HAL
open science

Chapter 13 Survival Analyses

Sarah Cubaynes, Simon Galas, Myriam Richaud, Ana Sanz, Roger Pradel, Giacomo Tavecchia, Fernando Colchero, Sebastien Roques, Richard Shefferson, Carlo-Giovanni Camarda

To cite this version:

Sarah Cubaynes, Simon Galas, Myriam Richaud, Ana Sanz, Roger Pradel, et al.. Chapter 13 Survival Analyses. Roberto Salguero-Gomez (ed.), Marlène Gamelon (ed.). Demographic Methods Across the Tree of Life, Oxford University Press, inPress, 9780198838609. 10.1093/oso/9780198838609.003.0013 . hal-03153205

HAL Id: hal-03153205

https://hal.science/hal-03153205v1

Submitted on 9 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapter 12 Survival Analyses

Sarah Cubaynes, Simon Galas, Myriam Richaud, Ana Sanz, Roger Pradel, Giacomo Tavecchia, Fernando Colchero, Sebastien Roques, Richard Shefferson, Carlo Giovanni Camarda

Introduction

Fitness differences among individuals are the bedrock of ecological and evolutionary dynamics (Stearns 1992). Survival is without doubt one major component of fitness, which makes of survival analyses a key tool for demographers, ecologists and evolutionary biologists (Metcalf and Pavard 2007).

Assessing survival is not always an easy task. This is because individuals die only once in a lifetime, therefore precluding repeated measurements on the same individual, and death is rarely directly observed. However, a diversity of monitoring techniques exists to gather survival data for species across the tree of life, from short-living lab organisms, plants exhibiting dormancy, to long-living and elusive wild vertebrates (see chapter 4). Each monitoring technique leading to a peculiar type of survival data, a plethora of methods exist to estimate survival, from simple non-parametric estimators, to more or less complex semi-parametric and fully parametric models fitted in a continuous or in a discrete manner (reviewed in e.g. Wienke 2010, Miller 2011, Klein et al. 2016, Canudas-Romo et al. 2018, Cox 2018). The choice of method is guided by the type of survival data (e.g. collecting age at death or monitoring live individuals with perfect or imperfect detection), the species life-history (e.g. single or numerous stages or ages) and the environment it experiences (e.g. controlled conditions versus variable environments).

The aim of this chapter is not to provide another extensive review of the existing techniques for survival analyses, but rather to illustrate and contrast the most commonly used

methods to estimate survival across the tree of life. The focus is placed on seven case studies of survival analyses in lab organisms, free-ranging animal and plant populations and in human populations. While it is relatively straightforward to gather survival data and assess survival using simple models under controlled lab conditions (Klein 2016; see section 2), monitoring survival in free-ranging populations often require more sophisticated capture-mark-recapture (CMR) techniques to deal with imperfect detection of individuals (Williams et al. 2002, see section 3). Indeed, survival data are often 'incomplete' in free-ranging populations, timing and cause of death can be hard to assess, and multiple environmental factors are at play in influencing survival (see section 3, see also chapter 4 and 5). Humans are an exception, with the existence of several consequent databases with perfect knowledge of age and cause of death for several human populations, such as the Human Mortality Database (2019). A great diversity of mortality models has been developed specifically to analyse age at death data in human demography (e.g. Canudas-Romo et al. 2018, see section 4).

1. What is survival analysis and why do we need it?

1.1. Time matters

Survival analysis is used to analyze the time up to a specific event is going to occur. The event is generally death, but can be of another kind (e.g. reproduction, migration, or exposure to a pathogen). Survival data usually involve following a set of individuals over a specific period of time and recording the time of occurrence of the event of interest.

1.2. Censoring and missing data

By nature, survival data have some particularities: survival cannot take a negative value, survival data are often censored, and missing data are frequent. Right censoring occurs when an individual remains alive after the end of the study or drops out of the study, e.g. due to permanent emigration (Klein et al. 2006). Left-censoring occurs when we cannot observe the

time when the event occurred, e.g. an individual was already exposed to a pathogen before the study started (Klein et al. 2006). Other types of missing data are common, because individuals cannot always be observed at all occasions (see chapter 4 and section 3). We need survival analyses to avoid introducing bias, because ordinary linear regression cannot effectively handle the censoring of observations and missing data (but see section 3).

1.3. Known fate data versus imperfect detection

In presence of data referred to as 'know-fate' data, survival can be modelled in a continuous manner using continuous distributions (e.g. Gompertz, Weibull, Makeham, Siler). The survival function *S(t)* is the probability that an individual survives up to a certain time t. The hazard function *h(t)* represents the instantaneous event rate for an individual who has already survived to time t. Both functions are related, the hazard relates to the death rate, while survival reflects its cumulative non-occurrence. One might choose to model survival when the data involve counts of individuals alive at different points in time while mortality models can be preferred to model age at death data.

Alternatively, CMR models are used to estimate survival when the data involve individuals missing at certain monitoring occasions, and unknown time at death (Lebreton et al. 1992, see chapter 4). CMR models involve estimating survival probability in between consecutive occasions as a function of a set of covariates, while accounting for the imperfect detection of individuals at each occasion (e.g. a marked individual might be present in the study area at a given occasion, but missed). Fully parametric or semi-parametric approaches are available to model survival, transition, or detection in CMR models (examples in section 3).

In both cases, survival or mortality can be modelled using a non-parametric (see example in section 2), semi-parametric (see example in section 4) or fully parametric model (see examples in section 3 and 4).

2. Survival analyses in the lab: when it's (almost) all under control

2.1. Monitoring survival in the lab

In the lab, gathering survival data usually involve counting the number of individuals alive at different times or simply collecting the exact time at death. Collecting survival data is generally easier than in the wild because individuals are available for monitoring at all times during the study, and the influence of external variables is limited and often controlled. A common objective is to analyse differences in survival functions between groups of individuals that present different characteristics, e.g. various genotypes; groups of individuals exposed to various treatments; age, sex, and life history differences.

2.2. Kaplan-Meier estimator and log-rank tests

Laboratory experiments generate data sometimes referred to as 'known fate' data, in which the probability of detection is 1 and thus does not need to be accounted for when estimating survival. The survival function $S(t)$, is usually estimated using the non-parametric Kaplan-Meier (KM) estimator (Kaplan and Meier 1958). For each time interval, it is calculated as the number of individuals surviving divided by the number of individuals at risk. Censored individuals who have dropped out of the study are not counted in the denominator. The cumulative survival probability is calculated by multiplying probabilities of surviving from one interval to the next. A plot of the KM survival probability against time provides a useful summary that can be used to estimate parameters such as median survival time (Figure 1). The smaller the time intervals, the smoother the survival curve.

[Figure 1 about here]

Graphically, a vertical gap between the survival curves of different groups means that at a specific time point, one group has a greater fraction of individuals surviving while a horizontal gap means that it took longer for one group to experience a certain fraction of deaths (Figure 1). The log-rank test (Peto et al, 1977) is often used to compare the survival

distributions of two or more groups. It is a non-parametric test based on a chi-square statistic, which makes no assumptions about the survival distributions. It assumes that the groups have the same survival as the null hypothesis. When the log-rank statistic is large, it is evidence for a significant difference in the survival times between the groups.

