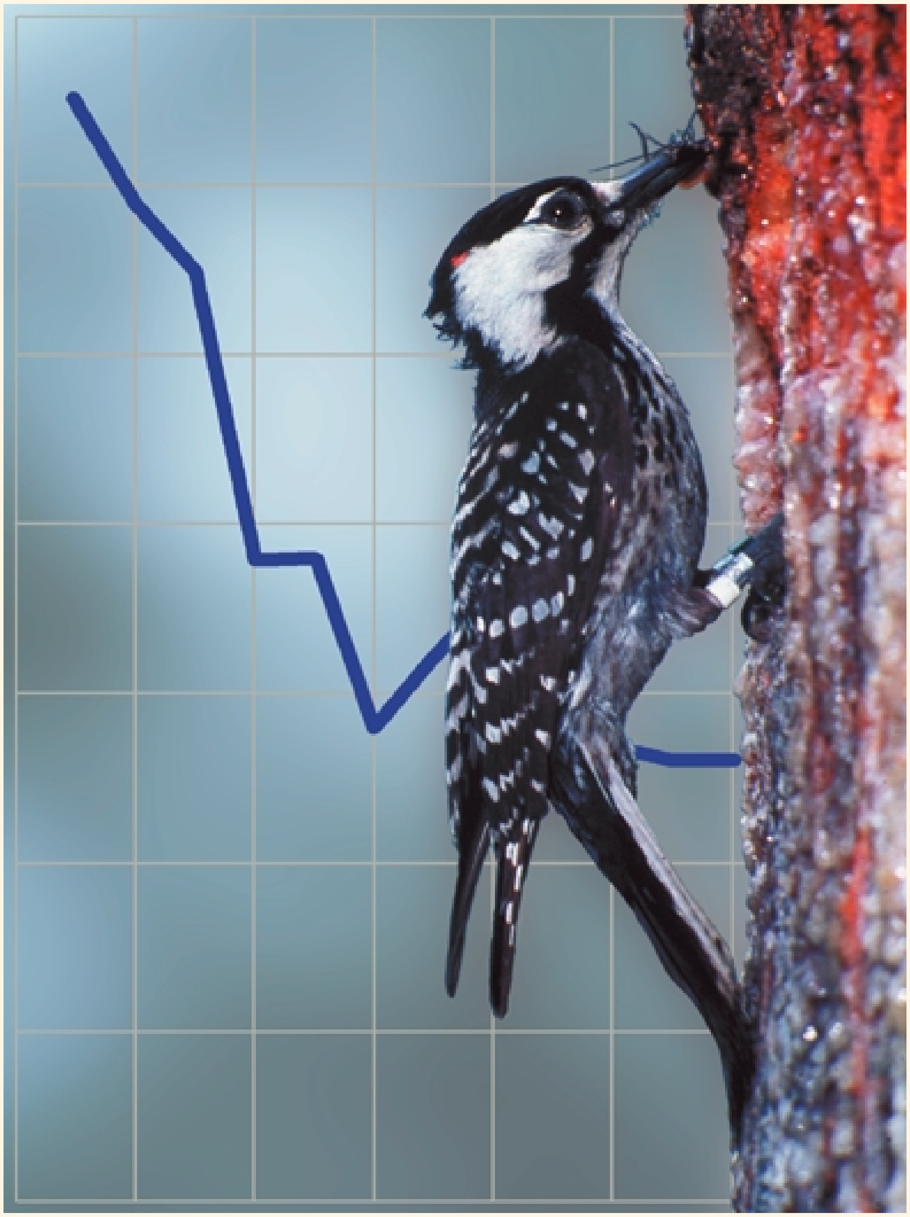
A Practical Handbook for Population Viability Analysis



William Morris, Daniel Doak, Martha Groom, Peter Kareiva,
John Fieberg, Leah Gerber, Peter Murphy, and Diane Thomson

The Nature Conservancy
Saving the Last Great Places

A Practical Handbook for Population Viability Analysis

William Morris[1,2], Daniel Doak[3,4], Martha Groom[5], Peter Kareiva[5],
John Fieberg[6], Leah Gerber[7], Peter Murphy[2], and Diane Thomson[4].

Evolution, Ecology and Organismal Biology Group[1] and Dept. of Zoology[2]
Duke University, Box 90235, Durham, NC 27708-0325

Biology Board[3] and Environmental Studies Board[4]
University of California, Santa Cruz, CA 95064

Dept. of Zoology[5]
Box 351800, University of Washington, Seattle, WA 98195-1800

Biomathematics Program[6]
North Carolina State University, Raleigh, NC 27695

Dept. of Ecology, Evolution, and Marine Biology[7]
University of California Santa Barbara, Santa Barbara, CA 93106

April 1999

The Nature Conservancy®
Saving the Last Great Places

A Practical Handbook for Population Viability Analysis

ISBN: 0-9624590-4-6

On the cover: photograph copyright S. Maka/VIREO.

ACKNOWLEDGMENTS

The authors thank Craig Groves for helping to organize the workshop out of which this handbook grew, the staff of the National Center for Ecological Analysis and Synthesis in Santa Barbara, CA, for logistical assistance, and Nicole Rousmaniere for designing the final version. The workshop was generously supported by grants from the Moriah Fund and the David B. Smith Foundation. We also thank all the workshop participants for providing data sets, and for the enthusiasm and dedication with which they approach the task of preserving biological diversity. Finally, we thank Christopher Clampitt, Doria Gordon, Craig Groves, Elizabeth Crone, Larry Master, Wayne Ostlie, Lance Peacock, Ana Ruesink, Cheryl Shultz, Steve Sutherland, Jim Thorne, Bob Unnasch, Douglas Zollner, and students in W. Morris's Fall 1998 graduate seminar at Duke University for reading and commenting on drafts of the handbook.

TABLE OF CONTENTS

CHAPTER ONE

What is Population Viability Analysis, and Why This Handbook?

