

Selection of species and sampling areas: the importance to inference

Paul Stephen Corn

23.1 Introduction

Inductive inference, the process of drawing general conclusions from specific observations, is fundamental to the scientific method. Platt (1964) termed conclusions obtained through rigorous application of the scientific method as “strong inference” and noted the following basic steps: generating alternative hypotheses; devising experiments, the results of which will exclude one or more hypotheses; conducting the experiment to get a “clean result”; and repeating the process with revision based on the information obtained. Every student is exposed to these basics in introductory courses, and a considerable proportion of a modern graduate education in the sciences is devoted to acquiring the analytic (statistical) skills necessary to apply the scientific method. Not even considering the field of mathematical statistics or applied statistics in disciplines such as social sciences, library shelves groan under the weight of texts on applied statistics, ranging from introductory (Hayek and Buzas 1997) to advanced (Williams *et al.* 2002), for conducting research in ecology, and new works are published every year. Much effort is currently devoted to the mechanisms of analysis and the issues involved in choosing among statistical methods; specifically, traditional hypothesis testing versus information-theoretic or Bayesian approaches (Hobbs and Hilborn 2006). This chapter does not address these topics, but instead discusses some of the issues related to selection of study sites and species necessary to obtain a “clean result.” Regardless of the method of analysis, successful inference relies on correctly designed data collection, meaning that the observations represent the population of interest. As Anderson (2008, p. 7) puts it, valid

inference requires, “some type of probabilistic sampling of the well-defined population.”

Experimental design and sampling are well-developed topics in applied statistics. There are numerous books available (see Anderson 2001 for a sample of representative works), and many universities offer their graduate students specific courses in these areas. Most biologists have had the requisite education and are familiar with and employ statistically valid designs. However, good study designs are not universal. Anderson (2008, p. 143) stated:

... there is no excuse for collecting data that are fundamentally flawed. Still, I see data collected from convenience sampling where any valid inductive inference is precluded. In some cases the population of interest is not even defined.

Convenience sampling is a broad term that describes non-random study-site selection (Hayek 1994; Anderson 2001). Inadequate design and sampling is often justified on logistical grounds, or, to put it more crudely, that it is unrealistic to satisfy the demands of a statistician, who is usually office-bound in an ivory tower. As a field biologist and not a statistician, I sometimes find myself sympathetic to such rationalizations. However, the solution is not to politely thank the statistician and then go ahead and sample as originally planned (see the first phrase in the Anderson quote above). Instead, more planning is necessary. Either a means must be devised for obtaining a valid sample, or the project's objectives should be revised so that a population can be defined that will allow a valid sample, or, in extreme cases, the project might need to be abandoned.

In this chapter, I provide a brief overview of sampling design, but this is a large, complex topic and it is impossible to do more than scratch the surface here. For more thorough instruction and discussion, including statistics specific to different designs, consult one of the many texts that have been written for that purpose (e.g. Cochran 1977; Thompson 1992; Hayek and Buzas 1997; Williams *et al.* 2002). The US Geological Survey (www.pwrc.usgs.gov/monmanual/) and National Park Service (<http://science.nature.nps.gov/im/monitor/>) also have extensive and very useful web-based resources to aid in study design for inventory and monitoring studies. Following this introduction to sampling, I will discuss issues related to selection of study sites and effects of estimating abundance on inference. This discussion will use a few selected cases to illustrate problems that can arise from either poor study design or interpretation of results that cannot be supported by the design. These examples include issues arising from both selection of study sites and collection and analysis of data from individual species, and describe some of the methods that can be used to minimize bias.

23.2 Sampling

In all but a tiny number of exceptional cases, it is impossible to survey every possible habitat or catch every individual in a population. Therefore, inference depends on statistics generated from a subset of individuals or habitats drawn from the population. Ideally, a sample has three qualities (Williams *et al.* 2002): it comprises separate individual units, and these share the same underlying distribution and are statistically independent. These criteria are more difficult to satisfy in inventory and monitoring studies than in a carefully designed experiment, but a good study design should try to come as close as possible.

