

## Quantifying abundance: counts, detection probabilities, and estimates

---

*Benedikt R. Schmidt and Jérôme Pellet*

### **25.1 Background: imperfect detection in amphibian ecology and conservation**

Understanding temporal and spatial variation in distribution and abundance has been, and will remain, a central goal in amphibian ecology and conservation. Even though these two quantities, distribution and abundance, are so fundamental, we usually cannot observe them directly in the field. We rarely observe all individuals, all populations, or all species in an area of interest. Imperfect detection is the rule rather than the exception and is a characteristic that all field studies share.

Imperfect detection is a trivial fact and herpetologists are often aware of it. For example, Hairston and Wiley (1993) attributed all fluctuations in salamander counts to variation in weather conditions (salamanders tend to remain underground during cold weather) and motivation of students (a class of highly motivated students found an exceptionally high number of salamanders). This implies that salamander detection was imperfect and, equally important, variable among years (Hyde and Simons 2001).

Nevertheless, imperfect detection is all too often ignored when estimating abundance or when temporal and spatial trends are analysed. In this chapter we pinpoint patterns and the consequences of imperfect detection and show how to deal with it in general (Chapter 24 introduces many of the estimators required to deal with imperfect detection). We focus on the effects of imperfect detection on population abundance estimation but we call attention to the fact that imperfect detection also affects other ecological quantities, such as distribution and species richness (e.g. Schmidt 2004; Royle and Dorazio 2006; Mazerolle *et al.* 2007). We focus on how the issue of imperfect detection should be incorporated into the

study of amphibians populations; that is, how to sample amphibian populations and how to estimate abundance. By doing so, we assume that clear study objectives have been formulated in advance (Yoccoz *et al.* 2001) and sites appropriately selected (see Chapter 23).

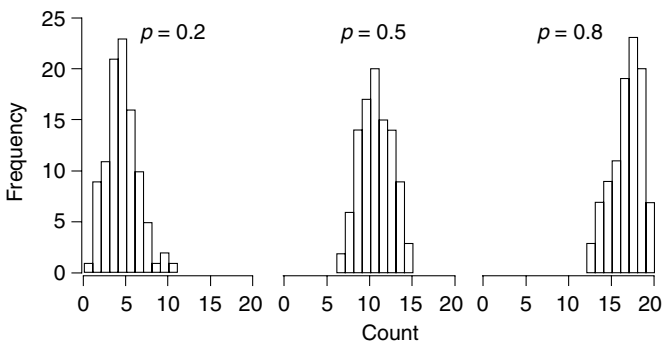
## 25.2 Imperfect detection

It is convenient to use a simple equation to conceptualize imperfect detection:

$$E(C) = Np$$

where  $N$  is the true value of the parameter of interest (i.e. number of individuals, density, number of populations in an area, or species richness) and  $p$  is the detection probability (Gill 1985; Yoccoz *et al.* 2001; Pollock *et al.* 2002; Schmidt 2004).  $E()$  denotes a statistical expectation. The expectation  $E(C)$  is the average of the count  $C$  over repeated realizations of the sampling process. This equation has three major implications that we discuss below.

Sampling a population should be viewed as a stochastic process because it involves a probability of detection. This is why there is a statistical expectation  $E(C)$  in the above equation. Even under identical conditions we should not expect to obtain the same result if we sample the same population multiple times. We should therefore expect variability in the counts. Variability in counts ( $C$ ) does not imply variation in abundance ( $N$ ) or detection probability ( $p$ ); it can simply be random variation. Technically, counts are random variables. This is illustrated in Figure 25.1. The figure shows that, as expected under



**Fig. 25.1** Expected variation in counts of individuals under identical conditions for 100 repeated counts, a true population size  $N = 20$  and detection probabilities  $p = 0.2, 0.5,$  and  $0.8$ . Data were simulated using the R code `rbinom(n = 100, size = 20, p = x)`, where  $x$  had the values 0.2, 0.5, and 0.8, respectively).

binomial sampling, variability of the counts is greatest if  $p = 0.5$ . The range of likely counts is large no matter what the detection probability is.

