement by anchoring causes fast and unpredictable mortality on larger noble pen shells, and selective mortality of large individuals might have important consequences for the future of the population.

In these first two applications (chapters 1 and 2) I have noticed that, despite a great flexibility, this approach has an important limitation in that the distribution of (body) size has to be known or assumed. Indeed, posteriors have the same distribution fixed by the one of the priors. In this case, I have assumed a normal distribution of sizes, but population structure might be skewed toward large or small individuals and have to be symmetrical, while there is no reason to assume that small and large animals are equally represented. On the other hand, using a flexible distribution (e.g. Gamma distribution) would limit the capability of drawing samples from the posterior distribution.

In Chapter 3 I showed how hierarchical models can be used as a formal way of combining information coming from different studies. A joint analysis of data on demographic parameters and population size allows a comprehensive assessment of the state and the dynamics of a system. This brings other advantages as well. The precision of parameter estimates is increased, meaning that we have more power to detect an effect, which is particularly useful with small sample sizes. Moreover, parameters that are not estimable otherwise may become estimable (*Schaub and Kéry*, 2012; *Tavecchia et al.*, 2009).

In my case, I integrated radio-tracking data, nest and fledglings counts and time series of birds found dead by the general public into a single analysis to obtain consensus estimates and explore the demographic impact of poisoning on an endangered raptor. The joint analysis delivered more precise estimates, which incorporated all available information and led to a more accurate assessment of the importance of specific causes of mortality and short-term population forecasting. Moreover, the integrated analysis allowed the estimation of the breeding and non-breeding sector of the population, an elusive parameter that can be measured only under special circumstances and usually through a pronounced survey. Additionally, it is interesting to note that the raw frequency of animals reported to have died from a specific cause is an information frequently available from local authorities but generally considered as too coarse to yield any meaningful estimate of mortality impact. In my case, even if this information was scarce, it contributed to improve parameter estimates.

The aim of this work was to measure the demographic consequences of illegal poisoning for a threatened raptor. My analysis confirmed the extreme sensitivity of this extinction-threatened species to poisoning, which resulted to be the most important cause of mortality of Red Kites in the Island of Mallorca and was able to suppress population growth by 20%. Nevertheless, a high fecundity rate seemed helpful to counter-sustain population viability. Unfortunately, it was not possible to evaluate to what degree this could be ascribed to the beneficial effect of artificial feeding sites, and further research is needed to understand their role as a population management tool.

It is important to note that no goodness-of-fit (GOF) test is currently available for integrated population models (*Schaub and Abadi*, 2010). At the moment the best solution is to assess the goodness of fit of individual likelihoods to single data set. If these tests are satisfactory, we may assume that integrated model is also a satisfactory description of the complete data (*Kéry and Schaub*, 2011). The development of GOF approach is very crucial for establishing credibility and confidence of the integrated population model (*Abadi*, 2010). Model comparison might be possible and relatively easy to implement when the number of models is limited, and the models concerned differ only in the form of linear predictors.

Integrated modelling can be of great help in comprehensive studies where information is scattered across multiple sources. However, they have a conceptual and practical limitation. To allow the communal likelihood, data should be, at least partially, independent (*McCrea et al.*, 2010). Independent data are likely to come from independent populations or from different biological levels (i.e. when national census are joined to local survival parameters). In this case it might be conceptually wrong to assume that they describe the same system or process as they incorporate too much variability due to the ecological scale. The practical problem is that integrated analyses provide consensual estimates combining models for the different parts of information available. A lack of fit in one or more of these models would automatically be compensated by the integrated analysis because the multiple variance components are not independent. For example, assuming a wrong model for let's say fertility, might result in an inflated error in the census data (*Tavecchia et al.*, 2009).

A communal problem and, so far, limitation of the Bayesian approach is the numerical and conceptual difficulty in variable and model selection. This is a very important topic in ecological modelling and well developed in the frequentist approach (*Burnham and Anderson*, 2002). It is however still *in fieri* for the Bayesian analysis as the available methods are numerically difficult to implement and their performance depends on the set of models considered. I tackled this problem in Chapter 4 where I provided an example to address the problems of both Bayesian variable and model selection using two Gibbs sampler based strategies, taken from the wider survey of methods already implemented in the BUGS software. After a theoretical description I illustrated the Gibbs variable selection approach with an example of logistic regression with random effects, and the product space method with the comparison of two non-nested hierarchical models of different distributional form. For their conceptual simplicity and ease of implementation the proposed methods can represent potentially useful tools for ecologists and conservation biologists that want to start addressing formal comparison of hierarchical models in a Bayesian framework. Particular emphasis has been placed on the difficulties related to the maximization of sampling efficiency and the sensitivity of inference (i.e. the Bayes factors) to the choice of priors on parameters. Note that the prior has a fundamental importance in the Bayesian approach (e.g. *Gelman et al.*, 2003), but its flexibility has been often a source of criticisms from those that are sceptic about incorporating previous information on the analysis (see Appendix A for further details).

Finally, in Chapter 5 I presented a novel application of the Bayesian state-space formulation of dynamic occupancy models (*Royle and Kéry*, 2007). To my knowledge, this is the first attempt to model spatial-temporal dynamics of site colonization and extinction probabilities

in migrating birds. By exploiting data derived from a medium-term monitoring program held in the souther Alps, I looked at possible altitudinal trends in the dynamic components of occupancy probability. I adopted a general and flexible formulation of dynamic occupancy models that provides a clear segregation of parameters governing the ecological and the observation process, yielding to an analytical framework that can deal even with strongly unbalanced data by accounting for different sources of detection bias. Moreover, I related a site-specific covariate (elevation) to the dynamic components of occupancy, by means of auto-logistic model. However, the usefulness of the Bayesian state-space formulation is currently limited by the issue of selecting between competing models (see Chapter 4). As already mentioned in the introductive part this contribution has to be considered as a work in progress, with objectives that include (i) the exploration of more complex spatial-temporal relationships in local extinction and colonization probabilities, (ii) the extension of inference to other migratory species, and (iii) the implementation of a model selection procedure through a Gibbs sampler.

All applications presented are unified by the hierarchical formulation of models for various ecological systems. Even if in the ecological literature Bayesian inference and hierarchical modelling are often thought and presented together, the utility of the hierarchical approach, to describe complex systems, is independent of the statistical framework. However, we cannot forget the usefulness of the Bayesian approach to statistics. Its utility mainly derives from its simplicity, which allows the investigation of far more complex models than can be handled by the tools in the classical toolbox (*Link and Barker*, 2009).

In this thesis I provided an empirical illustration on some of the current analytical capabilities available for population analysis, which is fundamental for conservation and management. As a matter of fact, a rigorous scientific approach to wildlife management and conservation should be based on quantitative evidence and rely on the best available assessment of demographic quantities, along with a full assessment in their uncertainty (*Kéry and Schaub*, 2011). However, recurring themes in population analysis are e.g. small sample sizes, data collected under non-standardized sampling designs, multiple levels uncertainty and stochasticity. As I meant to show throughout the thesis, a hierarchical model formulation within a Bayesian framework can be extremely functional for these challenges.

Besides the quantitative evidence of conservation concern, the contributions in this thesis provide examples of methodological novelties for some classes of models. More specifically, I extended inference based on individual covariate models to account for group and covariate effects, and I addressed the issue of Bayesian variable and model selection which is still an open question between ecologists. In the meanwhile, as always happens, many other interesting questions have arisen from this experience, ranging from the implementation of a goodness-of-fit test for individual covariate models, to the exploration of the relationships between latent parameters and the lack-of-fit in integrated population models. I hope to address at least some of them in the near future.

Appendix A

The Bayesian paradigm in brief

It would be impossible to explain here in detail the distinction between classical statistics, based on frequentist inference, and the Bayesian paradigm, together with the basic principles of Bayesian inference. What follows is based on some main references (e.g. *Congdon*, 2006; *Clark and Gelfand*, 2006; *Gelman et al.*, 2003; *King*, 2009; *McCarthy*, 2007; *Royle and Dorazio*, 2008), aiming to provide the reader with some hints about these complex and vast topics.

The main difference between the classical statistical theory and the Bayesian approach is that the latter consider parameters as random variables that are characterized by a prior distribution, on the same footing as the data. Before data are collected, it is the *prior* distribution that describe the variation in the parameters. After the collection of data, this prior distribution is combined with the traditional likelihood to obtain the *posterior* distribution of the parameter of interest on which the statistical inference is based. No optimization is involved.

In classical statistics one does not condition on the observed data but rather entertains the notion of replicate realizations (i.e. hypothetical data) and evaluates properties of estimators by averaging over these unobserved things. Conversely, in the Bayesian paradigm, the Bayesian conditions on the data, since it is the only thing known for certain. So the frequentist will evaluate some procedure, say an estimator, $\hat{\theta}$, that is a function of y, say $\hat{\theta}(y)$, by averaging over realizations of y. The nature of $\hat{\theta}$ becomes somewhat important to the frequentist way of life, and there are dozen of rules and procedures for cooking up various flavours of $\hat{\theta}$. On the other hand, the Bayesian will fix y, and base inference on the conditional probability distribution of θ given y, which is called the posterior distribution of θ. For this reason, the Bayesian approach is (conceptually) completely objective, with inference always based on the posterior distribution. But, therein lies also the conflict. To compute the posterior distribution, the Bayesian has to prescribe a prior distribution for θ, and this is a model choice. Fortunately, in practice this is usually not so difficult to do in a reasonably objective fashion. As such, this is viewed as a minor cost for being able to exploit probability calculus to yield a coherent framework for modelling and inference in many situations.

The main practical distinction between Bayesian and non-Bayesian treatments of hierarchical models comes down to how latent variables (random effects) are treated. Bayesian put

prior distributions on all unknown quantities and use a basic probability calculus in conjunction with simulation methods (see below section A.0.2) to characterize the posterior distribution of parameters and random effects by Monte Carlo simulation. The non-Bayesian removes the random effects from the model by integration.

The Bayesian approach follows from a simple application of Bayes' Theorem which provides a posterior distribution for all of the model parameters jointly. If we are interested in only a particular parameter (or a subset of parameters) we want to estimate the *marginal* posterior distribution for that parameter alone (or for each parameter in the subset). In order to obtain the marginal posterior distribution we have to integrate the *joint* posterior distribution. It is here that the difficulties arise, due to the intractabilities involved in the calculation of the marginal posterior distribution. The optimization of classical analysis has been replaced by integration in the Bayesian framework. The modern approach to this problem is not to try to integrate the joint posterior distribution analytically, but instead to employ special simulation procedures (known as Markov chain Monte Carlo, MCMC, *Robert and Casella*, 2004) which results in samples from the posterior distribution. Having simulated values from the joint posterior distribution means that one then naturally has simulated values form the marginal posterior distribution of the parameter of interest.

A.0.1 Bayes' Theorem

The Bayesian approach is based upon the idea that the experimenter begins with some prior beliefs about the system under study and then updates these beliefs on the basis of the observed data, $\boldsymbol{y}$. As said above, in order to make probability statements about a parameter θ given $\boldsymbol{y}$, we must begin with a model providing a joint probability distribution for θ and $\boldsymbol{y}$. The joint probability mass or density function can be written as a product of two densities, the prior distribution $p(\theta)$ and the sampling (or data) distribution $p(\boldsymbol{y}|\theta)$:

$$p(\theta, \boldsymbol{y}) = p(\theta)\, p(\boldsymbol{y}|\theta). \tag{A.1}$$

Simply conditioning on the known value of the data $\boldsymbol{y}$, using the basic property of conditional probability known as Bayes' rule, yields the posterior density:

$$p(\theta|\boldsymbol{y}) = \frac{p(\theta, \boldsymbol{y})}{p(\boldsymbol{y})} = \frac{p(\theta)\, p(\boldsymbol{y}|\theta)}{p(\boldsymbol{y})}, \tag{A.2}$$

where $p(\boldsymbol{y}) = \sum_\theta p(\theta)\, p(\boldsymbol{y}|\theta)$, and the sum is over all possible values of θ (or $p(\boldsymbol{y}) = \int p(\theta)\, p(\boldsymbol{y}|\theta)\, d\theta$ in the case of continuous θ). Note that we obtain a posterior *distribution* for the parameter, rather than simply a single point estimate. An equivalent form of eq. A.1 omits the factor $p(\boldsymbol{y})$, which does not depend on θ and, with fixed $\boldsymbol{y}$, can thus be considered a constant, yielding to the *unnormalized* posterior density:

$$p(\theta|\boldsymbol{y}) \propto p(\theta)\, p(\boldsymbol{y}|\theta). \tag{A.3}$$

The omitted term $p(\boldsymbol{y})^{-1}$ is essentially simply the normalising constant, which is needed so that the posterior distribution $p(\theta|\boldsymbol{y})$ is a proper distribution (and integrates to unity).

These expressions encapsulate the technical core of Bayesian inference: we must first define an appropriate probability model $p(\theta, \boldsymbol{y})$ and perform the necessary computations to summarize $p(\theta|\boldsymbol{y})$ in appropriate ways. Plugging the $p(\boldsymbol{y}|\theta)$ and $p(\theta)$ terms into Bayes' Theorem, we obtain the posterior distribution $p(\theta|\boldsymbol{y})$. This is a new distribution for the model parameter, corresponding to our new beliefs, which formally combine our initial beliefs (represented by $p(\theta)$) with the information gained from the data through the model $p(\boldsymbol{y}|\theta)$.

This approach is in contrast to the classical approach, which simply does not consider the prior and focuses on the model term $p(\boldsymbol{y}|\theta)$. Rather than inverting $p(\boldsymbol{y}|\theta)$ through Bayes' Theorem to obtain a function of the model parameter (which requires a prior), the classical approach simply sets

$$L(\theta|\boldsymbol{y}) \equiv p(\boldsymbol{y}|\theta), \tag{A.4}$$

where $L(\theta|\boldsymbol{y})$ is referred to as the likelihood function. Unlike the posterior distribution, the likelihood in classical statistics is not a probability distribution and, rather than reflecting the experimenter's beliefs, the likelihood refers to the relative likelihood of the observed data under the model, given the associated parameter(s). As we know, the classical approach typically seeks to identify the value (previously denoted by $\hat{\theta}$) of each parameter that maximises the likelihood, and a point estimate is obtained for that value. Differently, as we have seen, with the Bayesian approach the parameters are assumed to have a distribution rather than a single value. Note that, perhaps a little confusingly, the term *likelihood* is also widely used for the probability density (or mass) function $p(\boldsymbol{y}|\theta)$ within the Bayesian framework, since it is of the same mathematical form.

A.0.2 Markov chain Monte Carlo

At the beginning of the 1990s two groups of statisticians had (re)discovered Markov chain Monte Carlo (MCMC) methods (*Gelfand and Smith*, 1990). Physicists were familiar with MCMC methodology from the 1950s. Implementation of MCMC methods in combination with a rapid evolution of personal computers made the new computational tool popular within a few years. This provided an enormous impetus to the development of modern Bayesian methods. Using MCMC we can now set up and estimate complicated models that describe and solve problems that could not be solved with traditional methods. The development of MCMC methodology had also promoted the implementation of random effects and hierarchical models. MCMC approach has two components: Monte Carlo integration (the second MC) and Markov chains (the first MC).

Monte Carlo Integration

As we said, Bayesian inference requires integration of the posterior density. Usually this entails the problem of evaluating an integral which is too complex to calculate explicitly. We can replace the integration problem by a sampling problem by means of Monte Carlo integration. This simulation technique provides an estimate of a given integral, and hence the posterior expected value. The method is based upon drawing observations from the distribution of the variable of interest and simply calculating the empirical estimate of the

expectation. If we want to estimate the posterior expectation of a function $\xi(.)$ of a parameter θ, given the observed data $\boldsymbol{y}$

$$\mathbb{E}_p[\xi(\theta)] = \int \xi(\theta) p(\theta|\boldsymbol{y}) d\theta,$$

with Monte Carlo integration technique we obtain a sample of observations, $\theta^{(1)}, \ldots, \theta^{(n)} \sim p(\theta|\boldsymbol{y})$, and we can estimate the integral

$$\mathbb{E}_p[\xi(\theta)] = \int \xi(\theta) p(\theta|\boldsymbol{y}) d\boldsymbol{y},$$

by the average

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^{n} \xi(\theta^{(i)}). \tag{A.5}$$

That is, we draw sample from the posterior distribution and we calculate the sample mean. For independent samples, the Law of Large Numbers ensure that

$$\bar{\xi}_n \to \mathbb{E}_p[\xi(\theta)] \qquad \text{as} \qquad n \to \infty.$$

However, obtaining independent samples from $p(\theta|\boldsymbol{y})$ may be difficult, but equation A.5 may still be used if we generate samples (not independently) via some other methods. Note that in general $p(\theta|\boldsymbol{y})$ represents a high-dimensional and complex distribution from which samples would usually be difficult to obtain. Moreover, large sample sizes are often required and so powerful computers are needed to generate these samples. The solution to the issue of obtaining a large sample from the posterior distribution comes from the use of Markov chains.

Markov chains

A Markov chain is a stochastic sequence of numbers where each value in the sequence depends only upon the last. For example, a sequene can be represented by $\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(n)}$, where the value $\theta^{(0)}$ is chosen from some arbitrary starting distribution. If we want to simulate a Markov chain, we generate the new state of the chain, $\theta^{(k+1)}$, from some density, dependent only on $\theta^{(k)}$

$$\theta^{(k+1)} \sim \mathcal{K}(\theta^{(k)}, \theta) \qquad (\equiv \mathcal{K}(\theta|\theta^{(k)})),$$

where $\mathcal{K}$ denotes the transition kernel [1] for the chain. The transition kernel uniquely describes the dynamics of the chain.

The distribution over the states of the Markov chain will converge to a *stationary* distribution under certain conditions (aperiodic and irreducible chain) that are always assumed met. For our purposes, if a Markov chain has reached a stationary distribution ($\pi(\theta)$) then this distribution will be the posterior distribution of interest (i.e. $\pi(\theta) \equiv p(\theta|\boldsymbol{y})$).

[1] The kernel of a probability density function (pdf) or probability mass function (pmf) is the form of the pdf or pmf in which any factors that are not functions of any of the variables in the domain are omitted. In the example of a Normal distribution, the pdf is $p(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$ and the associated kernel is $p(y|\mu, \sigma^2) \propto e^{-\frac{(y-\mu)^2}{2\sigma^2}}$. Note that the factor in front of the exponential has been omitted, even though it contains the parameter σ^2, because it is not a function of the domain variable y.