Before a sample can be taken, the sampling units must be defined and be available for selection. Sample units can be individual animals, artificially bounded areas (quadrats), or natural habitat features, such as ponds, streams, or drainage basins. For population studies, where the individual is the sample unit, the availability for selection is often assumed, but this assumption is tested when the data are analyzed. For inventory or monitoring studies, site delineation and selection has become easier with the increasing availability of tools such as satellite imagery and Geographic Information Systems (GIS) software. A good example is the study by Kroll *et al.* (2008), which took advantage of detailed GIS data layers to select stream reaches for sampling. However, detailed GIS coverage exists mainly where there is an economic use for the data (e.g. the commercial forests surrounding the streams studied by Kroll *et al.* 2008), and even when GIS data are available the amphibian habitats may be poorly described. Temporary wetlands are notoriously underrepresented in aquatic data layers, and studies of lentic-breeding amphibians often will have more success in defining sampling units if areas are used instead of specific habitats, for example, 1 km² blocks (Johannson *et al.* 2006) or drainage basins (Corn *et al.* 2005).

The most basic form of probabilistic sampling is a simple random sample. All the available sample units are put into a metaphorical hat (the mechanics of choosing a sample now usually involve a computer) and the desired number of units are selected in random order. The number of sample units should be sufficient to estimate parameters of interest with sufficient precision, but samples that are too large should be avoided. Determining the desired sample size requires knowing the variability of the parameter to be estimated, the magnitude of the effect to detect (e.g. a 10% difference or a 5% annual trend), and the strength of the inference. The formula for determining sample size depends on the sampling scheme. Hayak and Buzas (1997) and Williams *et al.* (2002) describe the details for determining sample size.

If, as is usually the case, habitats are not uniform, species are patchily distributed, or the number of samples is small relative to the area of interest, simple random sampling may be inadequate to characterize the variability of the system being studied. There are several slightly more complex designs that can be used to achieve a more representative sample. In the case where the number of samples is relatively small, strictly random selection can result in sample sites clumped together instead of dispersed throughout the study area. If the study area is not uniform, this is not desirable, and a common modification is to employ a systematic random sample (Figure 23.1), in which, after a random start, every n th unit is selected, where the number of available units divided by n equals the desired sample size. Williams *et al.* (2002) cautioned that systematic sampling risks a biased result if the units are arranged in such a way that environmental gradients are correlated with the order of the sampling units.



Fig. 23.1 Sample locations for monitoring amphibian populations in Glacier National Park, MT, USA. Twenty small drainage basins (shaded) composed a systematic random sample from a frame of 235 basins (outlined). Areas without sufficient surface water, where the elevation precluded amphibian occurrence, or where extremely rugged terrain prevented safe access were excluded from the frame. Inference about status of amphibians in the park is limited to the frame.

If a study area is not homogenous and differences within the area can be described (e.g. altitudinal gradients, different types of wetlands, differences in ability to access potential sample units) then a stratified design is usually preferable to a simple random or systematic random sample. Once the strata and sampling units within them are delineated, then simple or systematic random samples can be drawn from each stratum. The formulas for calculating various statistics vary among sampling designs. See Williams *et al.* (2002) for a concise description of these. For long-term monitoring studies, strata should be based on features that are not expected to change significantly during the course of the study. For example, strata based on geological differences would be preferable to strata based on vegetative land cover.

This section is not a comprehensive treatment of sampling design. Advanced schemes, such as adaptive sampling (Thompson 1992; Williams *et al.* 2002), are beyond the scope of this chapter, but may be necessary or desirable in many cases. Sources listed above should be consulted when beginning the design of any study.

Any valid sampling design must incorporate replication of sample units. A study that compared toad populations in only two ponds could only describe the differences in abundance between the two ponds. Tests of hypotheses of the causes of the differences are not appropriate because there is no replication (Underwood 1998). Most inventory or monitoring studies include numerous sample units, but care must be taken to ensure that external factors are dealt with in the design stage (e.g. through stratification). If variation among sample units can be attributed to something other than random processes, then the study would suffer from pseudoreplication (Hurlbert 1984). Large field studies are vulnerable to temporal pseudoreplication (sample units vary in some systematic fashion over time), because it is seldom possible to visit all sample units simultaneously. This is a particular problem for studies of amphibians. For example, surveys that use breeding activity are affected by the changing composition of breeding choruses over time, both among and within species, and external factors, such as weather, that influence behavior of individuals. Studies that focus on larval stages must deal with growth and development and changes in abundance that may influence detection. It is always a good practice to minimize the time span of field surveys where possible.

23.3 Study sites and consequences of convenience sampling

If study sites or sample units are not selected using a probalistic design, the result is a convenience sample. The motivations for convenience sampling are several,

but common reasons are to choose study sites that are most accessible, or where the target species are most abundant. The latter case poses particular problems for analysis of trends in abundance, because there may be a built-in bias for detecting declining populations (Alford and Richards 1999; but see Green 2003).