### 25.2.1 Counts underestimate abundance

Because  $p$  can take any value between 0 and 1, counts almost always underestimate true abundance (or other biological parameters such as species richness or occupancy probabilities). Detection probabilities vary depending on the methods used to sample a population. In studies using drift fences, detection probabilities can be very close to 1 (Bailey *et al.* 2004a), meaning that very few individuals escape detection. When using other methods, such as hand capture, trapping, netting, or cover-board surveys, detection probabilities are usually far below 1 (see, for example, case studies and estimates of detection probabilities in Mazerolle *et al.* 2007). Consequently, the discrepancy between counts and true abundances increases with decreasing detection probability. Table 25.1 shows the correlations between counts and abundance where the true number of individuals was known. Counts and true abundances are somewhat correlated but the proportion of variance explained can be quite low.

Pellet *et al.* (2007) analysed the relationship between the chorus counts, number of captures and mark-recapture estimates of abundance in detail (Table 25.2). Their study of the European treefrog *Hyla arborea* showed that calling males represented a variable fraction of the total male population present at breeding ponds. In two sites studied over 3 years, the proportion of males calling varied between 0.32 and 0.65, suggesting that there is no solid link between chorus activity and (estimated) population size. Similarly, their capture effort (total number of males captured) did not truly reflect actual population size,

**Table 25.1** Results of linear regressions between counts ( $C$ ), population size estimates ( $\hat{N}$ ), and censuses for tadpoles of two anuran species.  $R^2$  and  $F$  tests were calculated using PROC GLM in SAS. Asterisks indicate significance at  $\alpha = 0.05$ . The Lincoln–Peterson estimator was used to estimate abundance. Table adapted from Schmidt (2004). Data are from Jung *et al.* (2002).

Species	Studied in	Intercept	Slope	$R^2$	$F$
<i>Between count and census</i>					
<i>Hyla arenicolor</i>	Natural ponds	-13.92	1.73	0.66	34.21*
<i>Scaphiopus couchii</i>	Mesocosms	12.44	1.17	0.96	283.1*
<i>Between estimate and census</i>					
<i>Hyla arenicolor</i>	Natural ponds	18.35	0.87	0.97	1005.0*
<i>Scaphiopus couchii</i>	Mesocosms	15.25	0.94	0.99	1509.0*

**Table 25.2** Counts and estimates of male European tree frogs (*Hyla arborea*) in two breeding aggregations. Maximum and mean chorus size, as well as number of males captured, correlate only weakly with actual (estimated) male population size, demonstrating the effects of imperfect detection on population size estimates. Adapted from Pellet et al. (2007).

	Year	Maximum chorus size	Mean chorus size	Total males captured	Modeled male population size $\pm$ SE	Proportion of calling males
Males in population	2002	27	11.4	35	57.9 $\pm$ 9	47%
	2003	18	6.8	34	49.6 $\pm$ 6.9	36%
Camp Romain	2004	20	12.5	75	62.1 $\pm$ 15	32%
Males in population	2002	25	6.9	29	38.5 $\pm$ 5.7	65%
Les Mossières	2003	15	7.2	30	30.8 $\pm$ 1.3	49%
	2004	20	7.3	45	46.8 $\pm$ 2.2	43%

thus demonstrating that population trends estimated from counts (callers or captures) are likely to reflect a mixture of population and detectability trends. In addition, Pellet *et al.* (2007) compared mean chorus counts and maximum chorus counts. Mean counts were only 25–63% of the maximum counts. Such variability probably arises from the fact that a maximum count is an extreme value that is likely to be highly variable spatially and temporally. We argue that the maximum count is not a suitable metric if the goal of a study is a comparison between years and/or sites. If anything, mean counts would be better comparable across sites and years than maxima (but see Schmidt and Pellet 2005 for a study where maximum counts predicted population persistence).

Funk *et al.* (2003) compared the efficiency of abundance estimates based on visual encounter surveys, distance sampling, and mark–recapture methods for monitoring population trends of forest-floor-dwelling *Eleutherodactylus* frogs. They found that mark–recapture methods were best at estimating abundance and the method had the greatest power to detect population declines. Like in the examples above, the message is clear: estimates clearly outperform counts in reflecting true population abundances.

### 25.2.2 Per-visit and cumulative detection probabilities

Detection probabilities come in two flavors: per visit and cumulative. Whereas per-visit detection probabilities may be low, cumulative detection probabilities are usually much higher. Cumulative detection probability,  $p_c$ , is given by

$$p_c = 1 - (1 - \bar{p})^n$$

where  $\bar{p}$  is the average per-visit detection probability and  $n$  is the number of visits or capture events. Per-visit (i.e. per day) detection probabilities of the frog *Colostethus stepheni* in the study of Funk and Mills (2003) ranged from 0 to 1 with a mean of 0.58 (W.C. Funk, personal communication). Because Funk and Mills (2003) had six capture events during a short time period when the population did not change, the cumulative capture probabilities were greater than 0.987. This implies that Funk and Mills (2003) achieved an almost complete census of the frogs and suggests that multiple capture events are an obvious way to deal with imperfect detection because cumulative detection probabilities can be quite high.