MCMC in brief

MCMC methods allows the construction of a sequence of values whose distribution converges to the posterior distribution of interest, given an arbitrarily large sample. This sequence is a dependent sample form the joint posterior distribution, since each realization depends directly on its predecessor. Once the chain has converged to the stationary distribution we can use the sequence of values of the chain in order to obtain empirical estimates of any posterior summary of interest, such as the marginal mean, median or standard deviation. Note that the updating procedure remains relatively simple, no matter how complex the posterior distribution. Thus we can do the required integration by sampling from the posterior distribution, and we can sample from the posterior by generating a Markov chain. We can simulate Markov chains by means of different algorithms, such as the Metropolis-Hastings algorithm or the Gibbs sampler.

A.0.3 Gibbs sampling

The Gibbs sampler is one of the most widely used algorithms for sampling posterior distribution. It is a special case of the Metropolis-Hastings algorithm and generates a multidimensional Markov chain by splitting vector of random variables $\boldsymbol{\theta}$ into subvectors (often scalars) and sampling each subvector in turn, conditional on the most recent value of all other elements of $\boldsymbol{\theta}$. Let's illustrate the Gibbs sample for a model that includes three parameters, i.e. $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. Given the data $\boldsymbol{y}$, we want to compute an arbitrary large sample from the joint posterior $p(\theta_1, \theta_2, \theta_3|\boldsymbol{y})$. We assume that the full conditional distributions

$$
\begin{aligned}
&p(\theta_1|\theta_2, \theta_3, \boldsymbol{y}) \\
&p(\theta_2|\theta_1, \theta_3, \boldsymbol{y}) \\
&p(\theta_3|\theta_1, \theta_2, \boldsymbol{y})
\end{aligned}
$$

of the model's parameters are relatively easy to sample. The Gibbs sampler begins by choosing arbitrary initial values for each parameter

$$
\begin{aligned}
\theta_1 &= \theta_1^{(0)} \\
\theta_2 &= \theta_2^{(0)} \\
\theta_3 &= \theta_3^{(0)}
\end{aligned}
$$

where the superscript in parentheses denotes the order in the sequence of random draws. The Gibbs sampler proceeds as follows:

$$
\begin{aligned}
&\textbf{Step 1:} \quad \text{Sample} \quad \theta_1^{(1)} \sim p(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, \boldsymbol{y}) \\
&\textbf{Step 2:} \quad \text{Sample} \quad \theta_2^{(1)} \sim p(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \boldsymbol{y}) \\
&\textbf{Step 3:} \quad \text{Sample} \quad \theta_3^{(1)} \sim p(\theta_3|\theta_1^{(1)}, \theta_2^{(1)}, \boldsymbol{y}).
\end{aligned}
$$

This completes one iteration of the Gibbs sampler and generates a new realization of the vector of unknowns, $\boldsymbol{\theta}^{(1)}$. Steps 1 to 3 are repeated many times, always conditioning on

the most recent values of the other parameters, to obtain a sequence of realizations of the vector of unknowns, $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(T)}$, where T is always of the order of many thousands. The stationary distribution of this Markov chain is equivalent to the joint posterior $p(\boldsymbol{\theta}|\boldsymbol{y})$ provided a set of technical, regularity conditions are satisfied (this implies the discard of the beginning of the chain). To obtain a random sample of approximately independent draws from $p(\boldsymbol{\theta}|\boldsymbol{y})$, the remainder of Markov chain must be subsampled (choosing every kth draw where $k > 1$) because successive draws $\boldsymbol{\theta}^{(t)}$ and $\boldsymbol{\theta}^{(t+1)}$ are not independent.

A.0.4 Monte Carlo simulations using BUGS

BUGS stands for **B**ayesian inference **U**sing **G**ibbs **S**ampling, reflecting the basic computational technique originally adopted. The BUGS project began in 1989, and has always been featured by an explicit attempt to separate what was known as the "knowledge base", encapsulating what was assumed about the state of the world, from inference engine "used to drawn conclusions" in specific circumstances. The knowledge base naturally makes use of a "declarative" form of programming, in which the structure of the "model" for the world is described using a series of local relationships that can often be conveniently expressed as a graph. There is a clear separation between the BUGS language for specifying Bayesian models and the various programs that might be used for actually carrying out the computations. The beauty of BUGS language stands in its power to describe almost arbitrarily complex models using a very limited syntax. The BUGS language comprises syntax for functions and distributions which allow a series of logical or stochastic local relationships between nodes (i.e. variables) and its parents (variables that directly influence other variables) to be expressed. By "chaining" these relationship together, a full joint distribution over all unknown quantities can be expressed (*Lunn et al.*, 2012). Currently, there are three BUGS engines in widespread use: WinBUGS (*Lunn et al.*, 2000), OpenBUGS (*Lunn et al.*, 2009b), and JAGS (*Plummer*, 2003). For further details about BUGS language and engines see *Lunn et al.* (2012).

Appendix B

Population abundance in an endemic lizard: Supporting Information

B.0.5 Model selection

We performed Bayesian model selection using the method developed by Kuo and Mallick (1998) and followed by Royle (2009) to calculate posterior model probabilities in a very similar case. Examples of implementation are reported in Royle (2008) and Royle and Dorazio (2009). For a performance comparison of this and other variable selection methods see O'Hara and Sillanpää (2009). For selecting the four (sex-independent) models, we specified as follows a set of two indicator variables, w_j for the jth effect, having a Bernoulli(0.5) prior distribution:

$$w_k = \begin{cases} 1 \text{ if covariate } k \text{ is included in the linear predictor} \\ 0 \text{ if covariate } k \text{ is not included in the linear predictor} \end{cases}$$

with the linear predictor specified as:

$$logit(p_{ij}) = \alpha_0 \, (1 - x_{1,ij}) + w_1 \, \alpha_1 \, x_{1,ij} + w_2 \, \beta \, x_{2,i}.$$

Posterior model probabilities, for each of the four possible models, were computed using the MCMC samples, and taking the ratio between the number of iterations giving this model over the total number of iterations. We tackled the well known issue of sensitivity of posterior model probabilities to the prior specification as in Royle (2009) and Royle and Dorazio (2008). We therefore fixed the prior for parameters at the posterior distribution under the full model, using the mean and precision of the posterior distribution of each regression coefficient. Furthermore, we explored this sensitivity to priors on parameters repeating the model selection for three different priors for all regression coefficients assumed drawn from a $N(0, \sigma^2)$ with $\sigma^2 \in \{10, 100, 1000\}$.

As before, we estimated posterior model probabilities for the two candidate models with sex-dependent parameters, expanding the linear predictor with the indicator variable w_1 as:

$$logit(p_{iju}) = tr \, \alpha_{1,u} \, (1 - x_{1,ij}) + \alpha_{1,u} \, x_{1,ij} + w_1 \, \beta \, x_{2,i} + (1 - w_1) \, \beta_{1,u} \, x_{2,i}.$$

Thus, the model with a sex-independent effect of body size on p occurs if $w_1 = 1$, while a sex-specific effect of body size is selected if $w_1 = 0$. We tested the influence of parameter priors on the posterior model probabilities as before, using the three different values for variance of the parameter priors reported above.
Posterior model probabilities were calculated from three independent Markov chains run 1,000,000 times, with a burn-in of 500,000 and a thinning rate of 20.

B.0.6 Posterior summaries of model parameters for the simulated datasets

Simulated datasets were derived from different detection probabilities (indicated in the first column of the tables) both in presence (p_0, p_1) or absence (p) of behavioural response, but always with individual heterogeneity $\beta = 0.5$, and starting from a population size $N = 200$.

Table B.1: Posterior summaries of parameters for models containing only standardized body length as a covariate on detection probability (p) or the latter in addition with a permanent behavioural response (p_0, p_1). All detection probability values are given in probability scale.

Simulated data	Model	Mean	SD	2.5%	Median	97.5%
$p = 0.30$	N	200.677	24.829	165.000	197.000	259.000
	p	0.303	0.037	0.232	0.302	0.377
	β	0.547	0.174	0.210	0.546	0.894
	N	394.589	113.280	209.000	387.000	608.000
	p_0	0.139	0.053	0.072	0.127	0.266
	p_1	0.372	0.047	0.282	0.372	0.459
	β	0.514	0.181	0.155	0.513	0.870
$p = 0.60$	N	216.777	9.972	201.000	216.000	240.000
	p	0.547	0.027	0.494	0.547	0.598
	β	0.570	0.112	0.353	0.569	0.793
	N	210.720	11.688	194.000	209.000	239.000
	p_0	0.583	0.049	0.487	0.586	0.671
	p_1	0.511	0.022	0.477	0.506	0.564
	β	0.564	0.110	0.356	0.562	0.785
$p_0 = 0.30$	N	423.164	88.363	276.000	414.000	605.000
$p_1 = 0.15$	p	0.118	0.026	0.078	0.115	0.176
	β	0.544	0.199	0.133	0.552	0.905
	N	213.544	52.106	154.000	201.000	356.000
	p_0	0.294	0.071	0.158	0.294	0.432
	p_1	0.095	0.025	0.051	0.093	0.149
	β	0.617	0.195	0.238	0.614	1.002
$p_0 = 0.60$	N	274.121	26.114	234.000	271.000	336.000
$p_1 = 0.30$	p	0.325	0.029	0.268	0.325	0.383
	β	0.422	0.140	0.154	0.419	0.706
	N	199.559	9.410	186.000	198.000	221.000
	p_0	0.570	0.049	0.467	0.572	0.660
	p_1	0.276	0.028	0.221	0.276	0.331
	β	0.398	0.121	0.164	0.394	0.640

Table B.2: Posterior summaries of parameters for models with a constant detection probability (p) or a permanent behavioural response (p_0, p_1). Unlike models reported in Table B1, individual heterogeneity (β) was here never modelled. All detection probability values are given in probability scale.

Simulated data	Model	Mean	SD	2.5%	Median	97.5%
$p = 0.30$	N	178.354	13.472	156.000	177.000	208.000
	p	0.353	0.033	0.289	0.353	0.419
	N	333.552	111.607	184.000	307.000	592.000
	p_0	0.175	0.067	0.077	0.167	0.325
	p_1	0.516	0.010	0.505	0.514	0.541
$p = 0.60$	N	204.002	5.548	194.000	203.000	216.000
	p	0.557	0.025	0.507	0.558	0.607
	N	200.746	6.719	191.000	200.000	217.000
	p_0	0.587	0.046	0.496	0.589	0.672
	p_1	0.522	0.026	0.483	0.518	0.582
$p_0 = 0.30$	N	339.229	60.149	246.000	331.000	480.000
$p_1 = 0.15$	p	0.154	0.028	0.102	0.152	0.213
	N	202.908	56.296	149.000	187.000	370.000
	p_0	0.319	0.084	0.134	0.324	0.468
	p_1	0.141	0.027	0.092	0.140	0.198
$p_0 = 0.60$	N	252.685	16.598	224.000	251.000	288.000
$p_1 = 0.30$	p	0.340	0.028	0.287	0.340	0.397
	N	195.698	7.677	184.000	194.000	214.000
	p_0	0.568	0.048	0.468	0.570	0.654
	p_1	0.293	0.028	0.241	0.293	0.350

B.0.7 Posterior distributions

Posterior distributions of total and sex-specific population size, plus posterior distribution of trap response parameter under different priors.

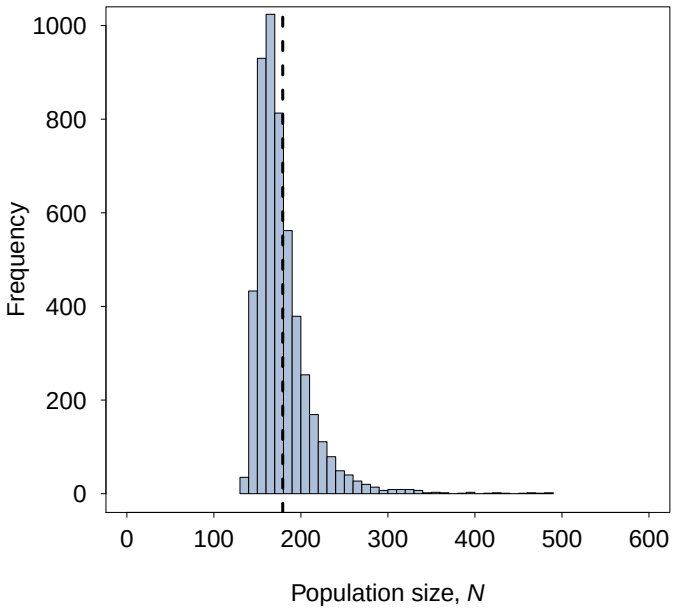

Figure B.1: Posterior distribution of population size N and its mean value (dashed line) for data of all sampled lizards.

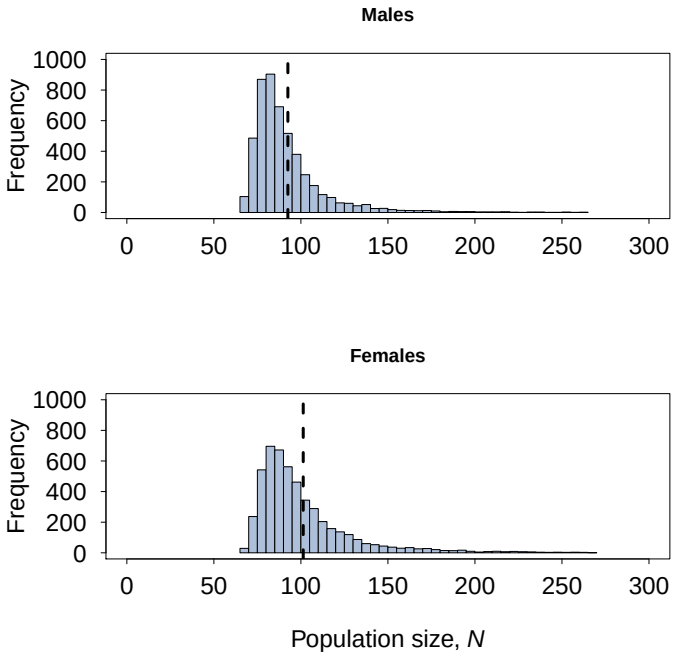

Figure B.2: Posterior distribution of population size and its mean value (dashed line) for data of sexed lizards.

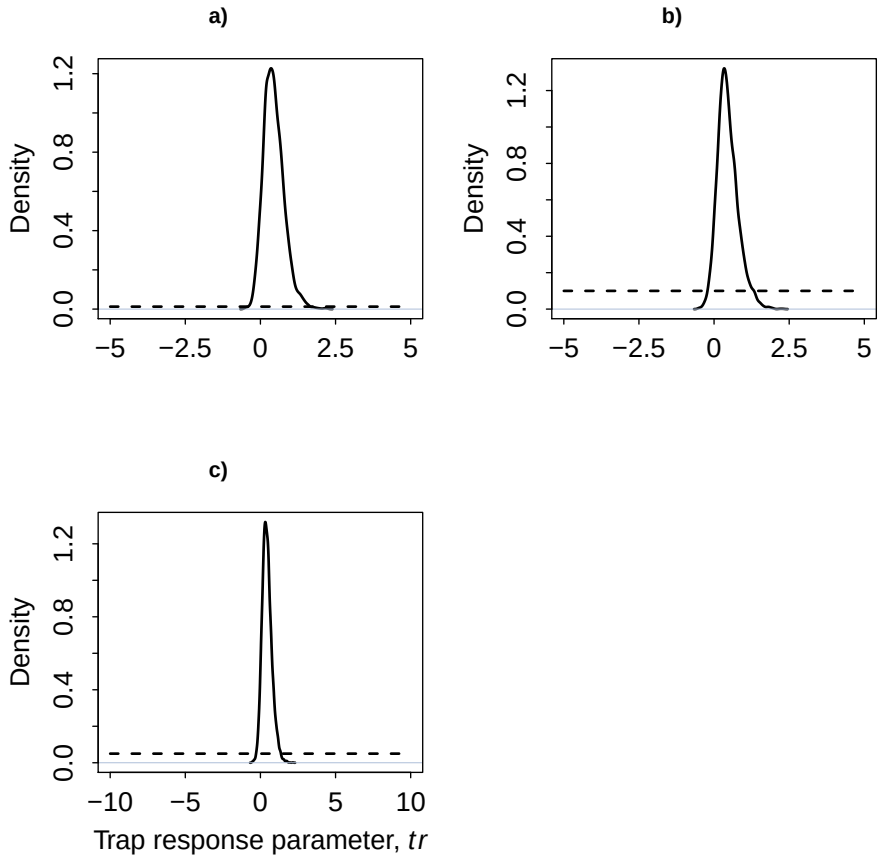

Figure B.3: Prior (dashed line) and posterior (solid line) distributions of parameter for trap response (tr). A N(0,1000) (panel a), U(-5,5) (panel b), and a U(-10,10) (panel c) were used to model data on sexed lizards. The probabilities that tr was positive are 0.925, 0.935, and 0.936 respectively.