The 1997 document *Conservation by Design: A Framework for Mission Success* states that the conservation goal of The Nature Conservancy is "the long-term survival of all viable native species and community types through the design and conservation of portfolios of sites within ecoregions." In an ideal world, conservation organizations like TNC would seek to preserve every location that harbors a rare, threatened, or endangered species. But in the real world, financial considerations make this strategy impossible, especially given the number of species whose status is already cause for concern. Thus it is an inescapable fact that for all but the rarest of species, TNC will need to focus on preserving only a subset of the known populations, and upon this choice will rest the success of the entire mission. To make this choice, Conservancy staff require the means to find answers, at the very least qualitative and conditional ones, to two critical questions. First, what is the likelihood that a known population of a species of conservation concern will persist for a given amount of time? Second, how many populations must be preserved to achieve a reasonable chance that at least one of them will avoid extinction for a specified period of time? The goal of this handbook is to introduce practical methods for seeking *quantitative* answers to these two questions, methods that can provide some guidance in the absence of highly detailed information that is unlikely to be available for most rare species. The use of such methods has come to be known as population viability analysis (PVA).

Broadly defined, the term "population viability analysis" refers to the use of quantitative methods to predict the likely future status of a population or collection of populations of conservation concern. Although the acronym PVA is now commonly used as though it signified a single method or analytical tool, in fact PVAs range widely both in methods and applications. Among the most influential PVAs to date is one of the original analyses of Northern Spotted Owl data (Lande 1988). This work relied upon quite simple demographic data, and its main points were that logging could result in owl population collapse and that the data available at that time were insufficient to determine how much forest was needed for the owl population to persist. This second point is important, as it emphasizes that PVAs can be highly useful even when data are sparse. Another influential PVA (Crouse et al. 1987) used a more complex size-structured model to assess the status of loggerhead sea turtles and to ask whether protecting nestlings on beaches or preventing the death of older turtles in fishing trawls would have a greater effect on enhancing population recovery. This single PVA played a critical role in supporting legislation to reduce fishing mortality of turtles (Crowder et al. 1994). More recent PVAs have involved yet more complex spatial models, for example of individual Lead-beater's Possums (Lindemeyer and Possingham 1994). Furthermore, while most PVAs are ultimately concerned with assessing extinction risks, they are often motivated by the need to address specific problems, for example sustainable traditional use levels of forest palms (Ratsirarson et al. 1996), the risks posed by different poaching techniques to wild ginseng populations (Nantel et al. 1996), or loss of movement corridors (Beier 1993). The uniting theme of PVAs is simply that they all are *quantitative* efforts to assess population health and the factors influencing it.

This handbook grew out of a workshop held at the National Center for Ecological Analysis and Synthesis in Santa Barbara, CA, in February, 1998, in which ecologists from four universities (the authors of this handbook) and TNC practitioners came together to explore how quantitative methods from the field of population biology might be used to inform TNC decision making. Prior to the workshop, TNC participants were asked to supply data sets that exemplify the types of information that TNC or Heritage employees and volunteers would collect about species of conservation concern. In Chapter 2, we classify the data sets into 3 categories, which we then use as a starting point to identify a few quantitative methods that we describe in detail in the subsequent chapters. In Chapters 3 and 4, we review methods for assessing viability of single populations when the data represent census counts or demographic information about individuals, respectively. In Chapter 5, we address the question of how to assess regional viability when a species is distributed across multiple populations of varying size and "quality".

We begin with two important caveats. First, this handbook does not attempt to review the field of population viability analysis as a whole, but instead focuses on the subset of all available PVA methods that we deemed, through our interactions with TNC biologists, to be the most practical given the types of data typically available. Second, population viability analyses, because they are typically based upon limited data, MUST be viewed as tentative assessments of current population risk based upon what we now know rather than as iron-clad predictions of population fate. Thus, as we will argue repeatedly below, we should not put much faith in the exact predictions of a single viability analysis (e.g. that a certain population will have a 50% chance of persisting for 100 years). Rather, a better use of PVA in a world of uncertainty is to gain insight into the *range* of likely fates of a single population based upon 2 or more different analyses (if possible), or the *relative* viability of 2 or more populations to which the same type of analysis has been applied. When data on a particular species are truly scarce, performing a PVA may do more harm than good. In such cases, basing conservation decisions on other methods (e.g. the known presence/absence of a species at a suite of sites, or its known habitat requirements) makes far better sense. We discuss the question of when NOT to perform a PVA in greater detail in the final chapter of this handbook. Thus, while we view PVA as a potentially useful tool, we do not see it as a panacea.

While data scarcity is a chronic problem facing all decision making in conservation, we should also recognize that it is often feasible to collect additional data to better inform viability assessments. Indeed, TNC and Heritage personnel are constantly collecting new information in the course of monitoring sites for rare and threatened species. Simple counts of the number of individuals of a certain species at a site over a number of years are often made with other purposes in mind, but they can also serve as grist for a population viability analysis. We hope that awareness of the possible use of monitoring data in PVA will lead TNC/Heritage biologists to consider ways that their monitoring schemes can maximize the usefulness of monitoring data for future viability assessments, without entailing costly changes in existing monitoring protocols. In Chapter 6 of this handbook, we make easy-to-follow recommendations for how the design of monitoring strategies can best meet the data requirements of PVA.

Before proceeding to the consideration of typical TNC data sets, we say a brief word about the structure of this handbook. To illustrate the application of each method, we provide step-by-step examples, usually using one of the TNC/Heritage data sets. These worked examples are featured in Key Boxes that are set aside from the background text of the handbook. We also use Key Boxes to highlight key assumptions or caveats about each of the methods we review. While the Key Boxes emphasize the methods we have found to be the most practical, it is also important to point out that more complex population viability analyses may be possible in cases in which more data are available. Because we do not have the space to thoroughly review these more complex (and therefore less frequently useful) analyses in a handbook of this length, we have also included Optional Boxes that give a brief overview of other methods and provide references that will allow the interested reader to learn more about them.

Finally, we make one further point of clarification. In this handbook, we aim to quantify the likelihood of persistence of a population (that is the collection of individuals of a single species living in a prescribed area) or a set of populations over a specified time period. We use the terms “population” and “element occurrence” (or “EO”) interchangeably. Thus we use “EO” to refer to a population of a single species, which we realize is a more restricted usage of the term than the one used by TNC/Heritage biologists, which defines elements as “viable native species AND communities” (see *Conservation by Design*). We emphasize that the methods we review are NOT intended to be used to determine the long-term viability of communities. However, we note that PVAs targeted at populations of dominant or characteristic species in a particular community type may serve as useful tools for evaluating the viability of community occurrences.