### 25.2.3 Temporal and spatial variation in detection probabilities

Counts of animals are usually made to assess temporal or spatial variation in population size. Such comparisons are only valid if  $E(p)$  remains constant in time or space; however, this is hardly ever the case (MacKenzie and Kendall 2002), even under strictly standardized methods. If  $p$  varies temporally and/or spatially then variation in  $p$  and  $N$  are confounded. In the form of an equation, this may be expressed as

$$E(T) = C_1/C_2 = N_1 p_1 / N_2 p_2$$

Spatial or temporal variation in  $p_i$  can lead to the detection of spurious temporal or spatial trends  $E(T)$ . The contrary may also be true: one may miss true patterns in abundance or (under- or) overestimate the effects of variables that are thought to explain variation in abundance (Mazerolle *et al.* 2005). Obviously, the absolute values of  $p$  matter. If  $p$  is high, then bias will be low. If  $p$  is low, however, then most variation in  $C$  is likely due to variation in  $p$ . For example, if  $p$  varies between 80 and 90%, then variation in  $C$  will be relatively small. However, if a low  $p$  varies by the same absolute amount, say between 10 and 20%, then variation in  $C$  will be relatively greater.

Temporal and spatial variation is the rule rather than the exception. Tacitly, Hairston and Wiley (1993) attributed variation in salamander counts entirely to variation in detection probability. In the study of Funk and Mills (2003),  $p$  varied from 0 to 1 depending on prevailing weather conditions (W.C. Funk, personal communication). Daily means (across sites) ranged from 0.43 to 0.75 and different sites had different means (across days) ranging from 0.43 to 0.90 (W.C. Funk, personal communication). Most variation in detection probabilities of the European treefrog *H. arborea* in the study of Pellet *et al.* (2007) was between sites whereas Schmidt *et al.* (2007) found that detection probabilities of *Salamandra salamandra* were low in autumn and high in spring. Bailey *et al.* (2004a) used a drift fence to

capture salamanders (*Ambystoma tigrinum*) on their way to and from the pond. Detection probabilities were high in most years (greater than 90%) but inexplicably lower in one year (76%). Although detection probabilities of egg masses of *Rana sylvatica* and *Ambystoma maculatum* were generally high (usually > 80%), Grant *et al.* (2005) documented substantial spatial and temporal variation in detection probabilities that could bias estimates of population trends based on unadjusted counts.

### 25.3 Components of imperfect detection

There are many reasons why detection is imperfect and it makes sense to decompose detection probabilities into its components (Pollock *et al.* 2004; Nichols *et al.* 2008):

$$E(C) = Npe$$

where  $e$  is exposure to sampling (also referred to as availability for sampling). In this equation,  $p$  should best be called detection probability *given exposure*. Nichols *et al.* (2008) describe how detection probability and exposure to sampling can be further decomposed. Detection probability ( $p$ ) and exposure to sampling ( $e$ ) are easy to distinguish at the conceptual level. In practice, the distinction may not always be obvious.

Under different names, exposure to sampling is a well-known phenomenon of mark–recapture studies dealing with survival estimation. The two cases are transients and temporary emigrants (Williams *et al.* 2002). Transients are animals that show up only once in the study area and then leave. Temporary emigrants are animals that leave the study area temporarily during the sampling period and then return (e.g. animals that skip a breeding season). In other words, these animals are only partly exposed to sampling and this can cause negative bias in survival estimates.

Incomplete exposure to sampling also affects abundance estimation (Kendall 1999; Bailey *et al.* 2004b; Royle and Dorazio 2006). Non-exposure to sampling can result from both biological and methodological reasons. Biological reasons include brooding salamanders that are underground and not at the surface during the time of survey or frogs that do not breed in a particular year and hence do not migrate to the study site which may be a breeding site (e.g. pond). Another biological reason may be that only a proportion of all males present at a pond may be calling at a particular time. This would explain why there is a difference between mean and maximum chorus counts (Pellet *et al.* 2007). Methodological reasons may play a part when funnel or minnow traps

are placed along the edge of a pond and some individuals spend all their time near the center of the pond. Non-exposure to sampling may also result from breeding phenology. Some individuals may breed early and some may breed later in the season (e.g. Sinsch 1988). If the population is sampled only early in the season, then many individuals will not be exposed to sampling. This might be a case where both biology and method cause non-exposure to sampling.