B.0.8 R and BUGS codes

R script with the WinBUGS model specification for the sex-specific model. The directories in the code need to be customized.

```
##############################################
#                      #
# MODEL WITH SEX-SPECIFIC PARAMETERS    #
#                      #
##############################################

# example of structure of starting data
str(data)
#'data.frame':  115 obs. of  7 variables:
# $ TG    : Factor w/ 115 levels "111","122","124",..: 1 3 4 5 7 13 17 19 105 107 ...
# $ d1    : int  0 1 0 0 1 1 1 1 0 0 ...
# $ d2    : int  0 0 1 1 1 0 0 1 1 1 ...
# $ d3    : int  1 1 0 0 0 0 0 1 1 1 ...
# $ male  : int  1 1 1 1 1 1 1 1 1 1 ...
# $ female: int  0 0 0 0 0 0 0 0 0 0 ...
# $ l     : num  6.6 7.7 7.4 6.9 7.4 7.3 6.8 6.8 6.9 6.7 ...

# Starting variables
# TG: individual ID number
# d1, d2, d3: capture histories for the three capture events
# male, female: sex of each individual
# l: body length (cm)

dat_males <- data[which(data$male==1),]
dat_females <- data[-(which(data$male==1)),]

y_males <- as.matrix(dat_males[,2:4])
attr(y_males, "dimnames") <- NULL
y_females <- as.matrix(dat_females[,2:4])
attr(y_females, "dimnames") <- NULL

size_not_centered_males <- dat_males[,7]
size_not_centered_females <- dat_females[,7]

nind_males <- nrow(y_males)
nind_females <- nrow(y_females)
nind <- nind_males + nind_females
J <- ncol(y_males)

# centring covariate
size_centered_males <- size_not_centered_males - mean(size_not_centered_males, na.rm=
    TRUE)
size_centered_females <- size_not_centered_females - mean(size_not_centered_females,
    na.rm=TRUE)

# prevcap, previous capture
first_males <- rep(NA,nind_males)
prevcap_males <- matrix(0,nrow=nind_males,ncol=J)
for(i in 1:nind_males){
  first_males[i]<- min((1:J)[y_males[i,]==1])
  if(first_males[i]<J)
  prevcap_males[i,(first_males[i]+1):J]<-1
}

```

```
first_females <- rep(NA,nind_females)
prevcap_females <- matrix(0,nrow=nind_females,ncol=J)
for(i in 1:nind_females){
  first_females[i]<- min((1:J)[y_females[i,]==1])
  if(first_females[i]<J)
  prevcap_females[i,(first_females[i]+1):J]<-1
}


# data augmentation
nz_males <- 200
nz_females <- 200
nz <- nz_males + nz_females
y_males <- rbind(y_males,matrix(0, nrow=nz_males,ncol=J))
y_females <- rbind(y_females,matrix(0, nrow=nz_females,ncol=J))

size_males <- c(size_centered_males,rep(NA,nz_males))
size_females <- c(size_centered_females,rep(NA,nz_females))

prevcap_males <- rbind(prevcap_males, matrix(0,nrow=nz_males,ncol=J))
prevcap_females <- rbind(prevcap_females, matrix(0,nrow=nz_females,ncol=J))


# Aggregate augmented data
y <- rbind(y_males, y_females)
size <- c(size_males, size_females)
prevcap <- rbind(prevcap_males, prevcap_females)
sex <- c(rep(1, dim(y_males)[1]), rep(2, dim(y_females)[1]))


library("R2WinBUGS")
wb<-"/path/to/WinBUGS14"
setwd("/path/to/working/directory")


# WinBUGS model (sex-independent coefficient for body length effect)
sink("M_sexes.txt")
cat("model {

for(u in 1:2){
  psi[u]~dunif(0,1)
  alpha2[u]~dunif(-10,10)
  av.p1[u] <- 1 / (1 + exp(-alpha1[u]))
  alpha1[u] <- tr * alpha2[u]
  av.p2[u] <- 1 / (1 + exp(-alpha2[u]))
}

mu~dunif(-10,10)
tau~dgamma(.01,.01)
sigma<-sqrt(1/tau)
beta~dnorm(0,.001)
tr~dunif(-5,5)

for(i in 1:(nind+nz)){
  size[i] ~ dnorm(mu, tau)I(-10,10)
  z[i] ~ dbern(psi[sex[i]])
  for(t in 1:J){
    p[i,t] <- 1 / (1 + exp(-lp[i,t])) # lp is p on the logit scale
    lp[i,t] <- tr * alpha2[sex[i]] * (1-prevcap[i,t]) + alpha2[sex[i]] * prevcap[i,t]
        + beta * size[i]
```

```
    muY[i,t] <- p[i,t]*z[i]
    y[i,t] ~ dbern(muY[i,t])
  }
}

N1 <- sum(z[1:(nind_males+nz_males)])
N2 <- sum(z[(1+nind_males+nz_males):(nind+nz)])
}
",fill=TRUE)
sink()

# Data
win.data <- list(y = y, size = size, nind = nind, nz = nz, nind_males = nind_males, nz
    _males = nz_males, J = J, prevcap = prevcap, sex = sex  )

# Initial values
inits <- function() list(alpha2=rnorm(2), tr=runif(1, -4, 4), beta=rnorm(1),  z=rbinom
    (nind+nz,1,.5), mu=rnorm(1), tau=runif(1,1,2)  )

# Define parameters
params <- c("N1", "N2", "alpha1", "alpha2", "tr", "beta", "psi", "mu", "sigma", "av.p1
    ", "av.p2")

# MCMC settings
ni <- 100000
nt <- 30
nb <- 50000
nc <- 3

# run the model
out <- bugs(win.data, inits, params, "M_sexes.txt", n.thin = nt, n.chains = nc, n.
    burnin = nb, n.iter = ni,
            debug=TRUE, bugs.directory=wb, working.directory= getwd(), useWINE=TRUE)
```

End of code.

Appendix C

Recreational tourism and population structure of an endangered bivalve: R and BUGS codes

R and BUGS codes for modelling anchoring effect on Noble Pen Shell density. The directories in the code need to be customized.

```
1  ######################################
2  # Code to model anchoring effect  #
3  ######################################
4
5  # import transects data
6  dat <- read.csv("/path/to/folder/transects.csv", header = TRUE)
7
8  str(dat)
9  #'data.frame':  356 obs. of  18 variables:
10 # $ date             : int  43 43 43 43 43 43 43 43 43 43 ...
11 # $ day              : int  12 12 12 12 12 12 12 12 12 12 ...
12 # $ month            : int  2 2 2 2 2 2 2 2 2 2 ...
13 # $ year             : int  2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
14 # $ site             : Factor w/ 6 levels "CalaOr","EsCargol",..: 1 1 1 1 1 1 1 1 1 1
       ...
15 # $ method           : Factor w/ 2 levels "Quadrants","Transect": 2 2 2 2 2 2 2 2 2 2
       ...
16 # $ time_ded_h       : int  1 1 1 1 1 1 1 1 1 1 ...
17 # $ area_sampled_team: int  240 285 240 240 285 240 240 240 285 285 ...
18 # $ width            : num  13.2 12.3 16.6 19.9 10.7 15.1 18.2 16.6 12.3 12.4 ...
19 # $ id               : Factor w/ 446 levels "1","10","1000",..: 67 82 112 114 125 131
       134 151 162 166 ...
20 # $ team             : int  4 5 4 4 5 4 4 4 5 5 ...
21 # $ first_pass       : int  1 1 1 1 0 1 0 1 1 1 ...
22 # $ second_pass      : int  0 0 0 0 1 1 0 1 1 1 ...
23 # $ meadow_cov_1D    : num  95.7 95.7 95.7 95.7 95.7 ...
24 # $ meadow_cov_2D    : num  78.2 78.2 78.2 78.2 78.2 ...
25 # $ shoot_density    : num  384 384 384 384 384 384 384 384 384 384 ...
26 # $ anchoring        : int  0 0 0 0 0 0 0 0 0 0 ...
27 # $ season           : num  1 1 1 1 1 1 1 1 1 1 ...
28
```

```
# number of sampled individuals for each site, in relation to presence/absence of anchoring
# dat$anchoring=0 denotes no anchoring; dat$anchoring=1 denotes anchoring
table(dat$site, dat$anchoring)
#              0   1
#  CalaOr     69   0
#  EsCargol    0  24
#  EsCastell  86   0
#  Magalluf    0  22
#  Pollenca    0  24
#  StMaria   131   0

# convert anchoring from 0/1 to 1/2 for BUGS
dat$anchoring2 <- dat$anchoring + 1

# split the dataset
# no anchoring
dat_1anchoring <- dat[which(dat$anchoring2==1),]
# anchoring
dat_2anchoring <- dat[which(dat$anchoring2==2),]

# arrange data
y_1anchoring <- cbind(dat_1anchoring$first_pass, dat_1anchoring$second_pass)
y_2anchoring <- cbind(dat_2anchoring$first_pass, dat_2anchoring$second_pass)

size_not_centered_1anchoring <- dat_1anchoring$width
size_not_centered_2anchoring <- dat_2anchoring$width

nind_1anchoring <- nrow(y_1anchoring)
nind_2anchoring <- nrow(y_2anchoring)
nind <- nind_1anchoring + nind_2anchoring
J <- ncol(y_1anchoring)

# centring covariate
size_centered_1anchoring <- size_not_centered_1anchoring - mean(size_not_centered_1anchoring, na.rm=TRUE)
size_centered_2anchoring <- size_not_centered_2anchoring - mean(size_not_centered_2anchoring, na.rm=TRUE)

y_1anchoring <- apply(y_1anchoring,1,sum)
attr(y_1anchoring, "names") <- NULL
y_2anchoring <- apply(y_2anchoring,1,sum)
attr(y_2anchoring, "names") <- NULL

# data augmentation
nz_1anchoring <- 850
nz_2anchoring <- 210
nz <- nz_1anchoring + nz_2anchoring
y_1anchoring <- c(y_1anchoring, rep(0,nz_1anchoring))
y_2anchoring <- c(y_2anchoring, rep(0,nz_2anchoring))

size_1anchoring <- c(size_centered_1anchoring,rep(NA,nz_1anchoring))
size_2anchoring <- c(size_centered_2anchoring,rep(NA,nz_2anchoring))

# Aggregate augmented data
y <- c(y_1anchoring, y_2anchoring)
size <- c(size_1anchoring, size_2anchoring)

anchoring <- c(rep(1,length(y_1anchoring)), rep(2,length(y_2anchoring)))

```

```
library("R2WinBUGS")
wb<-"/path/to/folder/WinBUGS14"

# BUGS model
sink("model_anchoring.txt")
cat("model {
# Priors
for(u in 1:2){
  psi[u]~dunif(0,1)
  mu[u]~dnorm(0,.001)
  tau[u]~dgamma(.001,.001)
  sigma[u]<-sqrt(1/tau[u])
}

alpha~dnorm(0,.001)
av.p <- 1 / (1 + exp(-alpha))
beta~dnorm(0,.001)

# Likelihood
for(i in 1:(nind+nz)){
  size[i] ~ dnorm(mu[anch[i]], tau[anch[i]])
  z[i] ~ dbern(psi[anch[i]])
  p[i] <- 1 / (1 + exp(-lp[i])) # lp is detection probability (p) in logit scale
  lp[i] <- alpha + beta * size[i]
  muy[i] <- z[i]*p[i]
  y[i] ~ dbin(muy[i],J)
}

# average density for all sites
N_1anch <- sum(z[1:length_1anch])
N_2anch <- sum(z[(length_1anch+1):(nind+nz)])
D_1anch <- N_1anch/47.55      # NO anchoring (area in 100 m^2)
D_2anch <- N_2anch/59.10      # YES anchoring

# difference between densities
D_diff <- D_1anch - D_2anch
}
",fill=TRUE)
sink()

# Data
win.data <- list(y = y, size = size, nind = nind, nz = nz, J = J, anch = anchoring,
    length_1anch = length(y_1anchoring)  )

# Initial values
inits <- function() list(alpha=rnorm(1), beta=rnorm(1), psi=runif(2), z=rbinom(length(
    size),1,.5), mu=rnorm(2), tau=runif(2,1,2)  )

# Define parameters
params <- c("N_1anch", "N_2anch", "D_1anch", "D_2anch", "D_diff", "alpha", "beta", "
    psi", "mu", "sigma", "av.p")

# MCMC settings
ni <- 100000
nt <- 30
nb <- 50000
nc <- 3

# Run WinBUGS (in Linux via WINE)
out <- bugs(win.data, inits, params, "model_anchoring.txt", n.thin = nt, n.chains = nc
```

```
    , n.burnin = nb, n.iter = ni, debug=TRUE, bugs.directory=wb, working.directory=
    getwd(), useWINE=TRUE)
```

End of code.

Appendix D

Demographic cost of illegal poisoning: data and codes for the integrated population model

