

Introduction

Ecology is frequently defined as the study of the distribution and abundance of plants and animals (e.g., Andrewartha and Birch 1954; Krebs 1972). Consequently, the practice of counting animals in order to draw inferences about their numbers and distribution has a long tradition in animal ecology and management. In his classic book, *Animal Ecology*, Charles Elton (1927:173) wrote: "The study of numbers is a very new subject, and perfect methods of recording the numbers and changes in the numbers of animals have yet to be evolved." Elton then devoted 6 pages to the topic of animal "census" methods. In his equally influential classic, *Game Management*, Aldo Leopold (1933:139) listed "Census" as the first of four steps required to initiate game management on any piece of land. He then devoted a 30-page chapter to "game census" and another 25 pages to "measurement and diagnosis of productivity," a chapter that focused on assessing vital rates and population change. Methods for counting animals have indeed evolved over the last 70 years, and animal ecologists and managers now have an impressive methodological toolbox for estimating parameters associated with animal abundance and with the vital rates that produce changes in abundance (e.g., Seber 1973, 1982; Williams *et al.* 2002).

Today, biologists interested in understanding and managing animal populations and communities include some individuals who make full use of the methods available for drawing inferences about variation in animal numbers over time and space, and many others who do not appear to recognize the importance of appropriate inferential procedures. Because of those scientists and managers who do not take advantage of available estimation methods, the fields of animal population and community ecology, wildlife management, and conservation biology include numerous examples of substantial field efforts that do not produce reliable conclusions. These disciplines suffer not only from the failure of animal ecologists and managers to utilize the range of available methods for drawing inferences about animal abundance and associated vital rates but also from the lack of rigorous methods for estimating other quantities that may be biologically relevant. For example, other variables that could be used to quantify the current status of a community or population (we refer to these as *state variables*) include species richness (number of species) and occupancy (proportion of an area occupied by a species or fraction of landscape units where the species is present). Scant attention has been devoted to estimation of these latter state variables, with the result that there is a great need for methodological development.

In this book, we emphasize the need for estimation methods that permit inference about occupancy based on so-called presence-absence data and report results of our initial efforts to develop a set of such methods. We begin this chapter by providing brief operational definitions for some important terms, then move on to an outline of general principles for sampling animal populations, focusing on the why, what, and how of such sampling. This outline is followed by a more detailed look at the critical step of using field data to discriminate among competing hypotheses about system response to environmental variation and management actions. We note different field designs that are used to generate system dynamics for such discrimination and comment on the different strengths of inference resulting from these designs. The chapter concludes with a more detailed statement of book objectives and contents.

1.1. OPERATIONAL DEFINITIONS

The methods presented in this book should be useful to biologists involved in either science or management of biological populations. Both endeavors use the following three constructs: *hypothesis*, *theory*, and *model*. These terms are not always used consistently in the literature, and therefore we provide our own operational definitions for use in this book (also see Nichols 2001). We view a *hypothesis* simply as a plausible explanation (i.e., a "story") about

how the world, or part of it, works. For example, we would deem density-dependent recruitment for mid-continent, North American mallards (*Anas platyrhynchos*) as a hypothesis, with density-independent recruitment as an alternative, competing hypothesis (Johnson *et al.* 1997). Once a hypothesis has withstood repeated efforts to falsify it, to the extent that we have some faith in predictions deduced from it, the hypothesis may become a *theory* (e.g., Einstein's theory of relativity). A theory can still be disproved in the future given new data or the expansion of the part of the world to which the theory is thought to be applicable (e.g., Newtonian physics).

Very generally, we view a *model* as an abstraction of a real-world system, which can be used to describe observed system behavior and predict how the system may respond to changes or perturbations. Within this broad definition we recognize many different kinds of models (Nichols 2001), three of which are especially useful within the context of this book. A *conceptual model* is a set of ideas about how the system of interest works, and may include one or more hypotheses or theories about the system. A *verbal model* is created by translating these ideas into words. Finally, a *mathematical model* results from translating a conceptual or verbal model into a set of mathematical equations, using defined parameters to symbolize the key processes of the system. In this book we derive mathematical expressions from our conceptual ideas about the processes that occur when collecting occupancy field data, placing particular attention on using the collected data to estimate the parameters of these models.

By following the logical progression above, note that a mathematical model is ultimately a representation of one or more hypotheses or theories about the system. Therefore, competing hypotheses can be formulated into competing mathematical models. Applying each model to the same set of available data, it may be possible to formally determine which model (and therefore which hypothesis) has a greater degree of support given the data at hand. Essentially this is an exercise in model selection. We advocate and use such an approach throughout this book.

1.2. SAMPLING ANIMAL POPULATIONS AND COMMUNITIES: GENERAL PRINCIPLES

It is our belief that many existing programs for sampling animal populations and communities are not as useful as they might be because investigators have not devoted adequate thought to fundamental questions associated with establishment of such programs. These failures have greatly reduced the value of efforts ranging from individual scientific investigations to large-scale monitoring programs. These latter programs are especially troubling, because they

can require nontrivial fractions of the total funding and effort available for the conduct of science and management of animal populations and communities. Here we present some opinions about the sort of thinking that should precede and underlie good animal sampling programs. These opinions are structured around three basic questions to be addressed during the design of an animal sampling program (see Yoccoz *et al.* 2001): *Why? What? and How?*

WHY?

Efforts to sample animal populations are generally associated with one of two main classes of endeavor, science or conservation and management (or possibly both). Science can be viewed as a process used to discriminate among competing hypotheses about system behavior, that is, discriminating among different ideas about how the world, or a part of it, works (e.g., whether recruitment to a population is density dependent). This process typically involves mathematical models. For example, a mathematical model that could be used to represent the number of recruits to a population (r) that assumes no density dependence would be $r = N_F b$, where N_F is the number of breeding females in the population and b is the average number of female births per adult female and is viewed as a constant (with respect to current breeding female population size). A different model that conceptualizes the effect of density-dependent recruitment would be $r = N_F b(N_F)$, where $b(N_F)$ specifies a functional relationship, such that number of recruits per female is a function of total female abundance. The primary use of models is to project the consequences of hypotheses, that is, to deduce predictions about system behavior (e.g., Nichols 2001). In the case of our example, the model is used to predict the number of recruits at different levels of population density.