Non-exposure to sampling can have profound consequences for abundance estimation (Kendall 1999). Imagine that the goal of a study is to estimate abundance in a particular area. If amphibians do not move in and out of the study area, then an abundance estimator such as the Lincoln–Peterson estimator will provide an unbiased estimate the number of amphibians in the area. However, if some individuals move randomly in and out of the area (i.e. if they are not always exposed to sampling), then the very same estimator will estimate a different quantity. It will now estimate the size of the superpopulation where the superpopulation is defined as the total number of amphibians exposed to sampling at least once. This includes all amphibians that are residents in the study area, but also all individuals that move in and out or that move through the study area. Bailey *et al.* (2004b) encountered a related problem in their study of salamanders in the Appalachian Mountains, USA. They argued that short-term studies where salamander movement was negligible yielded estimates of the “surface” population (i.e. salamanders exposed to sampling only). Long-term studies, where salamanders had time to move from the surface to deeper ground and vice versa, gave estimates of the total number of salamanders in the sampled area. Kinkead and Otis (2007) describe a similar situation with breeding and non-breeding ambystomatid salamanders that were sampled at the breeding site.

If many individuals are not exposed to sampling (low  $e$ ), then the mismatch between the spatial or temporal scale at which a population is sampled and the desired temporal or spatial scale of inference is likely to be great. In conclusion, study design (both the spatial and the temporal scales) and species behavior can jointly determine which biological entity is quantified.

## 25.4 How to deal with imperfect detection

### 25.4.1 Estimation of abundance

Evidently, an elegant way to deal with imperfect detection is to estimate detection probability ( $\hat{p}$ ). An estimate of detection probability can then be used to correct counts ( $C$ ) and estimate abundance ( $\hat{N}$ ):

$$\hat{N} = C / \hat{p}$$

This equation is the conceptual basis for all kinds of abundance estimators, be they mark–recapture, distance sampling, point count, removal, or other methods (Williams *et al.* 2002; Mazerolle *et al.* 2007; see also Chapter 24 in this volume). When non-exposure to sampling is a problem, then  $\hat{N}$  can only be estimated if one knows the fraction of the population that is exposed to sampling as well as the probability of detecting exposed individuals, populations, or species:

$$\hat{N} = C / \hat{p}\hat{e}$$

The best tool to estimate abundance in the presence of non-exposure to sampling is the robust design (see Williams *et al.* 2002) which has been successfully used with plethodontid and ambystomatid salamanders (Bailey *et al.* 2004b; Kinkead and Otis 2007). The robust design allows estimating both abundance and non-exposure (known as temporary emigration in this particular case). In the previously cited case of the European treefrog (*H. arborea*), the robust design also allowed demonstrating that annual male temporary emigration was negligible (i.e. that males rarely skipped breeding seasons; Pellet *et al.* 2007). In contrast, Bailey *et al.* (2004a) showed that temporary emigration in *A. tigrinum* was substantial.

Other approaches to dealing with non-exposure were described by Royle and Dorazio (2006) and Condit *et al.* (2007). Royle and Dorazio (2006) describe a method for point counts that allows dealing with non-exposure to sampling that arises from a mismatch between the scale at which data was collected and the desired scale of inference. They describe a case where quadrats have only been partially searched (in that study, only a fraction of the area of 1 km<sup>2</sup> quadrats was surveyed). Their method works if there is a suitable covariate (such as transect length) that can statistically link the exposed population to the true population.

Condit *et al.* (2007) developed a method for estimating the size of a population when individuals are asynchronously present. Their method may be particularly relevant for pond-breeding amphibians where the breeding season of the population is longer than the breeding season of an individual (e.g. in the natterjack toad *Bufo calamita* and the European treefrog *H. arborea*; Sinsch 1988; Friedl and Klump 2005).