R and BUGS codes including the data to fit the integrated population model. The directories in the code need to be customized.

```
# The R and WinBUGS codes including the data to fit the integrated population model

# capture-recapture data (2000-2010: 11 capture events)
CH <- matrix(c(1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 2, 2, 7, 7, 7, 7, 2, 7, 1, 1, 1, 1, 1, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3,
7, 0, 0, 0, 0, 7, 0, 2, 0, 0, 2, 7, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 0, 0,
0, 0, 0, 0, 0, 3, 0, 0, 3, 0, 2, 8, 8, 2, 2, 2,
```

```
2, 2, 8, 2, 7, 8, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 11, 0, 0, 0, 0, 3, 3, 3, 3,
3, 0, 3, 0, 0, 2, 2, 2, 2, 2, 2, 2, 2, 8, 2, 9,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 11, 11, 11, 0, 0,
0, 11, 0, 0, 3, 3, 3, 3, 3, 3, 3, 3, 0, 8, 0,
9, 7, 8, 7, 2, 2, 2, 2, 0, 2, 2, 2, 2, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
15, 0, 0, 4, 4, 4, 4, 4, 4, 4, 11, 0, 0, 0, 0,
0, 0, 0, 3, 3, 0, 0, 0, 3, 3, 3, 3, 2, 2, 2, 2,
8, 8, 8, 2, 8, 7, 9, 2, 0, 0, 8, 2, 2, 2, 2, 2,
2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 5, 5, 5, 5, 5, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 4, 7, 0, 0, 0, 4, 4, 4, 4, 7, 7, 8, 7, 0, 0,
0, 7, 0, 0, 0, 0, 0, 0, 0, 3, 0, 3, 3, 3, 3, 3,
3, 3, 2, 8, 7, 8, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 0, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
13, 13, 13, 13, 13, 13, 13, 0, 0, 0, 0, 0, 0, 0,
0, 7, 0, 0, 0, 0, 12, 12, 12, 12, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 4, 4, 4, 4, 4,
4, 4, 4, 7, 0, 0, 0, 7, 3, 3, 3, 3, 0, 3, 3, 3,
3, 3, 3, 3, 0, 3, 0, 0, 0, 2, 2, 2, 2, 2, 2, 8,
0, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 13, 13, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 12, 12, 12, 12, 12,
12, 12, 12, 12, 0, 0, 0, 0, 0, 4, 4, 4, 4, 0, 4,
4, 4, 4, 4, 4, 4, 0, 4, 0, 0, 0, 3, 3, 3, 3, 0,
3, 0, 0, 3, 3, 3, 3, 3, 3, 3, 9, 8, 2, 2, 9, 9,
```

```
9, 8, 2, 2, 2, 2, 7, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 13, 0, 0, 0, 0, 0, 0, 0, 0, 14, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 13, 0, 0, 0,
0, 0, 0, 0, 13, 0, 0, 0, 0, 0, 7, 5, 0, 0, 0,
5, 5, 12, 0, 12, 0, 12, 0, 5, 0, 0, 0, 4, 4, 4,
4, 11, 7, 0, 0, 11, 4, 4, 8, 4, 8, 4, 0, 0, 10,
3, 0, 10, 0, 0, 10, 10, 3, 10, 0, 8, 2, 2, 2, 7,
2, 2, 2, 2, 2, 2, 8, 8, 7, 2, 2, 2, 9, 2), nrow = 142, byrow = FALSE)

# First capture occasion
get.first <- function(x) min(which(x!=0))
f <- apply(CH, 1, get.first)

# Recode CH observations
# 16 = not seen, not detected
rCH <- CH
rCH[rCH==0] <- 16

# read matrix of known latent states
zCH <- matrix(c(NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, 2, 2, 7, 7, 7, 7, 2, 7, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 3,
7, 16, 16, 16, 16, 7, 16, 2, NA, NA, 2, 7, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 4,
16, 16, 16, 16, 16, 16, 16, 3, NA, NA, 3, 16, 2, 8,
8, 2, 2, 2, 2, 2, 8, 2, 7, 8, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
16, 16, 16, 16, 16, 16, 16, NA, NA, NA, 11, 16, NA,
16, 16, 3, 3, 3, 3, 3, 16, 3, 16, 16, 2, 2, 2,
```

```
2, 2, 2, 2, 2, 8, 2, 9, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, 16, 16, 16, 16, 16, 16, 16, NA, NA, NA, NA,
16, NA, 16, 16, 11, 11, 11, NA, NA, 16, 11, 16, 16,
3, 3, 3, 3, 3, 3, 3, 3, 16, 8, NA, 9, 7, 8, 7,
2, 2, 2, 2, NA, 2, 2, 2, 2, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, 16, 16, 16, 16, 16, 16, 16,
NA, NA, NA, NA, 16, NA, 16, 16, NA, NA, NA, NA, NA, 16, 15,
16, 16, 4, 4, 4, 4, 4, 4, 4, 11, 16, 16, NA, NA,
16, 16, 16, 3, 3, NA, NA, NA, 3, 3, 3, 3, 2, 2, 2,
2, 8, 8, 8, 2, 8, 7, 9, 2, NA, NA, 8, 2, 2, 2, 2,
2, 2, 2, 2, 2, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 16, 16, 16, 16,
16, 16, 16, NA, NA, NA, NA, 16, NA, 16, 16, NA, NA, NA,
NA, NA, 16, 16, 16, 16, 5, 5, 5, 5, 5, NA, NA, NA, 16,
16, NA, NA, 16, 16, 16, 4, 7, NA, NA, NA, 4, 4, 4, 4,
7, 7, 8, 7, 16, 16, 16, 7, 16, 16, NA, NA, NA, NA,
16, 3, NA, 3, 3, 3, 3, 3, 3, 3, 2, 8, 7, 8, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, NA, 2, 2, 2,
2, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, 16, 16, 16, 16, 16, 16, 16, NA, NA, NA, NA, 16, NA,
16, 16, NA, NA, NA, NA, NA, 16, 16, 16, 16, 13, 13, 13,
13, 13, 13, 13, NA, 16, 16, NA, NA, 16, 16, 16, 7,
16, NA, NA, NA, 12, 12, 12, 12, 16, 16, 16, 16, 16,
16, 16, 16, 16, 16, NA, NA, NA, NA, 16, 4, 4, 4, 4,
4, 4, 4, 4, 4, 7, 16, 16, 16, 7, 3, 3, 3, 3, NA,
3, 3, 3, 3, 3, 3, 3, NA, 3, NA, NA, NA, 2, 2, 2, 2,
2, 2, 8, NA, 2, 2, 2, 2, 2, 2, 2, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 16, 16, 16,
16, 16, 16, 16, NA, NA, NA, NA, 16, NA, 16, 16, NA, NA,
NA, NA, NA, 16, 16, 16, 16, 13, 13, NA, NA, NA, NA, NA,
NA, 16, 16, NA, NA, 16, 16, 16, 16, 16, NA, NA, NA, NA,
NA, NA, NA, 16, 16, 16, 16, 16, 16, 16, 16, 16, 16,
NA, NA, NA, NA, 16, 12, 12, 12, 12, 12, 12, 12, 12,
12, 16, 16, 16, 16, 16, 4, 4, 4, 4, NA, 4, 4, 4,
4, 4, 4, NA, 4, NA, NA, NA, 3, 3, 3, 3, NA, 3, 16,
NA, 3, 3, 3, 3, 3, 3, 3, 9, 8, 2, 2, 9, 9, 9, 8,
2, 2, 2, 2, 7, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, 16, 16, 16, 16, 16,
16, 16, NA, NA, NA, NA, 16, NA, 16, 16, NA, NA, NA, NA,
NA, 16, 16, 16, 16, NA, 13, NA, NA, NA, NA, NA, NA, 16,
16, 14, NA, 16, 16, 16, 16, 16, NA, NA, NA, NA, NA, NA,
NA, 16, 16, 16, 16, 16, 16, 16, 16, 16, 16, NA, NA,
NA, NA, 16, 13, NA, NA, NA, NA, NA, NA, NA, 13, 16, 16,
```

```
16, 16, 16, 7, 5, NA, NA, NA, 5, 5, 12, NA, 12, NA,
12, NA, 5, NA, NA, NA, 4, 4, 4, 4, 11, 7, 16, NA, 11,
4, 4, 8, 4, 8, 4, NA, 16, 10, 3, NA, 10, NA, 16, 10,
10, 3, 10, 16, 8, 2, 2, 2, 7, 2, 2, 2, 2, 2, 2,
8, 8, 7, 2, 2, 2, 9, 2), nrow = 142, byrow = FALSE)

# Number of years
nyears <- 12

# Population count data
popcount <- c(4,5,2,5,8,7,13,10,14,12,13,17)

# Reproductive success data
nestlings <- c(8,8,5,12,14,11,26,20,18,17,24,19) # number of offspring
R <- c(4,5,2,5,8,7,13,10,14,12,13,17)        # number of surveyed broods

# Unmarked birds found dead by poison or other causes
rec.dead.poison <- c(0,0,1,0,0,0,0,0,0,1,1,1)
rec.dead.other <- c(1,1,0,0,0,0,1,1,0,2,3,2)

# Load R2WinBUGS package
library("R2WinBUGS")


# WinBUGS MODEL:
sink("ipm_redkite.txt")
cat("
model {
#############################################################################
#
#  Integrated population model for the red  kite population of Mallorca (Spain)
#  - Age structured, female-based model with 3 age classes:
#  - Pre-breeding census
#  - Age at first breeding: 2 years old
#
#############################################################################

#################################################################
# 1. Priors for the parameters
#################################################################

# Probabilities for telemetry data
# Priors and constraints
for (t in 1:(n.occasions-1)){
   s[t] <- mean.s
   betajuv[t] <- mean.betajuv
   beta1y[t] <- mean.beta1y
   beta2my[t] <- mean.beta2my
   alpha1[t] <- mean.alpha1
   alpha2[t] <- mean.alpha2
   alpha3[t] <- mean.alpha3
   p[t] <- mean.p
   c[t] <- mean.c
   d1[t] <- mean.d1
   d2[t] <- mean.d2
   }

mean.s ~ dunif(0, 1)
mean.betajuv ~ dunif(0, 1)
mean.beta1y ~ dunif(0, 1)
```

```
mean.beta2my ~ dunif(0, 1)
mean.alpha1 ~ dunif(0, 1)
mean.alpha2 ~ dunif(0, 1)
mean.alpha3 ~ dunif(0, 0.1)
mean.p ~ dunif(0, 1)
mean.c ~ dunif(0, 1)
mean.d1 ~ dunif(0, 0.1)
mean.d2 ~ dunif(0, 1)

# Initial population sizes
N1[1] ~ dnorm(10, 0.001)I(1,)       # 1-year old individuals
N2[1] ~ dnorm(10, 0.001)I(1,)   # 2-year old individuals
Nad[1] ~ dnorm(10, 0.001)I(1,)      # Adults > 2 years

# Proportion of adult breeders
SBad ~ dunif(0.5, 1)

# Mean fecundity (on log scale)
l.mfec ~ dnorm(0, 0.0001)I(-10,10)

# Fecundity temporal variability
sig.fec ~ dunif(0, 10)
tau.fec <- pow(sig.fec, -2)

for (t in 1:(nyears-1)){
   epsilon.fec[t] ~ dnorm(0, tau.fec)I(-15,15)
   }

# Initial population size of unmarked birds dead by poison
Ndp[1] ~ dnorm(1, 0.01)I(0,)

# Initial population size of unmarked birds dead by other causes
Ndo[1] ~ dnorm(1, 0.01)I(0,)


#################################################################
# 2. Constrain fecundity
#################################################################

for (t in 1:(nyears-1)){
   log(b[t]) <- l.mfec + epsilon.fec[t]
   }

#################################################################
# 3. Derived and fixed parameters
#################################################################

# Mean fecundity (on real scale)
mfec <- exp(l.mfec)

# Population growth rate
for (t in 1:(nyears-1)){
   lambda[t] <- Ntot[t+1] / Ntot[t]
   logla[t] <- log(lambda[t])
   }
mlam <- exp((1/(nyears-1))*sum(logla[1:(nyears-1)]))   # Geometric mean

# Survival probability in absence of illegal poisoning
Snp.juv <- 1 - (1-mean.s) * (1-mean.betajuv)
Snp.1y <- 1 - (1-mean.s) * (1-mean.beta1y)
```

```
Snp.ad <- 1 - (1-mean.s) * (1-mean.beta2my)

# Mean probability of death due to poisoning
mean.betahat <- (mean.betajuv + mean.beta1y + mean.beta2my) / 3

# Probability of dying because of poisoning
mdp <- (1-mean.s) * mean.betahat

# Probability of dying for causes other than poison
mdo <- (1-mean.s) * (1-mean.betahat)

# Complementary probability for (mdp+mdo)
msurv <- 1-mdp-mdo

#############################################################################
# 4. The likelihoods of the single data sets
#############################################################################

#############################################################################
# 4.1. Likelihood for population population count data (state-space model)
#############################################################################

   ##############################
   # 4.1.1 System process
   ##############################
      for (t in 2:nyears){
      mean1[t] <- 0.5 * b[t-1] * mean.s * (N2[t-1] * 0.10 + Nad[t-1] * SBad)
      N1[t] ~ dpois(mean1[t]) # total number of 1y

      N2[t] ~ dbin(mean.s, N1[t-1]) # total number of 2y

      tot2ad[t-1] <- N2[t-1] + Nad[t-1]
      Nad[t] ~ dbin(mean.s, tot2ad[t-1])  # total number of adults

      Npop[t-1] <- N1[t-1] + N2[t-1] + Nad[t-1]

      # unmarked birds found dead
      # Multinomial distribution with unknown order N was coded as a sequence of
          conditional univariate binomial distributions
      # For further details see WinBUGS manual (Spiegelhalter 2007)
      Ndp[t] ~ dbin(mdp, Npop[t-1])       # number of birds dead by poison
      Ndo[t] ~ dbin(mdo, Npop_minus_dp[t-1])  # number of birds dead by other causes (
          CONDITIONAL on Ndp)
      Npop_minus_dp[t-1] <- Npop[t-1] - Ndp[t]  # *(see comment below)
      Nsurv[t] ~ dbin(msurv, Npop_surv[t-1])  # number of survived birds
      Npop_surv[t-1] <- Npop[t-1] - Ndp[t] - Ndo[t] # *this makes the three
          conditional binomial as a multinomial distribution
      }                        #

   ##############################
   # 4.1.2 Observation process
   ##############################
   for (t in 1:nyears){
      Ntot[t] <- (N2[t] * 0.10 + Nad[t] * SBad)
      census[t] ~ dpois(Ntot[t])
      deaths_p[t] ~ dbin(mean.d1, Ndp[t])
      deaths_o[t] ~ dbin(mean.d2, Ndo[t])
      }

#############################################################################
```

```
# 4.2 Likelihood for capture-recapture data: multistate model (6 age classes)
################################################################################

#####################################################
# Parameters:
#
# alpha1: probability of radio signal retention during the first 3 years of life
# alpha2: probability of radio signal retention during the fourth year of life
# alpha3: probability of radio signal retention from the fifth year of life onwards
# betajuv: probability of death due to poisoning given an animal has died as juvenile
    (between the birth and the first spring)
# beta1y: probability of death due to poisoning given an animal has died as 1-year old
     (between the first and the secod spring)
# beta2my: probability of death due to poisoning given an animal has died as 2 or more
     year old
# c: probability of encounter of an animal alive without the radio signal
# p:  probability of encounter of an animal with a functioning radio
# d1: probability of encounter of an animal dead by poisoning and without the radio
    signal
# d2: probability of encounter of an animal dead by other causes and without the radio
     signal
# s: survival probability
#
#####################################################
# States (S)
# 1 alive as juvenile
# 2 alive as 1-year old with tag
# 3 alive as 2-year old with tag
# 4 alive as 3-year old with tag
# 5 alive as 4-year old with tag
# 6 alive as 5-year old or more with tag
# 7 dead by poison with tag
# 8 dead by other causes with tag
# 9 alive as 1-year old without tag
# 10 alive as 2-year old without tag
# 11 alive as 3-year old without tag
# 12 alive as 4-year old without tag
# 13 alive as 5-year old or more without tag
# 14 dead by poison without tag
# 15 dead by other causes without tag
# 16 unobserved dead
#
# Observations (O)
# 1 detected alive as juvenile
# 2 detected alive as 1-year old with tag
# 3 detected alive as 2-year old with tag
# 4 detected alive as 3-year old with tag
# 5 detected alive as 4-year old with tag
# 6 detected alive as 5-year old or more with tag
# 7 detected dead by poison with tag
# 8 detected dead by other causes with tag
# 9 detected alive as 1-year old without tag
# 10 detected alive as 2-year old without tag
# 11 detected alive as 3-year old without tag
# 12 detected alive as 4-year old without tag
# 13 detected alive as 5-year old or more without tag
# 14 detected dead by poison without tag
# 15 detected dead by other causes without tag
# 16 not detected
#####################################################
```

```
444
445
446 # Define parameters
447 for (i in 1:nind){
448    # Define probabilities of state S(t+1) given S(t); see Figure S1
449    for (t in f[i]:(n.occasions-1)){  # loop over time
450       # First index = states at time t-1, last index = states at time t
451       ps[1,i,t,1] <- 0
452       ps[1,i,t,2] <- s[t] * alpha1[t]
453       ps[1,i,t,3] <- 0
454       ps[1,i,t,4] <- 0
455       ps[1,i,t,5] <- 0
456       ps[1,i,t,6] <- 0
457       ps[1,i,t,7] <- (1-s[t]) * betajuv[t] * alpha1[t]
458       ps[1,i,t,8] <- (1-s[t]) * (1-betajuv[t]) * alpha1[t]
459       ps[1,i,t,9] <- s[t] * (1-alpha1[t])
460       ps[1,i,t,10] <- 0
461       ps[1,i,t,11] <- 0
462       ps[1,i,t,12] <- 0
463       ps[1,i,t,13] <- 0
464       ps[1,i,t,14] <- (1-s[t]) * betajuv[t] * (1-alpha1[t])
465       ps[1,i,t,15] <- (1-s[t]) * (1-betajuv[t]) * (1-alpha1[t])
466       ps[1,i,t,16] <- 0
467
468       ps[2,i,t,1] <- 0
469       ps[2,i,t,2] <- 0
470       ps[2,i,t,3] <- s[t] * alpha1[t]
471       ps[2,i,t,4] <- 0
472       ps[2,i,t,5] <- 0
473       ps[2,i,t,6] <- 0
474       ps[2,i,t,7] <- (1-s[t]) * beta1y[t] * alpha1[t]
475       ps[2,i,t,8] <- (1-s[t]) * (1-beta1y[t]) * alpha1[t]
476       ps[2,i,t,9] <- 0
477       ps[2,i,t,10] <- s[t] * (1-alpha1[t])
478       ps[2,i,t,11] <- 0
479       ps[2,i,t,12] <- 0
480       ps[2,i,t,13] <- 0
481       ps[2,i,t,14] <- (1-s[t]) * beta1y[t] * (1-alpha1[t])
482       ps[2,i,t,15] <- (1-s[t]) * (1-beta1y[t]) * (1-alpha1[t])
483       ps[2,i,t,16] <- 0
484
485       ps[3,i,t,1] <- 0
486       ps[3,i,t,2] <- 0
487       ps[3,i,t,3] <- 0
488       ps[3,i,t,4] <- s[t] * alpha1[t]
489       ps[3,i,t,5] <- 0
490       ps[3,i,t,6] <- 0
491       ps[3,i,t,7] <- (1-s[t]) * beta2my[t] * alpha1[t]
492       ps[3,i,t,8] <- (1-s[t]) * (1-beta2my[t]) * alpha1[t]
493       ps[3,i,t,9] <- 0
494       ps[3,i,t,10] <- 0
495       ps[3,i,t,11] <- s[t] * (1-alpha1[t])
496       ps[3,i,t,12] <- 0
497       ps[3,i,t,13] <- 0
498       ps[3,i,t,14] <- (1-s[t]) * beta2my[t] * (1-alpha1[t])
499       ps[3,i,t,15] <- (1-s[t]) * (1-beta2my[t]) * (1-alpha1[t])
500       ps[3,i,t,16] <- 0
501
502       ps[4,i,t,1] <- 0
503       ps[4,i,t,2] <- 0
```

```
504       ps[4,i,t,3] <- 0
505       ps[4,i,t,4] <- 0
506       ps[4,i,t,5] <- s[t] * alpha2[t]
507       ps[4,i,t,6] <- 0
508       ps[4,i,t,7] <- (1-s[t]) * beta2my[t] * alpha2[t]
509       ps[4,i,t,8] <- (1-s[t]) * (1-beta2my[t]) * alpha2[t]
510       ps[4,i,t,9] <- 0
511       ps[4,i,t,10] <- 0
512       ps[4,i,t,11] <- 0
513       ps[4,i,t,12] <- s[t] * (1-alpha2[t])
514       ps[4,i,t,13] <- 0
515       ps[4,i,t,14] <- (1-s[t]) * beta2my[t] * (1-alpha2[t])
516       ps[4,i,t,15] <- (1-s[t]) * (1-beta2my[t]) * (1-alpha2[t])
517       ps[4,i,t,16] <- 0
518
519       ps[5,i,t,1] <- 0