KM curves and log-rank tests are very useful in assessing whether a categorial covariate (e.g. treatment A vs. treatment B; males vs. females) affects survival. However, it does not allow investigating the effects of multiple or continuous covariates (e.g. weight, age) and to know how much more at risk one group is than another.

2.3. Cox's proportional hazards to adjust for covariates

An alternative method is the Cox's proportional hazards regression analysis (Cox, 1972). It is a semi-parametric method, which can be used to assess simultaneously the effect of several risk factors (both categorial and continuous covariates) on survival time. It is analogous to a multiple regression in which the response variable is the hazard measuring the instantaneous rate of the event. It assumes a constant proportional hazard across groups over time, i.e. the ratio of the risk of dying at a particular point in time in one group over another group is constant over time. Proportionality tests are used to evaluate this assumption (e.g. Miller 2011, see Supplementary Information S1 for an example). Models allowing for different assumptions, such as accelerated failure time (AFT) models, are discussed for example in Kirkwood and Sterne (2003) and Klein et al. (2016).

2.4. Example of a stress assay in an unconventional resistant organism: the tardigrade

Tardigrades hold their own phylum that stands in between the phylum Arthropoda and Nematoda (Image 1). They are renowned organisms for their ability to cope with the harshest environments, such as exposure to organic solvents, extreme temperatures (rom -272 to 151°C) or high radiation doses, at any stage of their life (e.g. Jönsson et al, *2005*). Most notably, some

tardigrades have survived a ten days flight in the vacuum of space, and at a depth of 180 km below the surface of the earth (Ono et al. 2016). However, it seems that the genetic "toolbox" that ensures the uncommon resistance of these organisms to extreme stress is not shared by all tardigrade species. The resistance of tardigrade species can be assessed and compared under controlled conditions in the lab.

[Image 1 about here]

In this example, we illustrate the use of KM survival curve and log-rank test to evaluate the effect of the chemical stressor on tardigrades survival time. We use a Cox Proportional Hazard model to assess the effect of the chemical stressor while taking into account tardigrades age. Full procedure including data and R script to run the analysis using the R package *survival*, are provided in Supplementary Information S1.

3. Survival analyses in the wild: dealing with uncertainty and variable environments

In free-ranging animal and plant populations, CMR field methods are generally used to gather survival data (procedure described in chapter 4). A plethora of CMR models exist to estimate demographic parameters (including survival, access to reproduction, dispersal) while accounting for imperfect detection of individuals (e.g. Lebreton et al. 1992, Williams et al. 2002, Schaub et al. 2004, Pradel et al. 2005).

Hereafter, we introduce principle of CMR analysis and the main types of CMR models. We then provide an overview of their use in animal and plant demography with step-by-step study cases.

3.1. Overview of CMR models

3.1.1. Principle

Within a typical CMR protocol designed to estimate survival probabilities, individuals are sampled on discrete occasions (often ≥3 occasions), at which they may be detected or not. Data

collected in the field is then encoded into encounter histories organized by individual or cohort (see details in White and Burnham 1999). The simplest way to codify the encounter histories is by using binary codes “1” and “0” specifying if the individual has been detected or not. However, when additional information reflecting the state of the individual (e.g. breeder vs. non-breeder, different resighting sites, epidemiological states) is collected, encounter histories may include additional codes (see multistate and multievent capture-recapture frameworks below). Moreover, individual covariates can also be recorded and included at the end of the individual encounter histories either in order to categorize individuals (e.g. males and females, age class, different populations) or by indicating a particular quantitative individual trait (e.g. size) (White and Burnham 1999).

Models for the analysis of capture-mark-recapture data are classically based on multinomial distributions that describe on one hand the biological processes and on the other hand the observational processes, conditional to the biological ones. In the simplest case, the probabilities involved are, ϕ_i, probability of surviving the time interval i, $i+1$ and p_i the probability to detect a live individual at occasion i. More complex models (see section 3.4.2 and 3.4.3 below) actuate multinomial rather than binomial biological processes and observational processes but the basic structure remains the same.

3.1.2. Single-state models

The Cormack-Jolly-Seber (CJS) model was the first CMR model (with an earlier formulation: the Jolly-Seber (JS) model; Jolly 1965, Seber 1965) to allow the estimation of demographic parameters under the assumption of an open population (i.e. open to birth, immigration, death, and emigration, a.k.a. B.I.D.E. models). Under the CJS approach, apparent survival (ϕ_i) is the probability that an individual alive at time i will be alive at time i+1, while resighting (p_i) is the probability that an individual alive and present just before time i is seen (and marked) on that occasion (Lebreton *et al.* 1992). Apparent survival is generally not

referred to simply as ‘survival’ because it may be confounded by permanent emigration of marked individuals out of the study site, and by long-term vegetative dormancy in plants. Further, apparent survival is the probability to survive during the interval between two monitoring occasions and thus not at a specific monitoring occasion. In contrast, resighting is an estimator of detection of previously seen individuals during the monitoring occasion.

CMR models rely on several assumptions, the most important of which are that marked individuals are independent, tags are not lost, and the past history does not influence the fate of the individual (i.e. no trap response or negative effect of capture on survival). Pollock et al. (1990) developed a series of goodness-of-fit tests based on contingency tables for the CJS model to assess the validity of the assumptions. Later, directional tests for the detection of specific effects were derived (trap-dependence, Pradel 1993; transients, Pradel et al. 1997). The availability of informative goodness-of-fit tests makes the CJS model a common ‘umbrella’ model in model selection procedure (Lebreton et al. 1992). The CJS model can be easily expanded to age-dependent parameters when new animals are released at different ages or in multiple cohorts, or reduced assuming parameters are constant over time (examples in Lebreton et al. 1992).

The CJS model, and its extensions, have been, and still are, extensively used in the ecological literature as they provide a suitable analytical framework to address multiple questions that tackle variability in survival over time (Lebreton et al. 1992), with external covariates (e.g. Grobois et al. 2008), changes in recruitment probabilities (e.g. Pradel and Lebreton 1999), recapture processes (e.g. Sanz-Aguilar et al. 2010) or evolutionary trade-offs (Tavecchia et al. 2001). However, the CJS model is based on capture-recapture data obtained from a single population and cannot explicitly frame observations of animals moving across multiple sites or between relevant biological states.

3.1.3. Multi-state models

Mutli-state models extend the CJS model by making the parameters state-specific (Arnason 1973; Schwartz 1993). Thus, apparent survival (ϕ_i^{jk}) is now the probability that an individual alive in state *j* on monitoring occasion *i* survives to occasion *i*+1 and, during the latter occasion, transits to state *k*. This parameter may be decomposed into two parameters unseen in the CJS model: state-specific survival (S_i^j) and state-transition (Ψ_i^{jk}). Here, state-specific survival (S_i^j) is the probability that an individual alive in state *j* at monitoring occasion *i* survives to monitoring occasion *i*+1, irrespective of the state of the individual at the latter occasion. State-transition (Ψ_i^{jk}) is the probability, conditional on survival, of changing or moving from state *j* at time *i* to state *k* at time *i*+1 assuming that the individual survived the interval between the two occasions. Additionally, multi-state models estimate resighting (p_i^j), defined as the probability that an individual alive and in state *j* at monitoring occasion *i* is also observed in that occasion.