The key step in science, then, involves the confrontation of these model-based predictions with the relevant components of the real-world system (Hilborn and Mangel 1997; Williams *et al.* 2002). Faith and confidence increase for those models (and hence those underlying hypotheses) whose predictions match observed system behavior well and decrease for models that do a poor job of predicting. However, for most practical situations involving animal populations and communities, true system behavior cannot be directly observed, but must be estimated from data collected from sampling programs. Thus, sampling programs constitute a key component of scientific research.

In the conduct of management and conservation, estimates of state variables for animal populations and communities serve three distinct roles (Kendall 2001). First, estimates of system state are needed in order to make state-dependent management decisions (e.g., Kendall 2001; Williams *et al.* 2002). For example, the decision of which management action to take fre-

quently depends upon the current population size. Second, system state is frequently contained in the objective functions (precise, usually mathematical, statements of management objectives) for managing animal populations and communities. Evaluation of the objective function is an important part of management, addressing the question "To what extent are management objectives being met?" Finally, good management requires either a single model thought to be predictive of system response to management actions or a set of models with associated weights reflecting relative degrees of faith in their validity. The process of developing faith in a single model or weights for members of a model set involves the confrontation of model predictions with estimates of true system response. This confrontation is the scientific component of informed management and requires animal sampling programs that provide reliable estimates of state variables and associated vital rates.

Despite the importance of being explicit about why a program for sampling animal populations or communities is needed, we believe that many studies suffer from a failure to clearly articulate specific study objectives. This is especially evident in many large-scale monitoring programs (Yoccoz *et al.* 2001). For example, the following objectives statements from a report on ecological monitoring programs in the United States (LaRoe *et al.* 1995:3, 4) are fairly typical: "The goal of inventory and monitoring is to determine the status and trends of selected species or ecosystems"; "Inventory and monitoring programs can provide measures of status and trends to determine levels of ecological success or stress." The second statement implies an interest in management and conservation, but without specification of available management actions and hypotheses about system response to those actions, the statement provides little basis for monitoring program design. Thus, we advocate clear specification of monitoring program objectives.

Objective specification is facilitated by the recognition that monitoring of animal populations and communities is not a stand-alone activity of great inherent utility, but is more usefully viewed as a component of the processes of science and/or management. This recognition leads naturally to detailed consideration of exactly how the monitoring program results are to be used in the conduct of science or management or both. Such considerations lead directly to decisions about monitoring program design, whereas vague objectives that fail to specify use of program data and estimates provide little guidance for program design and can lead to endless debate about design issues.

WHAT?

The selection of what state variable(s) and associated vital rates to estimate will depend largely on the answer to the initial question of "Why?" The selec-

tion of state variables for scientific programs will depend on the nature of the competing hypotheses and specifically on the quantities most likely to lead to discrimination among the hypotheses (i.e., for what quantities are predictions of competing hypotheses most different?). The selection of state variables for management programs will depend on the most relevant characterization of system state, on management objectives, and on the ability to discriminate among competing hypotheses about system response to management actions. Practicality must also be considered in both cases as, most likely, logistical resources will be limited.

When dealing with single species, the most commonly used state variable is abundance or population size. Estimation of abundance frequently requires substantial effort, but it is a natural choice for state variables in studies of population dynamics and management of single-species populations. Some studies of animal abundance focus directly on changes in abundance, frequently expressed as the ratio of abundances in two sampling periods (e.g., two successive years) and termed the finite rate of population increase or population growth rate, λ . In scientific studies, mechanistic hypotheses frequently concern the vital rates responsible for changes in abundance, rates of birth (reproductive recruitment), death, and movement in and out of the population. In management programs, effects of management actions on animal abundance must also occur through effects on one or more of these vital rates. Thus, many animal sampling programs involve efforts to estimate abundance and rates of birth, death, and movement for animals inhabiting some area(s) of interest.

We believe that another useful state variable in single-species population studies is occupancy, defined as the proportion of area, patches, or sample units that is occupied (i.e., species presence). Sampling programs designed to estimate occupancy tend to require less effort than programs designed to estimate abundance (e.g., Tyre *et al.* 2001; MacKenzie *et al.* 2002; Manley *et al.* 2004). In the case of rare species, it is sometimes practically impossible to estimate abundance, whereas estimation of occupancy is still possible (MacKenzie *et al.* 2004a, 2005). Thus, for reasons that include expense and necessity, occupancy is sometimes viewed as a surrogate for abundance. However, there are also a number of kinds of questions for which occupancy would be the state variable of choice regardless of the effort involved in sampling. For example, metapopulation dynamics (e.g., Hanski and Gilpin 1997; Hanski 1999) are frequently described by patch occupancy models. So-called incidence functions (e.g., Diamond 1975a; Hanski 1994a) relate patch occupancy to patch characteristics such as size, distance to mainland or some source of immigrants, habitat, etc. Occupancy is the natural state variable for use in studies of distribution and range (e.g., Brown 1995; Scott *et al.* 2002) and should also be useful in the study of animal invasions and even disease dynamics. Patch occu-

pancy dynamics may be described using the rate of change in occupancy over time, and the vital rates responsible for such change are patch-level probabilities of extinction and colonization. Historical, current, and proposed uses of patch occupancy as a state variable for science and management will be discussed in more detail in Chapter 2.

