

On the Versatility of Capture-Recapture Modeling: Counting What We Don't See



James D. Nichols

Abstract Initial development of capture-recapture modeling occurred almost exclusively within the disciplines of wildlife management and animal ecology. Virtually all methods for surveying animals “miss” individuals; i.e., some unknown fraction of animals present in surveyed areas goes undetected. In order to draw inferences about all animals actually present, we must deal with this nondetection. In addition, we sometimes misclassify animals as to species, sex, reproductive condition, etc., requiring us to deal with probabilities of misclassification. Capture-recapture models differ from many other kinds of statistical models in that they incorporate parameters that deal with both the process being studied (e.g., population size, survival rate, recruitment rate) and the sampling process giving rise to the data (e.g., capture or detection probability, correct classification probability). Many other disciplines face these same kinds of counting errors, nondetection and misclassification. These disciplines include epidemiology, medicine, social sciences, paleobiology, remote sensing, military imaging, philately, space exploration, quality control, and software development. This chapter includes a brief history of capture-recapture modeling, an introduction to the logic underlying basic models, a discussion of nontraditional uses of these models, and recommendations for additional potential uses.

1 Introduction

Wildlife biologists and animal ecologists realized early on that their methods for surveying animal populations did not provide accurate counts. Animals are missed by virtually all survey methods, and biologists were forced to develop methods that produced not only counts of animals detected but also estimates of those present, but not detected. Several clever approaches have been developed to deal with this

J. D. Nichols (✉)

Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL, USA

issue of nondetection (Seber, 1982; Williams et al., 2002; Seber & Schofield, 2019), and they have seen wide use in ecology and wildlife biology for decades. Another problem faced when surveying animal populations and communities is misclassification. Some of the same factors that cause animals to be difficult to detect (e.g., reliance on partial sightings or auditory cues) can also result in misclassification. For example, an auditory bird survey may result in species misidentification. A visual survey of a sexually monomorphic bird species may result in sex being misclassified. A sighting survey of reproduction in manatees may result in a reproductive female being misclassified as nonreproductive because her young is too far away from her or obscured by her body. Methods for dealing with nondetection have thus been modified to incorporate misclassification as well (Pradel, 2005; MacKenzie et al., 2018).

These same problems in counting and classifying characterize other scientific disciplines as well, but with much less corresponding effort to deal with them. Here, I first provide a brief history of both capture-recapture and closely related occupancy modeling. Then, I survey uses of these models in disciplines other than ecology and wildlife biology and identify opportunities for even greater use.

2 Capture-Recapture

In this section, I provide a brief introduction to capture-recapture modeling by describing the basic ideas underlying several classes of capture-recapture models. Data are typically summarized as capture histories depicting whether or not an individual was captured or detected at each sampling occasion of the study. Rather than developing the full likelihoods for these different classes of model, I define parameters and then write out probability structures for sample capture histories as an abbreviated way of explaining the thinking that underlies these models.

2.1 2-Sample, Closed Population, Single State

The most basic capture-recapture estimator is based on the recognition that the proportion of a specific type of individual or entity in a representative sample from a population should be roughly equal to that in the population itself. Let M be the known number of animals of a certain type in a total population of N individuals. Then define m as the number of animals of that type in a sample of n individuals. If the sample is representative of the population, in the sense of a similar proportion of marked animals, then we expect:

$$\frac{m}{n} \approx \frac{M}{N}. \quad (1)$$

We can rearrange (1) to obtain the following estimator for N :

$$\hat{N} = \frac{nM}{m}. \quad (2)$$

Expression (2) is known as the Lincoln-Petersen estimator, and the prototypical study to which it applies entails two sampling occasions separated by a short time interval over which the sampled population is assumed to be “closed,” with no animals entering or departing. On occasion 1, M animals are captured and marks are applied to them. On occasion 2, another sample of n animals is captured, m of which are found to be marked. This estimator has been independently derived a number of times, initially by Laplace (1786), who used it to estimate the human population of France, and later by Lincoln (1930) to estimate the number of waterfowl in late summer to autumn in North America.

The estimator in (2) can also be viewed as a precursor to the general Horvitz and Thompson (1952) estimator for a population total. M is the number of animals sampled on occasion 1, and m/n estimates p_1 , the probability that a member of the population of size N is caught in occasion 1:

$$\hat{N} = \frac{M}{\hat{p}_1}. \quad (3)$$

The estimator of (2) and (3) and its associated variance have been derived using both hypergeometric (fixed sample size) and binomial (random sampling) likelihoods (reviewed by Seber, 1982, Williams et al., 2002, Seber & Schofield, 2019).

The statistics used in capture-recapture modeling are most frequently written as the number of animals exhibiting each possible capture or detection history. For a 2-sample study, there are only three such statistics denoted as x_{ij} , where $i = 1$ if caught on occasion 1 and 0 if not caught then, and $j = 1$ if caught on occasion 2, and 0 if not.

$$\begin{aligned} x_{11} &= m = \text{number of animals caught on occasions 1 and 2,} \\ x_{10} &= M - m = \text{number of animals caught on occasion 1 but not on occasion 2,} \\ x_{01} &= n - m = \text{number of animals caught on occasion 2 but not on occasion 1.} \end{aligned}$$

The number of animals not captured at either sampling occasion, x_{00} , is unknown, and the problem of estimating total abundance, N , is equivalent to the problem of estimating x_{00} .

Given that an animal is a member of the sampled population, the probabilities of it exhibiting each observable capture history are then:

$$\begin{aligned} \Pr(11) &= p_1 p_2, \\ \Pr(10) &= p_1(1 - p_2), \\ \Pr(01) &= (1 - p_1)p_2, \end{aligned}$$

where p_t is the capture probability for sampling occasion t . Under this 2-sample model, N , p_1 , and p_2 can all be estimated. It is termed a “single state” model because all animals are assumed to have the same probabilities of appearing in a sample (i.e., no stratification by age, sex, size, etc.)

2.2 >2-Sample, Closed Population, Single State

This approach was extended to multiple sampling occasions (Schnabel, 1938; Darroch, 1958) for closed populations. Capture histories were modeled as in the 2-sample case. For a study with five sampling occasions, the probability of an animal in the sampled population showing a capture history of 01010 is:

$$\Pr(01010) = (1 - p_1)p_2(1 - p_3)p_4(1 - p_5).$$

Each possible capture history has an associated probability such as above, and we know how many animals exhibited the history, so we can develop a corresponding likelihood and estimate the capture probabilities and abundance. Subsequent developments included consideration of behavioral response of animals to initial capture, heterogenous capture probabilities, and other generalizations (Otis et al., 1978; Chao and Huggins, 2005a 2005b; Seber & Schofield, 2019).

2.3 >2 Samples, Open Populations, Single State

Capture-recapture methods were extended to “open” populations as well, where sampling occasions could be separated by long time intervals such that gains and losses to the population could occur between occasions (e.g., Jackson, 1933, 1939; Cormack, 1964; Jolly, 1965; Seber, 1965). These models were also based on multinomial likelihoods and required additional parameters for survival of an animal from one sampling occasion t to the next. For example, let p_t denote capture probability for sampling occasion t , and let ϕ_t denote the probability that an animal alive at sampling occasion t survives until occasion $t + 1$ and remains in the sampled population. The conditional probability associated with capture history x_{01010} in a 5-occasion study is:

$$\Pr(01010|\text{release in 2}) = \phi_2(1 - p_3)\phi_3p_4(1 - \phi_4p_5). \quad (4)$$

The last terms in parentheses include both the possibility that the animal survived until 5 but was not caught and the possibility that the animal did not survive.

Likelihoods conditional on new releases in each sampling period can be used to estimate capture probabilities and survival probabilities and, assuming that animals that were and were not previously captured exhibit the same capture probabili-

ties, occasion-specific abundance (see expression 3). Subsequent parameterizations include different ways of modeling the entry of new animals into the sampled population (e.g., Crosbie & Manly, 1981; Pradel, 1996; Schwarz & Arnason, 1996), as opposed to simply conditioning on entries as in (4).

2.4 >2 Samples, Open Populations, Multiple States

Arnason (1972, 1973) introduced the concept of multiple states in which an animal could be captured, where states initially represented different locations and were later generalized to characteristics of individual animals such as age, reproductive condition, body mass, etc. The first multistate models to be widely used allowed capture and survival parameters to depend on age, for studies in which sampling occasions were separated by time intervals that corresponded to the exact interval required for an animal to make the transition from 1 age class to the next (Manly & Parr, 1968; Pollock, 1981; Stokes, 1984). These age-specific models are much simpler than the general models of Arnason (1972, 1973) because of the deterministic, unidirectional nature of age transitions.

In the general multistate models of Arnason (1972, 1973), state transitions are stochastic, necessitating additional new parameters for transitions between states (also see Hestbeck et al., 1991; Brownie et al., 1993; Schwarz et al., 1993). Define θ_t^{rs} as the probability that an animal in state r at occasion t that survives until occasion $t + 1$ is in state s at $t + 1$. Define S_t^r as the probability that an animal in state r at sampling occasion t is still alive and in the sampled population at occasion $t + 1$, and p_t^r as the probability that an animal in state r at occasion t is captured at t . Capture histories must now indicate the state of the animal at each capture. In a study area with two locations, 1 and 2, a capture history of 0102 would indicate an animal first captured in state/location 1 at sampling occasion 2, not captured at occasion 3, and captured in state 2 at occasion 4. The number of animals showing this history is denoted as x_{0102} , and the probability that an animal released in occasion 2, state 1, will exhibit this history and thus appear in this statistic is:

$$\begin{aligned} & \Pr(0102|\text{release in state 1 at occasion 2}) \\ &= S_2^1[(1 - \theta_2^{12})(1 - p_3^1)S_3^1\theta_3^{12} + \theta_2^{12}(1 - p_3^2)S_3^2(1 - \theta_3^{21})]p_4^2. \end{aligned} \quad (5)$$

The portion of expression (5) in brackets reflects the state uncertainty of the animal at occasion 3 and can be viewed as a mixture model incorporating the possibilities that the animal was in state 1 or state 2. Likelihoods are conditional on new releases in each state in each sampling occasion.

These multistate models assume the ability to classify an animal to its appropriate state without error at each capture, and they have been generalized to deal with state uncertainty and misclassification (Kendall et al., 2003, 2004; Nichols et al., 2004; Pradel, 2005). These generalizations include additional classification parameters

and sometimes use ancillary data to reduce uncertainty in modeling the capture and classification processes.