The effects of convenience sampling on inference can be subtle. Results of ecological experiments are often interpreted to demonstrate the generality of effects or even causality, but experiments by themselves are insufficient to explain complex ecological phenomena; such an effort requires integrating observation and theory with experimentation (Werner 1998). Broad inference is limited when experiments, even those that are internally well designed, are conducted at locations that are convenient for the researcher. The results may be useful for investigating possible mechanisms, but cannot be generalized without making the unsupportable assumption that the study sites are an unbiased representation of the habitats in question. For example, Blaustein *et al.* (1994) found that ambient ultraviolet-b (UV-B) radiation caused higher mortality of amphibian embryos than when embryos were shielded from UV-B at four lakes in the Cascade Mountains in Oregon, USA. This and subsequent research formed the basis of a theory of how climate change, UV-B, and pathogens might play a significant role in global amphibian declines (Kiesecker *et al.* 2001; Pounds 2001; Blaustein and Kiesecker 2002). The issue of amphibian declines and UV-B has been controversial with a large literature that I will not delve into here. Relevant to this chapter, the sites used for the UV-B studies were not chosen with respect to the potential for exposure to UV-B, but were a convenience sample. The primary study site turned out to have water much more transparent to UV-B transmission than most amphibian breeding habitats in the Pacific Northwest (Palen *et al.* 2002), severely restricting the generality of the proposed hypothesis.

Convenience sampling is almost always done without any intention to produce a biased result. The water chemistry and UV-B transmission at the study sites in the Oregon UV-B studies were not known beforehand. These sites were known to the researchers to have suitable amphibian populations and included locations where long-term studies had been conducted. Similarly, Corn and Bury's (1989) study of the effects of logging on stream amphibians in western Oregon did not include conscious bias in selection of streams to be sampled, but it was none-the-less based on a convenience sample. In this study, Bruce Bury and I identified likely sample locations on topographic maps beforehand, but the decision to sample was made in the field after inspecting each stream. We attempted to select typical streams and sample reaches based on our knowledge of the characteristics of headwater streams in the region. However, we did not begin with a well-defined population of streams, and we did not apply

any probabilistic sampling. As stated by Hayek and Buzas (1997, p. 113), “If the basis for inclusion in a sample is judgment, regardless of how expert, we will not have a reproducible measure of our field study’s usefulness.” Corn and Bury (1989) found strong differences in abundance and diversity of amphibians between streams in logged and unlogged forest. The paper was an early influence on what has become a spate of research on stream amphibians and forest management in the Pacific Northwest. Although the original results are largely supported by subsequent work (Olson *et al.* 2007), the convenience sampling design employed limits the scope of the conclusions and Corn and Bury (1989) should be viewed as hypothesis-generating work rather than a definitive demonstration of differences between streams in managed and unmanaged forests.

Cautions about convenience sampling apply equally to the sample frame, the pool from which the sample is drawn, as to individual study sites. Study areas, meaning the regions containing the sample frames, are almost never chosen at random. Study areas may be defined by a relevant management question (e.g. what is the status of amphibians in a National Park?), or they might be contain habitats or species of interest, yet be located conveniently near a researcher’s institution. Probabilistic methods can be used properly to select study sites, but if the frame is defined by convenience, then inference is restricted to the study area and the generality of the results may be limited. Two recent studies, conducted less than 200 km apart in Switzerland, illustrate this point. Pellet *et al.* (2004) found that urbanization and roads had a strong negative influence on presence of European treefrogs (*Hyla arborea*). Conversely, Van Buskirk (2005) found only weak support for landscape variables (including urbanization) to explain occurrence and abundance of treefrogs. Both studies were well designed, and although different methods of analysis were used (logistic regression versus information theoretic analysis) the different conclusions likely resulted from intrinsic differences between the study areas.