520       ps[5,i,t,2] <- 0
521       ps[5,i,t,3] <- 0
522       ps[5,i,t,4] <- 0
523       ps[5,i,t,5] <- 0
524       ps[5,i,t,6] <- s[t] * alpha3[t]
525       ps[5,i,t,7] <- (1-s[t]) * beta2my[t] * alpha3[t]
526       ps[5,i,t,8] <- (1-s[t]) * (1-beta2my[t]) * alpha3[t]
527       ps[5,i,t,9] <- 0
528       ps[5,i,t,10] <- 0
529       ps[5,i,t,11] <- 0
530       ps[5,i,t,12] <- 0
531       ps[5,i,t,13] <- s[t] * (1-alpha3[t])
532       ps[5,i,t,14] <- (1-s[t]) * beta2my[t] * (1-alpha3[t])
533       ps[5,i,t,15] <- (1-s[t]) * (1-beta2my[t]) * (1-alpha3[t])
534       ps[5,i,t,16] <- 0
535
536       ps[6,i,t,1] <- 0
537       ps[6,i,t,2] <- 0
538       ps[6,i,t,3] <- 0
539       ps[6,i,t,4] <- 0
540       ps[6,i,t,5] <- 0
541       ps[6,i,t,6] <- s[t] * alpha3[t]
542       ps[6,i,t,7] <- (1-s[t]) * beta2my[t] * alpha3[t]
543       ps[6,i,t,8] <- (1-s[t]) * (1-beta2my[t]) * alpha3[t]
544       ps[6,i,t,9] <- 0
545       ps[6,i,t,10] <- 0
546       ps[6,i,t,11] <- 0
547       ps[6,i,t,12] <- 0
548       ps[6,i,t,13] <- s[t] * (1-alpha3[t])
549       ps[6,i,t,14] <- (1-s[t]) * beta2my[t] * (1-alpha3[t])
550       ps[6,i,t,15] <- (1-s[t]) * (1-beta2my[t]) * (1-alpha3[t])
551       ps[6,i,t,16] <- 0
552
553       ps[7,i,t,1] <- 0
554       ps[7,i,t,2] <- 0
555       ps[7,i,t,3] <- 0
556       ps[7,i,t,4] <- 0
557       ps[7,i,t,5] <- 0
558       ps[7,i,t,6] <- 0
559       ps[7,i,t,7] <- 0
560       ps[7,i,t,8] <- 0
561       ps[7,i,t,9] <- 0
562       ps[7,i,t,10] <- 0
563       ps[7,i,t,11] <- 0
```

```
564       ps[7,i,t,12] <- 0
565       ps[7,i,t,13] <- 0
566       ps[7,i,t,14] <- 0
567       ps[7,i,t,15] <- 0
568       ps[7,i,t,16] <- 1
569
570       ps[8,i,t,1] <- 0
571       ps[8,i,t,2] <- 0
572       ps[8,i,t,3] <- 0
573       ps[8,i,t,4] <- 0
574       ps[8,i,t,5] <- 0
575       ps[8,i,t,6] <- 0
576       ps[8,i,t,7] <- 0
577       ps[8,i,t,8] <- 0
578       ps[8,i,t,9] <- 0
579       ps[8,i,t,10] <- 0
580       ps[8,i,t,11] <- 0
581       ps[8,i,t,12] <- 0
582       ps[8,i,t,13] <- 0
583       ps[8,i,t,14] <- 0
584       ps[8,i,t,15] <- 0
585       ps[8,i,t,16] <- 1
586
587
588       ps[9,i,t,1] <- 0
589       ps[9,i,t,2] <- 0
590       ps[9,i,t,3] <- 0
591       ps[9,i,t,4] <- 0
592       ps[9,i,t,5] <- 0
593       ps[9,i,t,6] <- 0
594       ps[9,i,t,7] <- 0
595       ps[9,i,t,8] <- 0
596       ps[9,i,t,9] <- 0
597       ps[9,i,t,10] <- s[t]
598       ps[9,i,t,11] <- 0
599       ps[9,i,t,12] <- 0
600       ps[9,i,t,13] <- 0
601       ps[9,i,t,14] <- (1-s[t]) * beta1y[t]
602       ps[9,i,t,15] <- (1-s[t]) * (1-beta1y[t])
603       ps[9,i,t,16] <- 0
604
605       ps[10,i,t,1] <- 0
606       ps[10,i,t,2] <- 0
607       ps[10,i,t,3] <- 0
608       ps[10,i,t,4] <- 0
609       ps[10,i,t,5] <- 0
610       ps[10,i,t,6] <- 0
611       ps[10,i,t,7] <- 0
612       ps[10,i,t,8] <- 0
613       ps[10,i,t,9] <- 0
614       ps[10,i,t,10] <- 0
615       ps[10,i,t,11] <- s[t]
616       ps[10,i,t,12] <- 0
617       ps[10,i,t,13] <- 0
618       ps[10,i,t,14] <- (1-s[t]) * beta2my[t]
619       ps[10,i,t,15] <- (1-s[t]) * (1-beta2my[t])
620       ps[10,i,t,16] <- 0
621
622       ps[11,i,t,1] <- 0
623       ps[11,i,t,2] <- 0
```

```
624       ps[11,i,t,3] <- 0
625       ps[11,i,t,4] <- 0
626       ps[11,i,t,5] <- 0
627       ps[11,i,t,6] <- 0
628       ps[11,i,t,7] <- 0
629       ps[11,i,t,8] <- 0
630       ps[11,i,t,9] <- 0
631       ps[11,i,t,10] <- 0
632       ps[11,i,t,11] <- 0
633       ps[11,i,t,12] <- s[t]
634       ps[11,i,t,13] <- 0
635       ps[11,i,t,14] <- (1-s[t]) * beta2my[t]
636       ps[11,i,t,15] <- (1-s[t]) * (1-beta2my[t])
637       ps[11,i,t,16] <- 0
638
639       ps[12,i,t,1] <- 0
640       ps[12,i,t,2] <- 0
641       ps[12,i,t,3] <- 0
642       ps[12,i,t,4] <- 0
643       ps[12,i,t,5] <- 0
644       ps[12,i,t,6] <- 0
645       ps[12,i,t,7] <- 0
646       ps[12,i,t,8] <- 0
647       ps[12,i,t,9] <- 0
648       ps[12,i,t,10] <- 0
649       ps[12,i,t,11] <- 0
650       ps[12,i,t,12] <- 0
651       ps[12,i,t,13] <- s[t]
652       ps[12,i,t,14] <- (1-s[t]) * beta2my[t]
653       ps[12,i,t,15] <- (1-s[t]) * (1-beta2my[t])
654       ps[12,i,t,16] <- 0
655
656       ps[13,i,t,1] <- 0
657       ps[13,i,t,2] <- 0
658       ps[13,i,t,3] <- 0
659       ps[13,i,t,4] <- 0
660       ps[13,i,t,5] <- 0
661       ps[13,i,t,6] <- 0
662       ps[13,i,t,7] <- 0
663       ps[13,i,t,8] <- 0
664       ps[13,i,t,9] <- 0
665       ps[13,i,t,10] <- 0
666       ps[13,i,t,11] <- 0
667       ps[13,i,t,12] <- 0
668       ps[13,i,t,13] <- s[t]
669       ps[13,i,t,14] <- (1-s[t]) * beta2my[t]
670       ps[13,i,t,15] <- (1-s[t]) * (1-beta2my[t])
671       ps[13,i,t,16] <- 0
672
673       ps[14,i,t,1] <- 0
674       ps[14,i,t,2] <- 0
675       ps[14,i,t,3] <- 0
676       ps[14,i,t,4] <- 0
677       ps[14,i,t,5] <- 0
678       ps[14,i,t,6] <- 0
679       ps[14,i,t,7] <- 0
680       ps[14,i,t,8] <- 0
681       ps[14,i,t,9] <- 0
682       ps[14,i,t,10] <- 0
683       ps[14,i,t,11] <- 0
```

```
684       ps[14,i,t,12] <- 0
685       ps[14,i,t,13] <- 0
686       ps[14,i,t,14] <- 0
687       ps[14,i,t,15] <- 0
688       ps[14,i,t,16] <- 1
689
690       ps[15,i,t,1] <- 0
691       ps[15,i,t,2] <- 0
692       ps[15,i,t,3] <- 0
693       ps[15,i,t,4] <- 0
694       ps[15,i,t,5] <- 0
695       ps[15,i,t,6] <- 0
696       ps[15,i,t,7] <- 0
697       ps[15,i,t,8] <- 0
698       ps[15,i,t,9] <- 0
699       ps[15,i,t,10] <- 0
700       ps[15,i,t,11] <- 0
701       ps[15,i,t,12] <- 0
702       ps[15,i,t,13] <- 0
703       ps[15,i,t,14] <- 0
704       ps[15,i,t,15] <- 0
705       ps[15,i,t,16] <- 1
706
707       ps[16,i,t,1] <- 0
708       ps[16,i,t,2] <- 0
709       ps[16,i,t,3] <- 0
710       ps[16,i,t,4] <- 0
711       ps[16,i,t,5] <- 0
712       ps[16,i,t,6] <- 0
713       ps[16,i,t,7] <- 0
714       ps[16,i,t,8] <- 0
715       ps[16,i,t,9] <- 0
716       ps[16,i,t,10] <- 0
717       ps[16,i,t,11] <- 0
718       ps[16,i,t,12] <- 0
719       ps[16,i,t,13] <- 0
720       ps[16,i,t,14] <- 0
721       ps[16,i,t,15] <- 0
722       ps[16,i,t,16] <- 1
723
724       # Define probabilities of O(t) given S(t); see Figure S2
725       # First index = states at time t, last index = observations at time t
726
727       po[1,i,t,1] <- 0
728       po[1,i,t,2] <- 0
729       po[1,i,t,3] <- 0
730       po[1,i,t,4] <- 0
731       po[1,i,t,5] <- 0
732       po[1,i,t,6] <- 0
733       po[1,i,t,7] <- 0
734       po[1,i,t,8] <- 0
735       po[1,i,t,9] <- 0
736       po[1,i,t,10] <- 0
737       po[1,i,t,11] <- 0
738       po[1,i,t,12] <- 0
739       po[1,i,t,13] <- 0
740       po[1,i,t,14] <- 0
741       po[1,i,t,15] <- 0
742       po[1,i,t,16] <- 1
743
```

```
744         po[2,i,t,1] <- 0
745         po[2,i,t,2] <- p[t]
746         po[2,i,t,3] <- 0
747         po[2,i,t,4] <- 0
748         po[2,i,t,5] <- 0
749         po[2,i,t,6] <- 0
750         po[2,i,t,7] <- 0
751         po[2,i,t,8] <- 0
752         po[2,i,t,9] <- 0
753         po[2,i,t,10] <- 0
754         po[2,i,t,11] <- 0
755         po[2,i,t,12] <- 0
756         po[2,i,t,13] <- 0
757         po[2,i,t,14] <- 0
758         po[2,i,t,15] <- 0
759         po[2,i,t,16] <- 1-p[t]
760
761         po[3,i,t,1] <- 0
762         po[3,i,t,2] <- 0
763         po[3,i,t,3] <- p[t]
764         po[3,i,t,4] <- 0
765         po[3,i,t,5] <- 0
766         po[3,i,t,6] <- 0
767         po[3,i,t,7] <- 0
768         po[3,i,t,8] <- 0
769         po[3,i,t,9] <- 0
770         po[3,i,t,10] <- 0
771         po[3,i,t,11] <- 0
772         po[3,i,t,12] <- 0
773         po[3,i,t,13] <- 0
774         po[3,i,t,14] <- 0
775         po[3,i,t,15] <- 0
776         po[3,i,t,16] <- 1-p[t]
777
778         po[4,i,t,1] <- 0
779         po[4,i,t,2] <- 0
780         po[4,i,t,3] <- 0
781         po[4,i,t,4] <- p[t]
782         po[4,i,t,5] <- 0
783         po[4,i,t,6] <- 0
784         po[4,i,t,7] <- 0
785         po[4,i,t,8] <- 0
786         po[4,i,t,9] <- 0
787         po[4,i,t,10] <- 0
788         po[4,i,t,11] <- 0
789         po[4,i,t,12] <- 0
790         po[4,i,t,13] <- 0
791         po[4,i,t,14] <- 0
792         po[4,i,t,15] <- 0
793         po[4,i,t,16] <- 1-p[t]
794
795         po[5,i,t,1] <- 0
796         po[5,i,t,2] <- 0
797         po[5,i,t,3] <- 0
798         po[5,i,t,4] <- 0
799         po[5,i,t,5] <- p[t]
800         po[5,i,t,6] <- 0
801         po[5,i,t,7] <- 0
802         po[5,i,t,8] <- 0
803         po[5,i,t,9] <- 0
```

```
804         po[5,i,t,10] <- 0
805         po[5,i,t,11] <- 0
806         po[5,i,t,12] <- 0
807         po[5,i,t,13] <- 0
808         po[5,i,t,14] <- 0
809         po[5,i,t,15] <- 0
810         po[5,i,t,16] <- 1-p[t]
811
812         po[6,i,t,1] <- 0
813         po[6,i,t,2] <- 0
814         po[6,i,t,3] <- 0
815         po[6,i,t,4] <- 0
816         po[6,i,t,5] <- 0
817         po[6,i,t,6] <- p[t]
818         po[6,i,t,7] <- 0
819         po[6,i,t,8] <- 0
820         po[6,i,t,9] <- 0
821         po[6,i,t,10] <- 0
822         po[6,i,t,11] <- 0
823         po[6,i,t,12] <- 0
824         po[6,i,t,13] <- 0
825         po[6,i,t,14] <- 0
826         po[6,i,t,15] <- 0
827         po[6,i,t,16] <- 1-p[t]
828
829         po[7,i,t,1] <- 0
830         po[7,i,t,2] <- 0
831         po[7,i,t,3] <- 0
832         po[7,i,t,4] <- 0
833         po[7,i,t,5] <- 0
834         po[7,i,t,6] <- 0
835         po[7,i,t,7] <- p[t]
836         po[7,i,t,8] <- 0
837         po[7,i,t,9] <- 0
838         po[7,i,t,10] <- 0
839         po[7,i,t,11] <- 0
840         po[7,i,t,12] <- 0
841         po[7,i,t,13] <- 0
842         po[7,i,t,14] <- 0
843         po[7,i,t,15] <- 0
844         po[7,i,t,16] <- 1-p[t]
845
846         po[8,i,t,1] <- 0
847         po[8,i,t,2] <- 0
848         po[8,i,t,3] <- 0
849         po[8,i,t,4] <- 0
850         po[8,i,t,5] <- 0
851         po[8,i,t,6] <- 0
852         po[8,i,t,7] <- 0
853         po[8,i,t,8] <- p[t]
854         po[8,i,t,9] <- 0
855         po[8,i,t,10] <- 0
856         po[8,i,t,11] <- 0
857         po[8,i,t,12] <- 0
858         po[8,i,t,13] <- 0
859         po[8,i,t,14] <- 0
860         po[8,i,t,15] <- 0
861         po[8,i,t,16] <- 1-p[t]
862
863
```

```
      po[9,i,t,1] <- 0
      po[9,i,t,2] <- 0
      po[9,i,t,3] <- 0
      po[9,i,t,4] <- 0
      po[9,i,t,5] <- 0
      po[9,i,t,6] <- 0
      po[9,i,t,7] <- 0
      po[9,i,t,8] <- 0
      po[9,i,t,9] <- c[t]
      po[9,i,t,10] <- 0
      po[9,i,t,11] <- 0
      po[9,i,t,12] <- 0
      po[9,i,t,13] <- 0
      po[9,i,t,14] <- 0
      po[9,i,t,15] <- 0
      po[9,i,t,16] <- 1-c[t]

      po[10,i,t,1] <- 0
      po[10,i,t,2] <- 0
      po[10,i,t,3] <- 0
      po[10,i,t,4] <- 0
      po[10,i,t,5] <- 0
      po[10,i,t,6] <- 0
      po[10,i,t,7] <- 0
      po[10,i,t,8] <- 0
      po[10,i,t,9] <- 0
      po[10,i,t,10] <- c[t]
      po[10,i,t,11] <- 0
      po[10,i,t,12] <- 0
      po[10,i,t,13] <- 0
      po[10,i,t,14] <- 0
      po[10,i,t,15] <- 0
      po[10,i,t,16] <- 1-c[t]

      po[11,i,t,1] <- 0
      po[11,i,t,2] <- 0
      po[11,i,t,3] <- 0
      po[11,i,t,4] <- 0
      po[11,i,t,5] <- 0
      po[11,i,t,6] <- 0
      po[11,i,t,7] <- 0
      po[11,i,t,8] <- 0
      po[11,i,t,9] <- 0
      po[11,i,t,10] <- 0
      po[11,i,t,11] <- c[t]
      po[11,i,t,12] <- 0
      po[11,i,t,13] <- 0
      po[11,i,t,14] <- 0
      po[11,i,t,15] <- 0
      po[11,i,t,16] <- 1-c[t]

      po[12,i,t,1] <- 0
      po[12,i,t,2] <- 0
      po[12,i,t,3] <- 0
      po[12,i,t,4] <- 0
      po[12,i,t,5] <- 0
      po[12,i,t,6] <- 0
      po[12,i,t,7] <- 0
      po[12,i,t,8] <- 0
      po[12,i,t,9] <- 0
```

```
924         po[12,i,t,10] <- 0
925         po[12,i,t,11] <- 0
926         po[12,i,t,12] <- c[t]
927         po[12,i,t,13] <- 0
928         po[12,i,t,14] <- 0
929         po[12,i,t,15] <- 0
930         po[12,i,t,16] <- 1-c[t]
931
932         po[13,i,t,1] <- 0
933         po[13,i,t,2] <- 0
934         po[13,i,t,3] <- 0
935         po[13,i,t,4] <- 0
936         po[13,i,t,5] <- 0
937         po[13,i,t,6] <- 0
938         po[13,i,t,7] <- 0
939         po[13,i,t,8] <- 0
940         po[13,i,t,9] <- 0
941         po[13,i,t,10] <- 0
942         po[13,i,t,11] <- 0
943         po[13,i,t,12] <- 0
944         po[13,i,t,13] <- c[t]
945         po[13,i,t,14] <- 0
946         po[13,i,t,15] <- 0
947         po[13,i,t,16] <- 1-c[t]
948
949         po[14,i,t,1] <- 0
950         po[14,i,t,2] <- 0
951         po[14,i,t,3] <- 0
952         po[14,i,t,4] <- 0
953         po[14,i,t,5] <- 0
954         po[14,i,t,6] <- 0
955         po[14,i,t,7] <- 0
956         po[14,i,t,8] <- 0
957         po[14,i,t,9] <- 0
958         po[14,i,t,10] <- 0
959         po[14,i,t,11] <- 0
960         po[14,i,t,12] <- 0
961         po[14,i,t,13] <- 0
962         po[14,i,t,14] <- d1[t]
963         po[14,i,t,15] <- 0
964         po[14,i,t,16] <- 1-d1[t]
965
966         po[15,i,t,1] <- 0
967         po[15,i,t,2] <- 0
968         po[15,i,t,3] <- 0
969         po[15,i,t,4] <- 0
970         po[15,i,t,5] <- 0
971         po[15,i,t,6] <- 0
972         po[15,i,t,7] <- 0
973         po[15,i,t,8] <- 0
974         po[15,i,t,9] <- 0
975         po[15,i,t,10] <- 0
976         po[15,i,t,11] <- 0
977         po[15,i,t,12] <- 0
978         po[15,i,t,13] <- 0
979         po[15,i,t,14] <- 0
980         po[15,i,t,15] <- d2[t]
981         po[15,i,t,16] <- 1-d2[t]
982
983         po[16,i,t,1] <- 0
```

```
      po[16,i,t,2] <- 0
      po[16,i,t,3] <- 0
      po[16,i,t,4] <- 0
      po[16,i,t,5] <- 0
      po[16,i,t,6] <- 0
      po[16,i,t,7] <- 0
      po[16,i,t,8] <- 0
      po[16,i,t,9] <- 0
      po[16,i,t,10] <- 0
      po[16,i,t,11] <- 0
      po[16,i,t,12] <- 0
      po[16,i,t,13] <- 0
      po[16,i,t,14] <- 0
      po[16,i,t,15] <- 0
      po[16,i,t,16] <- 1

      } #t
   } #i

# State-space model likelihood
for (i in 1:nind){
   z[i,f[i]] <- Y[i,f[i]]
   for (t in (f[i]+1):n.occasions){  # loop over time
      # State equation: draw S(t) given S(t-1)
      z[i,t] ~ dcat(ps[z[i,t-1], i, t-1,])

      # Observation equation: draw O(t) given S(t)
      Y[i,t] ~ dcat(po[z[i,t], i, t-1,])
      } #t
   } # i


#################################################################
# 4.3. Likelihood for reproductive data: Poisson regression
#################################################################

for (t in 1:(nyears-1)){
   J[t] ~ dpois(rho[t])
   rho[t] <- R[t] * b[t]
   }

} # End Model
",fill = TRUE)
sink()


# combine all data
bugs.data <- list(nyears = nyears, census = popcount,
                  deaths_p = rec.dead.poison, deaths_o = rec.dead.other,
                  J = nestlings[1:(nyears-1)], R = R[1:(nyears-1)],
                  Y = rCH,  f = f, n.occasions = dim(rCH)[2],
                  nind = dim(rCH)[1], z = zCH)

# Initial values
# Function to create initial values for unknown z (modified from [42])
ms.init.z <- function(ch, f){
  for (i in 1:dim(ch)[1]){
    for (c in (f[i]+1):dim(ch)[2]){
      if (is.na(ch[i,c])) {ch[i,c] <- sample(c(3,4,5,6,11,12,13), 1, replace = TRUE)}
      else {ch[i,c] <- NA}
```

```
          }
       }
    return(ch)
    }

inits <- function(){list(SBad = runif(1, 0.5, 1), l.mfec = rnorm(1, 0.2, 0.5), sig.fec = runif(1, 0.1, 10),
                         N1 = c(NA, round(runif((nyears-1), 8,12),0)), N2 = c(NA, round(runif((nyears-1), 6,8),0)),
                         Nad = c(NA, round(runif((nyears-1), 4,6),0)),
                         Ndp = c(NA, rep(4,nyears-1)), Ndo = c(NA, rep(4,nyears-1)),
                         mean.s = runif(1, 0.5, 0.6), mean.betajuv = runif(1, 0, 1), mean.beta1y = runif(1, 0, 1), mean.beta2my = runif(1, 0, 1),
                         mean.alpha1 = runif(1, 0.5, 1), mean.alpha2 = runif(1, 0, 0.5), mean.alpha3 = runif(1, 0, 0.1),
                         mean.p = runif(1, 0.5, 1),  mean.c = runif(1, 0, 0.5),
                         mean.d1 = runif(1, 0, 0.1), mean.d2 = runif(1, 0, 0.5),
                         z = ms.init.z(zCH, f)
                         )}

# Define parameters to be monitored
parameters <- c("b", "lambda", "mfec", "mlam", "sig.fec", "N1", "N2", "Nad", "Ntot",
                "mean.s", "mean.betajuv", "mean.beta1y", "mean.beta2my",
                "mean.alpha1", "mean.alpha2", "mean.alpha3",
                "mean.p", "mean.c", "mean.d1", "mean.d2",
                "SBad", "Snp.juv", "Snp.1y", "Snp.ad", "mean.betahat",
                "Npop", "Ndp", "mdp", "Ndo", "mdo" )

# MCMC settings
niter <- 1000000
nthin <- 20
nburn <- 500000
nchains <- 3

# Call WinBUGS from R (on a unix machine)
out <- bugs(bugs.data, inits, parameters, "ipm_redkite.txt", n.chains = nchains, n.thin = nthin, n.iter = niter, n.burnin = nburn, debug = TRUE, bugs.directory = wb, working.directory = getwd(), useWINE = TRUE)
```