When scientific or conservation attention shifts to the community level of organization, many possible state variables exist. The basic multivariate state variable of community ecology is the species abundance distribution, specifying the number of individuals in each species in the community. Many derived state variables are obtained by attributing different values or weights to individuals of different species (Yoccoz *et al.* 2001). Several common diversity indices are computed by providing a weight of 1 to every individual of each species (e.g., Pielou 1975; Patil and Taillie 1979), but it is also possible to give additional weight to individuals of species thought to be of special importance (e.g., endemic species or species of economic value) (Yoccoz *et al.* 2001). A state variable that is used commonly in community studies is simply species richness, the number of species within the taxonomic group of interest that is present in the community at any point in time or space. This state variable is used in scientific investigations (e.g., Boulinier *et al.* 1998b, 2001; Cam *et al.* 2002) and programs for management and conservation (e.g., Scott *et al.* 1993; Keddy and Drummond 1996; Wiens *et al.* 1996). The vital rates responsible for changes in species richness over time are rates of local species extinction and colonization.

In this book we focus largely on the state variable of occupancy, but note how these methods can also be applied where species richness-type metrics may be of interest (Chapter 9).

How?

Proper estimation of state variables and inferences about their variation over time and space require attention to two critical aspects of sampling animal populations: spatial variation and detectability (Fig. 1.1) (Lancia *et al.* 1994; Thompson *et al.* 1998; Williams *et al.* 2002). Spatial variation in animal abundance is important because in large studies and most monitoring programs investigators cannot directly survey the entire area of interest. Instead, investigators must select a sample of locations to which survey methods are applied, and this selection must be done in such a way as to accomplish two things. First, selection of study locations should be based on study objectives. In the case of scientific objectives, study locations should be selected to provide the best opportunity to discriminate among the competing hypotheses of interest (see Section 1.3 for further discussion). For example, in the case of an obser-

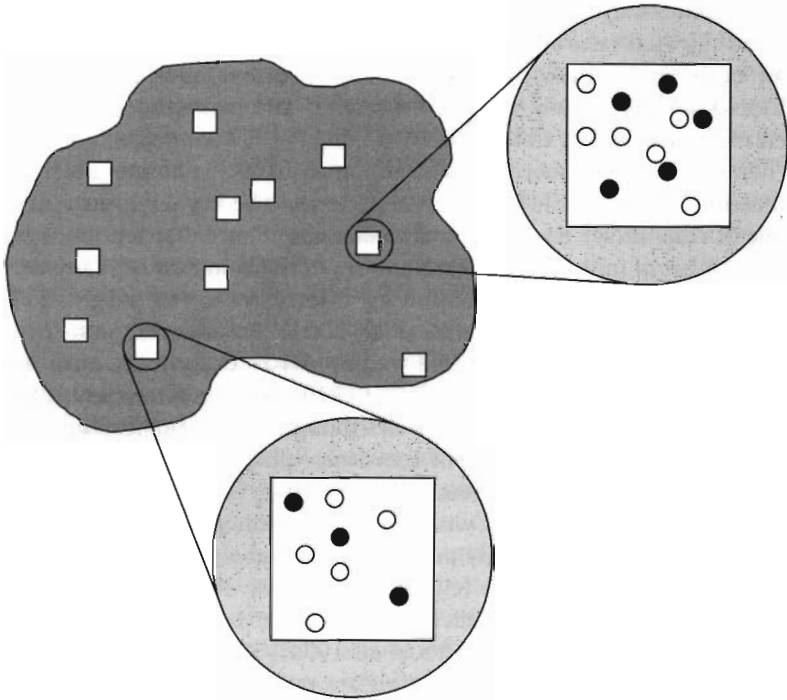


FIGURE 1.1 Illustration of the two critical aspects of sampling animal populations, spatial variation and detectability. The shaded region indicates the area or population of interest, with the small squares representing the locations selected for sampling. Within each sampling location, animals will be detected (filled circles) or undetected (hollow circles) during a survey or count.

vational study involving hypotheses about habitat variables, selected study locations might be extremes with respect to the variable(s) of interest or else might be locations at which changes in the variable(s) are anticipated. In the case of a management program, study locations should of course include the areas to which management actions are applied. Second, within larger areas selected based on study objectives, sample locations should be selected in a manner that permits inferences about the locations that are not surveyed and hence about the entire area(s) of interest. Approaches to sampling that accomplish this inferential goal include simple random sampling, unequal probability sampling, stratified random sampling, systematic sampling, cluster sampling, double sampling, and various kinds of adaptive sampling (e.g., Cochran 1977; Thompson 1992; Thompson and Seber 1996).

Detectability refers to the reality that, even in locations that are surveyed by investigators, it is very common for animals and even entire species to be missed and go undetected. Most animal survey methods yield some sort of count statistic. For example, when abundance is the quantity of interest, the count statistic might be the number of animals caught, seen, heard, or harvested. Let N_{it} be the true number of animals associated with an area or sample unit of interest, i , at time t , and denote as C_{it} the associated count statistic. This statistic can be viewed as a random variable whose expectation (basically the average value of the count if we could somehow conduct the count under the exact same conditions many times; see Chapter 3) is the product of the quantity of interest, abundance at the surveyed location, and the detection probability associated with the count statistic:

$$E(C_{it}) = N_{it}p_{it}, \quad (1.1)$$

where p_{it} is the detection probability (probability that a member of N_{it} appears in the count statistic, C_{it}). Estimation of N_{it} thus requires estimation of p_{it} :

$$\hat{N}_{it} = C_{it} / \hat{p}_{it}, \quad (1.2)$$

where the “hats” in this expression denote estimators (see Chapter 3). Expression (1.2) is very general and widely applicable. In fact, virtually all of the abundance estimation methods summarized and reviewed by Seber (1973, 1982), Lancia *et al.* (1994, 2005), Thompson *et al.* (1998), Williams *et al.* (2002), and Borchers *et al.* (2003) involve different approaches to the estimation of detection probability followed by (or integrated with) application of expression (1.2).