CHAPTER TWO

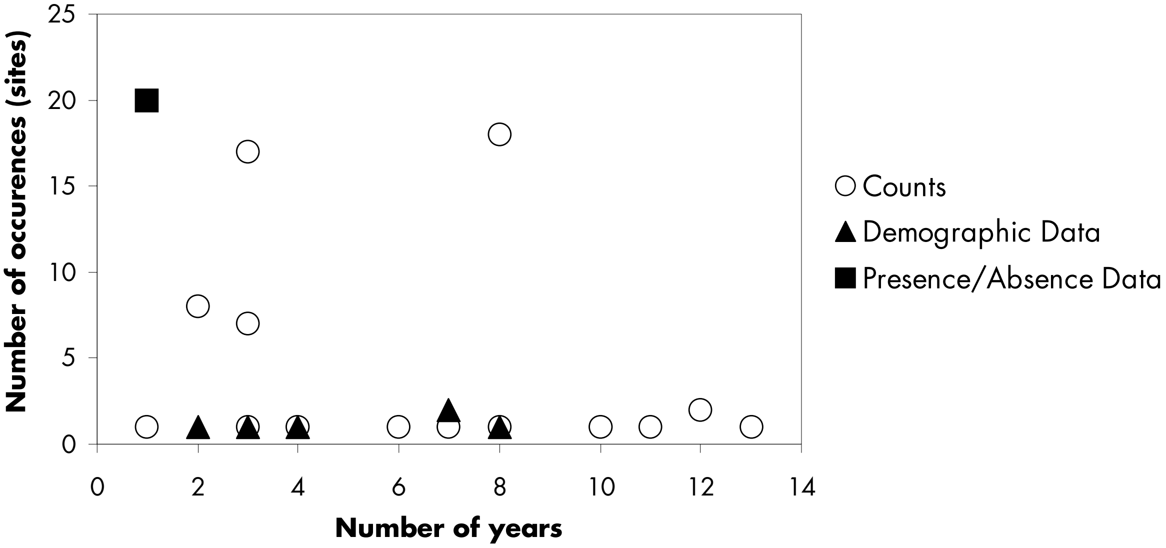
Letting the Data Determine an Appropriate Method for Population Viability Analysis

The first rule of population viability analysis is: "let the data tell you which analysis to perform". While population biologists have developed a vast array of complex and mathematically sophisticated population models, it is our view that when data are limited (as they almost always will be when we are dealing with the rare, seldom-studied species that are the typical concern of conservation planners) the benefits of using complex models to perform population viability analyses will often be illusory. That is, while more complex models may promise to yield more accurate estimates of population viability because they include more biological detail (such as migration among semi-isolated populations, the effects of spatial arrangement of habitat patches, and the nuances of genetic processes such as gene flow and genetic drift), this gain in accuracy will be undermined if the use of a more complex model requires us to "guess" at critical components about which we have no data. Instead, our philosophy is that the choice of models and methods in PVA should be determined primarily by the type of data that are available, and not the other way around.

With this philosophy in mind, and to get an idea of the kinds of data that TNC biologists will typically have at their disposal to perform population viability analyses, we asked workshop participants to provide us with data sets

FIGURE 2.1

Characteristics of 20 data sets on rare species considered in the PVA Workshop (see Table 2.1 for information on the species included)



that had been collected in conjunction with TNC field offices. We received 26 data sets, which included information about 25 species of conservation concern. We classified these data sets according to the type of data, the number of locations, and the number of years in which data were collected. By "type of data", we mean whether the persons who collected the data recorded the PRESENCE OR ABSENCE of the species at a location, COUNTS of individuals in one or more life stages, or DEMOGRAPHIC information about individual organisms (that is, whether each individual survived from one census to the next and if so, its size at each census and the number of offspring it produced in the time interval between the censuses).

This survey of data sets highlights four patterns (Table 2.1, Fig. 2.1). First, count data is the most common type of information in this sample of TNC data sets. Second, relatively long duration studies tended to focus on only a single site, while multi-site studies typically involved

TABLE 2.1 Data sets contributed to the TNC PVA workshop

Species	Type of Data	No. of sites	No. of years
Shale barren rockcress, *Arabis serotina*	Counts	1	6
Shale barren rockcress, *Arabis serotina*	Counts	17	3
Dwarf trillium, *Trillium pusillum*	Counts	1	4
Eriocaulon kornickianum	Counts	1	3
Mesa Verde cactus, *Sclerocactus mesae-verdae*	Counts	1	10
Mancos Milkvetch, *Astragalus humillimus*	Counts	1	8
Knowlton's cactus, *Pediocactus knowltonii*	Counts	1	11
Lesser prairie chicken, *Tympanuchus pallidicinctus*	Counts	1	13
Seabeach pinweed, *Amaranthus pumilus*	Counts	18	8
Golden Alexanders, *Zizia aptera*	Counts	1	7
Oenothera organensis	Counts	8	2
Arizona stream fish (7 species)	Counts	1	1
Red-cockaded woodpecker, *Picoides borealis*	Counts	2	12
Bog turtle, *Chlemmys muhlenbergii*	Counts	7	3
Kuenzler hedgehog cactus, *Echinocereus fendleri* var. *kuenzleri*	Demographic	1	2
Ornate box turtle, *Terrapene ornata*	Demographic	1	8
Larimer aletes, *Aletes humilis*	Demographic	2	7
Mead's milkweed, *Asclepias meadii*	Demographic	1	4
Trollius laxus	Demographic	1	3
Cave salamander, *Gyrinophilus palleucus*	Presence/Absence	20	1

only one or a few censuses, which is not surprising given the limited resources available to monitor populations of conservation concern. Only one of the 26 data sets included information from more than 8 sites in more than 3 years. Third, demographic data sets, because they are more difficult to collect, tend to include fewer years on average than do count data. Fourth, the data set that included the most sites comprised presence/absence data. The single example of presence/absence data here surely underestimates the true frequency of such data sets in Heritage data bases. While information about presence/absence of a species is critically important in identifying high-priority sites for acquisition or preservation (Church et al. 1996, Pressey et al. 1997), such data sets lack the population-level details required for a PVA, and we do not address them further in this handbook.