End of code.

Appendix E

Variable and Model selection codes

E.0.9 Product space code

R and BUGS codes for the product space method. The directories in the code need to be customized.

```
1
2  ##############################
3  #
4  # Product space method
5  #
6  ##############################
7
8
9  # read data
10 d <- read.csv("/path/to/folder/data_pinna.csv", header=TRUE)
11
12 library("R2jags")
13
14 sink("MS.txt")
15 cat("
16 data{
17 C <- 100
18 for (i in 1:n.ind) {
19   zeros[i] <- 0
20 }
21 }
22
23 model {
24
25 ######### MODEL INDEX
26 M ~ dcat(p[])
27 p[1] <- prior1
28 p[2] <- 1-prior1
29 postr1 <- 2-M
30 postr2 <- M-1
31
32
33 ######### PRIORS and PSEUDOPRIORS
34
35 # sigma of Normal model
```

```
sigmaN ~ dunif(PROP.MIN.sigmaN, PROP.MAX.sigmaN)
PROP.MIN.sigmaN <- (M-1) * N.sigma.min.ps
PROP.MAX.sigmaN <- (2-M) * 10 + (M-1) * N.sigma.max.ps

# sigma of Gamma model
tauG <- 1 / (sigmaG * sigmaG)
sigmaG ~ dunif(PROP.MIN.sigmaG, PROP.MAX.sigmaG)
PROP.MIN.sigmaG <- (2-M) * G.sigma.min.ps
PROP.MAX.sigmaG <- (M-1) * 10 + (2-M) * G.sigma.max.ps

# alpha of both models...
for (j in 1:n.sites){
  alphaN[j] ~ dnorm(mu.siteN, tau.siteN)
  alphaG[j] ~ dnorm(mu.siteG, tau.siteG)
}

# ... Normal model
mu.siteN ~ dnorm(PROP.MU.mu.siteN, PROP.TAU.mu.siteN)
PROP.MU.mu.siteN <- (M-1) * N.mu.site.mu.ps
PROP.TAU.mu.siteN <- (2-M) * 0.001 + (M-1) * N.mu.site.tau.ps

tau.siteN <- 1 / (sigma.siteN * sigma.siteN)

sigma.siteN ~ dunif(PROP.MIN.sigma.siteN, PROP.MAX.sigma.siteN)
PROP.MIN.sigma.siteN <- (M-1) * N.sigma.site.min.ps
PROP.MAX.sigma.siteN <- (2-M) * 10 + (M-1) * N.sigma.site.max.ps

# ... Gamma model
mu.siteG ~ dnorm(PROP.MU.mu.siteG, PROP.TAU.mu.siteG)
PROP.MU.mu.siteG <- (2-M) * G.mu.site.mu.ps
PROP.TAU.mu.siteG <- (M-1) * 0.001 + (2-M) * G.mu.site.tau.ps

tau.siteG <- 1 / (sigma.siteG * sigma.siteG)

sigma.siteG ~ dunif(PROP.MIN.sigma.siteG, PROP.MAX.sigma.siteG)
PROP.MIN.sigma.siteG <- (2-M) * G.sigma.site.min.ps
PROP.MAX.sigma.siteG <- (M-1) * 10 + (2-M) * G.sigma.site.max.ps


######### LIKELIHOOD
for (i in 1:n.ind) {
  zeros[i] ~ dpois(zeros.mean[i])
  zeros.mean[i] <- -loglik[M,i] + C

  # Normal model
  loglik[1,i] <- loglik1[i]
  loglik1[i] <- -0.5 * log(2 * 3.14159) - 0.5 * log(sigmaN * sigmaN) - 0.5 * pow((
      width[i]-mu[i]), 2) / (sigmaN * sigmaN)

  # Gamma model
  loglik[2,i] <- loglik2[i]
  loglik2[i] <- (mu[i]*tauG) * log(tauG) + ((mu[i]*tauG)-1) * log(width[i]) - tauG *
      width[i] - loggam((mu[i]*tauG))

  mu[i] <- (2-M) * alphaN[site[i]] + (M-1) * exp(alphaG[site[i]])   #NB: alphaG is
      equal to alphaN in log scale (same for mu.siteN and mu.siteG)
}

}
",fill=TRUE)
```

```
sink()

# load pseudopriors (import all objects from 'N.sigma.min.ps' to 'G.sigma.site.max.ps' reported in bugs.data)
load("/path/to/folder/Pseudo_Norm_REanova_trick.Rdata")
load("/path/to/folder/Pseudo_Gamma_REanova.trick.Rdata")

# Bundle data
bugs.data <- list(width = d$W, site = d$site, n.sites = length(unique(d$site)), n.ind = dim(d)[1],
                  N.sigma.min.ps = N.sigma.min.ps, N.sigma.max.ps = N.sigma.max.ps,
                  N.mu.site.mu.ps = N.mu.site.mu.ps, N.mu.site.tau.ps = N.mu.site.tau.ps,
                  N.sigma.site.min.ps = N.sigma.site.min.ps, N.sigma.site.max.ps = N.sigma.site.max.ps,
                  G.sigma.min.ps = G.sigma.min.ps, G.sigma.max.ps = G.sigma.max.ps,
                  G.mu.site.mu.ps = G.mu.site.mu.ps, G.mu.site.tau.ps = G.mu.site.tau.ps,
                  G.sigma.site.min.ps = G.sigma.site.min.ps, G.sigma.site.max.ps = G.sigma.site.max.ps,
                  prior1 = 1e-10
                  )

# Inits function
inits <- function(){ list(M = sample(c(1,2), 1), sigmaN = runif(1, N.sigma.min.ps, N.sigma.max.ps), sigmaG = runif(1, G.sigma.min.ps, G.sigma.max.ps),
                          mu.siteN = rnorm(1, N.mu.site.mu.ps, 1/sqrt(N.mu.site.tau.ps)), sigma.siteN = runif(1, N.sigma.site.min.ps, N.sigma.site.max.ps),
                          mu.siteG = rnorm(1, G.mu.site.mu.ps, 1/sqrt(G.mu.site.tau.ps)), sigma.siteG = runif(1, G.sigma.site.min.ps, G.sigma.site.max.ps)
                          )}


# Params to estimate
params <- c("postr1", "postr2")

# MCMC settings
ni <- 1001000
nt <- 50
nb <- 1000
nc <- 3

# Call JAGS
out <- jags(bugs.data, inits, params, "MS.txt", n.chains = nc, n.thin = nt, n.iter = ni, n.burnin = nb, working.directory = getwd())



#### Visualize prior and posterior model probabilities
Mpost <- out$BUGSoutput$sims.array[,,"postr1"]
M <- 2-Mpost
prob <- matrix(,2,5,dimnames=list(c("M1","M2"),c("pr","ps1","ps2","ps3","psc")));
prior <- c(bugs.data$prior1, 1-bugs.data$prior1)
prob[,1] <- prior

nchains <- 3
for(m in 1:2){
for(ch in 1:nchains){
```

```
  prob[m,ch+1]=mean(M[,ch]==m)
  }
}

postr <- rowMeans(prob[,2:4])
prob[,5] <- (postr/prior)/sum(postr/prior)

# print optimal prior probabilities (pr), observed posterior probabilities (ps1-ps3
    for three chains), and corrected posterior probabilities
prob

#### Log Bayes factors(12)
postr <- prob[,5]
logb <- log(postr[1])-log(postr[2])
logb
```

End of code.

E.0.10 GVS code

R and BUGS codes for the Gibbs Variable Selection. The directories in the code need to be customized.

```

##############################
#
# Gibbs variable selection
#
##############################

# read data
d <- read.csv("/path/to/folder/data_pinna.csv", header=TRUE)

library("R2jags")

# load pseudopriors (containing objects 'mu.beta.ps' and 'tau.beta.ps')
load("/path/to/folder/Pseudo_fullmodel.Rdata")

# model
sink("gvs.txt")
cat("
model {
### Priors
alpha ~ dnorm(0, 0.001)
beta ~ dnorm(beta.mean.prior, beta.tau.prior)
# GVS priors for beta
beta.mean.prior <- (1-g[1]) * mu.beta.ps
beta.tau.prior <- (1-g[1]) * tau.beta.ps + g[1] * 0.001

for(g in 1:n.sites){
  eps[g] ~ dnorm(0, tau.site)
}

tau.site <- 1/(sigma.site * sigma.site)
sigma.site ~ dunif(0, 5)

# Priors for variable indicators
```

```
g[1] ~ dbern(0.5)
g[2] ~ dbern(0.5)

### Likelihood
for (i in 1:n.ind){
  C[i] ~ dbern(p[i])
  p[i] <- 1 / (1 + exp(-lp[i]))
  lp[i] <- alpha + g[1] * beta * width[i] + g[2] * eps[site[i]]
}

### Model code
mdl <- 1 + g[1] * 1 + g[2] * 2

### Vector with model indicators
for (j in 1:nmodels) {
  pmdl[j] <- equals(mdl, j)
}

} # end model
",fill = TRUE)
sink()

# Bundle data
bugs.data <- list(C = d$CAPT, n.ind = dim(d)[1], site = d$site, width = (d$W-mean(d$W)
    )/sd(d$W), n.sites= length(unique(d$site)),
                  mu.beta.ps = mu.beta.ps, tau.beta.ps = tau.beta.ps, nmodels = 2^2)

# Function to generate initial values
inits <- function() list(alpha = runif(1, -2, 2), beta = runif(1, -2, 2), sigma.site =
     runif(1, 0, 2), g = rbinom(2,1,0.5) )

# Define parameters to be monitored
params <- c("alpha", "beta", "sigma.site",
            "g", "pmdl" )

# MCMC settings
ni <- 50000
nt <- 5
nb <- 25000
nc <- 3

# Call JAGS
out <- jags(bugs.data, inits, params, "gvs.txt", n.chains = nc, n.thin = nt, n.iter =
    ni, n.burnin = nb, working.directory = getwd())


# alternatively, getting posterior model probabilities from variable indicators
g_index <- as.data.frame(out$BUGSoutput$sims.list$g)
g_index$comb <- paste(g_index$V1, g_index$V2, sep="")
prop.table(table(g_index$comb))
```

End of code.

Appendix F

Dynamic occupancy models: advantages of the state-space formulation and model script

F.1 Advantages of the state-space formulation

The state-space formulation of dynamic occupancy models provides several important advantages over the likelihood-based approach. Following *Royle and Kéry* (2007) we want to highlight some of them. First, it yields a generic and flexible framework for modelling individual (i.e. "site") effects, or other latent structure in parameters (e.g. random year effects). Second, although the main object of inference may be the dynamic parameters (local colonization and extinction probabilities), sometimes summaries of the partial realization of the latent process itself will be of interest (e.g. the number of occupied sites from among those that were sampled). In this case sample quantities are typically more precisely estimated than population parameters. Third, the state-space approach generalizes recent advances in species distribution models, and it has recently been described for analysis of species distribution, where the true occupancy and detection in response to variations in sampling effort are dealt with different levels of the hierarchy of spatial models. These models outperformed more conventional, non-spatial and non-hierarchical models (GLM, GAM) in terms of the prediction of observed occurrences.

F.2 Dynamic occupancy model: R and BUGS codes

R script with the JAGS model specification for a dynamic occupancy model with the effect of elevation for both site survival (ϕ) and colonization (γ) probabilities, in addition to an additive effect of sampling effort and survey date on detection probability (p).

```
1 ### import detection/non detection data
```

```
y <- dget(file="/path/to/data/scorus_3Ddata.txt")

### import sampling effort
effort_1 <- dget(file="/path/to/data/scorus_3Deffort.txt")

### import site data
sites <- read.csv("/path/to/data/sites.csv", header=TRUE)

### structure of the imported datsets
str(y)
# int [1:32, 1:18, 1:14] NA 0 NA NA NA NA NA NA NA NA ...

str(effort_1)
# num [1:32, 1:18, 1:14] 0 4 0 0 0 0 0 0 0 0 ...

str(sites)
#'data.frame':  32 obs. of  5 variables:
# $ station: Factor w/ 32 levels "Balboutet","Bocca di Caset",..: 1 2 3 4 5 6 7 8 9 10
    ...
# $ site   : int  1 2 3 4 5 6 7 8 9 10 ...
# $ lat    : num  45 45.9 45.5 45.6 44.1 ...
# $ long   : num  7.02 10.68 10.83 9.73 7.73 ...
# $ elev   : int  1600 1608 330 190 1650 200 190 1774 1200 850 ...

### standardize elevation
stand_elev <- (sites$elev - mean(sites$elev))/sd(sites$elev)

### survey date
surv <- c(1:dim(y)[2])

### standardize survey date
stand_surv <- (surv-mean(surv))/sd(surv)

# standardize sampling effort
meaneffort_1 <- mean(effort_1, na.rm = TRUE)
sdeffort_1 <- sd(effort_1, na.rm = TRUE)
stand_effort_1 <- (effort_1 - meaneffort_1) / sdeffort_1


library("R2jags")

### Dynamic occupancy model: gamma(ELEVATION) phi(ELEVATION) p(EFFORT+TIME)

sink("occ_model.txt")
cat("
model {

for(i in 1:nsite){
  muZ[i,1] <- 0
}

### priors
psi~dunif(0,1)

alpha.p ~ dunif(-20,20)
beta1.p ~ dunif(-5,5)
beta2.p ~ dunif(-5,5)

for(t in 1:(nyear-1)){
  a[t] ~ dunif(-20,20)
```

```
  b[t] ~ dunif(-20,20)
}

beta1.muZ ~ dunif(-5,5)
beta2.muZ ~ dunif(-5,5)

### state process
for(i in 1:nsite){
  z[i,1] ~ dbern(psi)
    for(t in 2:nyear){
      # covariate effects with auto-logistic parameterisation
      muZ[i,t] <- 1 / (1 + exp(-lmuZ[i,t]))
      lmuZ[i,t] <- a[t-1] + b[t-1] * z[i,t-1] + beta1.muZ * ELEVATION[i] + beta2.muZ *
           ELEVATION[i] * z[i,t-1]
      z[i,t] ~ dbern(muZ[i,t])
    }
}

### observation process
for(i in 1:nsite){
  for(j in 1:nrep){
    for (t in 1:nyear){
      Py[i,j,t] <- z[i,t] * p[i,j,t]
      y[i,j,t] ~ dbern(Py[i,j,t])

      # covariate effect on detection probability
      p[i,j,t] <- 1 / (1 + exp(-lp[i,j,t]))
      lp[i,j,t] <- alpha.p + beta1.p * TIME[j] + beta2.p * EFFORT[i,j,t]
    }
  }
}

### additional, dervived parameters
# derive gamma and phi
for(t in 1:(nyear-1)){
  gamma[t] <- 1 / (1 + exp(-a[t]))
  phi[t] <- 1 / (1 + exp(-(a[t]+b[t])))
  eps[t] <- 1-phi[t]
}

# covariate effect on phi
effect_on_phi <- beta1.muZ + beta2.muZ

}
",fill=TRUE)
sink()

# Data
bugs.data <- list(y = y, nsite = dim(y)[1], nrep = dim(y)[2], nyear = dim(y)[3],
    ELEVATION = stand_elev, TIME = stand_surv, EFFORT = stand_effort_1)

# Initial values
zst <- apply(y, c(1, 3), function(x)sum(x, na.rm=TRUE))
zst1 = array(NA, c(dim(zst)[1], dim(zst)[2]))
for (i in 1:dim(zst)[1]){
  for(k in 1:dim(zst)[2]){
    if (zst[i,k]>0) {zst1[i,k] <- 1}
    else {zst1[i,k] <- zst[i,k]}
  }
}
```

```
119
120 inits <- function(){ list(z = zst1, alpha.p = runif(1, -5, 5), beta1.p = runif(1, -4,
        4), beta2.p = runif(1, -4, 4),
121                         a = runif((dim(y)[3]-1), -5, 5), b = runif((dim(y)[3]-1),
                                -5, 5),
122                         beta1.muZ = runif(1, -4, 4), beta2.muZ = runif(1, -4, 4))}
123
124 # Parameters
125 params <- c("gamma","phi", "eps", "beta1.p", "beta2.p", "beta1.muZ", "beta2.muZ", "
        effect_on_phi")
126
127 # MCMC settings
128 ni <- 10000
129 nt <- 10
130 nb <- 5000
131 nc <- 3
132
133 runif(1)  # JAGS work around for the "Error: object '.Random.seed' not found""
134
135 # Call JAGS from R
136 out <- jags(bugs.data, inits, params, "occ_model.txt", n.chains = nc, n.thin = nt, n.
        iter = ni, n.burnin = nb, working.directory = getwd())
137
138 out
```

End of code.

Bibliography

Abadi, F., Regulation of local bird populations: Deeper insights from integrated population models, Ph.D. thesis, University of Bern, 2010.

Abadi, F., O. Gimenez, R. Arlettaz, and M. Schaub, An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence, *Ecology*, *91*, 7–14, 2010a.

Abadi, F., O. Gimenez, B. Ullrich, R. Arlettaz, and M. Schaub, Estimation of immigration rate using integrated population models, *Journal of Applied Ecology*, *47*, 393–400, 2010b.