2.5 *Occupancy Models, Closed System, Single State*

Occupancy models extend the thinking underlying capture-recapture from individual animals to a set of locations or sites. The question for a single site is whether a focal species is present or not, and the objective of the modeling is to estimate occupancy, the probability that a site is occupied by the focal species. The motivation for these models is possible nondetection; surveys of sites sometimes “miss” detecting a species, despite presence of the species at the site. A key distinction between the occupancy problem for sites and the capture-recapture problem for individuals within a population is that the number of sites is known and sites can be sampled at every occasion, although the result of the sampling is still characterized by the uncertainty of possible nondetection. Early versions of occupancy models were developed by Geissler and Fuller (1987), Azuma et al. (1990), Nichols and Karanth (2002), and most current modeling is based on MacKenzie et al. (2002).

Sample units for occupancy studies may be naturally occurring units such as ponds or woodlots, or they may be cells in a grid superimposed on a continuous area. Each unit is surveyed on multiple occasions within a relatively short time period (e.g., 2 weeks) over which there are no changes in occupancy. Detection histories are analogous to capture histories and denote the sequence of detections and nondetections at each site. The statistics resulting from such a study are the numbers of sites exhibiting each possible detection history, e.g., x_{101} is the number of sites at which the species was detected on sampling occasions 1 and 3 of a 3-occasion study, but not occasion 2.

The modeling of the detection history data is similar to that for individual animal capture-recapture as well. Define p_t as the probability of detecting the focal species at a sample unit on sample occasion t , and ψ as the probability that a sample unit is occupied by the species. The probability that a surveyed sample unit shows detection history 101 is (MacKenzie et al., 2002):

$$\Pr(101) = \psi p_1 (1 - p_2) p_3. \quad (6)$$

The probability for a site at which the species was not detected in any of the three surveys is:

$$\Pr(000) = \psi (1 - p_1)(1 - p_2)(1 - p_3) + (1 - \psi). \quad (7)$$

We know the species was present for detection history 101, as it was detected, and we assume no false positives (Eq. 6). However, history 000 admits more uncertainty,

as there are two possibilities: the species was present and not detected or the species was absent (Eq. 7).

2.6 *Occupancy Models, Open System, Single State*

Open systems are those for which changes in occupancy status of sites may occur between some sample occasions. Define a primary sample occasion as a relatively short period (e.g., a specific month each year) during which occupancy status of a site is not likely to change. Multiple secondary samples (e.g., four survey days) occur within each primary period. However, the sites are permitted to be open to changes in occupancy between primary periods. For a study with three secondary occasions within each of two primary occasions, a detection history of 101,000 denotes a site with detections at secondary occasions 1 and 3 of primary occasion 1 and no detections in any of the three secondary occasions of primary occasion 2. Barbraud et al. (2003) developed an early model for such data, and the approach of MacKenzie et al. (2003) is the basis for most current modeling.

The modeling of detection probability requires an extra subscript for the two kinds of sampling occasions. Let p_{tk} denote the detection probability associated with secondary period k of primary period t . The possibility of changes in site occupancy requires two new parameters: ε_t is the probability that a site is unoccupied by the species at primary occasion $t + 1$, given that it was occupied at occasion t (local extinction); γ_t is the probability that a site is occupied by the species at occasion $t + 1$, given that it was not occupied in period t (local colonization). The probability associated with the above detection history is thus:

$$\Pr(101\ 000) = \psi_1(p_{11}(1 - p_{12})p_{13})[\varepsilon_1 + (1 - \varepsilon_1)(1 - p_{12})(1 - p_{22})(1 - p_{23})]. \quad (8)$$

The portion of (8) in brackets reflects the uncertainty about whether the species went locally extinct at the site or instead persisted but went undetected. The likelihood is then the product of these probabilities for the detection histories of all sites.

2.7 *Occupancy Models, Multiple States, False Positives*

Sometimes we may want to characterize occupied sites by "state," where state carries additional information about an occupied site. A common situation is where a site occupied by a species can be classified into multiple states that can be ordered by the degree of uncertainty characterizing the state classification (Royle, 2004; Royle & Link, 2005; Nichols et al., 2007; MacKenzie et al., 2009). For example, assume interest in a species and an associated pathogen, such that we designate state

0 as a site not occupied by the focal species, and state 1 as a site occupied by the species but where no individuals of the species have been infected by the pathogen. State 2 denotes occupancy, with pathogen infection of at least one member of the species. In addition to these three true states, we define three observation states that can apply to a site at any secondary occasion survey: 0 = no detection of the species; 1 = detection of the species, but no detection of the pathogen; and 2 = detection of both the species and pathogen. Observation state 0 admits the most uncertainty, as true state may be 0, 1, or 2. For observation state 1, true state may be 1 or 2. Under the assumption of no false-positive errors, observation state 2 is unambiguous, only occurring when true state = 2. The notation and modeling of multistate occupancy become increasingly complex (see MacKenzie et al., 2009, 2018).

The initial development of occupancy modeling assumed no false positives, where these refer to the investigator claiming to detect a species, when the species is actually absent from the sample unit. False positives typically occur when the investigator mistakes an individual or sign of one species for that of another. For example, the pugmark (track) or scat of a large leopard may be mistakenly recorded as that of a tiger. Royle and Link (2006) developed a general, single-season occupancy model that incorporates both nondetection and false positives. Miller et al. (2011) developed models that use two (or more) different detection methods to deal with false positives, and these have been extended to multiple designs (Chambert et al., 2015) and multiple seasons (Miller et al., 2013; MacKenzie et al., 2018).

2.8 Software

Computations for capture-recapture estimation of focal parameters and their variance-covariance structures are relatively complex, such that development of software has been critical to the use of these methods. Early software focused on specific parameterizations of capture-recapture models, whereas development of numerical differentiation algorithms has led to more flexible software, permitting inference for user-specified models. A variety of software packages now exists for implementing capture-recapture analyses. For example, one website (<https://www.capturerecapture.co.uk/software.html>) managed by R. McCrea provides links to a number of available capture-recapture packages.

Program MARK (White & Burnham, 1999; Cooch & White, 2022) implements closed and open capture-recapture models, occupancy models, and a variety of other models useful for inferences about demographic parameters. Program PRESENCE (Hines, 2006) was developed specifically for occupancy models. Historically, PRESENCE has incorporated new classes of occupancy models before other occupancy software. Program M-SURGE (Choquet et al., 2004) was developed to implement multistate capture-recapture models and is based on sufficient statistics, resulting in typically faster computation times than software such as MARK, which is based on individual capture history data. Program E-SURGE (Choquet et al.,

2012) provides a general analytic framework for implementing multistate models in the presence of state uncertainty.

2.9 Summary

Both capture-recapture modeling and closely related occupancy modeling have undergone substantial evolution since their initial development for relatively simple inference problems. Most of this development has been motivated by scientists investigating animal populations and has focused on extensions and generalizations to either estimate additional parameters (i.e., beyond abundance and occupancy) or relax restrictive assumptions.

3 Beyond Traditional Applications

There have been many nontraditional uses of capture-recapture thinking and methodology. Reviews of social science and medical applications include Bohning (2008), Chao (2014), Bird and King (2018), and Bohning et al. (2018). The applications discussed in this chapter are not exhaustive but are illustrative of the diverse estimation problems to which these methods have been applied. Most of these nontraditional uses begin with a focus on abundance of some focal entity, combined with a recognition that the entity is frequently undercounted using the standard survey methods of the discipline.

3.1 Human Health and Epidemiology

3.1.1 Population-Level Inferences

Uses of capture-recapture models for human health applications have a fairly long history, with key early contributions by Wittes and Sidel (1968), Fienberg (1972), Wittes (1974), Wittes et al. (1974), Hook et al. (1980), Hook and Regal (1982, 1992), LaPorte et al. (1992), McCarty et al. (1993), and LaPorte (1994) and useful reviews by IWGDMF (1995a, 1995b), Hook and Regal (1995, 1999), and Chao et al. (2001). Virtually, all of these epidemiological uses are based on data from incomplete lists.

Some capture-recapture applications focus on single lists consisting of frequency distributions of encounters. A single list might include the number of infected individuals for which there was a single recorded encounter (e.g., blood test result, hospital visit), the number with exactly two encounters, three encounters, etc., with the objective to estimate the number of infected individuals that were never

encountered. For example, Polonsky et al. (2018) used single list data to estimate the completeness, and thus effectiveness, of contact tracing.

Multiple list data are typically records of individuals infected with a particular disease or mortalities associated with a specific disease or other cause. In closed populations, the appearance of some individuals on one list and not another is clear evidence of nondetection, and early uses of multiple lists entailed first matching names that appear on multiple lists and then counting the total number of unique individuals. This approach does not include in the total count the number of individuals appearing on none of the lists, and inference about this number motivates the use of capture-recapture.

Multiple list data are encoded as individual capture or detection histories (Sects. 2.1 and 2.2), and the entire data set includes a detection history for every individual appearing on at least one list. For example, Hook et al. (1980) analyzed a data set consisting of three lists of individuals with spina bifida in New York state, 1969–1974. The lists were based on (1) birth certificates, (2) death certificates, and (3) medical rehabilitation records. Closed capture-recapture models were then used to estimate the total number of cases and disease “prevalence,” defined as the proportion of individuals in a population that is infected, or as the probability that a randomly selected individual in a population is infected. Numerous applications of capture-recapture to inferences about numbers of cases and prevalence now exist in the scientific literature.

Multiple list data differ from animal capture data in several ways that must be considered when selecting or developing capture-recapture models for epidemiological uses. The multiple lists are analogous to the multiple sampling periods of the animal ecologist, but unlike these animal sampling periods, there is frequently no natural temporal ordering of list data. Time-specificity of capture probabilities corresponds to list-specificity of detection probabilities. Certain kinds of behavioral response models in capture-recapture are based on temporal order of sampling occasions, and models (e.g., log-linear) for list data have been developed for more general kinds of dependence of detection probabilities for individuals among the different lists (see IWGDMF, 1995a, 1995b; Hook and Regal, 1995; Chao et al., 2001; Rivest and Lvesque, 2001).

Heterogeneous capture probabilities are sometimes associated with identifiable covariates, permitting inference based on a general Horvitz-Thompson approach (e.g., Huggins, 1989, 1991; also see Wang et al., 2006). Several approaches have been developed for the more difficult problem of heterogeneous capture probabilities that cannot be readily associated with covariates (e.g., Burnham & Overton, 1978; Chao, 1987; Norris & Pollock, 1995; Haas & Stokes, 1998; Dorazio & Royle, 2003; Haas et al., 2006). Problems deciding whether two similar records really match (represent the same individual) can occur when constructing detection histories from lists, and approaches for dealing with this problem (e.g., Seber et al., 2000; Lee et al., 2001) are similar, in some ways, to approaches for dealing with tag loss (e.g., Arnason & Mills, 1981; Kremers, 1988; Nichols & Hines, 1993) and misreading (e.g., McClintock et al., 2014).