This multisite-multistate formulation allows to address questions on survival and movement probabilities in metapopulation systems (e.g. Balkiz et al. 2010), but it also provides a suitable framework for the study of between-state transitions to study evolutionary trade-offs (Nichols and Kendall 1995), recruitment probability (e.g., Jenouvrier et al. 2008) or to mix information of different types, i.e. recoveries and recaptures (e.g. Lebreton et al. 1999). The multisite – multistate and robust design models have also been successfully applied to model dispersal to unobservable sites or transition to unobservable states, i.e. places or states through which animals can move but in which they cannot be seen, or unobservable dormancy states in plants (Kendall and Nichols 2002). In the unobservable site/state probability of detection is fixed to zero and for this reason they are often referred to as ‘ghost’ sites/states (Jenouvrier et al. 2008, Balkiz et al. 2010). Grosbois and Tavecchia (2003) applied this idea to unobservable transitions. They considered the probability of dispersal as a two-step process, one accounting for the probability of leaving a given site and a second, conditional to this, incorporating the

probability of settling into a new site. A similar approach was used by Schaub and Pradel (2004) to estimate the relative importance of different causes of death. However, multisite-multistate models have an important limitation: they assume that the state or the site in which an individual is observed is always certain.

3.1.4. Multievent models

State uncertainty is a general problem in CMR models, but it might be particularly relevant in some studies. For example, the study of the evolutionary trade-off between survival and reproduction relies on the fact that the breeding state of the individuals observed is always determined correctly. This can be true in most cases, but sometimes it can be difficult to determine with certainty whether an animal or plant is: breeding or not, is healthy or affected by a particular disease, or even if it is a male or a female. Pradel (2005) solved this problem by generalizing the multisite-multistate model into a multievent framework. In this new framework, individuals are still assumed to move across different states through survival and transition processes but a new parameter, the initial state probability, appears. Usually, field observations are not fully informative of biological processes. By formally separating the 'real' state process from the observational process-events, Pradel included a parameter to account for state uncertainty, i.e. the probability to not assign or erroneously assign a given state to an individual. Multievent models provide a solution to estimate sex and/or age dependent survival in species with cryptic or little age and/or sexual dimorphism (see section 3.4.1). Transients can be considered as a particular initial state, and modelled directly using multievent models (Genovart et al. 2012, Santidrian et al. 2017). More than uncertainty about the state of an individual, multievent models allow to explicitly model unobserved heterogeneity among individuals. Indeed, latent state mixture models can be easily implemented within the multievent framework, a model particularly suitable to frame individual heterogeneity or frailty (Gimenez, Cam and Gaillard 2018). Other approaches, e.g. using individual random effects

can also be used to implement frailty (e.g. Cam et al. 2016, Hamel et al. 2018 , see also section 3.2).

Multievent and multistate models can be used to address many other ecological and evolutionary questions. If individual states can change over time, multievent also allows modeling the transition dynamics between e.g. breeding states (Desprez et al. 2013) or epidemiological status (Benhaiem et al. 2018). Lagrange et al. (2014) developed multievent models able to study dispersal among numerous sites for birds and amphibians. Tavecchia et al. (2012), modelled mortality due to different causes of mortality of radiotagged individuals, while accounting for the loss of the radio signal (see section 3.4.2). Multievent models have also been used to model survival when marks identifying individuals are lost (even totally, see Badia-Boher et al. 2019). The multievent approach can be also used to exploit supplementary information and estimate survival, dispersal and/or recruitment in partially monitored populations (Tavecchia et al. 2016; Sanz-Aguilar et al. 2016). Finally, trap-responses and memory effects can be modeled into the multievent approach (Rouan et al. 2009; Pradel and Sanz-Aguilar 2012).

3.2. Environmental variability and individual heterogeneity in CMR models

Within a CMR model, survival, transition and/or detection parameters can be assumed to be a function of external covariates (Lebreton et al. 1992). This formulation allows the inclusion of environmental effects in a regression-like framework, by including e.g. climatic variables, or individual time-invariant characters such as genotype. Multifactorial effects of environmental covariates can also be modeled using hierarchical models (King et al. 2009), e.g. to study evolutionary processes in the wild (Cubaynes et al. 2012). However, time-varying individual covariates, such as body weight, are tricky because when an individual is not detected, the value of the covariate is unknown. Inference can be based on a conditional likelihood approach using only the observed covariate values (trinomial approach; Catchpole et al. 2008) or missing

values of can be imputed from an underlying distribution (e.g. multiple imputation; Worthington et al. 2015). However, methods of imputation are sensitive to the underlying model and the number of missing values (Langrock and King 2013). One possibility is to discretize the covariate and use a multistate model (Fernández-Chacón et al. 2015, Gimenez, Cam and Gaillard 2018).

Unfortunately, we often do not measure all covariates influencing demographic parameters. Individuals may also react in different ways to environmental variation depending on unobservable individual states. This leads to unobserved (latent) individual heterogeneity. Ignoring latent individual heterogeneity may lead to flawed inference about the ecological or evolutionary processes at hand (Cam et al. 2016, Hamel et al. 2018), such as senescence patterns (Cam and Monnat 2000, Service 2000, Peron et al. 2010). Latent individual heterogeneity can be framed using finite mixtures or as individual random effects (see Gimenez, Cam and Gaillard 2018 for a review about how to implement individual heterogeneity in CMR models).

3.3. Inference framework

Implementation of CMR models can be carried out either using a frequentist or Bayesian approach. A different philosophy stands behind each approach, and there is a long-standing debate about whether ecologists should use one or the other (e.g. Lele and Dennis 2009). While the frequentist approach may be faster, the Bayesian approach allows a great flexibility in the model writing which can be useful to tackle analytical complexity, such as choosing underlying distributions for model parameters, or fitting temporal random effects (e.g. Kery and Schaub 2011). Another appeal of the Bayesian approach is the possibility to include prior knowledge on biological parameters to facilitate the estimation, e.g. information on body weight or survival of a closely-related species (MacCarthy et al. 2005). The Bayesian approach of CMR models uses the state-space formulation (SSM) that clearly distinguishes the observation

process (detection) from the underlying demographic process of interest (transition between states; e.g. Gimenez et al. 2007, Royle 2008, Kery and Schaub 2011). The observation process being conditional to the state process. The SSM formulation therefore allows to easily implement complex multifactorial observation processes and combine multiple sources of information (e.g. Buoro et al. 2012, see Supplementary Information S5 for a SSM formulation of the Jolly-Seber model).