Methods to estimate abundance are not without problems. Detection probabilities can be low. This will have the effect that confidence intervals can be very wide (Williams *et al.* 2002). It is possible to make confidence intervals shrink with more effort or better capture techniques. That is, researchers should either use better methods that increase per-visit detection probabilities or



increase the number of repeat visits such that cumulative detection probabilities are increased. However, standard errors and confidence intervals are not a nuisance. Rather, they are an advantage of estimation methods. Standard errors and confidence intervals are a measure of uncertainty that allow an assessment of the estimates' reliability. Consequently, we view wide confidence intervals as an honest statement whether a particular estimate is or is not particularly reliable and useful. Wide confidence intervals are no reason to discard mark–recapture estimates and to prefer the simple counts ( $C$ ; e.g. Alford and Richards 1999). There is also uncertainty associated with counts (because detection probability  $p$  is unknown) but it is not explicit and is, in fact, unknowable.

Heterogeneity in detection probabilities among individuals can be a problem in mark–recapture studies (Link 2003). Heterogeneity usually leads to negative bias in abundance estimates and in the worst case it may not be possible to identify a best model that should be used for inference. We believe that amphibian ecologists and conservationists should attempt to minimize detection probability heterogeneity among individuals by adopting methods that account for variation in detectability among individuals (i.e. grouping individuals into homogeneous sets by sexes, colour morphs, age classes).

#### 25.4.2 Other approaches to dealing with imperfect detection

Amphibian ecologists have dealt with imperfect detection in many ways. Some authors simply did not analyse data from species where detection was uncertain and variable. For example, Pechmann *et al.* (1991) analysed data from ambystomatid salamanders that are unlikely to trespass a drift fence but they did not analyse data from treefrogs that could easily trespass a drift fence.

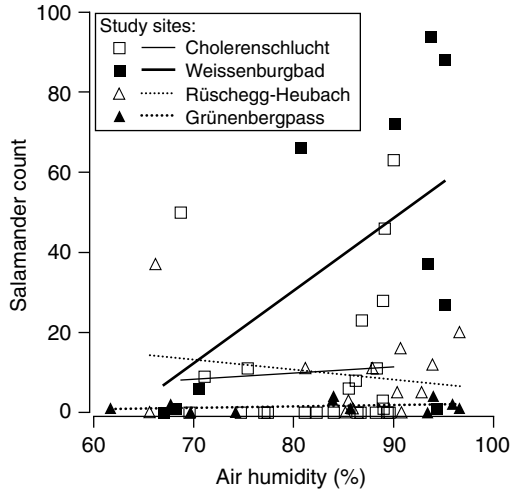
The most common objection to the use of methods that allow estimating population sizes and detection probabilities instead of counts is that they are demanding in terms of money, human resources, and statistical knowledge. This argument was not true in the detailed study of Funk *et al.* (2003) where different methods were compared. We discard this argument because every conservation action based on counts is likely to be biased to the point where they will be inefficient or, worse, counterproductive (Yoccoz *et al.* 2001). Moreover, it will be impossible to evaluate the success of actions taken because no data will be available to detect population trends accurately, and to adapt management actions accordingly.

Some authors have expressed the view that there is no need to estimate detection probabilities and to adjust counts accordingly (i.e. estimate abundance). One argument put forward in the context of long-term monitoring programs is that variability in detection probability does not matter as long as there is no

temporal trend in detection probability (Bart *et al.* 2004; tacitly, this is also the reasoning of Hairston and Wiley 1993). In such a situation, variation in detection probability likely causes extra variability (sampling variability) in the counts (in comparison to variation in absolute abundance which is the phenomenon of biological interest). Detectability-induced extra variation in the counts means that a monitoring program loses power to detect temporal trends. However, one should keep in mind that detection probabilities likely show temporal trends. Reasons include, but are not limited to, habitat succession or changes through time in the ability of the observer to detect the study species (Link and Sauer 1997).

One commonly held view is that field methods can be standardized to the extent where detection probability is constant. If this is the case, then counts or other estimates of relative abundance should serve well as proxies for absolute abundance. Unfortunately, variation in detection probabilities is the rule rather than the exception (MacKenzie and Kendall 2002). Whenever detection probabilities of amphibians have been estimated, they were found to be variable both within and across seasons (Bailey *et al.* 2004a, Kinkead and Otis 2007; Mazerolle *et al.* 2007; Pellet *et al.* 2007; Schmidt *et al.* 2007). This was the case even when researchers used standard(ized) methods; even when drift fences were used—where the assumption is that detection probability is 1—there was variation in detection probability (Bailey *et al.* 2004a). Pellet *et al.* (2007) used the same methods at two sites yet detection probabilities differed between sites by a factor of approximately two. Hyde and Simons (2001) showed that counts obtained from applying four standard methods gave results that were only weakly correlated. That is, the use of standardized methods does not guarantee that detection probabilities are constant. We believe that standardization of field methods is important because it can help to keep variation in detection probabilities within bounds, but it should certainly not be viewed as a panacea.