The North American Amphibian Monitoring Program (NAAMP; Weir and Mossman 2005), patterned after the North American Breeding Bird Survey (BBS; Peterjohn *et al.* 1995), also suffers from a sampling frame defined by convenience. NAAMP conducts manual calling surveys of breeding amphibians on prescribed road routes. Species are identified by their breeding calls, and data are collected mainly by volunteers. Survey routes are generated through a random process, but the goal of the program is to monitor trends in amphibian populations throughout the region where routes are conducted (Weir and Mossman 2005). Reliance on roadside observations limits the inference to those areas accessed by roads, or requires investigators to make an additional, untested assumption that the roadside amphibian population experiences the

same trends in abundance as those found in populations away from roads. The biases that potentially undermine this assumption can be related to both habitat and the observations themselves (Peterjohn *et al.* 1995). In Chapter 16 in this volume Dorcas *et al.* discuss assumptions about auditory observations. Habitat assumptions require that roadside wetlands reflect the same conditions found at wetlands away from roads, and that habitat condition changes over time in the same direction and at a similar pace alongside and away from roads. These assumptions are more likely to be violated than satisfied. For example, Keller and Scallan (1999) found that land cover types were similar near and away from BBS routes in Maryland and Ohio, but that in Maryland, urbanization was proceeding at a more rapid pace along roads. If urbanization is associated largely with existing road networks, then roadside habitats may diverge from distant habitats more rapidly in Maryland than in Ohio.

A more immediate and concrete difference between wetlands near and distant from roads (and also a difference between amphibians and birds) is that roads, especially paved roads with higher volumes of traffic, are a significant source of mortality of adult amphibians moving in and out of breeding ponds. Studies in North America and Europe have found negative relationships between traffic intensity and amphibian mortality and breeding activity (Fahrig *et al.* 1995), habitat occupancy (Pellet *et al.* 2004), and species richness and abundance (Eigenbrod *et al.* 2008). Similarly in New York, Karraker *et al.* (2008) found a twofold increase in density of egg masses of two amphibian species away from roads compared to roadside ponds. This may have been more an indirect road effect than from direct mortality. Demographic models showed significant negative effects on these species in roadside ponds resulting from the application of salt for de-icing (Karraker *et al.* 2008). Road de-icers epitomize a confounding variable contributing to differences that would be very difficult to model in analyses of NAAMP data. Application of de-icers varies among geographic regions, states, road types, and chemistry from year to year, and the data quantifying application are likely to be extremely difficult to obtain. The assumption that roadside habitats are equivalent to habitats away from roads is not supported by research to date.

The road studies illustrate the point that concerns about convenience sampling also apply to choosing variables for explaining patterns or trend. Data may be readily available in GIS data layers (e.g. road density), but less available data (e.g. traffic intensity, de-icer application) may be more important for generating well-supported models. Anderson (2008) emphasizes that considerable effort should be devoted to generating hypotheses before data are collected; this obviously applies to selection of the variables that make up the candidate models.

23.4 Abundance and inference

Issues of inference regarding species typically arise when deciding how to quantify abundance or the related estimate of habitat occupancy. These are important and somewhat controversial topics, and other chapters deal with estimating occupancy (Chapter 24) and abundance (Chapter 25) in detail. In this section, I focus on how uncertainty about numbers affects inference.

Studies often use an index, either simple counts of individuals observed or counts converted to relative abundance, to represent the abundance of a species (Hayek and Buzas 1997), and many standard field methods (e.g. in this volume or Heyer *et al.* 1994) are designed to obtain counts. However, inference about trends over time or differences among habitats requires that there is no trend in the detection probability (the relationship between the index and the true abundance). Bart *et al.* (2004) contended that this was a reasonable assumption for bird studies. Anderson (2001) regarded this as unlikely, and recently put it even more strongly (Anderson 2008, p. 20): "... the evidence is conclusive that they [index values] represent an amateur, unthinking approach and is not scientific" and "... index values are not data, they are just numbers." This viewpoint is not universal. Reliable estimation of detection probability may be difficult in many situations (Johnson 2008), and variation in detection probability, particularly among individuals, may result in unreliable estimates of abundance (Link 2003). Nevertheless, evidence from amphibian studies tends to support the idea that use of indices should be avoided, because estimates incorporating detection are more closely correlated to actual abundance than are counts (Schmidt 2004).

Welsh and Droege (2001) advocated use of count data from surveys of plethodontid salamanders to monitor forest condition and biological diversity. Terrestrial salamanders are associated with habitat features that are often disrupted by activities such as logging, and counts of salamanders in several studies sampled by a variety of techniques show relatively consistent numbers from year to year (Welsh and Droege 2001). Low inter-annual variation increases ability to detect trends. However, at any given time, the majority of *Plethodon* in a population are subterranean and unavailable to be captured (Bailey *et al.* 2004a), and counts of terrestrial plethodontids vary considerably among years, habitats, and sampling methods, violating many of the assumptions required for use of indices in monitoring (Hyde and Simons 2001; Dodd and Dorazio 2004). More critically, temporary immigration between surface and subterranean habitats varies among sampling occasions, so that the proportion of the population available to be sampled is not constant. Additionally, capture–recapture studies on