Open capture-recapture models are used to estimate survival rates, numbers of new recruits, and abundance for populations open to gains and losses across multiple sampling occasions. In the context of disease dynamics, multistate models for open populations can be especially useful, where states are defined, for example, as susceptible, infected, and recovered (SIR), a classification system used in classical compartmental disease models (Kermack & McKendrick, 1927; Bailey, 1975; Cooch et al., 2012). List data based on hospital visits or longitudinal data from studies with imperfect follow-up can be used to develop detection histories for such analyses, and estimated parameters include probabilities of state transition (e.g., the transition from susceptible to infected) and state-specific mortality rates. A feature of multistate capture-recapture models that is especially important for epidemiological uses is state-specific detection probabilities (e.g., infected individuals will typically have higher probabilities of detection for hospital lists than susceptible individuals). Uncertainty in state assignment (e.g., a false-negative or false-positive pathogen test result) led to the development of models to deal with this issue (reviewed by Lebreton et al., 2009), and the multi-event approach of Pradel (2005) provides a general approach to this problem (Conn & Cooch, 2009; Choquet et al., 2013; Benhaïem et al., 2018). Multistate capture-recapture models have been recommended for use in estimating epidemiological state transition probabilities and mortality rates (Jennelle et al., 2007; Cooch et al., 2012; Nichols et al., 2017), but such uses have been relatively rare (but see Viallefont & Auget, 1999) for human diseases.

Occupancy models have several potential uses for epidemiological studies. One use entails viewing individuals as the sample units and focusing on presence or absence of the disease organism (e.g., Bailey et al., 2014; MacKenzie et al., 2018). Multistate occupancy models can be used to estimate transitions (including infection rate) among SIR model states and state-specific mortality rates, as with multistate capture-recapture. Pathogen tests for a random or representative sample of individuals can be used with standard occupancy models to estimate prevalence in the case where false negatives (nondetection) are possible (e.g., Lachish et al., 2012; Nichols et al., 2021), and even infection intensity (Miller et al., 2012). Such testing programs should typically include a subset of individuals that receive multiple tests in order to deal with nondetection.

Occupancy models can also be used to model spatial dynamics of disease spread. Data are based on tests of individuals, but now the sample unit is a location (e.g., a county or city), and interest is in whether any infected individuals are present (McClintock et al., 2010; Bailey et al., 2014; MacKenzie et al., 2018). List data could come from hospital visits, and the replication required for the most general occupancy modeling could be obtained by treating each day or week as a sampling occasion. The ability to deal with imperfect detection is especially important in such studies, as detection probabilities are likely to vary among different locations (e.g., urban locations vs. rural locations far from medical centers).

3.1.2 Individual-Level Inferences

Decisions about individual treatment and quarantine depend on the same diagnostic test results that populate lists. When such tests admit false negatives and positives, it is useful to estimate the probability that a specific individual is infected, conditional on the test result(s). Define p_{lm} as the probability that a test result ($x_k = l$) indicates individual k to be in disease state l , given that true disease state is m . Define state $z_k = 1$ to mean that individual k is infected and state $z_k = 0$ to mean uninfected. Then p_{11} is the probability of correctly detecting infection when present, and its complement ($1 - p_{11}$) is the probability of a false negative. Similarly, p_{10} is the probability of a false positive, incorrectly declaring an individual in state $z_k = 0$ to be infected. Because these probabilities are conditional on the unknown true state of the tested individual, statements about the probability of true infection are also conditional on the underlying pathogen prevalence, ψ . All of the above parameters (detection/classification and prevalence) can be estimated directly using single-season occupancy models (Miller et al., 2011; Chambert et al., 2015; MacKenzie et al., 2018).

The conditional probability that an individual testing positive is actually infected (“positive predictive value”) can be written as:

$$\Pr(z_k = 1|x_k = 1) = \frac{\psi p_{11}}{\psi p_{11} + (1 - \psi)p_{10}}. \quad (9)$$

The conditional probability that an individual testing negative is truly not infected (termed “negative predictive value”) can be written similarly as:

$$\Pr(z_k = 0|x_k = 0) = \frac{(1 - \psi)(1 - p_{10})}{\psi(1 - p_{11}) + (1 - \psi)(1 - p_{10})}. \quad (10)$$

Note that if prevalence parameters are likely to differ for different groups of individuals (e.g., those exhibiting symptoms and those not), then group-specific prevalence parameters should be estimated and used. If the probabilities of an accurate test result [(9) and (10)] are thought to be too small for important decisions about individual treatment, then multiple tests can be used to increase them (e.g., Nichols et al., 2021).

An advantage of the occupancy approach over that frequently used by epidemiologists is that all of the relevant parameters can be estimated together in a joint likelihood. The probabilities of an individual being infected are computed directly as derived parameters, with the associated estimates of sampling variance properly accounting for the variances and covariances of the different parameter estimates.

3.2 *Social Sciences*

3.2.1 **Census**

Governments of most countries conduct periodic “censuses” of population size and distribution. However, direct counts are seldom possible, and virtually all census methods miss individuals (false negatives). Laplace (1786) was the first to derive the estimator (2) and used it to compute the human population size of France using two lists of citizens and their degree of overlap. Sekar and Deming (1949) appeared to derive the estimator (2) independently of (Laplace, 1786) and (Lincoln, 1930) and used it to draw inferences about the numbers of human births and deaths in a district near Calcutta, India. Application of capture-recapture methods (sometimes referred to as “multiple systems estimation” in the social science literature; Fienberg & Manrique-Vallier, 2009; Bird & King, 2018) to problems in the social sciences has increased in recent decades prompting methodological reviews (e.g., Bohning, 2008, Bird & King, 2018) and a book (Bohning et al., 2018).

Capture-recapture models for closed populations have been extended by scientists working with the United States Census Bureau and used with post-enumeration surveys to estimate the census undercount (Wolter, 1986, 1990; Cowan & Malec, 1986; U.S. Census, 2021). Evaluation of census coverage using post-enumeration surveys along with capture-recapture estimation has been recommended by the United Nations (Demographic and Social Statistics Branch, United Nations Statistics Division, 2009) and is being used by various countries (e.g., UK Abbott, 2009; Turkey, Ayhan & Ekni, 2003; Australia, Australian Bureau of Statistics, 2012). In addition to use with standard governmental censuses, capture-recapture approaches have been especially useful for providing inferences about “hidden” populations, groups of individuals that are especially difficult to count using conventional surveys, frequently because they do not wish to be counted (e.g., Sudman et al., 1988).

3.2.2 **Homeless**

Homeless persons are a problematic group for conventional governmental census methods, as they typically lack a mailing address and are not motivated to provide census information. Fisher et al. (1994) obtained list data for homeless persons from multiple sources including hospitals, local social service agencies, a healthcare center designated for homeless, and hostels and used capture-recapture to estimate the homeless population in an area of London. Their estimate was approximately three times larger than the number of list-identified individuals. Such multiple-list approaches have been used with capture-recapture modeling to estimate homeless populations elsewhere as well (e.g., Baltimore, Cowan et al., 1986; Budapest, David & Snijders, 2002). Berry (2007) used an observational approach to identify homeless individuals in Toronto on the street during multiple sampling occasions.

Closed capture-recapture models were used to estimate the homeless population, and detection probabilities of about 0.2 indicated the importance of dealing with nondetection.

An alternative approach to multiple lists is to insert some number of “marked” (M) or “planted” individuals into the focal homeless population and then survey the population directly, estimating detection probabilities as the proportion of planted individuals that is detected (Eq. 3). Laska and Meisner (1993) identified 103 sites frequented by homeless persons in a region of New York City and planted persons in a random sample of 41 of these sites. Census Bureau enumerators were then sent to directly survey homeless persons at these sites. Detection probability was estimated to be 0.48 and used to estimate the total number of homeless in the surveyed areas.

3.2.3 Problem Drug Users

Capture-recapture models have been used with list data on individuals to estimate numbers of problem drug users in various locations. For example, King et al. (2014) used four list sources, probation records, drug intervention program prison assessments, drug treatment facility records, and drug intervention program community assessments, to estimate the number of injecting drug users and heroin-associated deaths in England. They used a Bayesian approach to incorporate prior information into their capture-recapture modeling, obtaining estimates for England, as well as for specific regions within the country. Both prevalence of problem drug use and detection probabilities (probability that a problem drug user appears on at least one list) showed substantial regional variation. Approaches based on similar list data were used to estimate numbers of injecting drug users in Scotland (King et al., 2013). Capture-recapture approaches to inference about problem drug use are numerous and include inferences about the number of HIV-infected injecting drug users in Bangkok (Mastro et al., 1994), prevalence of opiate use in Dublin (Comiskey and Barry, 2001), prevalence of problem drug use in London (Hickman et al., 1999) and six French cities (Vaissade & Legleye, 2008), the risk of arrest of drug dealers and users in Quebec (Bouchard & Trembley, 2005), and the number of heroin users in the Australian Capital Territory (Larson et al., 1994).

3.2.4 Criminal Activities

Greene and Stollmack (1981) applied closed population capture-recapture methods to records from approximately 6000 males arrested at least once in Washington, D.C., 1974–1975. They estimated a total criminal population of about 30,000 individual criminals. Using these same data, Greene (1984) later applied an open population model permitting inferences about growth rate of the offender population, survival probabilities, and average criminal career length. Bouchard et al. (2019) used capture-recapture with arrest and rearrest record data from Quebec to estimate the number of criminals involved with illegal amphetamine-like stimulants.

They estimated that total arrests were only about 12% of those actually engaged in illegal activities and subject to arrest. Bouchard (2007) used capture-recapture methods with arrest data to estimate the number of criminal marijuana growers in Quebec, 1998–2002. Charette and van Koppen (2016) used capture-recapture methods to investigate selectivity in crime punishment, concluding that black male offenders were more likely to be arrested and punished than members of other demographic groups.

Cases of domestic violence in the Netherlands, 2006–2007, were estimated by van der Heijden (2014) using capture-recapture methods with police register records. Their estimates indicated that about 22% of offenders were actually observed and recorded by police. Silverman (2014) used capture-recapture modeling of multiple list data to estimate the number of victims of human trafficking in the UK, 2013. Data on individual victims came from six lists: local authority, police force, national government organization, nongovernment organization, National Crime Agency, and the general public. The estimated victim population was four to five times larger than the number of individuals detected.