Standardization of field methods is one solution to limit variation in detection probabilities. Another solution is to measure covariates that may affect detection probabilities and use these covariates at the analysis stage to adjust counts (Link and Sauer 1997, 1998). This approach may work well as long as the important covariates are known and has been successfully used in large-scale bird monitoring programs. However, it may be that the effect of a covariate on the counts varies from one site to the next. Lauber (2004) counted alpine salamanders (*Salamandra atra*) along fixed transects at four sites and tested whether weather covariates could be used to predict the salamander counts. An analysis of covariance found no main effect of air humidity on counts but there was



**Fig. 25.2** Geographic variation in the relationship between salamander (*Salamandra atra*) counts and air humidity. Salamanders were counted multiple times at four sites along transects. Analysis of covariance showed no significant main effect of air humidity on counts but there was a significant interaction between air humidity and site. Data were taken with permission from Lauber (2004).

a significant site-by-air-humidity interaction (Figure 25.2). Air humidity can thus be used to adjust counts at some sites but not at others; a likely explanation is that sites differ in overall humidity. Grant *et al.* (2005) also found that no single explanatory variable or set of variables best explained variation in detection probabilities across sites of egg masses of *Rana sylvatica* and *Ambystoma maculatum*. This implies that herpetologists must be very cautious when extrapolating results from one study site to another. At the planning stage, it also implies that one should try to replicate all experiments spatially and temporally to find out whether biological patterns are universally applicable.

In conclusion, we believe that the use of standard methods is always valuable. Because it does not always avoid variation in detection probabilities, it is better to rely on adjusting counts than on the strong assumption that detection probabilities are not showing any trends. We argue that one should assume *a priori* that detection probabilities are less than one and that they are variable in space and time. Thus, amphibian ecologists and conservationists should provide evidence for their studies that the counts that they report are indeed reliable indices of abundance. Because detection probabilities can vary from one study to the next even under apparently highly similar conditions (see examples in Mazerolle *et al.* 2007), the proof has to be provided every time anew.

## 25.5 Designing a sampling protocol

The adequate design of a sampling protocol is a fundamental aspect of research, whether it be on individuals, populations, or species. Once the research or monitoring question addressed has been explicitly formalized (Yoccoz *et al.* 2001) and sites were appropriately selected (see Chapter 23 in this volume), it is essential that aspects of imperfect detection are incorporated into the design. To do so, the procedures of data collection and data analysis must be identified in advance. Having a sound knowledge of your study species' ecology will be instrumental in determining which biological quantity (e.g. above-ground population size, breeders, super-population) your estimator will represent.

Importantly, one must be fully aware that detection probability ( $p$ ) is not only a species trait (see examples in Mazerolle *et al.* 2007); it also depends on methods, observers, year, site, and a myriad of other factors. Because  $p$  (and exposure to sampling,  $e$ ) are both variable in space and time, one cannot apply values obtained in one study to another. For these reasons we recommend that detection probability is explicitly integrated in all amphibian study protocols. Non-exposure to sampling is often hard to deal with at the analysis stage (except when using the robust design). We therefore recommend that researchers carefully plan a study such that all animals are exposed to sampling.

## 25.6 Software

There are many computer programs freely available to estimate population abundance while incorporating detection probability. The most versatile and widely used of them is program MARK (White and Burnham 1999). This software, available at [www.phidot.org/](http://www.phidot.org/), allows the analysis of a wide range of capture–recapture-based data sets. Every new version incorporates the latest development in capture–recapture and thus allows the user to choose from a wide range of models the one that will fit its data best. As the name implies, DISTANCE ([www.ruwpa.st and.ac.uk/distance/](http://www.ruwpa.st.and.ac.uk/distance/)) is the software tool that allows one to design and analyse distance sampling surveys. More recent developments have been integrated in statistical software such as R and WinBugs. These tools have the inconvenience of being less user-friendly than the previously listed programs, which benefit from a graphical user interface. There is software that can be used when planning a mark–recapture study (Devineau *et al.* 2006; Zucchini *et al.* 2007).