plethodontids have found low detection probabilities, often less than 0.05 (Jung *et al.* 2000; Welsh and Droege 2001; Dodd and Dorazio 2004). Low capture probabilities are not a problem if they are relatively uniform among habitats and observers (Welsh and Droege 2001), but there is considerable reason to doubt that this is true. Detection probabilities of plethodontids varied from 0.06 to 0.41 among years (Dodd and Dorazio 2004), and between 0.01 and 0.58 among sampling occasions in one model evaluated by Bailey *et al.* (2004a). Low detection magnifies any bias due to variation among observers, habitats, or years. For example, an increase in detection from 0.50 to 0.55 would result in a 10% increase in numbers of animals observed, but an increase from 0.06 to 0.38 would result in 6.5 times as many observations in the second count. This could produce a result similar to that illustrated in Schmidt (2004, figure 1), where counts were similar among years, but capture probability varied, with the result that counts did not reflect a large increase in actual abundance.

Concerns about use of count data apply more broadly than to just terrestrial salamanders. Johannson *et al.* (2006) used uncorrected counts of common frog (*Rana temporaria*) egg masses as an index to population size, and concluded that population size declined with increasing latitude and smaller populations had less genetic variability. Johannson *et al.* conceded that egg mass counts likely underestimated true abundance of breeding females but contended it was an unbiased index, because sampling was the same at all study sites. However, counts often fail to detect all egg masses present in a pond for a variety of reasons, such as differences in habitat complexity, weather conditions that might affect visibility, or variation in ability among observers. Grant *et al.* (2005) found detection probabilities of ranid egg masses to vary between 0.78 and 1.0. Variation in detection introduces uncertainty into conclusions about population size; this uncertainty is magnified if detection varies in a systematic manner across a study area.

Count data are incorporated into many indices of species diversity (Hayek and Buzas 1997), but calculation of these indices for amphibian assemblages are not appropriate, unless unbiased estimates of abundance are used instead of the raw counts. It has long been known that number of captures varies among sampling methods, so that a diversity index that included counts of species made using different techniques (for example, pitfall traps for one species and time-constrained searches for another) would not be valid (Corn 1994). Hyde and Simons (2001) demonstrated sampling efficiency varied among methods, but also among habitat types for some species of plethodontids when the same method was employed. Interpretations of diversity indices suffer from the same problems as interpretations about abundance.

The uncertainties about count data, the expense of obtaining unbiased estimates of population size, and high amount of year-to-year variation in abundance of many amphibian species prompted Green (1997) to suggest that tracking the changes in the occurrence of a species across the landscape—that is, whether or not a species occupied a given patch of habitat—was a more efficient way to monitor status and trend of amphibians. Such presence/absence (or, more accurately, detected/not detected) data used to be considered inferior to or less scientific than count data (MacKenzie *et al.* 2003). In part, this is because surveys to detect presence are typically at a reduced level of effort than surveys that include counts, and in most situations it is impossible to prove absence even with great effort. False negatives, or failures to detect species that are actually present, introduce bias that underestimates occupancy and can lead to errors in interpretation, such as incorrectly identifying the influence of habitat variables on occurrence (Mazerolle *et al.* 2005; Hossack and Corn 2007).

Occupancy estimation (MacKenzie *et al.* 2006; see also Chapter 24 in this volume) has been implemented in a variety of studies of amphibians, including effects of disturbance (e.g. Mazerolle *et al.* 2005; Hossack and Corn 2007), in addition to more general monitoring efforts (e.g. Corn *et al.* 2005; Schmidt 2005). Occupancy analysis has been recommended in the US Geological Survey's Amphibian Research and Monitoring Initiative (Muths *et al.* 2005), mainly through surveys for tadpoles or egg masses to indicate breeding populations. Tadpoles typically have high detection probabilities (Brown *et al.* 2007; Hossack and Corn 2007), which increases the precision of estimated occupancy. Occupancy analysis also shows promise as a means for monitoring terrestrial salamanders. Detection probabilities for occurrence of plethodontids on small plots were much higher than capture–recapture studies have found for detection of individual salamanders, and estimated occupancy was much less variable among years compared to estimated abundance (Bailey *et al.* 2004b).