To the extent that this informal sample gives a rough idea of the types of data accessible to TNC biologists, it suggests three themes about how PVA might best serve TNC decision making processes. First, our informal survey of data sets shows that counts of the number of individuals in one or more populations over multiple years will be the most common information upon which population viability analyses will need to be based, but that in some cases (most likely for umbrella or indicator species, and those for which particular reserves have been especially established) more detailed analyses based upon demographic information will be feasible. Second, while information will sometimes be available to perform PVAs on multiple local populations, most decisions about the number of occurrences needed to safeguard a species will require extrapolation from information collected at only one or a few populations at best. Third, the kinds of information that are *missing* from these data sets is also noteworthy. None of them include any information about genetic processes or, in the case of data sets that include multiple occurrences, about rates of dispersal of individuals among populations. Thus we conclude that more complex models that require this information

TABLE 2.2 A classification of PVA methods reviewed in this handbook

Number of populations or EOs included in the analysis:	Type of data collected:	Minimum number of years of data per population or EO:	PVA method:	Where to look in this handbook:
One	Counts	10 (preferably more)	Count-based extinction analysis	Chapter 3
One	Demographic information	2 or more	Projection matrix models	Chapter 4
More Than One	Counts	10 (preferably more) for at least one of the populations	Multi-site extinction analysis	Chapter 5

will not be justified in most cases. We reiterate these themes in the following chapters.

Thus Fig. 2.1 suggests three general classes of data sets that provide information that can be used to perform a PVA:

- Counts of individuals in a single population obtained from censuses performed over multiple years;
- Detailed demographic information on individuals collected over 3 or more years (typically at only 1 or 2 sites); and
- Counts from multiple populations, including a multi-year census from at least one of those populations.

Each of these classes require somewhat different methods for population viability analysis. Fortunately, population biologists have developed methods to deal with each of these situations. Table 2.2 summarizes the data requirements for PVA based upon each of these three classes of data, and points to where each type of PVA is presented in this handbook.

CHAPTER THREE

Using Census Counts Over Several Years to Assess Population Viability

As we saw in Chapter 2 (Fig. 2.1), the type of population-level data that is most likely to be available to conservation planners and managers is count data, in which the number of individuals in either an entire population or a subset of the population is censused over multiple (not necessarily consecutive) years. Such data are relatively easy to collect, particularly in comparison with more detailed demographic information on individual organisms (see Chapter 4). In this chapter, we review an easy-to-use method for performing PVA using count data. The method's simplicity makes it applicable to a wide variety of data sets. However, several important assumptions underlie the method, and we discuss how violations of these assumptions would introduce error into our estimates of population viability. We also point to other, similar methods that can be employed in the face of biological complexities that make the simpler method less appropriate.

In a typical sequence of counts from a population, the numbers do not increase or decrease smoothly over time, but instead show considerable variation around long-term trends (see examples in Fig. 3.1). One factor that is likely to be an important contributor to these fluctuations in abundance is variation in the environment, which causes the rates of birth and death in the population to vary from year to year. The potential sources of environmentally-driven variation are too numerous to list fully here, but they include inter-annual variation in factors such as rainfall, temperature, and duration of the growing season. Most populations will be affected by such variation, either directly or indirectly through its effects on interacting species (e.g. prey, predators, competitors, diseases, etc.). When we use a sequence of censuses to estimate measures of population viability, we must account for the pervasive effect of environmental variation that can be seen in most count data. To see how this is done, we first give a brief overview of population dynamics in a random environment, and then return to the question of how count data can be used to assess population viability.

Population dynamics in a random environment

Perhaps the simplest conceptual model of population growth is the equation

Equation 3.1

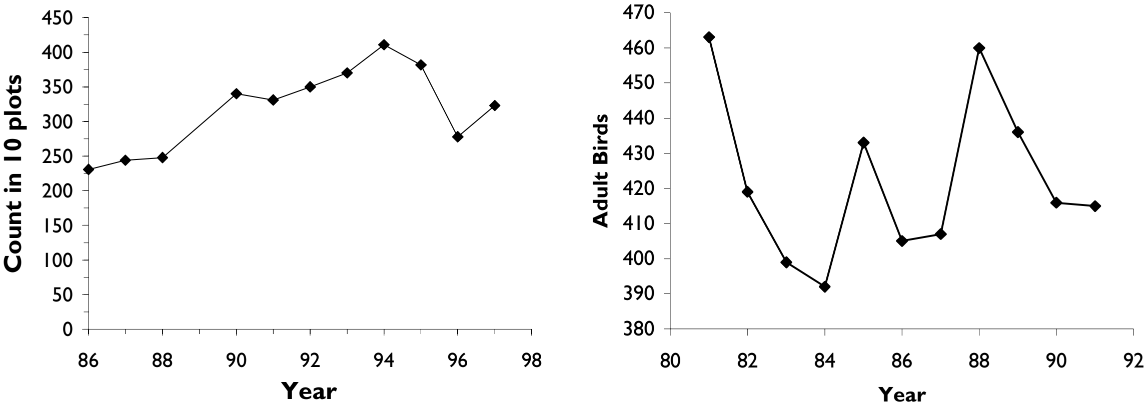
$$N(t+1) = \lambda N(t),$$

where N(t) is the number of individuals in the population in year t, and λ is the population growth rate, or the amount by which the population multiplies each year (the Greek symbol "lambda" is used here by tradition). If there is no variation in the environment from year to year, then the population growth rate λ is a constant, and only three qualitative types of population growth are possible (Fig. 3.2A): if λ is greater

(continued on page 11)