Frequently, interest will not be in abundance itself but in relative abundance, the ratio of abundances at two locations ($\lambda_{jit} = N_{it}/N_{jt}$, where i and j denote locations and t still denotes time), or in rate of population change, the ratio of abundances in the same location at two times ($\lambda_{it} = N_{it+1}/N_{it}$). Sometimes count statistics are treated as indices, and their ratio is used to estimate the true ratio of abundances. For example, consider the estimator $\hat{\lambda}_{it} = C_{it+1}/C_{it}$. The expectation of this estimator can be approximated using expression (1.1) as:

$$E(\hat{\lambda}_{it}) \approx \frac{N_{it+1}p_{it+1}}{N_{it}p_{it}} = \lambda_{it} \left(\frac{p_{it+1}}{p_{it}} \right). \quad (1.3)$$

As can be seen from (1.3), the ratio of counts estimates the product of the quantity of interest, λ_{it} , and the ratio of detection probabilities. If the detection probabilities are very similar for the two sample times, then the estimator will not be badly biased, but when detection probabilities differ, then the index-based estimator will be biased. If detection probability itself is viewed

as a random variable, then we still require $E(p_{it}) = E(p_{it+1})$ in order for a ratio of counts to be a reasonable estimator.

Proponents of the use of count statistics as indices for estimating relative abundance typically recommend standardization of survey methods as one means of trying to insure similar detection probabilities. Standardization involves factors that are under the control of the investigator (e.g., effort, trap type, bait, season and time of day of survey). While standardization of survey methods is usually a good idea, we believe that this approach is unlikely to produce equal detection probabilities, because there are always likely to be unidentified and uncontrollable factors that influence detection probabilities (Conroy and Nichols 1996). Sometimes it is possible to identify uncontrollable factors that could influence detection probability and incorporate them as covariates into analyses of count statistics. This approach is reasonable when dealing with factors that could only affect detection probability and not animal abundance itself. For example, differences in detection probabilities among observers are often incorporated into analyses of avian point count data (Link and Sauer 1997, 2002). However, it would not be wise to use a similar approach with habitat data, as habitat would be expected to influence not only detection probability but also animal abundance itself. Thus, "controlling" for habitat effects by incorporating them into analyses as covariates would not be appropriate. Of course, factors that we do not identify but still affect detection probability cannot be treated as covariates either.

Another common claim supporting the use of indices is that they are relatively assumption free, unlike the methods used to actually estimate abundance (e.g., Seber 1982; Williams *et al.* 2002). However, there are a large number of implicit assumptions to be made if the index is to be related to animal abundance. In fact, interpretation of an index as some indicator of true population size typically requires all the assumptions used to estimate abundance plus the assumption that a constant fraction of the population is counted each survey. Some uses of indices require the assumption that all animals are counted during each survey. As these assumptions are unlikely to be true, we believe that indices have a very limited use in good monitoring programs. We conclude that estimation of both absolute and relative abundance requires information about detection probability (also see Lancia *et al.* 1994; MacKenzie and Kendall 2002; Williams *et al.* 2002).

The importance of obtaining information about detection probability extends to other state variables as well. Investigations of species richness usually involve counts of the number of different species. Under some designs the counts are conducted at multiple locations within some large area to which inference is to apply, whereas other designs use counts conducted at multiple times (e.g., days) on a single area of interest (e.g., Nichols and Conroy 1996; Williams *et al.* 2002). In both designs, it is recognized that some species may

go undetected, and the replication (geographic or temporal) is used to estimate a species level detection probability, the probability that at least one individual of a species will be detected given that the species inhabits the area of interest. Efforts to estimate species richness from samples of animal communities are not new (Fisher *et al.* 1943; Preston 1948; Burnham and Overton 1979). Nevertheless, community ecologists have tended to ignore the issue of detection probabilities less than 1, and only recently has adequate attention been devoted to this estimation problem (e.g., Chao and Lee 1992; Bunge and Fitzpatrick 1993; Colwell and Coddington 1994; Walther *et al.* 1995; Chao *et al.* 1996; Nichols and Conroy 1996; Boulinier *et al.* 1998a; Cam *et al.* 2000; Williams *et al.* 2002; Dorazio and Royle 2005).

Detection probability is also very relevant to the estimation of occupancy. Define occupancy, ψ , as the probability that a randomly selected site or sampling unit in an area of interest is occupied by a species (i.e., the site contains at least one individual of the species). If x and s represent the number of occupied and total sites, respectively, then we can estimate occupancy as $\hat{\psi} = x/s$. However, x is not typically known. Instead, we will have a count of sites where the species has been detected, but this count will likely be smaller than x , because species will not always be detected in occupied sites (i.e., due to "false absences"). Thus, we must develop methods (e.g., based on multiple surveys of sites) to estimate detection probability and thus to estimate x . For example, we can use an analog of expression (1.2), where the count is the number of sites at which the species is detected, and the detection probability is the probability that the species is detected during sampling of an occupied site. Occupancy can then be estimated as:

$$\hat{\psi} = \frac{\hat{x}}{s}. \quad (1.4)$$

We have actually developed more direct ways to estimate occupancy (e.g., MacKenzie *et al.* 2002; Royle and Nichols 2003; Chapters 4, 5), but the basic rationale underlying these approaches is the same as outlined here.

Inferences about occupancy may be misleading when detection probability is not incorporated into the methods of data analysis. Not only will naïve approaches underestimate occupancy (as above), but indices intended to reflect relative occupancy also could be biased (MacKenzie 2006) and the effect of casual factors or variables may be underestimated (Tyre *et al.* 2003) or misidentified, particularly if detection probability covaries with the factors or variables thought to affect occupancy (Gu and Swihart 2004; MacKenzie 2006). Inferences about the dynamic processes that drive changes in occupancy may also be inaccurate (Moilanen 2002; MacKenzie *et al.* 2003). Indeed, an important theme of this book is that robust inference about occupancy and

related dynamics can only be made by explicitly accounting for detection probability.

1.3. INFERENCE ABOUT DYNAMICS AND CAUSATION

Chapter 2 will focus on the “what” of animal sampling programs and discuss the use of occupancy as a state variable. Much of the remainder of the book will then focus on the “how” question of sampling animal populations. That is, given interest in occupancy, how do we estimate this state variable and the vital rates responsible for its change in reasonable ways? Although we believe that this emphasis is justified by the absence of previous work and good guidance on drawing inferences about occupancy, we regret the need to abandon issues about “why” we sample animal populations. In our introductory discussion about why we might want to sample animal populations and communities, we emphasized that sampling programs are usefully viewed as components of the larger processes of science or management. In this section, we briefly discuss the manner in which results of animal sampling programs are used to draw the inferences needed for science or management. This discussion touches aspects of design that extend well beyond efforts to obtain reasonable estimates of state variables of interest.