Corlatti et al. (2019) studied illegal poaching of red deer in a park in the central Italian Alps, 2007–2017. They estimated age- and sex-specific mortality rates of deer associated with poaching and non-poaching sources using open, multi-event models with data for tagged red deer. Their modeling included parameters for tag loss and the possibility of misclassifying the cause of death (by poaching or not) and provided strong evidence of higher poaching mortality for older males than any other age-sex class.

Barber-Meyer (2010) proposed use of occupancy models with data on species (e.g., tiger parts and products) sold illegally at souvenir shops, traditional medicine stores, etc., within towns. Replication is provided by the multiple stores and shops within each town. Towns were the sample units, such that occupancy estimated the proportion of towns at which the focal species was illegally sold, and multiseason models could be used to estimate occupancy dynamics over time. Sharma et al. (2014) used reports of annual tiger poaching events reported by the Wildlife Protection Society of India, in conjunction with multiseason occupancy modeling, to estimate the prevalence of tiger poaching during periods of 3–7 years in 605 districts throughout India over a 40-year period. Results provided maps of tiger poaching crime and information about covariates associated with such crime.

Yeo et al. (2017) used eBay postings to estimate aspects of illegal elephant ivory trade dynamics in the UK. Each posting was identified by a description, item number, and seller identification, permitting identification of the item in subsequent postings. Postings were surveyed once per week for eight consecutive weeks, March–May 2014. Detection histories were developed for every item and used with open-population capture-recapture models to estimate numbers of items, as well as weekly survival (persistence in the eBay market) and entry probabilities. The authors concluded that a large fraction of illegal ivory sale items had very low probabilities of detection.

3.2.5 World Conflicts

Armed conflicts throughout the world result in numbers of persons being killed or disappearing, and “counts” of these victims are typically biased low. Capture-recapture methods have been used with casualty list data to estimate numbers of victims associated with conflicts in Peru (Ball et al., 2003; Manrique-Vallier et al., 2013) and Colombia (Lum et al., 2013); number of deaths in Kosovo, March–June 1999 (Ball & Asher, 2002); and the number of persons killed by state forces in Guatemala, 1981–1983 (Ball, 2000). For example, in the Guatemala analysis, lists of victims were provided by the following three sources, the Commission for Historical Clarification, the International Center for Human Rights Research, and the Catholic Church’s Interdiocesan Project for the Recuperation of Historical Memory. The estimated number of killings was about three times larger than the sum of victims identified via the three lists.

Social conflict events from some parts of the world are not well reported, such that counts of such events are typically biased low. Hendrix and Saleyhan (2015) used closed population capture-recapture models to estimate the number of social conflict events occurring across Africa in 2012. They obtained detection/nondetection data on 1443 events from the Social Conflict in Africa Database. They used data from two independent news agencies, Associated Press (AP) and Agence France-Presse (AFP), compiling statistics on numbers of events reported only by AP, only by AFP, and by both agencies. They concluded that these two news sources captured approximately 76% of all events in Africa and that the nondetection rate was predictably smaller for deadly events, events of a larger magnitude, and events associated with government repression.

3.3 Quality Control

Capture-recapture models have been used for several specific problems associated with quality control. Jewell (1985) noted that defects or errors can occur in production of various manufactured goods, in computer software, in manuscripts, etc. and recommended capture-recapture approaches for estimating numbers of them. Quality control efforts typically involve inspectors or proofreaders who examine products for defects or errors, but errors may go undetected. One approach to estimating number of errors/defects in the face of nondetection is to employ multiple inspectors or proofreaders. In the case of three inspectors, for example, each error detected by at least one inspector is represented by a row of three entries (one entry for each inspector), with a 1 denoting detection of the error by the particular inspector and a 0 denoting nondetection. Chao and Yang (1993) used this approach with computer code examined by multiple coders looking for errors and estimated the number of errors remaining (undetected). White et al. (1982) used this approach with multiple proofreaders of a large manuscript and estimated the number of undetected errors.

3.4 *Remote Sensing*

“Remote sensing” refers to use of aircraft or satellites to obtain information about the earth. Photography and video are typically used to provide images, which are then examined by individuals or computers in order to enumerate focal entities (e.g., wetlands, woodlots) or compute area measurements of specific cover types. However, such analyses of remote sensing images are usually characterized by two types of errors, nondetection and misclassification. These errors can sometimes be dealt with via replication provided by multiple observers in aircraft or multiple persons processing the same image. Capture-recapture methods are then used to estimate number of entities, for example, using the number of entities detected by just one observer, two observers, etc. (Magnusson et al., 1978; Cook and Jacobson, 1979).

More commonly, a sample of area covered by a survey is visited by ground observers providing direct counts and classifications, known as ground truthing. The number of ground truth entities that is correctly detected or classified via the remote images is then used with capture-recapture thinking to estimate detection and correct classification probabilities (see Maxim et al., 1981; Maxim & Harrington, 1982, 1983). Veran et al. (2012) focused on the question of land cover dynamics, noting that classification errors can be made at times t , $t + 1$, or both times, leading to large errors in estimates of land cover state transition probabilities. They proposed use of ground truth data with multistate capture-recapture models that included state misclassification as a means of directly estimating land cover transitions in the face of classification errors.

3.5 *Paleobiology*

Paleobiologists have long recognized that nondetection is an important issue for analyzing fossil data (e.g., Foote & Raup, 1996). Analyses that do not account for nondetection are subject to serious errors, as detection probabilities are thought not only to be substantial but also to vary across time and space (Brett, 1998). Paleobiological data consist of records of fossil taxa found via sampling at different strata (different geologic time horizons) and locations. Capture-recapture analyses typically treat each lower-level taxon (e.g., family) within a higher-level taxon (e.g., phylum) as an “individual.” Detection histories for each lower-level taxon can be developed using spatial samples (analogous to multiple lists) within some time stratum and area of interest, providing the data for estimation of total taxa using closed capture-recapture models. Detection histories for focal taxa developed from different geologic strata (time horizons) at the same sampling location, or even worldwide, can be used with open capture-recapture models to estimate number of taxa and rates of both local and global taxonomic origination and extinction for lower-level taxa (Nichols & Pollock, 1983). Capture-recapture models were

introduced to paleobiology over 40 years ago (Rosenzweig & Duek, 1979; Nichols & Pollock, 1983; Conroy & Nichols, 1984; Nichols et al., 1986), but they have seen only limited use (Connolly & Miller, 2001a, 2001b, 2002).

Occupancy models have multiple uses for fossil data as well, with focus on a specific taxon, rather than a group of lower-level taxa. Detection-nondetection data from replicate local samples can be used in conjunction with occupancy models to estimate geographic distribution (Liow & Nichols, 2010; Liow, 2013). Detection histories based on different time horizons from multiple locations can be used with occupancy modeling to estimate local probabilities of colonization and extinction as well (Liow & Nichols, 2010; Liow, 2013). Occupancy models were introduced to paleobiologists much more recently than capture-recapture models (Liow & Nichols, 2010; Liow, 2013; MacKenzie et al., 2018), and paleobiological use of occupancy approaches has been limited (but see Lawing et al., 2021).

3.6 *Miscellaneous Applications*

National databases for traffic accidents are maintained by law enforcement agencies in many countries, but accidents are thought to be underreported, leading to many efforts to estimate their true numbers using capture-recapture. Razzak and Luby (1998) compiled lists of police accident records and emergency ambulance service records over a 10-month period during 1994 in Karachi, Pakistan. Their estimates indicated that official records accounted for 56% of traffic accident deaths and only 4% of serious injuries. Capture-recapture inferences about traffic accidents have been used in various other locations including Nicaragua (Tercero & Andersson, 2004), Ethiopia (Abegaz et al., 2014), and Mali (Sango et al., 2016).

Beirne and Lambin (2013) studied volunteer “citizen scientists” who worked on a project to remove invasive mink from a large area of Scotland. Their objective was to draw inferences about volunteer retention (tendency to remain in the program) and the factors that affected it. They described the potential utility of open capture-recapture approaches, but collected data on volunteer activity data via telephone every 6 months and were thus able to use known-fate models (Pollock et al., 1989). They identified volunteer vocation and recent trapping and removal success as key determinants of retention in the program.

Interest in vocabulary size has prompted literary scholars to count the number of individual words that an author uses in her/his writing, but this number is likely smaller than the number actually known to the author. Efron and Thisted (1976) counted the number of words used once, twice, three times, etc., in samples of Shakespeare’s writing in order to estimate the total number of words that he knew. Words counted in the samples totaled 31,534, and capture-recapture estimators indicated that he knew about 35,000. Capture-recapture methods have also been used to estimate the song repertoire size of birds (Garamszegi et al., 2002).

An archaeological use of capture-recapture modeling was provided by Holst (1981), with subsequent reanalyses using different capture-recapture estimators by

Esty (1982, 1983) and Chao (1984). The problem was to estimate the number of different “dies” that produced a set of 204 coins in ancient India. The data were the number of dies that produced only a single coin in the sample, two coins in the sample, three, etc.

Herendeen and White (2013) collected data on appearances of specific rare stamps over the years from sources such as auction catalog, retail price lists, copies of expert certificates, and similar records. Each stamp has an individual identifying number. Herendeen and White (2013) viewed each year as a sampling occasion and used closed population models to estimate the total number of stamps still in existence.

Nichols et al. (2013a, 2013b) used capture-recapture thinking to estimate detection and classification probabilities for military imaging systems. Vessels of different classes (defined by size, and military vs. civilian status) were experimentally positioned at different distances from two new cameras. “No vessel” was one of the experimental possibilities as well. Resulting data were used to develop a model for detection and classification probabilities as a function of distance and vessel type. Capture-recapture model selection was used to infer that distance relationships were dependent on camera type but characterized by a common slope across vessel types (Nichols et al., 2013b).

K.H. Pollock (pers. comm.) used capture-recapture models to estimate the number of man-made objects orbiting earth. The field of astronomy is characterized by substantial nondetection, with detection probabilities a function of telescope type as well as distance, brightness, and size of focal object, and a number of potential uses of capture-recapture thinking can be envisaged.

4 Discussion

The problems of nondetection and misclassification characterize numerous types of count data. The various applications described in Sect. 3 have hopefully supported this assertion, and there are certainly many more applications that can be imagined. The adoption of capture-recapture thinking has not been as rapid as might be hoped for any of the disciplines of Sect. 3, and rate of adoption has varied among these disciplines. For example, my impression based on literature review is that epidemiological and human health applications are somewhat more common than those dealing with social sciences, whereas adoption within paleobiology has been very slow.