Prior to model fitting, goodness-of-fit tests are generally performed to check the validity of the assumptions behind a CMR model, e.g. using the R2ucare package (Gimenez et al. 2018). In the frequentist approach, model implementation can be carried out using program MARK and the widely used RMark package (White and Burnham 1999, Laake 2013), marked (Laake 2013) or E-SURGE (Choquet et al. 2009). Common tools for model comparisons include the Akaike Information Criterion and its variants (AICc, QAIC, wAIC; Burnham and Anderson 2002), which serves to rank the models, calculate weights of evidence for each of them or for a particular effect (Burnham and Anderson 2002). The Analysis of Deviance (Anodev) is also used to calculate the proportion of variance explained by a specific covariate (Grosbois et al. 2008). In the Bayesian approach, models can be implemented using program Jags (Plummer 2003), R packages such as rjags (Plummer et al. 2018) or BaSTA (Colchero et al. 2012). Posterior predictive checks can be used for performing model assessment (Chambert et al. 2014) and information criterion such as DIC or wAIC are often used for model comparison (Hooten and Hobbs 2015). Further details about implementation both in a frequentist and Bayesian framework can be found in McCrea and Morgan (2014).

Hereafter, we develop study cases in animal and plant demography showing how to implement CMR models in a frequentist framework using program E-SURGE and R package marked (Laake et al. 2013), and in a Bayesian framework using package BaSTA and rjags.

3.4. Study cases in animal demography

3.4.1. Estimating sex-dependent survival when sex assignment is uncertain: a multievent model of the Balearic Wall Lizard (*Podarcis lifordi*)

In this example, we consider the possibility of erroneous assignment of sex to a newly captured individuals, a situation common to the monitoring of species, for example, with little sexual dimorphisms. In CMR analyses, erroneously assigning sex at the beginning of the capture history leads to bias in the estimated survival difference between the sexes. Here, our aim is to estimate sex-specific survival rates in the Balearic wall lizards, a small species endemic of the Balearic archipelago, Spain. Immature males are sometimes difficult to be sexed in the field and can be confounded with mature females. Sex-specific survival rates can be estimated by accounting for the uncertainty on sex assignment using a multievent CMR model to separate the 'real' sex of the individuals (state) from the 'apparent' (observed) sex (event). In this model, we considered four events (type of observations) which code three states ('real' state of the individual). Data, together with step-by-step instructions to implement the models in E-SURGE and interpret the results, are provided in Supplementary Information S2.

3.4.2. Survival and the issue of tag-loss: study case of the Red Kite (*Milvus milvus)*

Tag-loss is a common issue in wildlife monitoring of marked individuals and can lead to underestimated survival (Arnason and Mills, 1981). To cope with this issue, ecologists have developed advanced methodological tools: from multiple-marking to advanced statistical methods to integrate tag-loss in the individual state (Cowen and Schwarz 2006, Tavecchia et al. 2012). Here is one case study to integrate the loss of a remote tracking device in the multievent modeling framework.

The use of remote tracking devices (radio-satellite and GSM/GPS transmitters) to collect detailed individual history data is increasingly common in the ecological literature (see Chapter 4). A problem in estimating survival from tracking data is that the lifespan of the remote signal is commonly shorter than the lifespan of the individual that carries the device. In this case the survival probability refers to the lifespan of the radio signal and not to the one of the animal. When animals are marked with tags or rings in addition to the radio device, their encounter history can follow the loss of the radio signal. In this example, we illustrate how multievent models can accommodate the loss of the signal and provide unbiased estimate of survival in the presence of radio-loss or radio-failure using a real dataset on Red Kite in the island of Mallorca. Data, together with step-by-step instructions to implement the models in E-SURGE, and interpret the results, are provided in Supplementary Information S3.

3.4.3. Bayesian implementation using the R package BaSTA

The R package Bayesian Survival Trajectory Analysis (BaSTA; Colchero et al. 2012) provides a set of tools that complement other CMR methods when users want to estimate age-specific mortality from CMR datasets where times of birth are known only for few individuals (or none). Several parametric mortality models are available in BaSTA, including the exponential, Gompertz, Weibull, logistic, Makeham and Gompert-Makeham models (Figure 2). In order to include all records in the analysis, BaSTA estimates the missing ages at birth and at death, which reduces the bias in the estimation of the mortality and cumulative survival functions. An example of implementation is provided in Supplementary information S4.

[Figure 2 about here]

3.4.4. Bayesian state-space formulation of the JS model to study stopover decisions of migratory birds using JAGS

Migratory birds cannot realize their journey between breeding and wintering areas in a single flight of thousands kilometres and usually stop-over at places where they can replenish their energy reserves. At these stopover places individuals are not easy to detect. Studying the stopover decisions of migratory birds is a typical case where the detectability need to be taken into account to be able to make strong ecological inference. In this example the survival between two capture occasions (ϕ_i) is considered as the remaining probability at the stopover place and thus $1-\phi_i$ is the departure probability between two occasions. As individuals may arrived in the stopover area before first capture, the model needs to not be conditional on the first capture (as in the CJS model). Thus, we can use the JS model parametrized with entry probabilities noted as η_i for the probability to entry in the stopover area between time i and $i+1$ if not previously entered (Schwarz and Arnason 1996). This SSM formulation of the JS model allows an easy implementation in the Bayesian framework and a straightforward computation of the stopover duration (Lyons et al. 2016). We can also easily incorporate the effect of a weather covariate on the departure probability. The implementation of this example with R and JAGS is provided in Supplementary S5.

3.5. Study cases in plant demography

Plants do not move, but certain aspects of their ecologies, such as vegetative dormancy, variable sprouting times, and complex growth patterns, can make them just as challenging to work with as animals.

3.5.1. Linear modeling of plant survival

The most common method to analyze survival using plant resighting datasets is using linear analysis under a logistic, generalized linear model (GLM), or generalized linear mixed model (GLMM) framework, with survival modeled assuming a binomial distribution (e.g. Salguero-Gómez et al. 2012). This method assumes that at the very least, the resighting of previously observed individuals is nearly perfect, because any phenomenon decreasing re-

detection would be observed as mortality and yield biased survival estimates. In cases where re-detection is not perfect, some have argued that this approach is still useful provided that monitoring datasets are particularly long and large, and that re-detection is still above 90% (Shefferson et al. 2018). However, in studies of vegetative dormancy-prone perennials, dormancy will increasingly be confused with mortality as study length decreases. In the final year of a study, there will be no ability to differentiate the dead from the dormant. This suggests that the final one to three years of data in a study should be used simply to assign dormancy within the remaining data, and that survival should not be estimated for those years. This loss of estimable years adds value to long datasets, particularly those over 10 years long.

We illustrate the use of linear modeling for survival analysis in plants using a case study on *Cypripedium parviflorum*, the North American small yellow lady's slipper, using data collected from 1994 to 2003 within a larger population from Illinois state in the United States. These data were previously used in e.g. Shefferson et al. (2018). Using the R packages lme4 for model fitting and MuMin for model comparison, we show that non-flowering plants have decreasing survival with increasing size, while flowering plants have increasing survival with increasing size. Full procedure including data and R script to run the analysis, are provided in Supplementary Information S6.