The key step in the scientific process involves a comparison of estimates of state variables with model-based predictions associated with competing hypotheses. Such comparisons also constitute an important management use of estimates from animal sampling programs, as the ability to predict consequences of different management actions is critically important to informed management. Scientific programs include interest in responses of animal populations and communities to a variety of factors (e.g., changes in predators, competitors, weather, habitat, disease, toxins/pesticides). Management programs focus not only on responses to management actions but also on other factors that might improve predictive abilities. We would like to discriminate among competing hypotheses about the relevance of different causal factors to system dynamics with the ultimate goal of being able to predict the magnitude of the state variable(s) at time $t + 1$, given the magnitude of the state variable at time t and knowledge of the causal factors operating between times t and $t + 1$ (Williams 1997; Williams *et al.* 2002).

GENERATION OF SYSTEM DYNAMICS

The scientific process usually includes some means of generating system dynamics so that estimated changes in state variables can be compared with

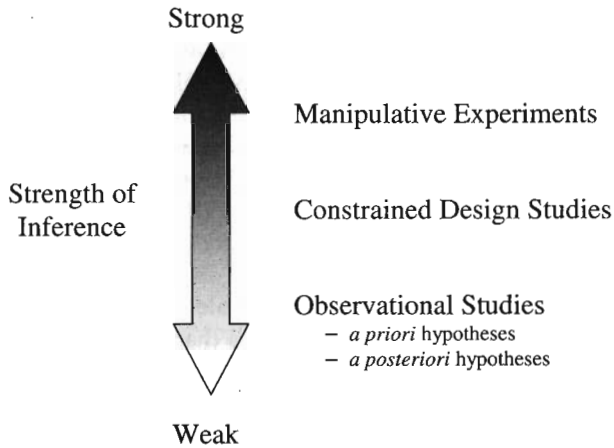


FIGURE 1.2 Strength of inference of different sampling designs that could be used to generate system dynamics.

the predictions of competing models. Multiple approaches are used to generate system dynamics in population and community ecology, and we classify these approaches broadly as true manipulative experiments, constrained designs or quasi-experiments, and observational studies (Romesburg 1981; Skalski and Robson 1992; Manly 1992; Williams *et al.* 2002). These approaches merit brief discussion here, as they provide different strengths of inference (Fig. 1.2).

Inferences are strongest when system dynamics are generated via the conduct of true manipulative experiments (see Fisher 1947; Hurlbert 1984; Skalski and Robson 1992; Manly 1995). Such experiments are characterized by replication, randomization in the assignment of different treatments (application of different hypothesized causal factors) to experimental units, and the use of a control or standard treatment group. In the context of population and community ecology, experimental units may be populations or communities occurring naturally or created as part of the experimental design. *Replication* refers to the application of treatments to multiple experimental units as a means of estimating the experimental error or error variance. The error variance reflects the variance among experimental units to be expected in the absence of treatment differences (i.e., the variance associated with all factors except the different treatments). *Randomization* refers to random assignment of treatments to experimental units. Randomization protects against systematic differences among experimental units receiving different treatments and represents an effort to insure that any systematic post-treatment differences among experimental units treated differently can be attributed to the treat-

ments themselves. One treatment type is typically designated as a control and is used to provide a baseline against which other treatments can be compared. The use of a control group is especially useful in attributing causation to different treatments and permitting estimation of treatment effects on response variables. Manipulative experiments thus seek to reduce potential sources of ambiguity to the extent possible, yielding strong inferences about causation.

True manipulative experiments are frequently difficult to perform on free-ranging animal populations and communities due to cost and practical field constraints. In many instances, we may be able to manipulate systems but may be required to do so using study designs that lack replication, randomization, or both of these features (see Green 1979; Skalski and Robson 1992; Williams *et al.* 2002). Inferences resulting from such constrained, or quasi-experimental, designs will typically not be as strong as those based on manipulative experimentation (see examples in Nichols and Johnson 1989).

Finally, the investigator may be unable to manipulate the system at all and may be forced to rely on natural variation to generate system dynamics. For example, large-scale animal monitoring programs may provide time series of estimated state variables, and retrospective analyses can be used to try to distinguish among competing hypotheses about system dynamics (Nichols 1991a). Two general approaches to observational studies are used, and they are distinguished by the existence of *a priori* hypotheses. The observational studies that tend to be most useful to science are those for which conditional *a priori* hypotheses are specified and used to guide monitoring program design (Nichols 2001; Williams *et al.* 2002). The hypotheses are conditional in the sense that changes in purported causal factors are not known *a priori*, as they are when the investigator imposes a manipulation. Instead, the different hypotheses predict different relationships between suspected causal factors and system state variables, and specific predictions then emerge as changes in the causal factors occur naturally and are observed. The initial specification of the hypotheses facilitates monitoring program design, as efforts can be devoted to monitoring changes in hypothesized causal factors as well.

The other approach to observational studies involves the development of *a posteriori* hypotheses to explain observed system dynamics. Monitoring programs may yield annual estimates of quantities such as population size over relatively long time periods (e.g., 20 years), and it is commonly thought that such trajectories lead directly to an understanding of underlying population dynamics. It is a common practice to use such data with correlation and regression analyses to investigate possible relationships between population size and various environmental and management variables. The problem with this approach is that it is unlikely to yield "reliable knowledge" (Romesburg 1981), because there will typically be multiple *a posteriori* hypotheses that provide

reasonable explanations for any observed time series (Nichols 1991a). Indeed, we tend to agree with Pirsig's (1974:107) assertion that The number of rational hypotheses that can explain any given phenomenon is infinite.