Occupancy analysis may not be the best choice in all situations. It includes assumptions about the relationship between occurrence on the landscape and abundance of a species (Royle and Nichols 2003), although combining occupancy and count data is topic of active development (Royle *et al.* 2005; see also Chapter 24). Occupancy analysis can be difficult to implement in habitats that are not discrete, for example extensive wetlands. It also may not be possible to obtain reliable estimates of occupancy for rare species with low detection probabilities (Bailey *et al.* 2004b). A reservation about use of occupancy for detecting trends in salamanders is that local populations must go extinct or be recolonized for change to be observed. This may be a common occurrence for

pond-breeding amphibians (Green 1997), but it is less likely that populations of terrestrial salamanders undergo the same dynamics.

One final note of caution: research can be well designed, but the resulting data and analysis can still be undermined if the investigators do not pay sufficient attention to the biology of the organisms they are studying. Kroll *et al.* (2008) recently conducted a study on stream amphibians that employed a rigorous statistically valid design for sample selection, estimated occupancy among streams for several species, and analyzed effects on detection and occupancy using information-theoretic methods. One of their findings was that detection of coastal tailed frogs (*Ascaphus truei*) declined as the field season progressed. This was likely a consequence of the timing of sampling, which was conducted from 19 July to 6 October. Although in parts of the Pacific Northwest, tadpoles of *A. truei* require 2 or more years to reach metamorphosis, many of the streams examined by Kroll *et al.* are in a region where tadpoles metamorphose in less than 1 year, beginning as early as late June (Bury and Adams 1999). During the time when Kroll *et al.* were sampling, only adult frogs or hatchling tadpoles (which tend to remain clustered for several weeks under rocks near the nest site) would have been present in many streams. Both of these life stages are less observable than older tadpoles. Reduced detection probability results in positive bias in the occupancy estimate. Studies that employ occupancy analysis should be designed so that detection does not vary among sample units in a systematic manner.

23.5 Conclusions

The path to strong inference leads through good study design that incorporates probabilistic sampling from a well-defined population. Inventory and, especially, monitoring studies stray from this path when scientific rigor is sacrificed to logistic constraints and convenience in data collection. Tension often exists between the field biologist and the consulting statistician regarding the requirements of good study design and the logistical realities of data collection. Having been on the field biologist's side of the argument, I can testify that the attitude summarized by "Yes, we realize valid sample selection is important, and it would be nice, but we have to collect data from the real world", is fairly common. Constraints in site selection can be incorporated into study design, such as by stratifying based on accessibility, and the resulting analysis can test hypotheses about whether populations that are easily accessible differ from those that are not.

The perils of convenience sampling also apply to choice of life stage to study or explanatory variables to incorporate in a model. The easiest life stage to study

may not be the same one that is most sensitive to external factors, and variables should not be included in a model simply because the data are available. There is no “magic bullet” for sampling amphibians. No single technique encompasses the variety of life histories of amphibians or the habitats in which they can be found. Occupancy analysis provides a useful tool for avoiding the pitfalls of using simple count data or the logistic difficulties of obtaining unbiased estimates of abundance, but it is not a panacea. Ultimately, the design that allows the strongest inference will be one that avoids convenience sampling and minimizes untested assumptions when the data are analyzed.

23.6 Acknowledgments

I thank Larissa Bailey, Blake Hossack, Benedikt Schmidt, and Susan Walls for suggestions that greatly improved the manuscript.

23.7 References

- Alford, R. A. and Richards, S. J. (1999). Global amphibian declines: a problem in applied ecology. *Annual Review of Ecology and Systematics*, **30**, 130–65.
- Anderson, D. R. (2001). The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin*, **29**, 1294–7.
- Anderson, D. R. (2008). *Model Based Inference in the Life Sciences: a Primer on Evidence*. Springer, New York.
- Bailey, L. L., Simons, T. R., and Pollock, K. H. (2004a). Estimating detection probability parameters for plethodon salamanders using the robust capture–recapture design. *Journal of Wildlife Management*, **68**, 1–13.
- Bailey, L. L., Simons, T. R., and Pollock, K. H. (2004b). Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecological Applications*, **14**, 692–702.
- Bart, J., Droege, S., Geissler, P., Peterjohn, B., and Ralph, C. J. (2004). Density estimation in wildlife surveys. *Wildlife Society Bulletin*, **32**, 1242–7.
- Blaustein, A. R. and Kiesecker, J. M. (2002). Complexity in conservation: lessons from the global decline of amphibian populations. *Ecology Letters*, **5**, 597–608.
- Blaustein, A. R., Hoffman, P. D., Hokit, D. G., Kiesecker, J. M., Walls, S. C., and Hays, J. B. (1994). UV repair and resistance to solar UV-B in amphibian eggs: a link to population declines? *Proceedings of the National Academy of Sciences USA*, **91**, 1791–5.
- Brown, G. W., Scroggie, M. P., Smith, M. J., and Steane, D. (2007). An evaluation of methods for assessing the population status of the threatened alpine tree frog *Litoria verreauxii alpina* in southeastern Australia. *Copeia*, **2007**, 765–70.
- Bury, R. B. and Adams, M. J. (1999). Variation in age at metamorphosis across a latitudinal gradient for the tailed frog, *Ascaphus truei*. *Herpetologica*, **55**, 283–91.
- Cochran, W. G. (1977). *Sampling Techniques*, 3rd edn. John Wiley & Sons, New York.