3.5.2. CMR survival analysis for plants

Plant population ecologists have long used field methods that may be considered in the same vein as mark-recapture methods in wildlife ecology. However, the application CMR methods to plant population ecology is very recent. In one of the first studies to use mark-recapture analyses in plants, Alexander et al. (1997) faced all of these problems in a population of the Mead's milkweed, *Asclepias meadii*. This population consists of plants that grow in high densities and do not always produce aboveground tissue in a growing season. Closed population mark-recapture analysis allowed them to produce estimates of population size

unbiased by these challenges. Expanding on this work, Shefferson et al. (2001) proposed the use of open population mark-recapture models to estimate annual survivorship in populations in which living individuals do not always sprout in a growing season. Since then, CMR studies have blossomed in plant population ecology, with extensions into the estimation of transition rates among life-history stages (e.g. Shefferson et al. 2003), estimation of the demographic impacts of herbivory (e.g. Kéry and Gregg 2004), investigations into relationships among life-history traits (e.g. Shefferson et al. 2003), tests of correlation with climatic factors (e.g. Shefferson and Tali 2007), the estimation of minimal recruitment levels necessary to sustain populations (Slade *et al.* 2003), and theoretical papers inspired by the problem of unobserved life stages (Kendall and Nichols 2002).

In this example, we also use the *Cypripedium parviflorum* dataset, in combination with the *R* package *marked* (Laake et al. 2013), to investigate costs of reproduction using CMR models. Full procedure including data and R script to run the analysis, are provided in Supplementary Information S7.

4. Mortality analysis in human populations

Modelling mortality in human populations is relatively easier than in non-human ones. For a given group of individuals, we often know their age-at-death, calendar year of the event and their sex. Thanks to these reliable data sources, methodological advances have been produced since De Moivre (1725) and Gompertz (1825). These long-standing demographic and statistical developments have been often drawn by political, military and economic reasons. Below, we present a brief overview of the most common models used to describe mortality patterns over age and/or time on human mortality data.

4.1. Human data and assumptions

For a given sex, we usually have deaths and exposures to the risk of death arranged in two matrices, whose rows and columns are classified by age at death, and year of death. The stochastic assumption behind mortality has a central role in modeling it. The most suitable distribution when we observe mortality data is the Poisson distribution. The aim of any mortality model is to seek for a parsimonious, yet satisfactory description of the so-called force of mortality $\mu_{i,j}$, given observed deaths $d_{i,j}$, and exposures $e_{i,j}$. One could estimate force of mortality in a fully non-parametric framework computing the death rates $\mu_{i,j} \approx \frac{d_{i,j}}{e_{i,j}}$. Simple plots of rates over age and/or time are good tools for a first presentation of mortality development. Commonly, rates are plotted on a logarithmic scale to better acknowledge differences (Figure 3). A relatively strong assumption behind is that within the Poisson distribution mean and variance are equal. When the observed variance is larger than the theoretical one, we often attribute this feature either to overdispersion, or to some hidden patterns in the data. Specific methods for coping with this issue have been proposed in the literature for all models below, though they will not be presented in this chapter. For a comprehensive overview of them see, e.g. Cameron and Trivedi (2013). Furthermore, binomial distribution could be used when we deal with probability of dying (deaths divided by persons-at-risk) and the multinomial distribution has been suggested for modeling mortality from a cohort perspective (Canudas Romo et al., 2018).

[Figure 3 about here]

4.2. Parametric models over age

Parametric models for a suitable representation of the variation of mortality over age have been proposed since Gompertz (1825). He observed that after a certain age, a “law of geometric

progression pervades, in an approximate degree, large portions of different tables of mortality" (Gompertz, 1825, p. 514). He thus suggested representing force of mortality as:

$\mu_i = ae^{bi}$. (1)

where a represents the mortality at time zero (usually age 30/40) and b is the rate of increase of mortality and is frequently used as a measure of the rate of aging.

Makeham (1860) extended Gompertz' equation by adding a constant, an age-independent term, c > 0, to account for risks of death that do not depend on age:

$\mu_i = c + ae^{bi}$. (2)

Human mortality often shows a levelling-off above certain ages (often 80) (Vaupel, 1997). Logistic models have been proposed to portray this feature. Perks (1932) was the first to proposed a logistic modification of the Gompertz-Makeham models. A logistic function to model the late-life mortality deceleration can be given by

$\mu_i = c + \frac{ae^{bi}}{1+\alpha e^{bi}}$. (3)

where α capture the mortality deceleration at oldest ages. This law of mortality can be derived when heterogeneity is assumed in a proportional hazard setting. Commonly called the Gamma-Gompertz model, (3) is the hazard at a population level when standard mortality is described as a Gompertz and frailty values are assumed to be Gamma distributed (Wienke, 2010).

A simplified version of the previous logistic law has been proposed by Kannisto (1992):

$\mu_i = c + \frac{ae^{bi}}{1+ae^{bi}}$. (4)

Heligman and Pollard (1980) derived a descriptive model, covering the whole age range. Here we propose a version for death rates:

$$\mu_i = A^{(i+B)^C} + De^{-E(lni-lnF)^2} + \frac{GH^i}{1+GH^i}. \quad (5)$$

where A, B . . . , H are the eight parameters in the model. Each component aims to describe mortality at childhood, during young-adult ages and at older ages. It is easy to see that such parameterization can cause difficulties in the estimation procedure. Moreover, it would be hard to disentangle the physical meaning of each parameter (Booth and Tickle, 2008).

Another three-component, competing-risk mortality model has been proposed by Siler (1983). Initially developed for animal survival data, this model has been recently used in human demography, especially for simulating possible scenarios in mortality developments (Canudas-Romo, 2018). This model aims at portraying the whole of the age range with five parameters:

$$\mu_i = a_1 e^{-b_i i} + a_2 + a_3 e^{b_3 i} \quad (6)$$

4.3. Over parametrized models: the example of the Lee-Carter

Moving to a two-dimensional perspective and given the wealth of data, traditional demographic methods for analyzing mortality surfaces, i.e. data on deaths and exposures cross-classified by age and year of occurrence, tend to apply a high number of parameters leading to all but parsimonious models. A typical example of this kind is the widely used model introduced by Lee and Carter (1992). In its original formulation, this approach reduces the complexity of the whole surface by introducing the following bi-linear model for the log-death-rates:

$$\ln(m_{ij}) = \alpha_i + \beta_i \kappa_j + \varepsilon_{ij}, \qquad i = l, \dots, m \qquad (7)$$

$$j = l, \dots, n$$

where αi, βi and κj are vectors of parameters to be estimated, and εij represents the error term. Interpretation of the parameters are straightforward: αi and βi describe the general shape of mortality and the fixed rate of mortality improvement at age i respectively, and κj is a time-varying index which captures the general level of mortality. The variance εij in Lee and Carter (1992) is assumed to be constant for all i and j. As presented, the Lee-Carter model (LC) is under-determined and requires additional constraints on the parameters to be successfully estimated. Usually, the model is centered by $\sum \beta_i = 1$, $\sum \kappa_j = 0$.