## 25.7 Outlook

Imperfect detection is the feature that the vast majority of all amphibian surveys have in common. Complete censuses where all amphibians that are present at the study site are captured are not impossible, but require a lot of work (drift fences: Bailey *et al.* 2004a; many capture events: Funk and Mills 2003; Pellet *et al.* 2007). We argue that detection probabilities can be low and highly variable among years and/or sites. While counts that are not adjusted for imperfect detection can certainly indicate negative population trends (Laurance *et al.* 1996), variability in imperfect detection can seriously bias inference from surveys. Amphibian ecologists and conservationists should therefore estimate detection probabilities as the best tool to calibrate a survey and use robust methods for abundance estimation (Williams *et al.* 2002; Chapter 24). Unfortunately, the use of such methods is not yet widespread (Alford and Richards 1999).

The number of methods available for estimation of abundance that account for imperfect detection has increased tremendously in the recent past. Existing methods are constantly being refined, while new methods are being developed (e.g. Royle 2004; Royle and Dorazio 2006). Still, all methods need to be used with care as sampling design, the behavior of the species, and the estimator used all determine which biological quantity is being estimated. Notwithstanding, the quality of inference from methods that adjust for imperfect detection will be stronger than inference from any other kind of method.

In the future, we ought to be able to estimate abundance with a precision and freedom from bias that was not achievable in the past. We should now be able to determine which factors influence abundance rather than study patterns of an inseparable combination of abundance, detectability, and exposure to sampling (such as counts). This will help us gain new insights into fundamental and applied aspects of amphibian ecology and conservation.

## 25.8 References

- Alford, R. A. and Richards, S. J. (1999). Global amphibian declines: a problem in applied ecology. *Annual Reviews of Ecology and Systematics*, **30**, 133–65.
- Bailey, L. L., Kendall, W. L., Church D. R., and Wilbur, H. M. (2004a). Estimating survival and breeding probability for pond-breeding amphibians: a modified robust design. *Ecology*, **85**, 2456–66.
- Bailey, L. L., Simons, T. R., and Pollock, K. H. (2004b) Comparing population size estimators for plethodontid salamanders. *Journal of Herpetology*, **38**, 370–80.
- Bart, J., Droege, S., Geissler, P., Peterjohn, B., and Ralph, C. J. (2004). Density estimation in wildlife surveys. *Wildlife Society Bulletin*, **32**, 1242–7.

- Condit, R., Le Boeuf, B. J., Morris, P. A., and Sylvan, M. (2007). Estimating population size in asynchronous aggregations: A Bayesian approach and test with elephant seal censuses. *Marine Mammal Science*, **23**, 834–55.
- Devineau, O., Choquet, R., and Lebreton, J. D. (2006). Planning capture-recapture studies: straightforward precision, bias, and power calculations. *Wildlife Society Bulletin*, **34**, 1028–35.
- Friedl, T. W. P. and Klump, G. M. (2005). Sexual selection in the lek-breeding European treefrog: body size, chorus attendance, random mating and good genes. *Animal Behaviour*, **70**, 1141–54.
- Funk, W. C. and Mills, L. S. (2003). Potential causes of population declines in forest fragments in an Amazonian frog. *Biological Conservation*, **111**, 205–14.
- Funk, W. C., Almeida-Reinoso, D., Nogales-Sornosa, F., and Bustamante, M. R. (2003). Monitoring population trends of *Eleutherodactylus* frogs. *Journal of Herpetology*, **37**, 245–56.
- Gill, D. E. (1985). Interpreting breeding patterns from census data: a solution to the Husting dilemma. *Ecology*, **66**, 344–54.
- Grant, E. H. C., Jung, R. E., Nichols, J. D., and Hines, J. E. (2005). Double-observer approach to estimating egg mass abundance of pool-breeding amphibians. *Wetlands Ecology and Management*, **13**, 305–20.
- Hairston, Sr, N. G. and Wiley, R. H. (1993). No decline in salamander (Amphibia: Caudata) populations: a twenty-year study in the southern Appalachians. *Brimleyana*, **18**, 59–64.
- Hyde, E. J. and Simons, T. R. (2001). Sampling plethodontid salamanders: sources of variability. *Journal of Wildlife Management*, **65**, 624–32.
- Jung, R. E., Dayton, G. H., Williamson, S. J., Sauer, J. R., and Droege, S. (2002). An evaluation of population index and estimation techniques for tadpoles in desert pools. *Journal of Herpetology*, **36**, 465–72.
- Kendall W. L. (1999). Robustness of closed capture-recapture methods to violations of the closure assumption. *Ecology*, **80**, 2517–25.
- Kinkead, K. E. and Otis, D. L. (2007). Estimating superpopulation size and annual probability of breeding for pond breeding salamanders. *Herpetologica*, **63**, 151–62.
- Lauber, A. (2004). *Methodenevaluation zum Monitoring der Alpensalamanderpopulation*. Diploma thesis, Eidgenössische Technische Hochschule Zürich (ETHZ), Zürich.
- Laurance, W. F., McDonald, K. R., and Speare, R. (1996). Epidemic disease and the catastrophic decline of Australian rain forest frogs. *Conservation Biology*, **10**, 406–13.
- Link, W. A. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics*, **59**, 1123–30.
- Link, W. A. and Sauer, J. R. (1997). Estimation of population trajectories from count data. *Biometrics*, **53**, 488–97.
- Link, W. A. and Sauer J. R. (1998). Estimating population change from count data: Application to the North American Breeding Bird Survey. *Ecological Applications*, **8**, 258–68.
- MacKenzie, D. I. and Kendall, W. L. (2002). How should detection probability be incorporated into estimates of relative abundance? *Ecology*, **83**, 2387–93.
- Mazerolle, M. J., Desrochers, A., and Rochefort, L. (2005). Landscape characteristics influence pond occupancy by frogs after accounting for detectability. *Ecological Applications*, **15**, 824–34.