- Corn, P. S. (1994). Straight-line drift fences and pitfall traps. In W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L.-A. C. Hayek, and M. S. Foster (eds), *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*, pp. 109–17. Smithsonian Institution Press, Washington DC.
- Corn, P. S. and Bury, R. B. (1989). Logging in western Oregon: responses of headwater habitats and stream amphibians. *Forest Ecology and Management*, **29**, 39–57.
- Corn, P. S., Hossack, B. R., Muths, E., Patla, D. A., Peterson, C. R., and Gallant, A. L. (2005). Status of amphibians on the Continental Divide: surveys on a transect from Montana to Colorado, USA. *Alytes*, **22**, 85–94.
- Dodd, Jr, C. K. and Dorazio, R. M. (2004). Using counts to simultaneously estimate abundance and detection probabilities in a salamander community. *Herpetologica*, **60**, 468–78.
- Eigenbrod, F., Hecnar, S. J., and Fahrig, L. (2008). The relative effects of road traffic and forest cover on anuran populations. *Biological Conservation*, **141**, 35–46.
- Fahrig, L., Pedlar, J. H., Pope, S. E., Taylor, P. D., and Wegner, J. F. (1995). Effect of road traffic on amphibian density. *Biological Conservation*, **73**, 177–82.
- Grant, E. M., Jung, R. E., Nichols, J. D. and Hines, J. E. (2005). Double-observer approach to estimating egg mass abundance of pool-breeding amphibians. *Wetland Ecology and Management*, **13**, 305–20.
- Green, D. M. (1997). Perspectives on amphibian population declines: defining the problem and searching for answers. In D. M. Green (ed.), *Amphibians in Decline: Canadian Studies of a Global Problem. Herpetological Conservation*, vol. 1, pp. 291–308. Society for the Study of Amphibians and Reptiles, St. Louis, MO.
- Green, D. M. (2003). The ecology of extinction: population fluctuation and decline in amphibians. *Biological Conservation*, **111**, 331–43.
- Hayek, L.-A. C. (1994). Research design for quantitative amphibian studies. In W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L.-A. C. Hayek, and M. S. Foster (eds), *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*, pp. 21–39. Smithsonian Institution Press, Washington DC.
- Hayek, L.-A. C. and Buzas, M. A. (1997). *Surveying Natural Populations*. Columbia University Press, New York.
- Heyer, W. R., Donnelly, M. A., McDiarmid, R. W., Hayek, L.-A. C., and Foster, M. S. (eds), (1994). *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*. Smithsonian Institution Press, Washington DC.
- Hobbs, N. T. and Hilborn, R. (2006). Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. *Ecological Applications*, **16**, 5–19.
- Hossack, B. R. and Corn, P. S. (2007). Responses of pond-breeding amphibians to wildfire: short-term patterns in occupancy and colonization. *Ecological Applications*, **17**, 1403–10.
- Hurlbert, S. N. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, **54**, 187–211.
- Hyde, E. J. and Simons, T. R. (2001). Sampling plethodontid salamanders: sources of variability. *Journal of Wildlife Management*, **65**, 624–32.
- Johannson, M., Primmer, C. R., and Merilä, U. (2006). History vs. current demography: explaining the genetic population structure of the common frog (*Rana temporaria*). *Molecular Ecology*, **15**, 975–83.

