# Chapter 1
# Introduction

This book was conceived as a companion volume to *Estimating Animal Abundance: Closed Populations* by Borchers et al. (2002), who gave a unifying framework for estimating the abundance of closed populations. We seek to do the same for estimating the abundance of open populations: populations subject to births, deaths and movement in and out of the population. We focus primarily on populations of large vertebrates, for which we will typically model dynamics within the framework of an annual cycle, and for which stochastic variability in the demographic processes is usually modest. We consider discrete time models in which animals can be assigned to discrete states such as age class, gender, population (within a metapopulation), or species (for multi-species models).

We go beyond estimation of abundance, however, as we are interested in understanding reasons for variation in abundance over time. Knowledge of abundances alone will seldom be sufficient for the ecologist or wildlife manager. For example, suppose one knew there were exactly 5,218, 6,319, and 7,438 individuals for three consecutive years. A natural question would be "*Why* are the abundances increasing?". More generally, the question is "What is causing, or driving, the population dynamics?". The answer will depend in some way on the underlying processes of birth or recruitment, survival and movement. These processes are typically of greater interest to ecologists and managers than abundance itself.

To quantify these processes requires the formulation and fitting of population dynamics models. The resulting fitted models will potentially yield both estimates of abundance and estimates of parameters characterizing the underlying processes, such as survival probabilities.

## 1.1 Background to the Book

Population dynamics models serve a wide variety of uses, including explaining previous population fluctuations as well as projecting future population abundances under hypothetical scenarios. We might for example wish to reintroduce an animal

K.B. Newman et al., *Modelling Population Dynamics: Model Formulation, Fitting and Assessment using State-Space Methods*, Methods in Statistical Ecology, DOI 10.1007/978-1-4939-0977-3_1,

that has become extinct locally. We then need to be confident that the reintroduced population is viable. Models are needed to make this assessment. Or perhaps we need to manage an exploited fish stock, and assess whether the level of take is sustainable. Another example is when we wish to manage a species of conservation concern, and perhaps reverse a declining trend. We would like models of the species and its environment to allow us to explore “what-if” scenarios on the computer before we experiment with the real population. If we can build a model that relates demographic processes to habitat, we can then explore options for modifying the habitat to the species’ benefit. Perhaps we wish to assess the response of a species or community to climate change. To do this, we must model the population processes, and how these depend on climate. Yet another example is when a population is too large for the environment to support, so that habitat degrades, to the detriment of biodiversity in the region. We may then need models to indicate levels of culling needed to sustain the environment and its diversity.

Population dynamics models play a central role in adaptive resource management (Williams et al. 2007) where managers are experimenting with different management actions to produce desired population responses. Models are used to predict the consequences of management actions such as habitat modifications or predator removal. After-the-fact comparisons of model predictions to observed outcomes are also made. The strength of evidence for alternative models, i.e. competing explanations for the effects of actions, can be assessed to determine which explanations are more consistent with the data.

In all of the above examples, we need to address risk. This means that we need to quantify uncertainty in our model predictions. It is not enough therefore to develop mathematical models of animal dynamics. We also need to allow for the main sources of uncertainty in these models: demographic and environmental stochasticity; observation or sampling errors in our data; and model or structural uncertainty (Williams et al. 2001). Without these elements, a mathematical model of population dynamics is merely a tool of theoretical ecologists, with little to offer wildlife managers. Similarly, purely empirical models for trend, fitted to time series of data, have little predictive power, because they ignore the population processes that drive change, so cannot predict change with any degree of confidence. In this book, we develop an approach that allows models for population dynamics to be embedded within a full inferential framework, giving the predictive and explanatory power of mathematical process models, coupled with statistical tools for quantifying uncertainty.

The main steps involved are as follows. First, we formulate a model of the population dynamics. We provide a building block strategy for this, allowing quite complex models to be assembled from simple components. For convenience of model formulation, we use matrices to define these building blocks, which allows us also to relate our approach to the well-developed field of matrix population models (Caswell 2001). However, our fitting algorithms are not restricted to such models, and we can readily extend our models in various ways. We develop the tools needed for model formulation, and show how they provide a more general framework than is possible with matrix models.

Second, we formulate the population dynamics model as a hidden process model. If the model is first-order Markov, it is termed a *state-space model*. The states might reflect gender, age, geographic region, or other characteristics that define different sub-populations that might be subject to different dynamics. The state equation defines how the population updates stochastically through time. These processes are generally hidden because we cannot observe the entire population—we typically have counts or estimates of the population or of some component of the population at certain points in time. Hidden process models give great flexibility in defining both the systematic and the random components of the population dynamics model.

Third, we define how our observations relate to the states, through an observation equation. The random component of this equation relates to observation error, while the systematic component defines a "mapping" between the observations and the states.

As an aside, we note that partitioning the population dynamics model into two components, namely the underlying hidden state process and the observation model, has practical advantages. One advantage is that theories about the dynamics can be kept separate from the sampling and estimation issues; thus, for example, the subject matter specialist can focus attention on the underlying dynamic processes of survival and birth, postulating alternative theories largely independent of the data collection procedures. Sample size determination can also be partitioned into problems to determine the number of years to sample and problems to determine the sampling intensity within a given year, and trade-offs between different combinations of sample sizes can be more clearly determined.