- Mazerolle, M. J., Bailey, L. L., Kendall, W. L., Royle, J. A., Converse, S. J., and Nichols, J. D. (2007). Making great leaps forward: accounting for detectability in herpetological field studies. *Journal of Herpetology*, **41**, 672–89.
- Nichols, J. D., Thomas, L., and Conn, P. B. (2008). Inferences about landbird abundance from count data: recent advances and future directions. In D. L. Thomson, E. G. Cooch, G. Evan, and M. J. Conroy (eds), *Modeling Demographic Processes in Marked Populations*, pp. 203–38. Springer Science+Business Media, New York.
- Pechmann, J. H. K., Scott, D. E., Semlitsch, R. D., Caldwell, J. P., Vitt, L. J., and Gibbons, J. W. (1991). Declining amphibian populations: the problem of separating human impacts from natural fluctuations. *Science*, **253**, 892–5.
- Pellet, J., Helfer, V., and Yannic, G. (2007). Estimating population size in the European tree frog (*Hyla arborea*) using individual recognition and chorus counts. *Amphibia-Reptilia*, **28**, 287–94.
- Pollock, K. H., Nichols, J. D., Simons, T. R., Farnsworth, G. L., Bailey, L. L., and Sauer, J. R. (2002). Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics*, **13**, 105–19.
- Pollock, K. H., Marsh, H., Bailey, L. L., Farnsworth, G. L., Simons, T. R., and Alldredge, M. W. (2004). Separating components of detection probability in abundance estimation: an overview with diverse examples. In W. L. Thompson (ed.), *Sampling Rare or Elusive Species*, pp. 43–58. Island Press, Washington DC.
- Royle, J. A. (2004). *N*-mixture models for estimating population size from spatially replicated counts. *Biometrics*, **60**, 108–15.
- Royle, J. A. and Dorazio, R. M. (2006). Hierarchical models of animal abundance and occurrence. *Journal of Agricultural, Biological, and Environmental Statistics*, **11**, 249–63.
- Schmidt, B. R. (2004). Declining amphibian populations: the pitfalls of count data in the study of diversity, distributions, dynamics and demography. *Herpetological Journal*, **14**, 167–74.
- Schmidt, B. R. and Pellet, J. (2005). Relative importance of population processes and habitat characteristics in determining site occupancy of two anurans. *Journal of Wildlife Management*, **69**, 884–93.
- Schmidt, B. R., Schaub, M., and Steinfartz, S. (2007). Apparent survival of the salamander *Salamandra salamandra* is low because of high migratory activity. *Frontiers in Zoology*, **4**, e19.
- Sinsch, U. (1988). Temporal spacing of breeding activity in the natterjack toad, *Bufo calamita*. *Oecologia*, **76**, 399–407.
- White, G. C. and Burnham, K. P. (1999). Program MARK: survival estimation from populations of marked animals. *Bird Study*, **46**, 120–38.
- Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.
- Yoccoz, N. G., Nichols, J. D., and Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution*, **16**, 446–53.
- Zucchini, W., Borchers, D. L., Erdelmeier, M., Rexstad, E., and Bishop, J. (2007). *WiSP 1.2.4*. Institut für Statistik und Ökonometrie, Georg-August-Universität Göttingen, Göttingen.