The potential for being misled by retrospective analysis of data exists for all kinds of observations (Platt 1964; Romesburg 1981) but is probably especially large for time series of estimates of population size and related variables. One reason for this is that population size is not observed but is estimated, often with large sampling variances and sometimes with bias. Temporal variation in point estimates of population size is thus not equivalent to temporal variation in the underlying population (Link and Nichols 1994). Another difficulty in drawing inferences from retrospective analyses of population trajectory data involves the stochastic nature of population processes. Death, for example, is typically viewed as a simple stochastic process. If a population has 100 animals at time t and if each of these animals has a probability of 0.2 of dying during the interval $(t, t + 1)$, then we do not expect exactly 80 animals to be alive at time $t + 1$. Instead, the number of survivors will be a binomial random variable with expected value 80, but with likely realized values of 78, 83, 75, etc. Reproductive processes and movement are also stochastic in nature, leading to the view of a population trajectory as a single realization of a (likely complicated) stochastic process. There is little reason for us to expect to be able to infer much about the nature of an unknown stochastic process based on a single realization of that process (Nichols 1991a). This is analogous to being handed a loaded coin, being permitted to flip it once, and then being asked to specify the probability of obtaining heads.

Another difficulty associated with inferences from retrospective studies of population monitoring data involves using correlation analysis to draw inferences about the functional relationship between variables represented by time series. A clear example of such problems involves the existence of trends and monotonicity in many environmental covariates that potentially influence animal populations. Metrics of human-related environmental variables such as habitat fragmentation, habitat degradation, and pollutant levels will frequently tend to show an increasing trend over time. Correlation analyses involving two variables, each of which shows a time trend, will tend to indicate association, although this may have nothing to do with any functional relationship between the variables. In fact, the problem of conducting association analyses of two time series extends well beyond the case of monotonic trends, and such analyses frequently lead to inappropriate inferences (Yule 1926; Barker and Sauer 1992).

These various considerations lead us to conclude that development of a *posteriori* hypotheses based on retrospective analyses of monitoring data is an approach that necessarily results in weak inferences. Certainly we do not claim that such retrospective analyses are without value, as they can sometimes

provide useful insights and ideas about system behavior. Instead, our recommendation is that such analyses be viewed primarily as an approach to hypothesis generation rather than as an inferential assessment of the hypothesis as an explanation for system dynamics. We thus recommend that observational studies be guided by *a priori* hypotheses, with exploratory retrospective analyses possibly used as a means of hypothesis generation.

As noted above, distinguishing among competing hypotheses about system response to management is an important component of an informed decision process. The term *adaptive management* (e.g., Holling 1978; Walters 1986; Hilborn and Walters 1992; Williams *et al.* 2002) typically applies to management that is state dependent and that incorporates learning about system response to management actions. It is this learning component that distinguishes adaptive management from other decision processes (Kendall 2001; Williams *et al.* 2002). Estimates of system state are used not only for the purpose of making state-dependent decisions but also as a means of confronting the predictions of competing models about system response for the purpose of discriminating among their associated hypotheses. Based on objectives, potential actions, an estimate of system state, and models (with associated probabilities reflecting relative degrees of faith), managers make the decision to take a particular action at time t . This action drives the system to a new state at time $t + 1$, and this state is identified via a monitoring program. Probabilities associated with degrees of faith in the various system models are then updated based on the distance between estimated system state and the predictions of the competing models (Kendall 2001; Nichols 2001; Williams *et al.* 2002). Although this approach to multimodel inference is used in the current applications of adaptive management with which we are most familiar (Nichols *et al.* 1995; Johnson *et al.* 1997; Williams *et al.* 2002), hypothesis-testing approaches are also possible and are also based on the distance between estimated system state and model-based predictions.

In the context of the previous discussion of approaches for generating system dynamics, the learning component of adaptive management will virtually always be manipulative, in that management actions will be imposed and system response then observed. However, attainment of management objectives is of primary importance in adaptive management, and learning is valued only to the extent that it is useful in better meeting objectives. Thus, in most applications with which we are familiar, the learning components of adaptive management exhibit the features of constrained designs. However, if management is of a spatially extended system and if different actions are to be taken on different spatial units of the system, then a manipulative experimental approach might be taken as well.

In summary, the conduct of science requires some means of generating system dynamics for comparison with predictions of competing hypotheses.

True manipulative experiments represent a study design that permits strong inferences about causation. Constrained or quasi-experimental designs involve manipulations, but the absence of either randomization or replication, or both features, does not permit the strength of inference of a true experiment. Finally, observational studies based on retrospective analyses of monitoring data involve no manipulation as part of study design and rely on natural variation in purported causal factors. These analyses tend to yield weaker inferences than manipulative studies. Within observational studies using retrospective analyses, those that test predictions of *a priori* hypotheses tend to yield stronger inferences than analyses used to generate *a posteriori* hypotheses. Adaptive management represents an informed decision process incorporating explicit efforts to learn about system responses to management actions. Because learning is not the sole objective of adaptive management, management manipulations typically follow some form of constrained design.

STATICS AND PROCESS VS. PATTERN

Inferences about causation emerge most naturally from studies of system dynamics. Scientists and managers estimate the state variable at time t , apply or observe purported causal factors operating between times t and $t + 1$, and then estimate the state variable again at time $t + 1$. However, because of the difficulties in applying manipulations to animal populations and communities and in properly estimating relevant state variables over time, animal ecologists have also tried to draw inferences about dynamics based on observations of spatial pattern at a single time, t . Brown (1995:10) describes "macroecology" as a research program in ecology with "emphasis on statistical pattern analysis rather than experimental manipulation." Inferences based on such efforts have been applied to each of the state variables described above—abundance, species richness, and occupancy.

Ecologists frequently use spatial variation in abundance of animals to draw inferences about habitat "quality," based on the commonsense idea that if animals are found in higher density in one habitat than others, then that habitat is likely of high quality. For such a statement to have meaning, "quality" must be defined. In their influential work on habitat selection, Fretwell and Lucas (1969; Fretwell 1972) defined habitat quality in terms of the fitness of organisms in that habitat. The two fundamental fitness components, survival probability and reproductive rate, are also primary determinants of population dynamics, so this definition is relevant to population ecologists and managers as well. Observations of spatial variation in animal density associated with habitat variation do not yield reliable inferences about individual fitness or dynamics of populations inhabiting such areas (e.g., van

Horne 1983; Pulliam 1988). Instead, such inferences require studies of system dynamics, in this case habitat-specific demography (e.g., Franklin *et al.* 2000), preferably in conjunction with habitat manipulations.