Adrover, J., A. Muñoz, and J. Riera, Situació de la milana *Milvus milvus* a Mallorca., *Anuari Ornitologic de les Balears*, *17*, 41–45, 2002.

Ando, T., *Bayesian Model Selection and Statistical Modeling*, Taylor & Francis US, 2010.

Backhurst, M., and R. Cole, Biological impacts of boating at Kawau Island, north-eastern New Zealand, *Journal of Environmental Management*, *60*, 239–251, 2000.

Balaguer, P., A. Diedrich, R. Sardá, M. Fuster, B. Cañellas, and J. Tintoré, Spatial analysis of recreational boating as a first key step for marine spatial planning in Mallorca (Balearic Islands, Spain), *Ocean and Coastal Management*, *54*, 241–249, 2011.

Begon, M., J. L. Harper, and C. R. Townsend, *Ecology. Individuals, Populations and Communities*, Blackwell, Cambridge, 1990.

Beissinger, S. R., and D. R. McCullough, *Population viability analysis*, University of Chicago Press, 2002.

Beissinger, S. R., and I. M. Westphal, On the use of demographic models of population viability in endangered species management, *Journal of Wildlife Management*, *62*, 821–841, 1998.

Belda, E. J., and A. Sanchez, Seabids mortality on longline fisheries in the western mediterranean: factors affecting bycatch and proposed mitigation measures., *Biological Conservation*, *98*, 357–363, 2001.

Berger, J. O., and L. R. Pericchi, The intrinsic bayes factor for model selection and prediction, *Journal of The American Statistical Association*, *91*, 109–122, 1996.

Berliner, L. M., Hierarchical bayesian time series models, in *Maximum Entropy and Bayesian Methods*, pp. 15–22, 1996.

Besbeas, P., S. N. Freeman, B. J. T. Morgan, and E. A. Catchpole, Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters, *Biometrics*, *58*, 540–547, 2002.

Besbeas, P., J. . Lebreton, and B. J. T. Morgan, The efficient integration of abundance and demographic data, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *52*, 95–102, 2003.

Bouma, T., S. Olenin, K. Reise, and T. Ysebaert, Ecosystem engineering and biodiversity in coastal sediments: posing hypotheses, *Helgoland Marine Research*, *63*, 95–106, 2009.

Brook, B. W., J. J. O'Grady, A. P. Chapman, M. A. Burgman, H. R. Akcakaya, and R. Frankham, Predictive accuracy of population viability analysis in conservation biology, *Nature*, *404*, 385–387, 2000.

Brooks, S. P., and A. Gelman, Alternative methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics*, *7*, 434–455, 1998.

Brooks, S. P., R. King, and B. J. T. Morgan, A Bayesian approach to combining animal abundance and demographic data, *Animal Biodiversity and Conservation*, *27*, 515–529, 2004.

Bruderer, B., *Animal migration, navigation and homing*, chap. Effects of alpine topography and winds on migrating birds, pp. 252–265, 1978.

Buckland, S. T., *Introduction to distance sampling: estimating abundance of biological populations*, Oxford University Press, 2001.

Buckley, L. B., and W. Jetz, Insularity and the determinants of lizard population density, *Ecology Letters*, *10*, 481–489, 2007.

Burnham, K. P., and D. R. Anderson, *Model selection and inference. A practical information-theoretic approach*, Springer, 2nd edition. New York, 2002.

Butchart, S. H. M., A. J. Stattersfield, L. A. Bennun, S. M. Shutes, H. R. Akçakaya, J. E. M. Baillie, S. N. Stuart, C. Hilton-Taylor, and G. M. Mace, Measuring global trends in the status of biodiversity: Red list indices for birds, *PLoS Biology*, *2*, e383–e383, 2004.

Cam, E., 'Each site has its own survival probability, but information is borrowed across sites to tell us about survival in each site': random effects models as means of borrowing strength in survival studies of wild vertebrates, *Animal Conservation*, *15*, 129–132, 2012.

Cardona, M., M. Colomer, M. Pérez-Jiménez, D. Sanuy, and A. Margalida, Modeling ecosystems using p systems: The bearded vulture, a case study, in *Membrane Computing*, edited by D. Corne, P. Frisco, G. Paun, G. Rozenberg, and A. Salomaa, vol. 5391 of *Lecture Notes in Computer Science*, pp. 137–156, Springer Berlin / Heidelberg, 2009.

Carlin, B. P., and S. Chib, Bayesian model choice via markov chain monte carlo methods, *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, pp. 473–484, 1995.

Carrete, M., J. A. Donázar, and A. Margalida, Density-dependent productivity depression in pyrenean bearded vultures: implications for conservation, *Ecological Applications*, *16*, 1674–1682, 2006.

Caswell, H., *Matrix population models: construction, analysis, and interpretation*, Sinauer Associates, 2001.

Cave, V. M., R. King, and S. N. Freeman, An integrated population model from constant effort bird-ringing data, *Journal of Agricultural, Biological, and Environmental Statistics*, *15*, 119–137, 2010.

Ceccherelli, G., D. Campo, and M. Milazzo, Short-term response of the slow growing seagrass posidonia oceanica to simulated anchor impact, *Marine Environmental Research*, *63*, 341–349, 2007.

Chao, A., H. Pan, and S. Chiang, The PetersenLincoln estimator and its extension to estimate the size of a shared population, *Biometrical Journal*, *50*, 957970, 2008.

Chipman, H., Bayesian variable selection with related predictors, *Canadian Journal of Statistics*, *24*, 17–36, 1996.

Choquet, R., A.-M. Reboulet, J.-D. Lebreton, O. Gimenez, and R. Pradel, *U_Care 2.2 : User's Manual http://ftp.cefe.cnrs.fr/biom/Soft-CR/.*, 2002.

CITTIB, Govern de les Illes Balears, Palma de Mallorca, *Tech. rep.*, ed. G.d.l.I.B. Conselleria de Turisme, Palma de Mallorca, 2009.

Clark, J. S., Why environmental scientists are becoming bayesians, *Ecology Letters*, *8*, 2–14, 2005.

Clark, J. S., and O. N. Bjørnstad, Population time series: Process variability, observation errors, missing values, lags, and hidden states, *Ecology*, *85*, 3140–3150, 2004.

Clark, J. S., and A. E. Gelfand, *Hierarchical Modelling for the Environmental Sciences: Statistical methods and applications*, Oxford University Press, UK, 2006.

Coleman, F. C., and S. L. Williams, Overexploiting marine ecosystem engineers: potential consequences for biodiversity, *Trends in Ecology & Evolution*, *17*, 40–44, 2002.

Coltman, D. W., P. O'Donoghue, J. T. Jorgenson, J. T. Hogg, C. Strobeck, and M. Festa-Bianchet, Undesirable evolutionary consequences of trophy hunting, *Nature*, *426*, 655–658, 2003.

Congdon, P., *Bayesian Statistical Modelling*, John Wiley & Sons, 2006.

Coulson, T., E. A. Catchpole, S. D. Albon, B. J. Morgan, J. M. Pemberton, T. H. Clutton-Brock, M. J. Crawley, and B. T. Grenfell, Age, sex, density, winter weather and population crashes in soay sheep., *Science*, *292*, 1528–1531, 2001a.

Coulson, T., G. M. Mace, E. Hudson, and H. Possingham, The use and abuse of population viability analysis, *Trends in Ecology & Evolution*, *16*, 219–221, 2001b.

Cramp, S., P. Shajahan, A. Doris, D. Nelson, and M. Taylor, *The Complete Birds of the Western Palearctic*, Oxford University Press, 2000.

De Valpine, P., and A. Hastings, Fitting population models incorporating process noise and observation error, *Ecological Monograph*, *72*, 57–76, 2002.

Del Hoyo, J., A. Elliott, and J. Sargatal, *Handbook of the Birds of the World. Vol 3. Hoatzin to Auks*, Lynx Edicions, Barcelona, 1996.

Dellaportas, P., J. Forster, and I. Ntzoufras, Bayesian variable selection using the gibbs sampler, in *Generalized Linear Models: A Bayesian Perspective*, edited by D. Dey, S. Ghosh, and B. Mallick, 5, pp. 273–286, Chemical Rubber Company Press, 2000.

Dellaportas, P., J. J. Forster, and I. Ntzoufras, On bayesian model and variable selection using mcmc, *Statistics and Computing*, *12*, 27–36, 2002.

DeLong, J. P., S. W. Cox, and N. S. Cox, A comparison of avian use of high- and low-elevation sites during autumn migration in Central New Mexico, *Journal of Field Ornithology*, *76*, 326–333, 2005.

Duarte, C. M., Marine biodiversity and ecosystem services: an elusive link, *Journal of Experimental Marine Biology and Ecology*, *250*, 117–131, 2000.

Duarte, C. M., The future of seagrass meadows, *Environmental Conservation*, *29*, 192–206, 2002.

EEC, Council directive on the conservation of natural habitats and of wild fauna and flora (the habitats and species directive), *Tech. rep.*, EEC, Official Journal of the European Communities, 1992.

Galán, P., Structure of a population of the lizard podarcis bocagei in northwest spain: variations in age distribution, size distribution and sex-ratio, *Animal Biology*, *54*, 57–75, 2004.

García-Fernández, A. J., P. María-Mojica, E. Martínez-López, D. Romero, I. Navas, A. Hernández-García, and P. Gómez-Ramírez, Aspectos clínicos y forenses del envenenamiento de aves silvestres: diferencias entre aldicarb y estricnina, *Revista de Toxicología*, *23*, 44–48, 2006.

García-March, J., L. Pérez-Rojas, and A. García-Carrascosa, Influence of hydrodynamic forces on population structure of Pinna nobilis L., 1758 (Mollusca: Bivalvia): The critical combination of drag force, water depth, shell size and orientation, *Journal of Experimental Marine Biology and Ecology*, *342*, 202–212, 2007.

Gardner, B., J. Reppucci, M. Lucherini, and J. A. Royle, Spatially explicit inference for open populations: estimating demographic parameters from camera-trap studies, *Ecology*, *91*, 3376–3383, 2010a.

Gardner, B., J. A. Royle, M. T. Wegan, R. E. Rainbolt, and P. D. Curtis, Estimating black bear density using dna data from hair snares, *The Journal of Wildlife Management*, *74*, 318–325, 2010b.

Gelfand, A. E., and A. F. M. Smith, Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, *85*, 398–409, 1990.

Gelman, A., Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis*, *1*, 1–19, 2006.

Gelman, A., and J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, 2006.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, CRC Press, 2003.

George, E. I., and R. E. McCulloch, Variable selection via gibbs sampling, *Journal of The American Statistical Association*, *88*, 881–889, 1993.

Gimenez, O., V. Rossi, R. Choquet, C. Dehais, B. Doris, H. Varella, J.-P. Vila, and R. Pradel, State-space modelling of data on marked individuals, *Ecological Modelling*, *206*, 431–438, 2007.

Gimenez, O., et al., The risk of flawed inference in evolutionary studies when detectability is less than one, *The American Naturalist*, *172*, 441–448, 2008.

Giorgi, F., and P. Lionello, Climate change projections for the mediterranean region, *Global and Planetary Change*, *63*, 90–104, 2008.

Godsill, S. J., On the relationship between markov chain monte carlo methods for model uncertainty, *Journal of Computational and Graphical Statistics*, *10*, 230–248, 2001.

González, L. M., A. Margalida, R. Sánchez, and J. Oria, Supplementary feeding as an effective tool for improving breeding success in the Spanish imperial eagle (*Aquila adalberti*), *Biological Conservation*, *129*, 477–486, 2006.

González, L. M., A. Margalida, S. Mañosa, R. Sánchez, J. Oria, J. I. Molina, J. Caldera, A. Aranda, and L. Prada, Causes and spatio-temporal variations of non-natural mortality in the vulnerable Spanish imperial eagle *Aquila adalberti* during a recovery period, *Oryx*, *41*, 495–502, 2007.

Green, P. J., Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, *82*, 711–732, 1995.

Guitart, R., M. Sachana, F. Caloni, S. Croubels, V. Vandenbroucke, and P. Berny, Animal poisoning in Europe. Part 3: Wildlife, *The Veterinary Journal*, *183*, 260–265, 2010.

Gutiérrez, J., et al., *Physical ecosystem engineers and the functioning of estuaries and coasts*, pp. 53–81, Oxford: Elsevier, 2012.

Gutiérrez, J. L., and C. G. Jones, *Encyclopedia of Life Sciences*, chap. Ecosystem Engineers, John Wiley & Sons, Ltd, 2008.

Halstead, B. J., G. D. Wylie, P. S. Coates, P. Valcarcel, and M. L. Casazza, Exciting statistics: the rapid development and promising future of hierarchical models for population ecology, *Animal Conservation*, *15*, 133–135, 2012.

Hanski, I., *Metapopulation ecology*, Oxfor University Press, Oxford, UK, 1999.

Hedley, S., S. Buckland, and D. Borchers, *Advanced Distance Sampling: Estimating abundance of biological populations*, chap. Spatial distance sampling models, pp. 48–70, OUP Oxford, 2004.

Hendriks, I., T. Sintes, T. Bouma, and C. Duarte, Experimental assessment and modeling evaluation of the effects of seagrass (*Posidonia oceanica*) on flow and particle trapping, *Marine Ecology Progress Series*, *356*, 163–173, 2008.

Hendriks, I., T. Bouma, E. Morris, and C. Duarte, Effects of seagrasses and algae of the Caulerpa family on hydrodynamics and particle-trapping rates, *Marine Biology*, *157*, 473–481, 2010, 10.1007/s00227-009-1333-8.

Hendriks, I., M. Cabanellas-Reboredo, T. Bouma, S. Deudero, and C. Duarte, Seagrass Meadows Modify Drag Forces on the Shell of the Fan Mussel *Pinna nobilis*, *Estuaries and Coasts*, *34*, 60–67, 2011, 10.1007/s12237-010-9309-y.

Hendriks, I., S. Deudero, and G. Tavecchia, Recapture probability underwater: predicting the detection of the threatened noble pen shell in seagrass meadows, *Limnology and Oceanography: Methods*, *10*, 824–831, 2012.

Hendriks, I. E., C. M. Duarte, and C. H. R. Heip, Biodiversity Research Still Grounded, *Science*, *312*, 1715–1715, 2006.

Hernández, M., and A. Margalida, Poison-related mortality effects in the endangered Egyptian vulture (*Neophron percnopterus*) population in Spain, *European Journal of Wildlife Research*, *55*, 415–423, 2009.

Hobert, J. P., and G. Casella, The effect of improper priors on gibbs sampling in hierarchical linear mixed models, *Journal of The American Statistical Association*, *91*, 1461–1473, 1996.

Hutto, R. L., Seasonal changes in the habitat distribution of transient insectivorous birds in Southeastern Arizona: Competition mediated?, *The Auk*, *102*, 120–132, 1985.

Igual, J. M., G. Tavecchia, S. Jenouvrier, M. G. Forero, and D. Oro, Buying years to extinction: Is compensatory mitigation for marine bycatch a sufficient conservation measure for seabirds?, *PloS ONE*, *4*, e4826, 2009.

Irish, K. E., and E. A. Norse, Scant Emphasis on Marine Biodiversity, *Conservation Biology*, *10*, 680–681, 1996.

IUCN, IUCN red list of threatened species version 2010.4, 2010.

Jenni, L., and M. Kéry, Timing of autumn bird migration under climate change: advances in longdistance migrants, delays in shortdistance migrants, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*, 1467–1471, 2003.

Jones, C., J. Lawton, and M. Shachak, Organisms as ecosystem engineers, *Oikos*, *69*, 373–386, 1994.

Jordà, G., N. Marbà, and C. Duarte, Mediterranean seagrass vulnerable to regional climate warming, *Nature Climate Change*, p. online, 2012a.

Jordà, G., M. Marcos, M. Gomez, and D. Gomis, Wind wave climate in the Western Mediterranean under climate change scenarios, *Global and Planetary Change*, p. submitted, 2012b.

Kass, R. E., and A. E. Raftery, Bayes factors, *Journal of The American Statistical Association*, *90*, 773–795, 1995.

Katsanevakis, S., Population ecology of the endangered fan mussel *Pinna nobilis* in a marine lake, *Emdangered Species Research*, *1*, 1–9, 2005.

Katsanevakis, S., Density surface modelling with line transect sampling as a toll for abundance estimation of marine benthic species: the *Pinna nobilis* example in a marine lake, *Marine Biology*, *152*, 77–85, 2007.

Katsis, A., and I. Ntzoufras, Bayesian hypothesis testing for the distribution of insurance claim counts using the gibbs sampler, *Journal of Computational Methods in Science and Engineering*, *5*, 201–214, 2005.

Kendall, W. L., Robustiness of closed capture-recapture methods to violations of the closure assumption, *Ecology*, *80*, 2517–2525, 1999.

Kenward, R. E., S. S. Walls, K. H. Hodder, M. Pahkala, S. N. Freeman, and V. R. Simpson, The prevalence of non-breeders in raptor populations: evidence from rings, radio-tags and transect surveys, *Oikos*, *91*, 271–279, 2000.

Kéry, M., and M. Schaub, *Bayesian Population Analysis Using WinBUGS: A Hierarchical Perspective*, Academic Press, 2011.

King, R., *Bayesian Analysis for Population Ecology*, CRC Press, 2009.

Kochin, B., and P. Levin, Lack of concern deepens the oceans' problems, *Nature*, *424*, 723–723, 2003.

Kuo, L., and B. Mallick, Variable selection for regression models, *Sankhya Series B*, *60*, 65–81, 1998.

Laneri, K., et al., Trawling regime influences longline seabird bycatch in the mediterranean: new insights from a small-scale fishery, *Marine Ecology Progress Series*, *420*, 241–252, 2010.

Lawler, J., et al., Conservation science: a 20-year report card, *Frontiers in Ecology and the Environment*, *4*, 473–480, 2006.

Le Galliard, J.-F., R. Ferriere, and J. Clobert, Effect of patch occupancy on immigration in the common lizard, *J Anim Ecology*, *74*, 241–249, 2005.

Lebreton, J.-D., T. Almeras, and R. Pradel, Competing events, mixtures of information and multistratum recapture models, *Bird Study*, *46*, 39–46, 1999.

Lehman, R. N., P. L. Kennedy, and J. A. Savidge, The state of the art in raptor electrocution research: A global review, *Biological Conservation*, *136*, 159–174, 2007.