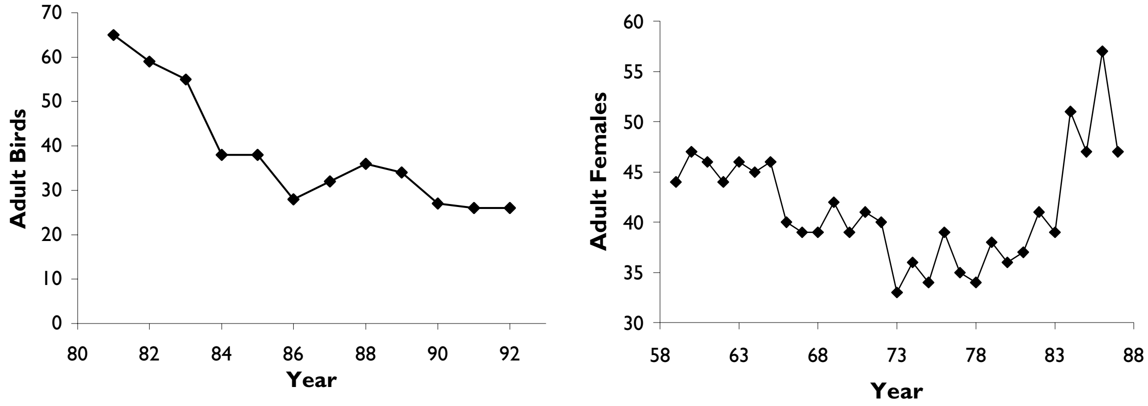
FIGURE 3.1

Examples of count data: A) Knowlton's cactus in New Mexico (data provided by R. Sivinski); B & C) Red-cockaded woodpecker in North Carolina and central Florida (data provided by J. Hardesty); D) Grizzly bears in the Greater Yellowstone Ecosystem (reproduced from Dennis et al. 1991); E) Lesser prairie chicken in Caprock Wildlife Management Area, NM (data provided by K. Johnson).



A. Knowlton's cactus

B. North Carolina Red-cockaded Woodpecker



C. Central Florida Red-cockaded Woodpecker

D. Yellowstone Grizzly Bear

E. Lesser Prairie Chicken

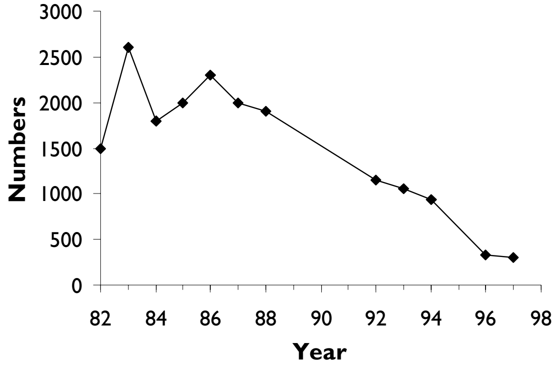
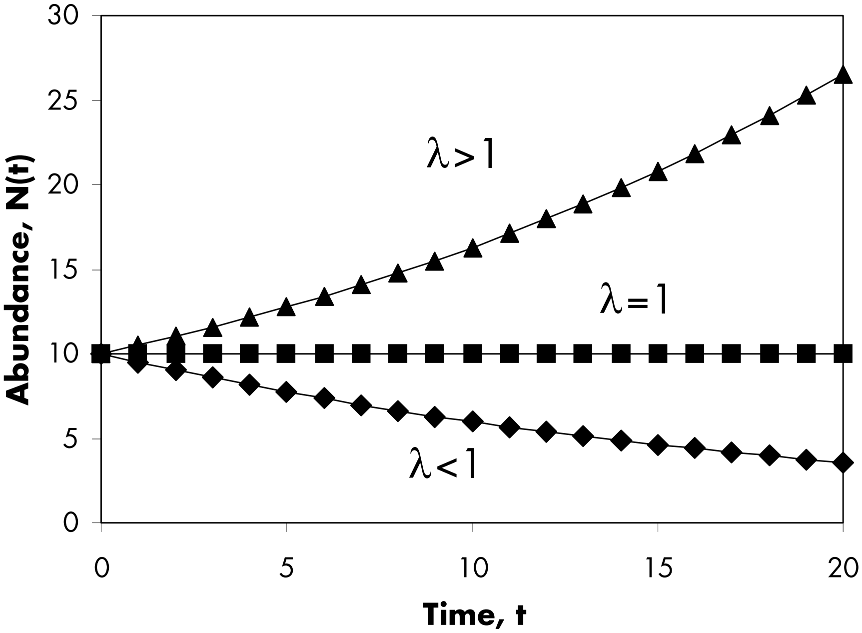


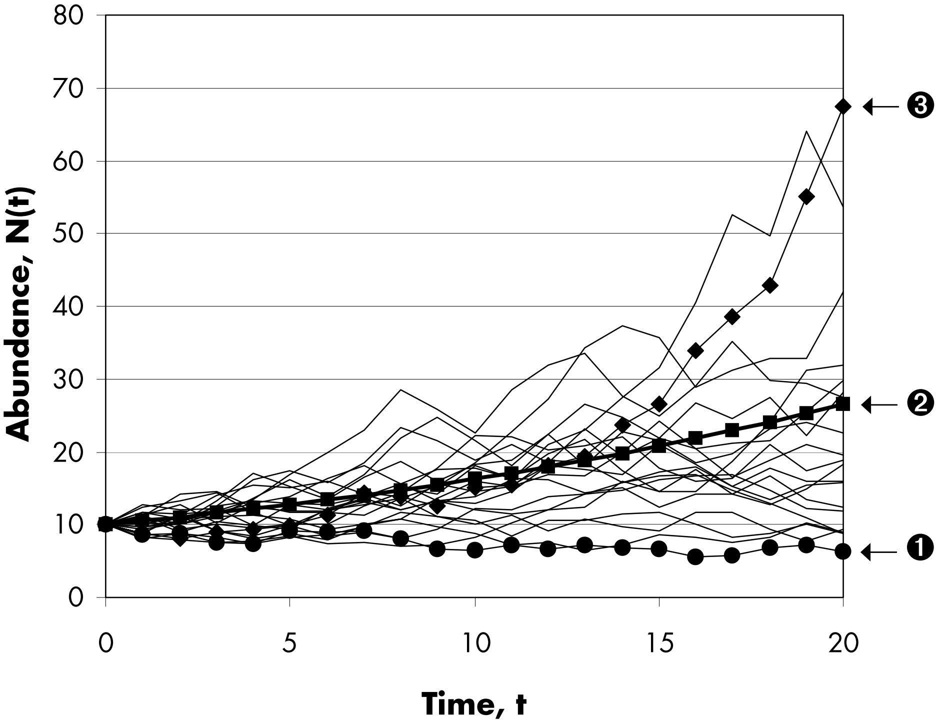
FIGURE 3.2

Population growth described by a geometric growth model $N(t+1) = \lambda N(t)$ in (A) a constant or (B) a stochastically-varying environment