- Johnson, D. H. (2008). In defense of indices: the case of bird surveys. *Journal of Wildlife Management*, **72**, 857–68.
- Jung, R. E., Droege, S., Sauer, J. R., and Landy, R. B. (2000). Evaluation of terrestrial and streamside salamander monitoring techniques at Shenandoah National Park. *Environmental Monitoring and Assessment*, **63**, 65–79.
- Karraker, N. E., Gibbs, J. P., and Vonesh, J. R. (2008). Impacts of road deicing salt on the demography of vernal pool-breeding amphibians. *Ecological Applications*, **18**, 724–34.
- Keller, C. M. E. and Scallan, J. T. (1999). Potential roadside biases due to habitat changes along Breeding Bird Survey routes. *The Condor*, **101**, 50–7.
- Kiesecker, J. M., Blaustein, A. R., and Belden, L. K. (2001). Complex causes of amphibian population declines. *Nature*, **410**, 681–4.
- Kroll, A. J., Risenhoover, K., McBride T., Beach E., Kernohan B. J., Light, J., and Bach, J. (2008). Factors influencing stream occupancy and detection probability parameters of stream-associated amphibians in commercial forests of Oregon and Washington, USA. *Forest Ecology and Management*, **255**, 3726–35.
- Link, W. A. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrika*, **59**, 1123–30.
- MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., and Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, **84**, 2200–7.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., and Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Academic Press, Burlington, MA.
- 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.
- Muths E., Jung, R. E., Bailey, L. L., Adams, M. J., Corn, P. S., Dodd, Jr, C. K., Fellers, G. M., Sadinski, W. J., Schwalbe, C. R., Walls, S. C. *et al.* (2005). The U.S. Department of Interior's Amphibian Research and Monitoring Initiative: a successful start to a national program. *Applied Herpetology*, **2**, 355–71.
- Olson, D. H., Anderson, P. D., Frissell, C. A., Welsh, Jr, H. H., and Bradford, D. F. (2007). Biodiversity management approaches for stream–riparian areas: perspectives for Pacific Northwest headwater forests, microclimates, and amphibians. *Forest Ecology and Management*, **246**, 81–107.
- Palen, W. J., Schindler, D. E., Adams, M. J., Pearl, C. A., Bury, R. B., and Diamond, S. A. (2002). Optical characteristics of natural waters protect amphibians from UV-B in the U. S. Pacific Northwest. *Ecology*, **83**, 2951–7.
- Pellet, J., Guisan, A., and Perrin, N. (2004). A concentric analysis of the impact of urbanization on the threatened European tree frog in an agricultural landscape. *Conservation Biology*, **18**, 1599–1606.
- Peterjohn, B. G., Sauer, J. R., and Robbins, C. S. (1995). Population trends from the North American Breeding Bird Survey. In T. E. Martin and D. M. Finch (eds), *Ecology and Management of Neotropical Migrants*, pp. 3–39. Oxford University Press, New York.
- Platt, J. R. (1964). Strong inference. *Science*, **146**, 347–53.
- Pounds, J. A. (2001). Climate and amphibian declines. *Nature*, **410**, 639–40.

- Royle, J. A. and Nichols, J. D. (2003). Estimating abundance from repeated presence–absence data or point counts. *Ecology*, **84**, 777–90.
- Royle, J. A., Nichols, J. D., and Kéry, M. (2005). Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, **110**, 353–9.
- 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. (2005). Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **15**, 681–92.
- Thompson, S. K. (1992). *Sampling*. John Wiley & Sons, New York.
- Underwood, A. J. (1998). Design, implementation, and analysis of ecological and environmental experiments. In W. J. Resetarits, Jr. and J. Bernardo (eds), *Experimental Ecology: Issues and Perspectives*, pp. 325–49. Oxford University Press, New York.
- Van Buskirk, J. (2005). Local and landscape influence on amphibian occurrence and abundance. *Ecology*, **86**, 1936–47.
- Weir, L. A. and Mossman, M. J. (2005). North American Amphibian Monitoring Program (NAAMP). In M. Lannoo (ed.), *Amphibian Declines: The Conservation Status of United States Species*, pp. 307–13. University of California Press, Berkeley, CA.
- Welsh, Jr, H. H. and Droege, S. (2001). A case for using plethodontid salamanders for monitoring biodiversity and ecosystem integrity of North American forests. *Conservation Biology*, **15**, 558–69.
- Werner, E. E. (1998). Ecological experiments and a research program in community ecology. In W. J. Resetarits, Jr and J. Bernardo (eds), *Experimental Ecology: Issues and Perspectives*, pp. 3–26. Oxford University Press, New York.
- Williams, B. K., Nichols, J. D., and Conroy M. J. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.