As pointed out in Section 4.1, a Poisson assumption is more suitable for estimating mortality models and a further development of the Lee-Carter model was devoted to this issue (Brouhns et al., 2002). In Supplementary Information S7 we will provide computational details for estimating the LC model within a Poisson framework.

In the last decades, further variants of this model have been proposed for enhancing several features of the model (e.g. Booth et al. 2002).

4.4. Semi-parametric models: the example of P-splines

A compromise between simple parametric and over-parametrized models could be found in the area of semi-parametric statistics. Without the aim of producing estimated parameters with physical meanings and searching for a good fit to the data, smoothing approaches have been lately developed in the study of nonlinear phenomena. Among numerous options, we will mention here (and illustrate with more details in Supplementary Information S7) a methodology which is particularly suitable for the analysis of mortality developments: B-splines with penalties, known as P-splines. In a unidimensional setting, applications to mortality have been proposed by Currie et al. (2004) and Camarda (2008). Estimation of a model for the complete mortality surface will be achieved by a specific R package:

MortalitySmooth (Camarda, 2012). It is noteworthy that, being an extremely flexible tool, this methodology can be easily implemented for the analysis of demographic phenomena other than mortality. The main idea of the P-splines is to construct an intentionally over-parametrized model and simultaneously to restrict, via a penalty, all redundant features for achieving a wisely parsimonious description of the data. Instead of smoothing a given structure, this approach used local supports such as equally-spaced B-splines over ages and/or years, and penalty term(s) on the associated coefficients.

4.5. Implementation example

We use data from the Human Mortality Database (2019) on Japanese females from year 1960 to 2016 and from age 0 to 100 to fit the models described above. Guidance with R script are provided in Supplementary Information S8.

Concluding remarks

This chapter has illustrated how the choice of a specific method for survival analysis is driven by the research question (e.g. comparing survival across groups vs assessing the effects of environmental predictors), the species life history (more or less easy to monitor) and its living environment (controlled vs. stochastic environments) which determines the type of survival data (e.g. proportion of individuals alive, CMR data, or age at death).

With the rising of new monitoring techniques allowing to gather more and more detailed data at the individual level (see chapter 4) has come the developments of advanced statistical tools for survival analyses, including multievent models and hierarchical models. More than simply assessing survival to feed population projection models (see chapter 8 and 9), modern survival analyses have addressed questions about evolutionary trade-offs (e.g. Nichols and Kendall 1995), static versus dynamic heterogeneity in demographic parameters (e.g. Cam et al. 2016; Gimenez et al. 2018), assessing senescence (e.g. Peron et al. 2010), or quantifying

heritability of demographic parameters in the wild (e.g. Papaïx et al. 2010). Addressing more and more complex questions has brought new challenges into the field of survival analyses. Among others, current methodological developments deal with methods to implement models with numerous states, predict age at death, or consider dependence among individuals when estimating survival.

References

Alexander, H.M., Slade, N.A. and Kettle, W.D. (1997). Application of mark-recapture models to estimation of the population size of plants. Ecology, 78, 1230–1237.

Arnason, A. N. (1973). The estimation of population size, migration rates and survival in a stratified population. Researches on Population Ecology, 15, 1-8.

Arnason, A. N., and Mills, K. H. (1981). Bias and loss of precision due to tag loss in Jolly–Seber estimates for mark–recapture experiments. Canadian Journal of Fisheries and Aquatic Sciences, 38(9), 1077-1095.

Badia-Boher, J. A., Sanz-Aguilar, A., de la Riva, M., Gangoso, L., van Overveld, T., García-Alfonso, M., Luzardo, O. P., Suarez-Pérez, A. and Donázar, J. A. (2019). Evaluating European LIFE conservation projects: Improvements in survival of an endangered vulture. Journal of Applied Ecology. in press.

Balkız, O., Béchet, A., Rouan, L., Choquet, R., Germain, C., Amat, J.A., Rendón-Martos, M., Baccetti, N., Nissardi, S., Ozesmi, U. and Pradel, R. (2010). Experience-dependent natal philopatry of breeding greater flamingos. Journal of Animal Ecology 79, 1045-1056.

Benhaiem, S., Marescot, L., Hofer, H., East, M.L., Lebreton, J.-D., Kramer-Schadt, S. and Gimenez, O. (2018). Robustness of eco-epidemiological capture-recapture parameter estimates to variation in infection state uncertainty. Frontiers in Veterinary Sciences 5, 197.

Booth, H., J. Maindonald, and L. Smith (2002). Applying Lee-Carter under conditions of variable mortality decline. Population Studies 56, 325–336.

Booth, H., and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. Annals of actuarial science, 3(1-2), 3-43.

Brouhns, N., M. Denuit, and J. K. Vermunt (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics 31, 373–393.

Buoro M., Prévost E. and Gimenez, O. (2012). Digging through model complexity: using hierarchical models to uncover evolutionary processes in the wild. Journal of Evolutionary Biology, 25(10): 2077–2090.

Burnham, K.P. and Anderson, D.R. (2002). Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer, New York

Cam, E., and Monnat, J. Y. (2000). Apparent inferiority of first-time breeders in the kittiwake: the role of heterogeneity among age classes. Journal of Animal Ecology, 69(3), 380-394.

Cam, E., Aubry, L. M., and Authier, M. (2016). The conundrum of heterogeneities in life history studies. Trends in Ecology and Evolution, 31(11), 872-886.

Camarda, C. G. (2008). Smoothing Methods for the Analysis of Mortality Development. Ph. D. thesis, Programa de Doctorado en Ingeniería Matemática. Universidad Carlos III, Departamento de Estadística, Madrid.

Camarda, C. G. (2012). MortalitySmooth: An R Package for Smoothing Poisson Counts with P-Splines. Journal of Statistical Software 50, 1–24. Available on www.jstatsoft.org/v50/i01.

Cameron, A. C., and Trivedi, P. K. (2013). Regression analysis of count data (Vol. 53). Cambridge university press.

Canudas-Romo, V., Mazzuco, S., and Zanotto, L. (2018). Measures and Models of Mortality. In Handbook of statistics (Vol. 39, pp. 405-442). Elsevier.

Catchpole, E. A., Morgan, B. J., and Tavecchia, G. (2008). A new method for analysing discrete life history data with missing covariate values. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 70(2), 445-460.

Chambert, T., Rotella, J. J., and Higgs, M. D. (2014). Use of posterior predictive checks as an inferential tool for investigating individual heterogeneity in animal population vital rates. Ecology and evolution, 4(8), 1389-1397.

Choquet, R., Rouan, L., and Pradel, R. (2009). Program E-SURGE: a software application for fitting multievent models. In Modeling demographic processes in marked populations (pp. 845-865). Springer, Boston, MA.

Cowen, L., and Schwarz, C. J. (2006). The Jolly–Seber model with tag loss. Biometrics, 62(3), 699-705.

Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202. Cox, D.R. (2018). Analysis of survival data. Chapman and Hall/CRC.