One possible explanation underlying this variation in rate of adoption involves the perceived severity of nondetection and misclassification (Nichols, 2019). For example, the development of capture-recapture thinking in the fields of wildlife and animal ecology is likely a natural response to the well-known nondetection problems associated with virtually all animal survey methods. Not only do surveys miss animals, but the fraction missed can be very large. In contrast, epidemiological and social science data based on counts of humans have historically been thought to

be closer to truth, although this perception is changing. Indeed, some of the detection probability estimates of Sect. 3 are quite small.

A second factor that may affect rate of adoption of capture-recapture is the cost of incorrect counts and estimates, and thus the scrutiny that analytic results receive. Results of epidemiological surveys and medical diagnoses are viewed as extremely important and are often carefully reviewed, as misleading inferences can have detrimental consequences that may be readily apparent. In contrast, inferences in the social sciences are certainly important, but not so highly scrutinized, and misleading inferences are less likely to be recognized. The greater the degree of scrutiny, the greater the expected attention to analytic details and inferential errors.

Another factor that may affect rate of adoption is the funding available to a discipline. Good funding helps ensure collaboration of statisticians, who are able to deal with the added complexity of modeling sampling processes. Epidemiology and human health are among the better funded scientific disciplines in most countries.

One more factor influencing methodological adoption is likely the familiarity of scientists with capture-recapture approaches. My search for uses of capture-recapture models for inferences about criminal activity produced a number of papers that used capture-recapture to study the criminal activity of animal poaching. I am guessing that the appearance of disproportionate numbers of applications for this particular type of crime resulted from prior familiarity of scientists investigating such crimes with capture-recapture approaches.

This relatively slow adoption of robust methods for dealing with nondetection and misclassification begs the question: what are the alternatives to modeling these components of the sampling process? The most common alternative appears to be to view the problems as so small and insignificant that they can be safely ignored. For example, this has been the case with remote sensing uses, as ground truthing data have provided clear evidence of nondetection and misclassification. However, estimates of these errors are frequently presented, claimed to be small, and then ignored in analysis (see discussion in Veran et al., 2012). I suspect that this alternative is also prevalent in disciplines where errors are not so readily estimated, but rather assumed or claimed to be small and thus not worthy of the effort to deal with nondetection.

A second alternative to the use of capture-recapture is to try to identify the key sources of variation in detection probability or misclassification, to develop models for each of these component processes separately, and then to combine these models to provide overall inferences about detection probability and the focal parameters that they influence. I encountered this approach at a 2011 workshop dealing with nondetection. A biostatistician working for the Centers for Disease Control and Prevention outlined this approach to inference about detection probability for a focal disease. Her strategy was to develop a model for each of 10–12 sources of variation in detection and to then combine these models into an overall model for the sampling process. At the time of the workshop, two of these models had been developed. In contrast to this incremental approach, capture-recapture

requires some form of replicate sampling (e.g., multiple lists) and then directly uses the information about nondetection available in detection history data. Detection probability, as well as focal parameters such as numbers of cases or disease prevalence, is estimated without the need for identification and modeling of all factors affecting nondetection. If potential influencing factors can be identified, then they can be treated as covariates in capture-recapture modeling, with the result that their influence can be formally tested and, if found to be important, included in the modeling of detection probability.

A problem related to slow adoption of capture-recapture approaches is the limited inferences to which they are applied. The early history of capture-recapture in wildlife and animal ecology was dominated by a focus on numbers. Capture-recapture estimators for closed populations were used to provide estimates of population size for specific locations and times. However, abundance is not necessarily interesting by itself, but rather is more usefully viewed as a state variable in studies of dynamical processes. The primary interests are sources of spatial and temporal variation in abundance, and the ability of human actions to influence population size. This recognition eventually led to increased interest in capture-recapture models for open populations that experience dynamical changes between sampling occasions. Today's capture-recapture studies of wildlife populations tend to focus on the processes of birth, death, and movement, and on transition probabilities governing changes of state within individuals. Similarly, occupancy estimates themselves are not viewed as especially interesting, and focus has shifted to the probabilities of local extinction and colonization that govern occupancy dynamics.

This review of nontraditional applications of capture-recapture models suggests to me that most of these studies are focusing on numbers of focal entities. As noted in Sect. 3.1.1, epidemiological list data from hospital visits and data from longitudinal studies with incomplete follow-up can be used with multistate capture-recapture models for open populations to draw inferences about state transition probabilities and state-specific mortality rates required by SIR (susceptible, infected, recovered) models. These models provide a way to deal with the state-specific detection probabilities likely to exist in longitudinal data. For example, an individual in the infected state at sampling occasion t is more likely to be found on a hospital list or re-encountered in a longitudinal study at that occasion than an individual in the susceptible or recovered state. Accompanying information on public health interventions or even individual treatments can be used with these models to directly test the efficacy of interventions and treatments. Despite the potential utility of multistate (Lebreton et al., 2009) and multi-event (Pradel, 2005) capture-recapture models, I saw little evidence that these approaches are being used in disciplines other than animal and wildlife ecology.

The primary interest of most capture-recapture applications in criminology (Sect. 3.2.4) was in numbers of criminals or victims, or in detection probability when this equated with probability of arrest. These studies were not focused on the influences of laws or enforcement interventions on criminal activity, or on the

effects of crimes on victims. Contrast this with a number of the studies of wildlife poaching crimes which included investigations of poaching-related mortality rates (Corlatti et al., 2019), and of relationships between dynamics of poaching activity and species distributions (Marescot et al., 2019; Moore et al., 2021), and between ranger (enforcement) activity and poaching activity (Moore et al., 2017). I suspect that a main reason for this difference in uses of capture-recapture for wildlife vs. other crimes stems from the familiarity of persons focused on wildlife crimes with these more complicated models and their utility.

With the exception of these investigations of wildlife crimes, the majority of the nontraditional uses of capture-recapture have focused on numbers. Such studies can be very useful when these numbers are incorporated into a larger sampling scheme designed to test hypotheses about system dynamics or effects of potential interventions. However, for most of the reviewed papers describing nontraditional uses of capture-recapture, this was not the case. I believe that studies that go beyond estimates of numbers to focus on system dynamics and key relationships (e.g., effects of interventions or treatments) are much more likely to be useful to both science and decision-making. Capture-recapture and occupancy models developed for open populations are especially useful for investigating underlying processes, and I would hope that we see the same increases in use of these models in nontraditional applications as we did in the fields of wildlife and animal population ecology.

In summary, studies in many disciplines are based on count data, yet counts are frequently inaccurate because of nondetection and misclassification. The fields of wildlife and animal ecology recognized these problems nearly a century ago and began to develop capture-recapture approaches to deal with them. Other disciplines began to adopt these methods and, in some cases, to modify them for their specific applications (e.g., log-linear models for closed populations to deal with list dependence). However, the integration of capture-recapture methods into the toolboxes of scientists of non-ecological disciplines has been incomplete and slower than might be hoped. In addition, the bulk of non-traditional uses of capture-recapture models has focused on estimation of totals for counted entities. I hope that use of capture-recapture models for nontraditional applications increases and that such uses better exploit the open-population models that permit inferences about dynamical processes.

Acknowledgments Thanks to Dainel F. Heitjan and H.K.T. Ng for the invitation to contribute to this volume, for useful comments on the initial manuscript, and for information about nontraditional uses of capture-recapture thinking. Constructive comments on the manuscript were also provided by an anonymous reviewer. Thanks to Ken Burnham and Ken Pollock for providing information on nontraditional uses of capture-recapture and to Evan Cooch for discussions of this topic. Thanks especially to Dr. Lynne S. Stokes for the 3 years that she spent working at a US Fish and Wildlife Service research lab early in her career. Lynne was a great teacher and collaborator and played a major role in increasing my statistical capabilities, as well as those of our entire research group.

References

- Abbott, O. (2009). 2011 UK Census coverage assessment and adjustment methodology. *Popular Trends*, 137, 2532.
- Abegaz, T., Berhane, Y., Worku, A., Assrat, A., & Assefa, A. (2014). Road traffic deaths and injuries are under-reported in Ethiopia: A capture-recapture method. *PLoS ONE* 9(7), e103001. <https://doi.org/10.1371/journal.pone.0103001>
- Arnason, A. N. (1972). Parameter estimates from mark-recapture experiments on two populations subject to migration and death. *Researches on Population Ecology*, 13, 97–113.
- 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., & 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, 1077–1095.
- Australian Bureau of Statistics. (2012). 2011 census of population and housing: details of undercount. Report 2940.0. Canberra: Australian Bureau of Statistics.
- Ayhan, H.Ö., & Ekni, S. (2003). Coverage error in population censuses: The case of Turkey. *Survival Methods*, 29, 155165.
- Azuma, D. L., Baldwin, J. A., & Noon, B. R. (1990). Estimating the occupancy of spotted owl habitat areas by sampling and adjusting for bias. USDA Gen. Tech. Rep. PSW-124, Berkeley, CA.
- Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases* (2nd ed.). Macmillan.
- Bailey, L. L., MacKenzie, D. I., & Nichols, J. D. (2014). Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5(12), 1269–1279.
- Ball, P. (2000). The Guatemalan commission for historical clarification: Intersample analysis. In P. Ball, H. Spierer & L. Spierer (Eds.), *Making the case: Investigating large scale human rights violations using information systems and data analysis* (pp. 259–285). American Association for the Advancement of Science.
- Ball, P., & Asher, J. (2002). Statistics and Slobodan: Using data analysis and statistics in the war crimes trial of former president Milosevic. *Chance*, 15, 17–24.
- Ball, P., Asher, J., Sulmont, D., & Manrique, D. (2003). How many Peruvians have died? An estimate of the total number of victims killed or disappeared in the armed internal conflict between 1980 and 2000 Washington, DC: Report to the Peruvian Commission for Truth and Justice (CVR).
- Barber-Meyer, S. M. (2010). Dealing with the clandestine nature of wildlife-trade market surveys. *Conservation Biology*, 24, 918–923.
- Barbraud, C., Nichols, J. D., Hines, J. E., & Hafner, H (2003). Estimating rates of extinction and colonization in colonial species and an extension to the metapopulation and community levels. *Oikos*, 101, 113–126.
- Beirne, C., & Lambin, X. (2013). Understanding the determinants of volunteer retention through capture-recapture analysis: Answering social science questions using a wildlife ecology toolkit. *Conservation Letters*, 6, 391401.
- Benhaïem, S., Marescot, L., Hofer, H., East, M. L., Lebreton, J.-D., Kramer-Schadt, S., Gimenez, O. (2018). Robustness of eco-epidemiological capture-recapture parameter estimates to variation in infection state uncertainty. *Frontiers in Veterinary Science*, 5. <https://doi.org/10.3389/fvets.2018.00197>
- Berry, B. (2007). A repeated observation approach for estimating the street homeless population. *Evaluation Review*, 31, 166–199.
- Bird, S. M., & King, R. (2018). Multiple systems estimation (or capture-recapture estimation) to inform public policy. *Annual Review of Statistics and Its Application*, 5, 95–118.
- Bohning, D. (2008). Editorial—recent developments in capture-recapture methods and their applications. *Biometrical Journal*, 50, 954956.