The relationship between species richness and area is one of the oldest and most-cited static relationships in ecology (e.g., Arrhenius 1921; Preston 1948). Hypotheses about the dynamic processes responsible for this relationship include habitat selection coupled with habitat heterogeneity (e.g., Williams 1964) and increased probabilities of local extinction in small areas (e.g., MacArthur and Wilson 1967). However, these two hypotheses yield similar species-area relationships, providing no basis for distinguishing between these or other mechanistic explanations (Connor and McCoy 1979).

Occupancy appears to be used more frequently in static analyses than either of the other discussed state variables, abundance and species richness. Static analyses of occupancy data in animal ecology can be illustrated with two common applications, single-species incidence functions and multiple-species co-occurrence patterns. Incidence functions involve efforts to model dichotomous spatial occupancy pattern (presence or absence) as a function of characteristics of the sampled locations or patches. Diamond (1975a) first described incidence functions in his studies of distributional ecology of birds inhabiting islands in the area of New Guinea. He grouped islands by such characteristics as land area and total avian species richness and then plotted the proportion of islands in each category (e.g., area, richness) that was occupied by a particular species. Diamond noted that some species tended to occur only on large, species-rich islands, whereas others were found only on remote, species-poor islands. Diamond (1975a:353) viewed the incidence function as a "fingerprint" of the distributional strategy of a species" and used these functions to draw inferences about such processes as dispersal, habitat selection, and competition (see below and Chapter 2). These inferences have been challenged based on the consistency of observed patterns with other processes (e.g., Connor and Simberloff 1979).

Hanski (1992) adapted the incidence function for use in describing and modeling metapopulation dynamics. He noted that in an equilibrium system of many patches of similar size, the fraction of occupied patches at any point in time can be written as an explicit function of patch probabilities of extinction and colonization. He then postulated functional forms for the relationships between extinction probability and patch area and between colonization probability and patch isolation. If metapopulation dynamics can be described as a stationary Markov process, then parameters of the extinction and colonization relationships can be estimated using occupancy data from a single point in time (e.g., Hanski 1992, 1994a,b, 1998, 1999). However, the difficulties of inferring process from pattern have been noted. For example, based on analyses of year-to-year changes in occupancy of pikas (*Ochotona princeps*),

Clinchy *et al.* (2002:351) recommended that "simple patch occupancy surveys should not be considered as substitutes for detailed experimental tests of hypothesized population processes."

Use of occupancy data from multiple species to draw inferences about species interactions also has a long history in ecology. Some of the first statistical analyses adapted by ecologists were used to test the null hypothesis of independence of species occurrence using occupancy data for two species (Forbes 1907; Dice 1945; Cole 1949). Non-independent occupancy patterns of multiple species on islands have been interpreted as evidence of competition (e.g., MacArthur 1972; Diamond 1975a). For example, the "assembly rules" of Diamond (1975a) include specification of species combinations that cannot exist for reasons of interspecific competition and are based on empirical observations of species distributions on different islands. However, such inferences about process based on observed patterns have been sharply criticized. Critics argued that rejection of predictions of neutral models developed from distributional null hypotheses should precede any attempt to develop more complicated explanatory hypotheses for static species distribution patterns (e.g., Connor and Simberloff 1979, 1986; Simberloff and Connor 1981). Neutral models themselves were then criticized by proponents of the original competitive hypotheses (Diamond and Gilpin 1982; Gilpin and Diamond 1984), neutral model proponents responded (Connor and Simberloff 1984, 1986), and the entire issue of inference based on species distribution patterns was hotly debated (Strong *et al.* 1984). Such debate is not surprising, as strong disagreement is a natural consequence of weak inference, which brings us back to Pirsig's (1974) assertion about the ability to develop large numbers of plausible hypotheses to explain any given pattern.

Each of the three quantities listed as state variables of potential interest in population ecology and management (abundance, occupancy, and species richness) has been investigated with respect to its distribution over space at one point in time. Identification of spatial patterns has then led to inferences about the dynamic processes that produced these patterns. However, these inferences are always very weak, as many alternative hypotheses can be invoked to explain most ecological patterns (Fig. 1.3). Our conclusions about drawing inferences about process based on snapshots of spatial pattern are simple and straightforward. First, inferences about system dynamics should be based on estimates and observations of those dynamics, and of the vital rates that produce them, whenever possible. Second, when ecologists do try to draw inferences about dynamics based on observations of static pattern, we believe that such inferences are much more likely to be useful if the specification of model-based predictions from competing or single hypotheses precedes the investigation of pattern (e.g., see Karanth *et al.* 2004). Brown (1995:18) stated, "Macroecology seeks to discover, describe, and explain the patterns of varia-

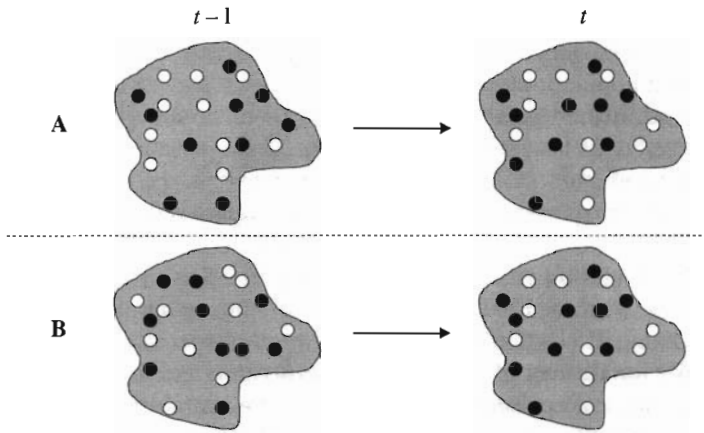


FIGURE 1.3 Illustration of how a pattern observed at time t may result from very different processes. Darkened circles represent occupied patches, and white circles represent unoccupied patches. The level of turnover between times $t-1$ and t is much greater in scenario B.

tion.” We recognize that such efforts can be useful, but we recommend that they be viewed as mechanisms for hypothesis generation rather than for inference and testing.