Colchero, F., Jones, O. R., and Rebke, M. (2012). BaSTA: an R package for Bayesian estimation of age-specific survival from incomplete mark–recapture/recovery data with covariates. Methods in Ecology and Evolution, 3(3), 466-470.

Cubaynes, S., Doutrelant, C., Grégoire, A., Perret, P., Faivre, B., and Gimenez, O. (2012). Testing hypotheses in evolutionary ecology with imperfect detection: capture–recapture structural equation modeling. Ecology, 93(2), 248-255.

Currie, I. D., M. Durbán, and P. H. C. Eilers (2004). Smoothing and Forecasting Mortality Rates. Statistical Modelling 4, 279–298.

De Moivre, A. (1725). Annuities Upon Lives: Or, The Valuation of Annuities Upon Any Number of Lives... NewsBank Readex.

Desprez, M., McMahon, C. R., Hindell, M. A., Harcourt, R., and Gimenez, O. (2013). Known unknowns in an imperfect world: incorporating uncertainty in recruitment estimates using multi-event capture–recapture models. Ecology and Evolution 3, 4658-4668.

Fernández-Chacón, A., Genovart, M., Álvarez, D., cano, J.M., Ojanguren, A.F., Rodriguez-Muñoz, R., Nicieza, A.G. (2015). Neighbouring populations, opposite dynamics: influence of body size and environmental variation on the demography of stream-resident brown trout (Salmo trutta). Oecologia, 78:379-89

Genovart, M., Pradel, R., Oro, D. (2012). Exploiting uncertain ecological fieldwork data with multi-event capture–recapture modelling: an example with bird sex assignment. Journal of Animal Ecology, 81, 970-977.

Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H., ... and Pradel, R. (2007). State-space modelling of data on marked individuals. Ecological Modelling, 206(3-4), 431-438.

Gimenez, O., Lebreton, J. D., Choquet, R., and Pradel, R. (2018). R2ucare: An r package to perform goodness-of-fit tests for capture–recapture models. Methods in Ecology and Evolution, 9(7), 1749-1754.

Gimenez, O., Cam, E., and Gaillard, J. M. (2018). Individual heterogeneity and capture–recapture models: what, why and how?. Oikos, 127(5), 664-686.

Gompertz, B. (1825). XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS. Philosophical transactions of the Royal Society of London, (115), 513-583.

Grosbois, V., Gimenez, O., Gaillard, J.-M., Pradel, R., Barbraud, C., Clobert, J., Møller, A. P. and Weimerskirch, H. (2008). Assessing the impact of climate variation on survival in vertebrate populations. Biological Reviews 83, 357-399.

Grosbois, V. and Tavecchia, G. (2003). Modeling dispersal with capture-recapture data: Disentangling decisions of leaving and settlement. Ecology 84, 1225-1236.

Hamel, S., Gaillard, J. M., Douhard, M., Festa-Bianchet, M., Pelletier, F., and Yoccoz, N. G. (2018). Quantifying individual heterogeneity and its influence on life-history trajectories: different methods for different questions and contexts. Oikos, 127(5), 687-704.

Heligman, L., and Pollard, J. H. (1980). The age pattern of mortality. Journal of the Institute of Actuaries, 107(1), 49-80.

Hooten, M. B., and Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. Ecological Monographs, 85(1), 3-28.

Jönsson KI, Harms-Ringdahl M, Torudd J (2005) Radiation tolerance in the eutardigrade Richtersius coronifer. Int J Radiat Biol 81(9):649–656.

Jenouvrier, S., Tavecchia, G., Thibault, J., Choquet, R. and Bretagnolle, V. (2008). Recruitment processes in long-lived species with delayed maturity: estimating key demographic parameters. Oikos 117, 620-628.

Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika, 52(1/2), 225-247.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American statistical association, 53(282), 457-481.

Kendall, W.L. and Nichols, J.D. (2002). Estimating state-transition probabilities for unobservable states using capture-recapture/resighting data. Ecology 83, 3276-3284.

Kéry, M. and Gregg, K.B. (2004). Demographic analysis of dormancy and survival in the terrestrial orchid Cypripedium reginae. Journal of Ecology, 92, 686–695.

Kéry, M., and Schaub, M. (2011). Bayesian population analysis using WinBUGS: a hierarchical perspective. Academic Press.

King, R., Morgan, B., Gimenez, O., and Brooks, S. (2009). Bayesian analysis for population ecology. Chapman and Hall/CRC.

Kirkwood B. R., Sterne J. A. C. (2003) Essential Medical Statistics 2 Edition Oxford, UK: Blackwell Science Ltd.

Klein, J. P., and Moeschberger, M. L. (2006). Survival analysis: techniques for censored and truncated data. Springer Science and Business Media.

Klein, J.P., Van Houwelingen, H.C., Ibrahim, J.G. and Scheike, T.H. eds. (2016). Handbook of survival analysis. CRC Press.

Laake, J. L., Johnson, D. S., and Conn, P. B. (2013). marked: an R package for maximum likelihood and Markov Chain Monte Carlo analysis of capture–recapture data. Methods in Ecology and Evolution, 4(9), 885-890.

Lagrange, P., Pradel, R., Bélisle, M. and Gimenez, O. (2014). Estimating dispersal among numerous sites using capture-recapture data. Ecology 95, 2316-2323.

Langrock, R., and King, R. (2013). Maximum likelihood estimation of mark–recapture–recovery models in the presence of continuous covariates. The Annals of Applied Statistics, 7(3), 1709-1732.

Lebreton, J.-D, Burnham, K.P., Clobert, J., Anderson, D.R. (1992). Modeling survival and testing biological hypotheses using marked animals - A unified approach with case-studies. Ecological Monographs 62, 67-118.

Lebreton, J.-D., Almeras, T. and Pradel, R. (1999). Competing events, mixtures of information and multistrata recapture models. Bird Study 46,S39-S46.

Lee, R. D., and Carter, L. R. (1992). Modeling and forecasting US mortality. Journal of the American statistical association, 87(419), 659-671.

Lele, S. R., and Dennis, B. (2009). Bayesian methods for hierarchical models: Are ecologists making a Faustian bargain?. Ecological Applications, 19(3), 581-584.

Lyons, J. E., Kendall, W. L., Royle, J. A., Converse, S. J., Andres, B. A., and Buchanan, J. B. (2016). Population size and stopover duration estimation using mark–resight data and Bayesian analysis of a superpopulation model. Biometrics, 72(1), 262-271.

Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. Journal of the Institute of Actuaries, 8(6), 301-310.

McCarthy, M.A. and Masters, P.I.P. (2005). Profiting from prior information in Bayesian analyses of ecological data. Journal of Applied Ecology, 42(6), pp.1012-1019.

McCrea, R. S., and Morgan, B. J. (2014). Analysis of capture-recapture data. Chapman and Hall/CRC.

Metcalf, C. J. E., and Pavard, S. (2007). Why evolutionary biologists should be demographers. Trends in Ecology and Evolution, 22(4), 205-212.

Miller Jr, R.G., 2011. Survival analysis (Vol. 66). John Wiley & Sons. Nichols, J.D. and Kendall, W.L. (1995). The use of multi-state capture-recapture models to address questions in evolutionary ecology. Journal of Applied Statistics 22, 835-846.