Fourth, we address how the hidden process model may be fitted to time series of observations. We consider three primary approaches. The Kalman filter allows us to fit a model quickly, at the expense of assuming that the errors in both the state equation and the observation equation are normally distributed, and that the expectation of the state vector at one time point is a linear function of the state vector at the previous time point. We also give two computer-intensive methods which can be implemented within a Bayesian framework to give much greater flexibility. The first of these is Markov chain Monte Carlo (MCMC), and the second is sequential Monte Carlo methods.

Finally, we discuss model selection, goodness-of-fit, model averaging, parameter redundancy, and other issues related to model assessment and uncertainty.

In subsequent chapters, we consider different methods of assessing wild animal populations, and how these can be combined with hidden process modelling to allow the dynamics of the populations to be modelled. We first address the case where closed-population assessment methods are used at each of several time points, and a population dynamics model fitted to the time series of abundance estimates. This may be viewed as an extension of the robust design approach of Pollock (1982). Next, we consider open population assessment methods, and how these may be integrated with hidden process models. Finally, we look at the case when there are multiple data sources. These methods are illustrated through examples, and `R` or `WinBUGS` code is provided for most examples, allowing users to modify the code for their own purposes.

## 1.2 Book Website

Computer code for our examples is available at the book website, www.creem.st-and.ac.uk/modpopdyn. The code can be used as a learning aid to re-run the analyses of this book, and can also be used as a template for the user to generate code to analyse his or her own data.

## 1.3 Related Books

As noted above, this book is a companion to Borchers et al. (2002). They give a unifying framework for methodologies for estimating the abundance of closed populations, and we seek to do the same for open populations. Population dynamics play a crucial role in open population modelling, yet methods for assessing the abundance of open populations have for the most part had no embedded model of these dynamics. Such embedded models are central to the methods of this book.

This book also complements the book on Bayesian analysis for population ecology by King et al. (2009). They focus on the underlying ideas associated with Bayesian analyses of ecological data and the corresponding methods to fit Bayesian models. There is an emphasis on mark-recapture-recovery data for open populations. Case studies are used to illustrate the methods, including state-space models, the use of covariate information (including dealing with missing data), multi-state data, integrated data analyses, random effects models and closed capture-recapture models. Computer codes are provided for examples, using both `WinBUGS` and `R`.

Williams et al. (2001) offer near encyclopedic coverage of many of the better known population dynamics models for animal populations as well as the wide variety of statistical procedures used to estimate abundance (for closed and open populations) and parameters of population dynamic processes. Integration of population process models with observation models is by and large not addressed, however.

Caswell (2001) uses matrices to model population dynamics, e.g., Leslie matrices to characterise survival and reproduction. Matrices, strictly speaking, imply linear models, although Caswell formulates non-linear extensions to the basic matrix model. We find matrix models to be useful for formulating state-space models and present many examples in Chap. 2.

Royle and Dorazio (2008) cover hierarchical models for ecological systems. In common with this book, they use both frequentist and Bayesian methods, and show how to combine explicit models of ecological system structure or dynamics with models of how ecological systems are observed. They apply the principles of hierarchical modelling to problems in population, metapopulation, community and metacommunity systems, whereas we focus on modelling population dynamics, and estimating population abundance.

Two classic fisheries science and management books, Hilborn and Walters (1992) and Quinn and Deriso (1999), describe a wide variety of methods and models for characterising fish population dynamics. These models and methods are often used for non-fish species. Both books also discuss the distinction between process variation and observation error, but at the time the books were written, the handling of both types of error simultaneously was not common practice.

Written for researchers and graduate students in statistics, ecology, demography, and the social sciences, McCrea and Morgan (2014) provide an up-to-date overview of capture-recapture methods. It covers model development and diagnostics. The authors use illustrative data sets drawn from a range of different areas and indicate available software for classical and Bayesian inference.

We use Bayesian methods for fitting many of the models of this book. Books that provide an introduction to Bayesian methods for ecologists are McCarthy (2007) and Link and Barker (2009). Kéry and Schaub (2012) focus on using `WinBUGS` to model population dynamics. They show how to fit Bayesian state-space models using mark-recapture or mark-recovery data, and they also provide example code for integrated population modelling, and for modelling counts.

## 1.4 Roadmap to the Book

After reading this chapter, we recommend that Chap. 2 be read next, which provides a path for those familiar with classic Leslie and Lefkovitch matrix models to a building-block approach to (a) constructing situation-specific matrix models and (b) formulating the skeleton of state-space models. Chapters 3–5 are the methodological core of the book giving general definitions of SSMs, methods for fitting SSMs, and then suggestions for formulation, or re-formulation, and assessment of SSMs. Chapters 6–9 are largely stand-alone chapters and can be read in any order, although there is some cross-reference between these chapters. Chapter 10 summarizes the book and can be read at any time.