1.4. DISCUSSION

We began this chapter by asserting that many animal sampling programs, including many large-scale monitoring programs, have deficiencies resulting from failure to adequately consider three basic questions: Why do we want to sample animal populations and communities? What quantities do we want to estimate? How should we estimate the quantities of interest? In answer to the “why” question, we suggested that animal sampling and monitoring programs should not be viewed as stand-alone activities but as components of the larger processes of science or management. This recognition forces consideration of exactly how resulting data are to be used in these processes, and this consideration leads to program designs that maximize utility of data. The answer to the “what” question will depend heavily on the answer to the “why” question, and we noted that abundance, occupancy, and species richness are reasonable state variables for a variety of objectives. Of these potential state variables, occupancy has received the least methodological attention. Indeed, our objective in this book is to provide a set of inference methods useful for investigating this state variable.

The answer to the "how" question depends on the answers to the previous two questions, but also requires attention to two basic issues. When the entire area of interest cannot be surveyed, space must be sampled in a manner that is maximally useful to study objectives and that permits inference about the entire areas of interest. Because this problem of spatial sampling characterizes a wide variety of applications in statistical inference, it has been addressed well elsewhere. Spatial sampling will be touched on throughout the other chapters but will not be emphasized in this book. The second issue involves imperfect detection, the likelihood that surveys of animal populations and communities will not result in complete counts of all individuals or species present in surveyed locations. We present a general conceptual framework that relates the various count statistics obtained in studies of animal populations and communities to the true state variables of interest. Until very recently, uses of occupancy as a state variable in animal ecology have simply not dealt with the issue that failure to find evidence of a species at a location does not necessarily mean that the species does not occupy the area. The suite of models, methods, and estimators that we develop in this book is basically designed to remedy this situation and permit inferences about occupancy that deal adequately with detection probabilities less than 1.

Because most of this book focuses on parameter estimation, we returned to the "why" question and the manner in which estimates of state variables are to be used in the conduct of science and management. We briefly addressed the general question of drawing inferences about system dynamics and causal factors responsible for these dynamics. Approaches to the generation of system dynamics for the purpose of conducting science include true manipulative experimentation, constrained design manipulative studies, and observational studies using retrospective analyses. Strength of inference is greatest for manipulative experiments and weakest for retrospective analyses of time series data from observational studies. Within the category of observational studies, those used to provide confrontations with predictions of *a priori* hypotheses are much more likely to be useful than those used solely to develop *a posteriori* hypotheses.

Finally, we noted that investigators sometimes try to draw inferences about system dynamics based on static looks at spatial patterns of state variables at single points in time. Such efforts to draw inferences about process based on observation of pattern have been used with all three state variables, abundance, occupancy, and species richness. However, such efforts suffer from the ability to develop many process-based hypotheses to explain the generation of any particular pattern. Previous uses of occupancy in animal ecology have relied heavily on inferences based on statics and pattern, and we note the shortcomings of this approach. In particular, we do not view the primary purpose of this book to be provision of methods for obtaining better estimates of static

occupancy patterns for use in drawing inferences about dynamic processes. Instead, we also provide methods for drawing inferences about occupancy dynamics based on data covering multiple time periods.

In Chapter 2, we consider both historical and proposed uses of occupancy as a state variable in studies of animal populations and communities. With each use, we emphasize the need to deal adequately with detection probabilities. Chapter 3 provides an elementary overview of the statistical concepts used throughout the book. Chapters 4–6 then deal with single-species occupancy studies in which multiple locations are surveyed during a single time period or “season.” The parameter of interest is the probability of occupancy of a site, given that occupancy cannot always be detected. Chapter 4 presents a basic model and estimators, and includes discussion of issues such as missing data, covariate modeling, goodness-of-fit tests, and consequences of violations of model assumptions. Chapter 5 focuses on the common assumption violation of heterogeneous detection probabilities. We present mixture models that allow for variation in detection probabilities that cannot be attributed to measured covariates. Animal abundance at a site is identified as one important source of heterogeneity in detection probability. The relationship between abundance and detection probability provides a basis for estimating abundance from occupancy survey data and for estimating occupancy itself in a manner that deals with this heterogeneity. Chapter 6 deals with the important topic of study design for single-season occupancy studies for a single species.

Chapter 7 then focuses on occupancy studies conducted over multiple years or seasons for the purpose of drawing inference about occupancy dynamics for a single species. Rate of change in occupancy over time is identified as a parameter of interest, and the vital rates responsible for such change, local probabilities of extinction and colonization, are also incorporated into estimation models. Estimation, covariate modeling, assumption violations, and study design are all considered. Chapter 8 shifts emphasis to multiple species and begins with inference procedures for two species in a single year or season. Methods permit inference about dependence in probabilities of occupancy given detection probabilities that are less than 1 and that may themselves exhibit dependence on presence or detection of the other species. These methods are then extended to multiple seasons, where the emphasis shifts to possible dependence of extinction and colonization probabilities of one species on the presence of the other species.

Chapter 9 includes some suggestions about potential uses of occupancy modeling in community-level studies. One approach exploits the analogy between the different species in a local species pool and the “locations” of typical occupancy studies in order to directly estimate the fraction of the pool that is present. If multiple time periods are available, then local extinction probability and turnover can be estimated directly using this basic approach

as well. The other potential use in community-level investigations involves the synthetic treatment of Dorazio and Royle (2005) in which species richness is estimated, as is the equitability component of many diversity metrics (based on relative occupancy). This work also provides a conceptual framework for considering species-area relationships. The concluding Chapter 10 contains several ideas for future work, as well as discussion linking spatial occupancy and abundance in a common framework that facilitates consideration of the relationship between range size and abundance.