Ono F, Mori Y, Takarabe, Fujii A, Saigusa M, Matsushima Y, Yamazaki D, Ito E, Galas S and Saini NL. (2016) Effect of ultra-high pressure on small animals, tardigrades and Artemia, Cogent Physics, 3:1

Papaïx, J., Cubaynes, S., Buoro, M., Charmantier, A., Perret, P., and Gimenez, O. (2010). Combining capture–recapture data and pedigree information to assess heritability of demographic parameters in the wild. Journal of evolutionary biology, 23(10), 2176-2184.

Perks, W. (1932). On some experiments in the graduation of mortality statistics. Journal of the Institute of Actuaries, 63(1), 12-57.

Peron, G., Crochet, P.A., Choquet, R., Pradel, R., Lebreton, J.-D. and Gimenez, O. (2010). Capture-recapture models with heterogeneity to study survival senescence in the wild. Oikos 119, 524-532.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., ... and Smith, P. G. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. British journal of cancer, 35(1).

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing, 124(125.10).

Plummer, M., Stukalov, A., Denwood, M., and Plummer, M. M. (2018). Package 'rjags'. update, 16, 1.

Pollock, K., Nichols, J., Brownie, C., and Hines, J. (1990). Statistical Inference for Capture-Recapture Experiments. Wildlife Monographs 107, 3-97.

Pradel, R. (1993). Flexibility in Survival Analysis from Recapture Data: Handling Trap-Dependence. In J.-D. Lebreton and P. M. North (Eds.), *Marked Individuals in the Study of Bird Population* (pp. 29–37). Basel, Switzerland: Birkhäuser Verlag.

Pradel, R., Hines, J. E., Lebreton, J.-D., and Nichols, J. D. (1997). Capture-recapture survival models taking account of transients. Biometrics 53, 60-72.

Pradel, R. and Lebreton, J.-D. (1999). Comparison of different approaches to the study of local recruitment of breeders. Bird Study 46, 74-81.

Pradel, R. (2005). Multievent: An extension of multistate capture-recapture models to uncertain states. Biometrics 61, 442-447.

Pradel, R. and Sanz-Aguilar, A. (2012). Modeling trap-awareness and related phenomena in capture-recapture studies. PloS One 7, e32666.

Rouan, L., Choquet, R. and Pradel, R. (2009). A general framework for modeling memory in capture-recapture data. Journal of Agricultural Biological and Environmental Statistics 14, 338-355.

Royle, J. A. (2008). Modeling individual effects in the Cormack–Jolly–Seber model: a state–space formulation. Biometrics, 64(2), 364-370.

Salguero-Gómez, R., Siewert, W., Casper, B.B. and Tielbörger, K. (2012). A demographic approach to study effects of climate change in desert plants. Philosophical Transactions of the Royal Society B: Biological Sciences, 367, 3100–3114.

Santidrián Tomillo, P., Robinson, N., Sanz-Aguilar, A., Spotila, J., Paladino, F. and Tavecchia, G. 2017. High and variable mortality of leatherback turtles reveal possible anthropogenic impacts. Ecology 98, 2170-2179.

Sanz-Aguilar, A., Tavecchia, G., Mínguez, E., Massa, B., Lo Valvo, F., Ballesteros, G.A., Barberá, G.G., Amengual, J.F., Rodríguez, A., McMinn, M. and Oro, D. (2010). Recapture processes and biological inference in monitoring burrowing nesting seabirds. Journal of Ornithology, 151, 133-146.

Sanz-Aguilar, A., Igual, J.M., Oro, D., Genovart, M. and Tavecchia, G. (2016). Estimating recruitment and survival in partially monitored populations. Journal of Applied Ecology, 53, 73-82.

Schaub, M., Gimenez, O., Schmidt, B. R., and Pradel, R. (2004). Estimating survival and temporary emigration in the multistate capture–recapture framework. Ecology, 85(8), 2107-2113.

Schaub, M. and Pradel, R. 2004. Assessing the relative importance of different sources of mortality from recoveries of marked animals. Ecology 85: 930-938.

Schwarz, C.J., Schweigert, J.F. and Arnason, A.N. 1993. Estimating migration rates using tag-recovery data. Biometrics, 49, 177-193.

Schwarz, C. J., and Arnason, A. N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. Biometrics, 860-873.

Service, P. M. (2000). Heterogeneity in individual mortality risk and its importance for evolutionary studies of senescence. The American Naturalist, 156(1), 1-13.

Seber, G. A. (1965). A note on the multiple-recapture census. Biometrika, 52(1/2), 249-259.

Shefferson, R.P., Sandercock, B.K., Proper, J. and Beissinger, S.R. (2001). Estimating dormancy and survival of a rare herbaceous perennial using mark-recapture models. Ecology, 82, 145–156.

Shefferson, R.P., Proper, J., Beissinger, S.R. and Simms, E.L. (2003). Life history trade-offs in a rare orchid: the costs of flowering, dormancy, and sprouting. Ecology, 84, 1199–1206.

Shefferson, R.P. and Tali, K. (2007). Dormancy is associated with decreased adult survival in the burnt orchid, Neotinea ustulata. Journal of Ecology, 95, 217–225.

Shefferson, R.P., Kull, T., Hutchings, M.J., Selosse, M.-A., Jacquemyn, H., Kellett, K.M., et al. (2018). Drivers of vegetative dormancy across herbaceous perennial plant species. Ecology Letters, 21, 724–733.

Siler, W. (1983). Parameters of mortality in human populations with widely varying life spans. Statistics in medicine, 2(3), 373-380.

Stearns, S. C. (1992). The evolution of life histories (No. 575 S81).

Tavecchia, G., Pradel, R., Boy, V., Johnson, A.R. and Cezilly, F. (2001). Sex- and age-related variation in survival and cost of first reproduction in Greater Flamingos. Ecology 82:165-174.

Tavecchia, G., Adrover, J., Navarro, A. M. and Pradel, R. (2012). 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.

Tavecchia, G., Sanz-Aguilar, A. and Cannell, B. (2016). Modelling survival and breeding dispersal to unobservable nest sites. Wildlife Research 43: 411-417.

Slade, N.A., Alexander, H.M. and Kettle, W.D. (2003). Estimation of population size and probabilities of survival and detection in Mead's milkweed. Ecology, 84, 791–797.

Vaupel, J. W. (1997). Trajectories of mortality at advanced ages. Between Zeus and the salmon: The biodemography of longevity, 17-37.

White, G.C. and Burnham, K.P. (1999). Program MARK: survival estimation from populations of marked animals. Bird Study 46: S120–S139.

Wienke, A. (2010). Frailty models in survival analysis. Chapman and Hall/CRC.

Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). Analysis and management of animal populations. Academic Press.

Worthington, H., King, R., and Buckland, S. T. (2015). Analysing mark–recapture–recovery data in the presence of missing covariate data via multiple imputation. Journal of Agricultural, Biological, and Environmental Statistics, 20(1), 28-46.