Link, W. A., and R. J. Barker, Model weights and the fundations of multimoldel inference, *Ecology*, *87*, 2626–2635, 2006.

Link, W. A., and R. J. Barker, *Bayesian Inference: With Ecological Applications*, Academic Press, 2009.

Link, W. A., and M. J. Eaton, On thinning of chains in mcmc, *Methods in Ecology and Evolution*, *3*, 112–115, 2012.

Liu, J. S., and Y. N. Wu, Parameter expansion for data augmentation, *Journal of The American Statistical Association*, *94*, 1264–1274, 1999.

Lodewyckx, T., W. Kim, M. D. Lee, F. Tuerlinckx, P. Kuppens, and E.-J. Wagenmakers, A tutorial on bayes factor estimation with the product space method, *Journal of Mathematical Psychology*, *55*, 331 – 347, 2011.

Lunn, D., N. Best, and J. Whittaker, Generic reversible jump mcmc using graphical models, *Statistics and Computing*, *19*, 395–408, 2009a.

Lunn, D., D. Spiegelhalter, A. Thomas, and N. Best, The BUGS project: Evolution, critique, and future directions, *Statistics in Medicine*, *28*, 3049–3067, 2009b, citation for OpenBUGS.

Lunn, D., C. Jackson, B. Nicky, A. Thomas, and D. Spiegelhalter, *The BUGS Book*, CRC Press Book, 2012.

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

MacKenzie, D. I., and W. L. Kendall, How should detection probability be incorporated into estimates of relative abundance?, *Ecology*, *83*, 2387–2393, 2002.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin, Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly, *Ecology*, *84*, 2200–2207, 2003.

Marbà, N., and C. Duarte, Mediterranean warming triggers seagrass (*Posidonia oceanica*) shoot mortality, *Global Change Biology*, *16*, 2366–2375, 2010.

Marbà, N., C. Duarte, E. Diaz-Almela, J. Terrados, E. Alvarez, R. Martiinez, R. Santiago, E. Gacia, and A. Grau, Direct evidence of imbalanced seagrass (*Posidonia oceanica*) shoot population dynamics in the Spanish Mediterranean, *Estuaries*, *28*, 53–62, 2005.

Marcos, M., G. Jordà, D. Gomis, and B. Pérez, Changes in storm surges in southern Europe from a regional model under climate change scenarios, *Global and Planetary Change*, *77*, 116–128, 2011.

Margalida, A., R. Heredia, M. Razin, and M. Hernández, Sources of variation in mortality of the bearded vulture *Gypaetus barbatus* in Europe, *Bird Conservation International*, *18*, 1–10, 2008.

Margalida, A., M. Á. Colomer, and D. Sanuy, Can wild ungulate carcasses provide enough biomass to maintain avian scavenger populations? an empirical assessment using a bio-inspired computational model, *PLoS ONE*, *6*, e20,248, 2011.

Martínez-Abraín, A., D. Oro, G. Tavecchia, J. Jiménez, M. Surroca, and H. Regan, Demographic response of a long-lived species to the synergistic effects of mortality in wind power plants and food scarcity after the european bovine spongiform encephalopathy crisis, *Journal of Applied Ecology*, *49*, 109–117, 2012.

Massot, M., J. Clobert, T. Pilorge, J. Lecomte, and R. Barbault, Density dependence in the common lizard: demographic consequences of a density manipulation, *Ecology*, *5*, 1742–1756, 1992.

Maunder, M. N., Population viability analysis based on combining Bayesian, integrated, and hierarchical analyses, *Acta Oecologica*, *26*, 85–94, 2004.

McCarthy, M. A., *Bayesian Methods for Ecology*, 1 ed., Cambridge University Press, 2007.

McCrea, R., B. Morgan, O. Gimenez, P. Besbeas, J.-D. Lebreton, and T. Bregnballe, Multi-site integrated population modelling, *Journal of Agricultural, Biological, and Environmental Statistics*, *15*, 539–561, 2010, 10.1007/s13253-010-0027-5.

M'Closkey, R. T., S. J. Hecnar, D. R. Chalcraft, and J. E. Cotter, Size distributions and sex ratios of colonizing lizards, *Oecologia*, *116*, 501–509, 1998.

Medina, R., Impactos sobre la costa española por efectos del cambio climático, *Tech. rep.*, Univ. de Cantabria para el Ministerio de Medio Ambiente, 2004.

Milazzo, M., F. Badalamenti, G. Ceccherelli, and R. Chemello, Boat anchoring on *Posidonia oceanica* beds in a marine protected area (Italy, western Mediterranean): effect of anchor types in different anchoring stages, *Journal of Experimental Marine Biology and Ecology*, *299*, 51–62, 2004.

Millar, R. B., Comparison of hierarchical bayesian models for overdispersed count data using dic and bayes' factors, *Biometrics*, *65*, 962–969, 2009.

Moore, F. R., P. Kerlinger, and T. R. Simons, Stopover on a gulf coast barrier island by spring trans-gulf migrants, *The Wilson Bulletin*, *102*, 487–500, 1990.

Moreteau, J., and N. Vicente, Evolution d'une population de Pinna nobilis l. (Mollusca, Bivalvia), *Malacologia*, *22*, 341–345, 1982.

Mougeot, F., and V. Bretagnolle, Breeding biology of the red kite *Milvus milvus* in Corsica, *Ibis*, *148*, 436–448, 2006.

Norse, E., and L. Crowder, *Marine Conservation Biology: The Science Of Maintaining The Seas Biodiversity*, Island Press, 2005.

Ntzoufras, I., Gibbs variable selection using bugs, *Journal of Statistical Software*, *7*, 1–19, 2002.

Ntzoufras, I., *Bayesian Modeling Using WinBUGS*, John Wiley and Sons, 2009.

O'Hara, R. B., and M. J. Sillanpää, A review of bayesian variable selection methods: What, how and which, *Bayesian Analysis*, *4*, 85–118, 2009.

Ormerod, S., Current issues with fish and fisheries: editor's overview and introduction, *Journal of Applied Ecology*, *40*, 204–213, 2003.

Oro, D., A. Margalida, M. Carrete, R. Heredia, and J. A. Donázar, Testing the Goodness of Supplementary Feeding to Enhance Population Viability in an Endangered Vulture, *PLoS ONE*, *3*, e4084, 2008.

Ortega, E., S. Mañosa, A. Margalida, R. Sánchez, J. Oria, and L. González, A demographic description of the recovery of the Vulnerable Spanish imperial eagle *Aquila adalberti*, *Oryx*, *43*, 113–121, 2009.

Park, T., and G. Casella, The bayesian lasso, *Journal of The American Statistical Association*, *103*, 681–686, 2008.

Pedrini, P., S. Tenan, and F. Spina (Eds.), *La migrazione post-riproduttiva degli Uccelli attraverso le Alpi italiane: fenologia ed andamenti – Post-nuptial bird migration across the Italian Alps: phenology and trends*, Museo delle Scienze, 2012.

Pérez-Mellado, V., J. Á. Hernàndez-Estévez, T. García-Díez, B. Terrassa, M. M. Ramón, J. Castro, J. Martín-Vallejo, and R. Brown, Population density in *Podarcis lilfordi* (Squamata, Lacertidae), a lizard species endemic to small islets in the Balearic Islands (Spain), *Amphibia-Reptilia*, *29*, 49–60, 2008.

Peterson, C., H. Summerson, and P. Duncan, The influence of seagrass cover on population structure and individual growth rate of a suspension-feeding bivalve, *Mercenaria mercenaria*, *Journal of Marine Research*, *42*, 123–138, 1984.

Pledger, S., Unified maximum likelihood estimates for closed CaptureRecapture models using mixtures, *Biometrics*, *56*, 434442, 2000.

Plummer, M., JAGS: A program for analysis of bayesian graphical models using gibbs sampling, in *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, 2003.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. Hines, Statistical inference for capture-recapture experiments, *Wildlife monographs*, *107*, 3–97, 1990.

Pradel, R., The stakes of capture-recapture models with state uncertainty, in *Modeling Demographic Processes In Marked Populations*, vol. 3 of *Environmental and Ecological Statistics*, pp. 781–795, Springer, 2009.

Procaccini, G., M. Buia, M. Gambi, M. Pérez, G. Pergent, C. Pergent-Martini, and J. Romero, *Word Atlas of Seagrasses*, chap. The seagrasses of the western Mediterranean, pp. 48–58, University of California Press, 2003.

Pullin, A. S., T. M. Knight, D. A. Stone, and K. Charman, Do conservation managers use scientific evidence to support their decision-making?, *Biological Conservation*, *119*, 245–252, 2004.

R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0.

R Development Core Team, R: A language and environment for statistical computing. URL http://www.r-project.org, 2011.

Rees, S. G., A. E. Goodenough, A. G. Hart, and R. Stafford, Testing the effectiveness of capture mark recapture population estimation techniques using a computer simulation with known population size, *Ecological Modelling*, *222*, 3291–3294, 2011.

Robb, G. N., R. A. McDonald, D. E. Chamberlain, and S. Bearhop, Food for thought: supplementary feeding as a driver of ecological change in avian populations, *Frontiers in Ecology and the Environment*, *6*, 476–484, 2008.

Robert, C., and G. Casella, *Monte Carlo Statistical Methods*, Springer, 2004.

Royle, J., and R. Dorazio, *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*, Academic Press, 2008.

Royle, J. A., Analysis of capture-recapture models with individual covariates using data augmentation, *Biometrics*, *65*, 267–274, 2009.

Royle, J. A., and R. Dorazio, Parameter-expanded data augmentation for bayesian analysis of capture-recapture models, *Journal of Ornithology*, *152 (Suppl. 2)*, S521–S537, 2011.

Royle, J. A., and M. Kéry, A Bayesian State-Space formulation of dynamic occupancy models, *Ecology*, *88*, 1813–1823, 2007.

Royle, J. A., and K. V. Young, A hierarchical model for spatial capture-recapture data, *Ecology*, *89*, 2281–2289, 2008.

Royle, J. A., R. M. Dorazio, and W. A. Link, Analysis of multinomial models with unknown index using data augmentation, *Journal of Computational and Graphical Statistics*, *16*, 67–85, 2007.

Sacchi, R., et al., Photographic identification in reptiles: a matter of scales, *Amphibia-Reptilia*, *31*, 489–502, 2010.

Salvador, A., Lagartija balear - Podarcis lilfordi, in *Enciclopedia Virtual de los Vertebrados Españoles*, Salvador, A., Marco, A. (Eds.). Museo Nacional de Ciencias Naturales, Madrid., 2009.

Sánchez-Camacho, M., Study of the spatio-temporal changes in the distribution of the seagrass *Posidonia oceanica* in Mallorca (Balearic Islands, Spain) between 1956 and 2001 by remote sensing and GIS methods, *Tech. rep.*, School of Ocean Sciences University of Wales, Bangor, 2003.

Sanz-Aguilar, A., et al., Recapture processes and biological inference in monitoring burrow-nesting seabirds, *Journal of Ornithology*, *151*, 133–146, 2010.

Sarrazin, F., and S. Legendre, Demographic approach to releasing adults versus young in reintroductions, *Conservation Biology*, *14*, 488–500, 2000.

Schaub, M., and F. Abadi, Integrated population models: a novel analysis framework for deeper insights into population dynamics, *Journal of Ornithology*, *152*, 227–237, 2010.

Schaub, M., and L. Jenni, Body mass of six long-distance migrant passerine species along the autumn migration route, *Journal für Ornithologie*, *141*, 441–460, 2000.

Schaub, M., and M. Kéry, Combining information in hierarchical models improves inferences in population ecology and demographic population analyses, *Animal Conservation*, *15*, 125–126, 2012.

Schaub, M., and R. Pradel, Assessing the relative importance of different sources of mortality from recoveries of marked animals, *Ecology*, *85*, 930–938, 2004.

Schaub, M., O. Gimenez, A. Sierro, and R. Arlettaz, Use of integrated modeling to enhance estimates of population dynamics obtained from limited data, *Conservation Biology*, *21*, 945–955, 2007.

Schaub, M., A. Aebischer, O. Gimenez, S. Berger, and R. Arlettaz, Massive immigration balances high anthropogenic mortality in a stable eagle owl population: Lessons for conservation, *Biological Conservation*, *143*, 1911–1918, 2010.

Schaub, M., T. S. Reichlin, F. Abadi, M. Kéry, L. Jenni, and R. Arlettaz, The demographic drivers of local population dynamics in two rare migratory birds, *Oecologia*, pp. 1–12, 2011.

Schoener, T. W., and A. Schoener, Densities, sex ratios, and population structure in four species of bahamian anolis lizards, *Journal of Animal Ecology*, *49*, 19–53, 1980.

Schwarz, C. J., and A. N. Anderson, Jolly-Seber models in MARK, in *MARK. A gentle introduction*, pp. 1–53, 2001.

Schwarz, C. J., and G. A. F. Seber, Estimating animal abundance: Review iii, *Statistical Science*, *14*, pp. 427–456, 1999.

Seber, G. A. F., *The Estimation of Animal Abundance and Related Parameters*, Blackburn Press, 1982.

Seber, G. A. F., A review of estimating animal abundance ii, *International Statistical Review / Revue Internationale de Statistique*, *60*, pp. 129–166, 1992.

Sergio, F., L. Marchesi, P. Pedrini, M. Ferrer, and V. Penteriani, Electrocution alters the distribution and density of a top predator, the eagle owl *Bubo bubo*, *Journal of Applied Ecology*, *41*, 836–845, 2004.

Sergio, F., J. Blas, M. Forero, N. Fernández, J. A. Donázar, and F. Hiraldo, Preservation of wide-ranging top predators by site-protection: black and red kites in Doñana National Park, *Biological Conservation*, *125*, 11–21, 2005.

Sisson, S. A., Transdimensional markov chains: A decade of progress and future perspectives, *Journal of The American Statistical Association*, *100*, 1077–1089, 2005.

Skalski, J. R., J. J. Millspaugh, and R. D. Spencer, Population estimation and biases in paintball, markresight surveys of elk, *Journal of Wildlife Management*, *69*, 1043–1052, 2005a.

Skalski, J. R., K. E. Ryding, and J. J. Millspaugh, *Wildlife Demography. Analysis of sex, age and count data*, Elsevier, London, 2005b.

Smart, J., A. Amar, I. M. Sim, B. Etheridge, D. Cameron, G. Christie, and J. D. Wilson, Illegal killing slows population recovery of a re-introduced raptor of high conservation concern - the red kite *Milvus milvus*, *Biological Conservation*, *143*, 1278–1286, 2010.

Soulé, M. E., The biological diversity crisis, *BioScience*, *35*, 727–734, 1985.

Spiegelhalter, D., A. Thomas, N. Best, and D. J. Lunn, *WinBUGS user manual, Version 1.4.3*, 2007.

Spiegelhalter, D. J., A. Thomas, N. G. Best, and W. R. Gilks, BUGS Examples Volume 2, Version 0.5, (version ii), *Tech. rep.*, MRC Biostatistics Unit, Cambridge, 1996.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde, Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*, 583–639, 2002.

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

Su, Y.-S., and M. Yajima, *R2jags: A Package for Running jags from R*, 2012, r package version 0.03-06.

Tavecchia, G., E. Minguez, A. De Leon, M. Louzao, and D. Oro, Living close, doing differently: small-scale asynchrony in demographic parameters in two species of seabirds, *Ecology*, *89*, 77–85, 2008.

Tavecchia, G., P. T. Besbeas, T. Coulson, B. J. Morgan, and T. H. Clutton-Brock, Estimating population size and hidden demographic parameters with state-space modelling, *The American Naturalist*, *173*, 722–733, 2009.

Tavecchia, G., J. Adrover, A. M. Navarro, and R. Pradel, Modelling mortality causes in longitudinal data in the presence of tag loss: application to raptor poisoning and electrocution, *Journal of Applied Ecology*, *49*, 297–305, 2012.

Tilman, D., R. M. May, C. L. Lehman, and M. A. Nowak, Habitat destruction and the extinction debt, *Nature*, *371*, 65–66, 1994.

Vaquer-Sunyer, R., and C. Duarte, Thresholds of hypoxia for marine biodiversity, *Proceedings of the National Academy of Sciences*, *105*, 15,452–15,457, 2008.

Vicente, N., J. Moreteau, and P. Escoubet, Etude de l'evolution d'une population de *Pinna nobilis* l. (mollusque eulamellibranche) au large de l'anse de la Palud (Parc National sous-marin de Port-Cros), *Trav. Scient. Parc Natn. Port-Cros*, *6*, 39–68, 1980.

Walker, D., R. Lukatelich, G. Bastyan, and A. McComb, Effect of boat moorings on seagrass beds near Perth, Western Australia, *Aquatic Botany*, *36*, 69–77, 1989.

Wanger, T., I. Motzke, S. Furrer, B. Brook, and B. Gruber, How to monitor elusive lizards: comparison of capturerecapture methods on giant day geckos (Gekkonidae, Phelsuma madagascariensis grandis) in the masoala rainforest exhibit, Zurich Zoo, *Ecological Research*, *24*, 345–353, 2009.

Waycott, M., et al., Accelerating loss of seagrasses across the globe threatens coastal ecosystems, *Proceedings of the National Academy of Sciences*, *106*, 12,377–12,381, 2009.

White, G. C., and K. P. Burnham, Program mark : survival estimation from populations of marked animals, *Bird Study*, *46 (suppl.)*, 120–129, 1999.

Whitfield, D. P., D. R. A. McLeod, J. Watson, A. H. Fielding, and P. F. Haworth, The association of grouse moor in Scotland with the illegal use of poisons to control predators, *Biological Conservation*, *114*, 157–163, 2003.

Williams, B. K., J. D. Nichols, and M. J. Conroy, *Analysis and management of animal populations*, Academic Press, San Diego, 2002.

Williams, T. C., J. M. Williams, P. G. Williams, and P. Stokstad, Bird migration through a mountain pass studied with high resolution radar, ceilometers, and census, *The Auk*, *118*, 389–403, 2001.

Wilson, S., and K. Martin, Songbird use of high-elevation habitat during the fall post-breeding and migratory periods, *Ecoscience*, *12*, 561–568, 2005.

Zavodnik, D., M. Hrs-Brenko, and M. Legac, *Les especes marines a protéger en Méditerranée. Marseille: GIS Posidonie*, chap. Synopsis on the fan shell *Pinna nobilis* L. in the Eastern Adriatic sea, pp. 169–178, 1991.