A. No Environmentally-Driven Variability



B. Populations in a Stochastic Environment



than one, the population grows geometrically; if λ is less than one, the population declines geometrically to extinction; and if λ exactly equals one, the population neither increases nor declines, but remains at its initial size in all subsequent years. But when variation in the environment causes survival and reproduction to vary from year to year, the population growth rate λ must also be viewed as varying over some range of values. Moreover, if the environmental fluctuations driving changes in population growth include an element of unpredictability (as factors such as rainfall and temperature are likely to do), then we must face the fact that we cannot predict with certainty what the exact sequence of future population growth rates will be. As a consequence, even if we know the current population size and both the average value and the degree of variation in the population growth rate λ, the best we can do is to make probabilistic statements about the number of individuals the population will include at some time in the future. To illustrate, Fig. 3.2B shows a hypothetical population governed by the same equation we saw above, but in which the value of the population growth rate λ in each year was generated on a computer so as to vary randomly around an average value. Each line in the figure can be viewed as a separate "realization" of population growth, or a possible trajectory the population might follow given a certain average value and degree of variability in λ.

Fig. 3.2B illustrates three important points about population growth in a random or "stochastic" environment. First, the possible realizations of population growth diverge over time, so that the farther into the future predictions about likely population size are made, the less precise they become. Second, the realizations do not follow very well the predicted trajectory based upon the average population growth rate. In particular, even though the average λ in this case would predict that the population should increase at a slow rate, a few realizations explode over the 20 years illustrated, while others decline (thus extinction is possible even though the average of the possible population trajectories increases). Third, the endpoints of the 20 realizations shown are highly skewed, with a few trajectories (such as ❸) winding up much higher than the average λ would suggest, but most (such as ❶) ending below the average. This skew is due in part to the multiplicative nature of population growth. Because the size of the population after 20 years depends on the product of the population growth rate in each of those years, a long string of chance "good" years (i.e. those with high rates of population growth) would carry the population to a very high level of abundance, while "bad" years tend to confine the population to the restricted zone between the average and zero abundance.

Skewness in the distribution of the likely future size of a population is a general feature of a wide variety of models of population growth in a stochastic environment. In fact, we can make the even more precise statement that for many such models, the endpoints of multiple independent realizations of population growth will lie approximately along a particular skewed probability distribution known as the log-normal, or equivalently that the natural log of population size will be normally

(continued on page 14)

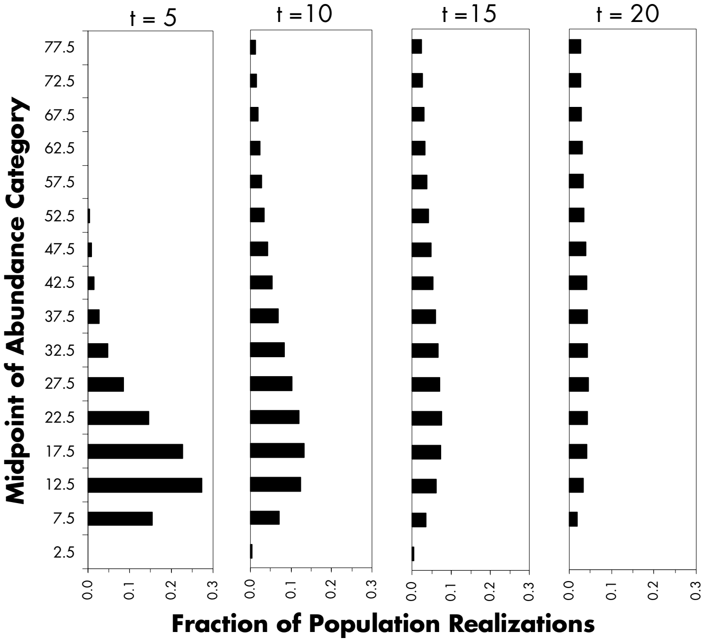
BOX 3.1 (Optional): Theoretical Underpinnings

Probability Distributions Describing Population Size in a Randomly Varying Environment

We saw in Fig. 3.2B that possible realizations of population growth in a stochastic environment become skewed, with a few high-abundance realizations outweighed by a large number of low-abundance realizations. If we were to simulate a large number of such realizations and then divide them into abundance "bins" at several different times, we would get the following sequence of histograms (Fig. 3.3), which clearly shows the skewness in population abundance. Note that with the passage of time, both the average value and the degree of spread in these histograms increases. If we make the size of the "bins" smaller and smaller, the histograms in Fig. 3.3 will come to resemble the skewed probability distribution known as the log-normal. If abundance has a log-normal distribution, then the natural log of abundance will have a normal distribution, whose mean and variance will also change over time (Fig. 3.4). Measures of population viability are derived directly from this shifting normal distribution. For example, the probability that the population lies below the threshold at a certain time is simply the area under the normal distribution below the threshold (shown in red in Fig. 3.4). The time until the threshold is first attained is also determined by the normal distribution (see Box 3.3). The shifting normal distribution is completely described by two parameters. The parameter μ determines how quickly the mean in-creases (if μ is greater than zero, Fig. 3.4A) or decreases (if μ is less than zero, Fig. 3.4B). The second parameter, σ^2, determines how quickly the variance in the normal distribution increases. Clearly if μ is less than zero, extinction is certain, but even if μ is positive (i.e. the population is expected to grow on average), there will be some chance that the population falls below the threshold, particularly if the variance increases rapidly (i.e. if σ^2 is large). Thus to measure a population's risk of extinction, we must know the values of both μ and σ^2.