- Bohning, D., van der Heijden, P. G. M., & Bunge, J. (Eds.). (2018). *Capture-recapture methods for the social and medical sciences*. CRC Press
- Bouchard, M. (2007). A capture-recapture model to estimate the size of criminal populations and the risks of detection in a marijuana cultivation industry. *Journal of Quantitative Criminology*, 23, 221–241.
- Bouchard, M., Morselli, C., MacDonald, M., Gallupe, O., Zhang, S., & Farabee, D. (2019). Estimating risks of arrest and criminal populations: Regression adjustments to capture-recapture models. *Crime & Delinquency*, 65, 1767–1797.
- Bouchard, M., & Trembley, P. (2005). Risks of arrest across drug markets: A capture-recapture analysis of “hidden” dealer and user populations. *Journal of Drug Issues*, 35, 733–754.
- Brett, C. E. (1998). Sequence stratigraphy, paleoecology, and evolution; biotic clues and responses to sea-level fluctuations. *Palaos*, 13, 241–262.
- Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H., & Hestbeck, J. B. (1993). Capture-recapture studies for multiple strata including non-Markovian transition probabilities. *Biometrics*, 49, 1173–1187.
- Burnham, K. P., & Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika*, 65, 625–633.
- Chambert, T., Miller, D. A. W., & Nichols, J. D. (2015). Modeling false positive detections in species occurrence data under different study designs. *Ecology*, 96, 332–339
- Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, 11, 265–270.
- Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, 43, 783–791.
- Chao, A. (2014). *Capture-recapture for human populations*. Wiley stats ref: Statistics reference Online. Wiley.
- Chao, A., & Huggins, R. M. (2005a). Classical closed-population capture-recapture models. In S. C. Amstrup, T.L. McDonald & B. F. J. Manly (Eds.), *Handbook of capture-recapture analysis* (pp. 22–35). Princeton Univ. Press.
- Chao, A., & Huggins, R. M. (2005b). Modern closed-population capture-recapture models. In S. C. Amstrup, T.L. McDonald & B. F. J. Manly (Eds.), *Handbook of capture-recapture analysis* (pp. 58–87). Princeton Univ. Press.
- Chao, A., Tsay, P. K., Lin, S. -H., Shau, W. -Y., & Chao, D. -Y. (2001). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20, 3123–3157.
- Chao, A., & Yang, M. C. K. (1993). Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika*, 80, 193201.
- Charette, Y., & van Koppen, V. (2016). A capture-recapture model to estimate the effects of extra-legal disparities on crime funnel selectivity and punishment avoidance. *Security Journal*, 29, 561–583.
- Choquet, R., Reboulet, A. M., Pradel, R., Gimenez, O., & Lebreton, J. D. (2004). M-SURGE new software specifically designed for multistate capture-recapture models. *Animal Biodiversity and Conservation*, 27(1), 207–215.
- Choquet, R., Rouan, L., & Pradel, R. (2012). Program E-SURGE: A software application for fitting multievent models. In D. L. Thomson, E. G. Cooch, M. J. Conroy (Eds.), *Modeling demographic processes in marked populations* (Vol. 3, pp. 845–865). Environmental and Ecological Statistics. <https://doi.org/10.1007/978-0-387-78151-8>
- Choquet, R., Carrie, C., Chambert, T., & Boulinier, T. (2013). Estimating transitions between states using measurements with imperfect detection: Application to serological data. *Ecology*, 94, 2160–2165.
- Comiskey, C. M., & Barry J. M. (2001). A capture-recapture study of the prevalence and implications of opiate use in Dublin. *European Journal of Public Health*, 11, 198–200.
- Conn, P. B., & Cooch, E. G. (2009). Multi-state capture-recapture analysis under imperfect state observation: An application to disease models. *Journal of Applied Ecology*, 46, 486492.
- Connolly, S. R., & Miller, A. I. (2001a). Global Ordovician faunal transitions in the marine benthos: Proximate causes. *Paleobiology*, 27, 779–795.

- Connolly, S. R., & Miller, A. I. (2001b). Joint estimation of sampling and turnover rates from fossil databases: Capture-mark-recapture methods revisited. *Paleobiology*, 27, 751–767.
- Connolly, S. R., & Miller, A. I. (2002). Global Ordovician faunal transitions in the marine benthos: Ultimate causes. *Paleobiology*, 28, 26–40.
- Conroy, M. J., & Nichols, J. D. (1984). Testing for variation in taxonomic extinction probabilities: A suggested methodology and some results. *Paleobiology*, 10, 328–337.
- Cooch, E. G., Conn, P. B., Ellner, S. P., Dobson, A. P., & Pollock, K. H. (2012). Disease dynamics in wild populations: Modeling and estimation: A review. *Journal of Ornithology*, 152, 485–509.
- Cooch, E. G., & White, G. C. (2022). *Program MARK – A Gentle Introduction* (22nd ed.). <http://www.phidot.org/software/mark/docs/book/>
- Cook, R. D., & Jacobson, J. O. (1979). A design for estimating visibility bias in aerial surveys. *Biometrics*, 35, 735–742.
- Corlatti, L., Sanz-Aguilar, A., Tavecchia, G., Gugiatti, A., & Pedrotti, L. (2019). Unravelling the sex- and age-specific impact of poaching mortality with multievent modeling. *Frontiers in Zoology*, 16, 20. <https://doi.org/10.1186/s12983-019-0321-1>
- Cormack, R. M. (1964). Estimates of survival from the sightings of marked animals. *Biometrika*, 51, 429–438.
- Cowan, C. D., & Malec, D. (1986). Capture-recapture models when both sources have clustered observations. *Journal of the American Statistical Association*, 81, 347–353.
- Cowan, C. D., Breakey, W. R., & Fischer, P. J. (1986). The methodology of counting the homeless. In *Proc. Surv. Res. Meth. Sect.* (pp. 170–175). American Statistical Association.
- Crosbie, S. F., Manly, B. F. J. (1981). Parsimonious modeling of capture-mark-recapture studies. *Biometrics*, 41, 385–398.
- Darroch, J. N. (1958). The multiple-recapture census. I. Estimation of a closed population. *Biometrika* 45, 343–359.
- David, B., & Snijders, T. A. B. (2002). Estimating the size of the homeless population in Budapest, Hungary. *Quality & Quantity*, 36, 291–303.
- Demographic and Social Statistics Branch, United Nations Statistics Division. (2009). *Manual on census evaluation. Post enumeration surveys*. United Nations.
- Dorazio, R. M., & Royle, J. A. (2003). Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics*, 59, 351–364.
- Efron, B., & Thisted, R. (1976). Estimating the number of unseen species: How many words did Shakespeare know? *Biometrika*, 63, 435–447.
- Esty, W. W. (1982). Confidence intervals for the coverage of low coverage samples. *The Annals of Statistics*, 10, 190–196.
- Esty, W. W. (1983). A normal limit law for a nonparametric estimator of the coverage of a random sample. *The Annals of Statistics*, 11, 905–912.
- Fienberg, S. E. (1972). The multiple recapture census for closed populations and incomplete 2^k contingency tables. *Biometrika*, 59, 591–603.
- Fienberg, S. E., & Manrique-Vallier, D. (2009). Integrated methodology for multiple systems estimation and record linkage using a missing data formulation. *Advances in Statistical Analysis*, 93, 49–60.
- Fisher, N., Turner, S. W., Pugh, R., & Taylor, C. (1994). Estimated numbers of homeless and homeless mentally ill people in north east Westminster by using capture-recapture analysis. *BMJ* 308, 27–30.
- Foote, M., & Raup, D. M. (1996). Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, 22, 121–140.
- Garamszegi, L. Z., Boulinier, T., Moller, A. P., Torok, J., Michl, G., & Nichols, J. D. (2002). The estimation of size and change of avian song repertoires. *Animal Behaviour*, 36, 623–630.
- Geissler, P. H., & Fuller, M. R. (1987). Estimation of the proportion of area occupied by an animal species. In *Proc. Sect. Surv. Res. Meth. Amer. Stat. Assoc.* (pp. 533–538).
- Greene, M. A. (1984). Estimating the size of a criminal population using an open population approach. In *Proc. Amer. Stat. Assoc. Surv. Res. Meth. Sec.* (pp. 8–13).

- Greene, M. A., & Stollmack, S. (1981) Estimating the number of criminals. In J. A. Fox (Ed.), *Models in quantitative criminology*. Academic Press.
- Haas, P. J., & Stokes, S. L. (1998). Estimating the number of classes in a finite population. *Journal of the American Statistical Association*, *93*, 1475–1487.
- Haas, P. J., Liu, Y., & Stokes, L. (2006). An estimator of number of species from quadrat sampling. *Biometrics*, *62*, 135–141.
- Hendrix, C. S., & Saleyhan, I. (2015). No news is good news: Mark and recapture for event data when reporting probabilities are less than one. *International Interactions*, *41*, 392–406.
- Herendeen, D. L., & White, G. C. (2013). Statistical estimates of rare stamp populations. *The History of Science and Technology at the Smithsonian*, *57*, 91–100.
- Hestbeck, J. B., Nichols, J. D., & Malecki, R. A. (1991). Estimates of movement and site fidelity using mark-resight data of wintering Canada geese. *Ecology*, *72*, 523–533.
- Hickman, M., Cox, S., Harvey, J. Howes, S., Farrell, M., Frischer, M., Stimson, G., Taylor, C., & Tilling, K. (1999). Estimating the prevalence of problem drug use in inner London: A discussion of three capture-recapture studies. *Addiction*, *94*, 1653–1662.
- Hines, J. E. (2006). *PRESENCE 3.1 Software to estimate patch occupancy and related parameters*. <http://www.mbr-pwrc.usgs.gov/software/presence.html>
- Holst, L. (1981). Some asymptotic results for incomplete multinomial or Poisson samples. *Scandinavian Journal of Statistics*, *8*, 243–246.
- Hook, E. B., & Regal, R. R. (1982). Validity of Bernoulli census, log-linear, and truncated binomial models for correction for underestimates in prevalence studies. *American Journal of Epidemiology*, *116*, 168–176.
- Hook, E. B., & Regal, R. R. (1992). The value of capture-recapture methods even for apparently exhaustive surveys: The need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. *American Journal of Epidemiology*, *135*, 1060–1067.
- Hook, E. B., & Regal, R. R. (1995). Capture-recapture methods in epidemiology: Methods and limitations. *Epidemiologic Reviews*, *17*, 243–264.
- Hook, E. B., & Regal, R. R. (1999). Recommendations for presentation and evaluation of capture-recapture estimates in epidemiology. *Journal of Clinical Epidemiology*, *52*, 917–933.
- Hook, E. B., Albright, S. G., & Cross, P. K. (1980). Use of Bernoulli census and log-linear methods for estimating the prevalence of spina bifida in livebirths and the completeness of vital record reports in New York State. *American Journal of Epidemiology*, *112*, 750–758.
- Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, *47*, 663–685.
- Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika*, *76*, 133–140.
- Huggins, R. M. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics*, *47*, 725–732.
- International Working Group for Disease Monitoring and Forecasting (IWGDMF). (1995a). Capture-recapture and multiple-record systems estimation. I: History and theoretical development. *American Journal of Epidemiology*, *142*, 1047–1058.
- International Working Group for Disease Monitoring and Forecasting (IWGDMF). (1995b). Capture-recapture and multiple-record systems estimation. II: Applications in human diseases. *American Journal of Epidemiology*, *142*, 1059–1068.
- Jackson, C. H. N. (1933). On the true density of tsetse flies. *Journal of Animal Ecology*, *2*, 204–209.
- Jackson, C. H. N. (1939). The analysis of an animal population. *Journal of Animal Ecology*, *8*, 238–246.
- Jennelle, C. S., Cooch, E. G., Conroy, M. J., & Senar, J. C. (2007). State specific detection probabilities and disease prevalence. *Ecological Applications*, *17*, 154–167.
- Jewell, W. S. (1985). Bayesian estimation of undetected errors. In J. M. Bernardo, M. H. DeGroot, D.V. Lindley, A.F.M. Smith (Eds.), *Bayesian statistics* (Vol. 2, pp. 663–671). Elsevier.
- Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, *52*, 225–247.

- Kendall, W. L., Hines, J. E., & Nichols, J. D. (2003). Adjusting multi-state capture-recapture models for misclassification bias: Manatee breeding proportions. *Ecology*, *84*, 1058–1066.
- Kendall, W. L., Langtimm, C. A., Beck, C. A., & Runge, M. C. (2004). Capture-recapture analysis for estimating manatee reproductive rates. *Marine Mammal Science*, *20*, 424–437.
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A*, *115*, 700–721.
- King, R., Bird, S. M., Overstall, A., Hay, G., & Hutchinson, S. J. (2013). Injecting drug users in Scotland, 2006: Listing, number, demography, and opiate-related death-rates. *Addiction Research & Theory*, *21*, 235–246.
- King, R., Bird, S. M., Overstall, A., Hay, G., & Hutchinson, S. J. (2014). Estimating prevalence of injecting drug users and associated heroin-related death-rates in England using regional data and incorporating prior information. *Journal of the Royal Statistical Society: Series A*, *177*, 1–28.
- Kremers, W. K. (1988). Estimation of survival rates from a mark-recapture study with tag loss. *Biometrics*, *44*, 117–130.
- Lachish, S., Gopalaswamy, A. M., Knowles, S. C. L., & Sheldon, B. C. (2012). Site-occupancy modelling as a novel framework for assessing test sensitivity and estimating wildlife disease prevalence from imperfect diagnostic tests. *Methods in Ecology and Evolution*, *2012*, 339–348.
- Laplace, M. (1786). Sur les naissances, les mariages et les mortes. *Historie Academie Royale Science. Ann'ee, 1783*, 693–702.
- LaPorte, R. E. (1994). Assessing the human condition: Capture-recapture techniques. *BMJ*, *308*, 5–6.
- LaPorte, R. E., McCarty, D. J., Tull, E. S., & Tajima, N. (1992). Counting birds, bees, and NCDs. *Lancet*, *339*, 494–495.
- Larson, A., Stevens, A., & Wardlaw, G. (1994). Indirect estimates of 'hidden' populations: Capture-recapture methods to estimate the numbers of heroin users in the Australian Capital Territory. *Social Science & Medicine* *39*, 823–831.
- Laska, E. M., & Meisner, M. (1993). A plant-capture method for estimating the size of a population from a single sample. *Biometrics*, *49*, 209–220.
- Lawing, A. M., Blois, J. L., Maguire, K. C., Goring, S. J., Wang, Y., & McGuire, J. L. (2021). Occupancy models reveal regional differences in detectability and improve relative abundance estimations in fossil pollen assemblages. *Quaternary Science Reviews*, *253*, 106747.
- Lebreton, J.-D., Nichols, J. D., Barker, R., Pradel, R., & Spendelov, J. (2009). Modeling individual animal histories with multistate capture-recapture models. *Advances in Ecological Research*, *41*, 87–173.
- Lee, A. J., Seber, G. A. F., Holden, J. K., & Huakau, J. T. (2001). Capture-recapture, epidemiology, and list mismatches: Several lists. *Biometrics*, *57*, 707–713.
- Lincoln, F. C. (1930). Calculating waterfowl abundance on the basis of banding returns. U.S. Dept. Agric., Circ. No. 118, Washington, D.C.
- Liow, L. H. (2013). Simultaneous estimation of occupancy and detection probabilities: An illustration using Cincinnatian brachiopods. *Paleobiology*, *39*, 193–213.
- Liow, L. H., & Nichols, J. D. (2010). Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: Capture-mark-recapture (CMR) approaches. In J. Alroy & G. Hunt (Eds.), *Quantitative methods in paleobiology* (pp. 81–94). The Paleontological Society.
- Lum, K., Price, M. E., & Banks, D. (2013). Applications of multiple systems estimation in human rights research. *American Statistician*, *67*, 191–200.
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., & Langtimm, C. A. (2002). Estimating site occupancy when detection probabilities are less than one. *Ecology*, *83*, 2248–2255.
- MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., & Franklin, A. B. (2003). Estimating site occupancy, colonization and local extinction probabilities when a species is not detected with certainty. *Ecology*, *84*, 2200–2207.

- MacKenzie, D. I., Nichols, J. D., Seamans, M. E., & Gutierrez, R. J. (2009). Dynamic models for problems of species occurrence with multiple states. *Ecology*, *90*, 823–835.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K.H., Bailey, L. A., & Hines, J. E. (2018). *Occupancy modeling and estimation* (2nd ed.). Academic Press.
- Magnusson, W. E., Caughley, G. J., & Grigg, G. C. (1978). A double-survey estimate of population size from incomplete counts. *Journal of Wildlife Management*, *42*, 174–176.
- Manly, B. F. J., & Parr, M. J. (1968). A new method of estimating population size, survivorship, and birth rate from capture-recapture data. *Transactions of the Society for British Entomology*, *18*, 81–89.
- Manrique-Vallier, D., Price, M. E., & Gohdes, A. (2013). Multiple systems estimation techniques for estimating casualties in armed conflicts. In T.B. Seybolt, J. D. Aronson, B. Fischhoff (Eds.), *Counting civilian casualties: An introduction to recording and estimating nonmilitary deaths in conflict* (pp. 165–182). Oxford Univ. Press.
- Marescot, L., Lyet, A., Singh, R., Carter, N., & Gimenez, O. (2019). Inferring wildlife poaching in southeast Asia with multispecies dynamic occupancy models. *Ecography*, *43*, 239–250.
- Mastro, T. D., Kitayaporn, D., & Weniger, B. G. (1994). Estimating the number of HIV-infected injection drug users in Bangkok: A capture-recapture method. *American Journal of Public Health*, *84*, 1094–1099.
- Maxim, L. D., & Harrington, L. (1982). Scale-up estimators with size-dependent detection. *Photogrammetric Engineering and Remote Sensing*, *48*, 1271–1287.
- Maxim, L. D., & Harrington, L. (1983). Aerial survey design: A systems-analytic perspective. *Photogrammetric Engineering and Remote Sensing*, *49*, 1425–1435.
- Maxim, L. D., Harrington, L., & Kennedy, M. (1981). A capture-recapture approach for estimation of detection probabilities in aerial surveys. *Photogrammetric Engineering and Remote Sensing*, *47*, 779–788.
- McCarty, D. J., Tull, E. S., Moy, C. S., Kwoh, C. K., & LaPorte, R. E. (1993). Ascertained corrected rates: Applications of capture-recapture methods. *International Journal of Epidemiology*, *22*, 559–565.
- McClintock, B. T., Nichols, J. D., Bailey, L. L., MacKenzie, D. I., Kendall, W. L., & Franklin, A. B. (2010). Seeking a second opinion: Uncertainty in wildlife disease ecology. *Ecology Letters*, *13*, 659–674.
- McClintock, B. T., Bailey, L. L., Dreher, B. P., & Link, W. A. (2014). Probit models for capture-recapture data subject to imperfect detection, individual heterogeneity, and misidentification. *The Annals of Applied Statistics*, *8*, 2461–2484.
- Miller, D. A. W., Nichols, J. D., McClintock, B. T., Grant, E. H. C., Bailey, L. L., & Weir, L. (2011). Improving occupancy estimation when two types of observational error occur: Nondetection and species misidentification. *Ecology*, *92*, 1422–1428.
- Miller, D. A. W., Talley, B. L., Lips, K. R., & Grant, E. H. C. (2012). Estimating patterns and drivers of infection prevalence and intensity when detection is imperfect and sampling error occurs. *Methods in Ecology and Evolution*, *2012*, 850–859.
- Miller, D. A. W., Nichols, J. D., Gude, J. A., Rich, L. N., Podruzny, K. M., Hines, J. E., & Mitchell, M. S. (2013). Determining occurrence dynamics when false positives occur: Estimating the range dynamics of wolves from public survey data. *PLoS One* *8*. <https://doi.org/10.1371/journal.pone.0065808>
- Moore, J. F., Mulindahabi, F., Masozera, M. K., Nichols, J. D., Hines, J. E., Turikunkiko, E., & Oli, M. K. (2017). Are ranger patrols effective in reducing poaching-related threats in protected areas? *Journal of Applied Ecology*, *2017*, 1–9.
- Moore, J., Uzabaho, E., Musana, A., Uwingeli, P., Hines, J., & Nichols, J. (2021). What is the effect of poaching activity on wildlife species? *Ecological Applications*, *31*, e02397.
- Nichols, J. D. (2019). Confronting uncertainty: Contributions of the wildlife profession to the broader scientific community. *Journal of Wildlife Management*, *83*, 519–533.
- Nichols, J. D., & Hines, J. E. (1993). Survival rate estimation in the presence of tag loss using joint analysis of capture-recapture and resighting data. In J.-D. Lebreton, P.M. North (Eds.), *The study of bird population dynamics using marked individuals* (pp. 229–243). Birkhauser Verlag.