FIGURE 3.3

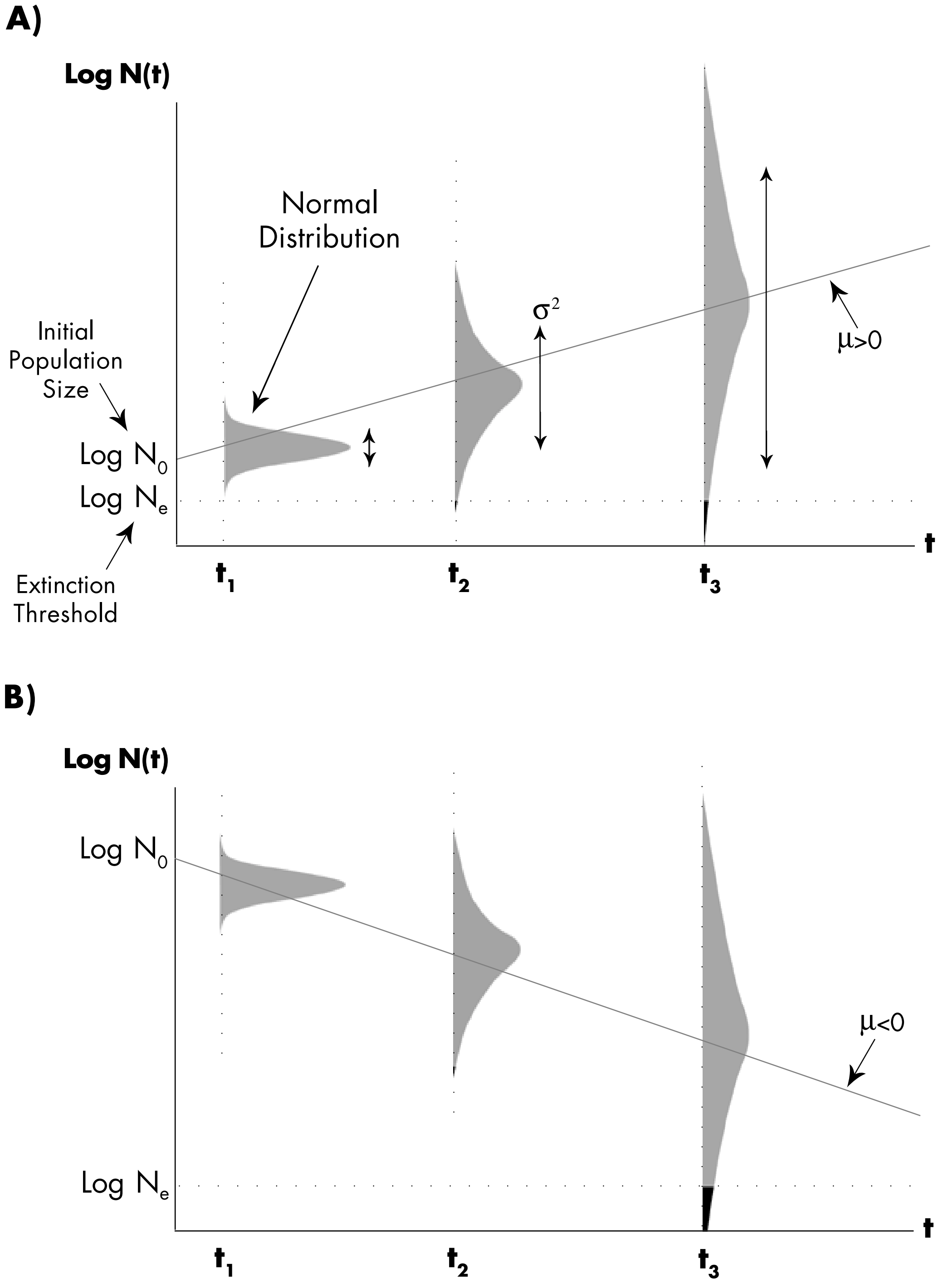
Log-normal distributions of abundance in a population growing exponentially in a stochastic environment



BOX 3.1 *(continued)*

FIGURE 3.4

Normal distributions of the log of population size, when the parameter μ is A) positive or B) negative. The size of the "black zone" gives the probability of extinction.



distributed (see Box 3.1). This important result means that we can use the normal distribution (whose properties are well understood, as it underlies much of modern statistical theory) to calculate measures of viability, such as the probability that the population will be above some threshold size a given number of years into the future, or the likely number of years before the population first hits the threshold. But before we can calculate these measures, we must first estimate two parameters that describe how the normal distribution of the log of population size will change over time: μ, which governs change in the mean of the normal distribution, and σ^2, which governs how quickly the normal distribution's variance will increase over time (Box 3.1). Both of these parameters will have important effects on measures of population viability, so we require a method to estimate their values using count data.

Using count data to estimate population parameters

Brian Dennis and colleagues (Dennis et al. 1991) have proposed a simple method for estimating μ and σ^2 from a series of population censuses. The method involves two easy steps: 1) calculating simple transformations of the counts and of the years in which counts were taken, and 2) performing a linear regression (Box 3.2). The results of the regression yield estimates of μ and σ^2.

Measures of viability based upon μ and σ^2

Once we have estimated the parameters μ and σ^2 from count data, we can calculate several measures of the viability of the population from which the counts were obtained (Box 3.3). One is the average value of the population growth rate, $\bar{\lambda}$. This value indicates whether the *average* of the possible population trajectories will tend to increase ($\bar{\lambda}>1$), decrease ($\bar{\lambda}<1$), or remain the same ($\bar{\lambda}=1$) over one census interval (thus $\bar{\lambda}$ describes ❷ in Fig. 3.2B). Keep in mind that some, or even most, realizations of population growth may decline even if their average increases (see Fig. 3.2B). The confidence interval for $\bar{\lambda}$ is also informative, because only if the entire confidence interval lies above or below the value 1 can we say (for example with 95% confidence) that the average of population trajectories will increase or decrease, respectively.

Because the average value of the population growth rate doesn't do a good job of predicting what most population realizations will do, two other viability measures, the mean time to extinction and the probability that extinction has occurred by a certain future time, may be calculated. These require us to specify an initial population size (typically the most recent count) and an "extinction" threshold. The "extinction" threshold need not be set at zero abundance. For a non-hermaphroditic species, we may wish to set the threshold at 1, at which point the population would be *effectively* extinct. It may be reasonable to set the threshold at even higher levels, such as the abundance at which genetic drift or demographic stochasticity reach a predetermined level of importance, or the lowest level of abundance at which it remains feasible to attempt intervention to prevent further decline. Once we arrive at an

(continued on page 17)

BOX 3.2 (Key): Methods of Analysis

Estimating Useful Parameters from a Series of Population Censuses

Dennis et al. (1991) outlined the following simple method to estimate the parameters μ and σ^2 from a series of counts from a population:

- First, choose pairs of counts N(i) and N(j) from adjacent censuses i and j performed in years t(i) and t(j).
- Second, calculate the transformed variables $x = \sqrt{t(j)-t(i)}$ and $y = \ln(N(j)/N(i))/\sqrt{t(j)-t(i)} = \ln(N(j)/N(i))/x$ for each pair.
- Third, use all the resulting pairs of x and y to perform a linear regression of y on x, *forcing the regression line to have a y-intercept of zero* (Fig. 3.5).