- Nichols, J. D., Hines, J. E., MacKenzie, D. I., Seamans, M. E., & Gutierrez, R. J. (2007). Occupancy estimation with multiple states and state uncertainty. *Ecology*, *88*, 1395–1400.
- Nichols, J. D., & Karanth, K. U. (2002). Statistical concepts: Assessing spatial distributions. In K.U. Karanth, J.D. Nichols (Eds.), *Monitoring tigers and their prey. A manual for wildlife managers, researchers, and conservationists* (pp. 29–38). Centre for Wildlife Studies.
- Nichols, J. D., & Pollock, K. H. (1983). Estimating taxonomic diversity, extinction rates and speciation rates from fossil data using capture-recapture models. *Paleobiology*, *9*, 150–163.
- Nichols, J. D., Morris, R. W., Brownie, C., & Pollock, K. H. (1986). Sources of variation in extinction rates, turnover and diversity of marine invertebrate families during the Paleozoic. *Paleobiology*, *12*, 421–432.
- Nichols, J. D., Kendall, W. L., Hines, J. E., & Spindelov, J. A. (2004). Estimation of sex-specific survival from capture-recapture data when sex is not always known. *Ecology*, *85*, 3192–3201.
- Nichols, J. M., Judd, K. P., Olsen, C. C., Waterman, J. R., & Nichols, J. D. (2013a). Estimating detection and identification probabilities for maritime target acquisition. *Applied Optics*, *52*, 2531–2545.
- Nichols, J. M., Hines, J. E., & Nichols, J. D. (2013b). Selecting among competing models of electro-optic, infrared camera range performance. *Optical Engineers*, *52*, 113108.
- Nichols, J. D., Hollman, T., & Grand, J. B. (2017). Monitoring for the management of disease risk in animal translocation programmes. *EcoHealth*, *14*, S156–S166.
- Nichols, J. D., Bogich, T. L., Howerton, E., Bjrnstad, O. N., Borcherling, R. K., Ferrari, M., Haran, M., Jewell, C., Pepin, K. M., Probert, W. J. M., Pulliam, J. R. C., Runge, M. C., Tildesley, M., Viboud, C., & Shea, K. (2021). Strategic testing approaches for targeted disease monitoring can be used to inform pandemic control and vaccine rollout programs. *PLoS Biology*, *19*, e3001307. <https://doi.org/10.1371/journal.pbio.3001307>
- Norris III, J. L., & Pollock, K. H. (1995). A capture-recapture model with heterogeneity and behavioral response. *Environmental and Ecological Statistics*, *2*, 305–313.
- Otis, D. L., Burnham, K. P., White, G. C., & Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs* *62*, 3–135.
- Pollock, K. H. (1981). Capture-recapture models for age-dependent survival and capture rates. *Biometrics*, *37*, 521–529.
- Pollock, K. H., Winterstein, S. R., Bunck, C. M., & Curtis, P. D. (1989). Survival analysis in telemetry studies: The staggered entry design. *Journal of Wildlife Management*, *53*, 7–15.
- Polonsky, J. A., Bhning, D., Keita, M., Ahuka-Mundeke, S., Nsio-Mbeta, J., Abedi, A. A., Mossoko, M., Estill, J., Keiser, O., Kaiser, L., Yoti, Z., Sangnawakij, P., Lerdsuwansri, R., & Vilas, V. J. D. R. (2018). Novel application of capture-recapture methods to estimate the completeness of contact tracing during a large outbreak of Ebola Virus Disease, Democratic Republic of Congo, 2018–2020. *International Journal of Infectious Diseases*, *116*, S98.
- Pradel, R. (1996). Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics*, *52*, 703–709.
- Pradel, R. (2005). Multievent: An extension of multistate capture-recapture models to uncertain states. *Biometrics*, *61*, 442–447.
- Razzak, J. A., & Luby, S. P. (1998). Estimating deaths and injuries due to road traffic accidents in Karachi, Pakistan, through the capture-recapture method. *International Journal of Epidemiology*, *27*, 866–870.
- Rivest, L.-P., & Lvesque, T. (2001). Improved log-linear model estimators of abundance in capture-recapture experiments. *Canadian Journal of Statistics*, *29*, 555–572.
- Rosenzweig, M. L., & Duek, J. L. (1979). Species diversity and turnover in an Ordovician marine invertebrate assemblage. In G.P. Patil & M.L. Rosenzweig (Eds.), *Contemporary quantitative ecology and related ecometrics. Statistical ecology series* (Vol. 12, pp. 109–119). International Cooperative Publishing House.
- Royle, J. A. (2004). Modeling abundance index data from anuran calling surveys. *Conservation Biology*, *18*, 1378–1385.
- Royle, J. A., & Link, W. A. (2005). A general class of multinomial mixture models for anuran calling survey data. *Ecology*, *86*, 2505–2512.

- Royle, J. A., & Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, *87*, 835–841.
- Sango, H. A., Testa, J., Meda, N., Conrand, B., Traoré, M. S., Stacini, P., & Lagarde, E. (2016). Mortality and morbidity of urban road traffic crashes in Africa: Capture-recapture estimates in Bamako, Mali, 2012. *PLoS ONE*, *11*(2), e0149070. <https://doi.org/10.1371/journal.pone.0149070>
- Schnabel, Z. E. (1938). The estimation of the total fish population of a lake. *The American Mathematical Monthly*, *45*, 348–352.
- Schwarz, C. J., & Arnason, A. N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics*, *52*, 860–873.
- Schwarz, C. J., Schweigert, J. F., & Arnason, A. N. (1993). Estimating migration rates using tag recovery data. *Biometrics*, *49*, 177–193.
- Seber, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika*, *52*, 249–259.
- Seber, G. A. F. (1982). *The estimation of animal abundance and related parameters*. MacMillan.
- Seber, G. A. F., & Schofield, M. R. (2019). *Capture-recapture: Parameter estimation for open animal populations*. Springer.
- Seber, G. A. F., Huakau, J. T., & Simmon, D. (2000). Capture-recapture, epidemiology, and list mismatches: Two lists. *Biometrics*, *56*, 1227–1232.
- Sekar, C. C., & Deming, W. E. (1949). On a method of estimating birth and death rates and the extent of registration. *Journal of the American Statistical Association*, *44*, 101–115.
- Sharma, K., Wright, B., Joseph, T., & Desai, N. (2014). Tiger poaching and trafficking in India: Estimating rates of occurrence and detection over four decades. *Conservation Biology*, *179*, 33–39.
- Silverman, B. (2014). Modern slavery: An application of multiple systems estimation Home Office. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/386841/Modern_Slavery_an_application_of_MSE_revised.pdf
- Stokes, S. L. (1984). The Jolly-Seber method applied to age-stratified populations. *Journal of Wildlife Management*, *48*, 1053–1059.
- Sudman, S., Sirken, M. G., & Cowan, C. D. (1988). Sampling rare and elusive populations. *Science* *240*, 991–996 (1988)
- Tercero, F., & Andersson, R. (2004). Measuring transport injuries in a developing country: An application of the capture-recapture method. *Accident Analysis and Prevention*, *36*, 13–20.
- U.S. Census Bureau. Post Enumeration Surveys. U.S. Census Bureau, Washington, DC (2021). <https://www.census.gov/programs-surveys/decennial-census/about/coverage-measurement/pes.html>
- Vaissade, L., & Legleye, S. (2008). Capture-recapture estimates of the local prevalence of problem drug use in six French cities. *European Journal of Public Health*, *19*, 32–37.
- van der Heijden, P. G. M., Cruyff, M., & Bohning, D. (2014). Capture recapture to estimate criminal populations. In G. J. N. Bruinsmaand, D. L. Weisburd (Eds.), *Encyclopedia of criminology and criminal justice* (pp. 267–278). Springer.
- Veran, S., Kleiner, K. J., Choquet, R., Collazo, J., & Nichols, J. D. (2012). Modeling habitat dynamics accounting for possible misclassification. *Landscape Ecology*, *27*, 943–956.
- Viallefont, A., & Auget, J. L. (1999). [Using capture-recapture models to estimate transition rates between states in interval-censored data] (published in French). *Revue D'épidemiologie et de Sante Publique*, *47*, 627–634.
- Wang, X., Lim, J., & Stokes, S. L. (2006). Forming post-strata via Bayesian treed capture-recapture models. *Biometrika*, *93*, 861–876.
- White, G. C., & Burnham, K. P. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study*, *46*(sup1), S120–S139. <https://doi.org/10.1080/00063659909477239>
- White, G. C., Anderson, D. R., Burnham, K. P., & Otis, D. L. (1982). *Capture-recapture and removal methods for sampling closed populations*. Los Alamos National Laboratory LA-8787-NERP, Los Alamos, New Mexico.

- Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and management of animal populations*. Academic Press.
- Wittes, J. T. (1974). Applications of a multinomial capture-recapture model to epidemiological data. *Journal of the American Statistical Association*, *69*, 93–97.
- Wittes, J. T., & Sidel, V. W. (1968). A generalization of the simple capture-recapture model with applications to epidemiological research. *Journal of Chronic Diseases*, *21*, 287–301.
- Wittes, J. T., Coulton, T., & Sidel, V. W. (1974). Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *Journal of Chronic Diseases*, *27*, 25–36.
- Wolter, K. M. (1986). Some coverage error models for census data. *Journal of the American Statistical Association*, *81*, 338–346.
- Wolter, K. M. (1990). Capture-recapture estimation in the presence of a known sex ratio. *Biometrics*, *46*, 157–162.
- Yeo, L. M., McCrear, R. S., & Roberts, D. L. (2017). A novel application of mark-recapture to examine behaviour associated with the online trade in elephant ivory. *PeerJ*, *5*, e3048. <https://doi.org/7717/peerj.3048>