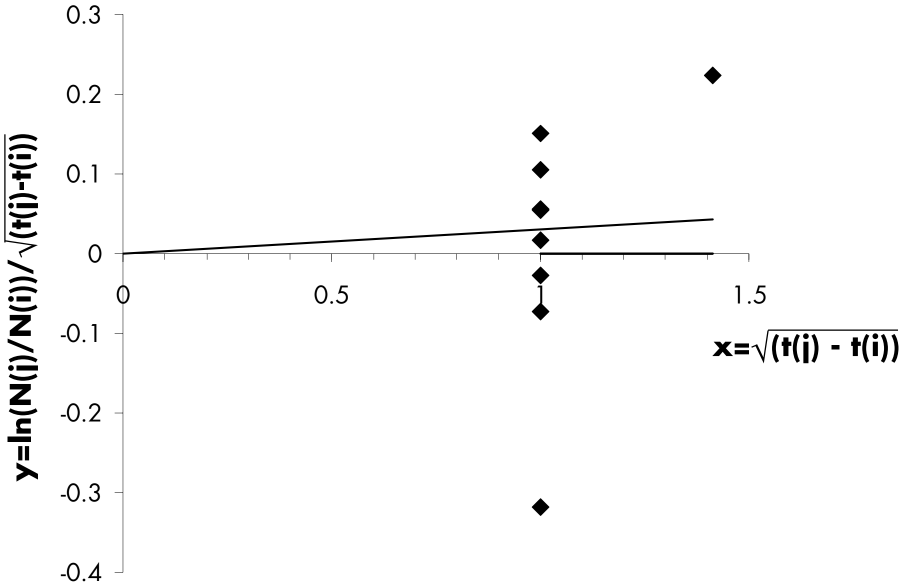
The slope of the resulting regression line is an estimate of the parameter μ. The mean squared residual, which can be read from the Analysis of Variance table associated with the regression, is an estimate of the parameter σ^2.

As an illustration of the method, the following data were collected in a monitoring study of the federally-listed Knowlton's cactus (*Pediocactus knowltonii*) made over 11 years by R.L. Sivinski at its only known location in San Juan County, New Mexico (see Fig. 3.1A). The data are summed counts of the number of individual plants in ten permanent 10 square meter plots (an eleventh plot was omitted because all the individuals were removed in 1996 by cactus poachers!). The transformed variables x and y are also shown. Note that an advantage of the Dennis et al. method is that it does not require that censuses be performed year after year without fail. For example, monitoring of

(continued on page 16)

FIGURE 3.5

The regression of y on x for the Knowlton's cactus data. The slope of the regression line is an estimate of μ, and the variance of the points around the line is an estimate of σ^2



BOX 3.2 *(continued)*

Knowlton's cactus was incomplete in 1989, and omitting that census results in adjacent counts in 1988 and 1990 that are 2 years apart. We simply use an x value of $\sqrt{2}$ = 1.414 and a y value of ln(340/248)/ $\sqrt{2}$ for that pair of counts in the regression.

Year	$x=\sqrt{t(j)-t(i)}$	**Count**	$y=\ln(N(j)/N(i)/x$
1986		231	
1987	$\sqrt{(1987-1986)} = 1$	244	ln(244/231)/1 = 0.054751
1988	1	248	0.016261
1990	1.414	340	0.223104
1991	1	331	-0.02683
1992	1	350	0.055815
1993	1	370	0.05557
1994	1	411	0.10509
1995	1	382	-0.07317
1996	1	278	-0.3178
1997	1	323	0.150031

Once x and y have been calculated, the linear regression can be performed by any statistical package or even by basic spreadsheet programs. The following output was produced using Microsoft Excel (with the Analysis Toolpak installed) to perform a linear regression on the transformed Knowlton's cactus data above (and forcing the y-intercept to be zero by checking the "**Constant is Zero**" option in the Regression window):

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.00432	0.00432032	0.20502671	0.662721138
Residual	9	0.189648	**0.021072** ❸		
Total	10	0.193968			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	**0.00000** ❶	#N/A	#N/A	#N/A	#N/A	#N/A
X Variable 1	**0.03048** ❷	0.04377	0.69631	0.50382	-0.06853	0.12949

First check to see that the circled number ❶ is zero; if not, you failed to check the "**Constant is Zero**" box, and must redo the regression. The circled number ❷ is the slope of the linear regression, which provides an estimate of the parameter μ. The circled number ❸ in the ANOVA table is the mean squared residual, which is an estimate of σ^2. For Knowlton's cactus, these estimates indicate that μ is positive and σ^2 is less than μ, as is expected given that the counts show an increasing trend without a

BOX 3.2 *(continued)*

great deal of inter-annual variability (Fig. 3.1A). In contrast, the estimated μ and σ^2 for the lesser prairie chicken (Fig. 3.1E) are -0.106 and 0.097, respectively; the negative value of μ reflects the sharp decline of this population.

Note that the last column of the ANOVA table above indicates that this is a non-significant regression ($p>0.05$). This should not deter us from using our estimated μ and σ^2 to calculate viability measures; we are using linear regression here to find the best-fit values of the parameters given the data, not to statistically test any particular hypotheses.

Regression methods also allow one to detect outliers (years of unusually high population growth or unusually steep decline) in the count data. If these outliers coincide with events such as a change in the census protocol or one-time human impacts (e.g. oilspills) that are unlikely to recur, we may wish to omit them when estimating μ and σ^2. One can also test statistically whether μ and σ^2 differed before and after a management strategy was instituted or a permanent change in the environment took place. Interested readers should consult Dennis et al. (1991).

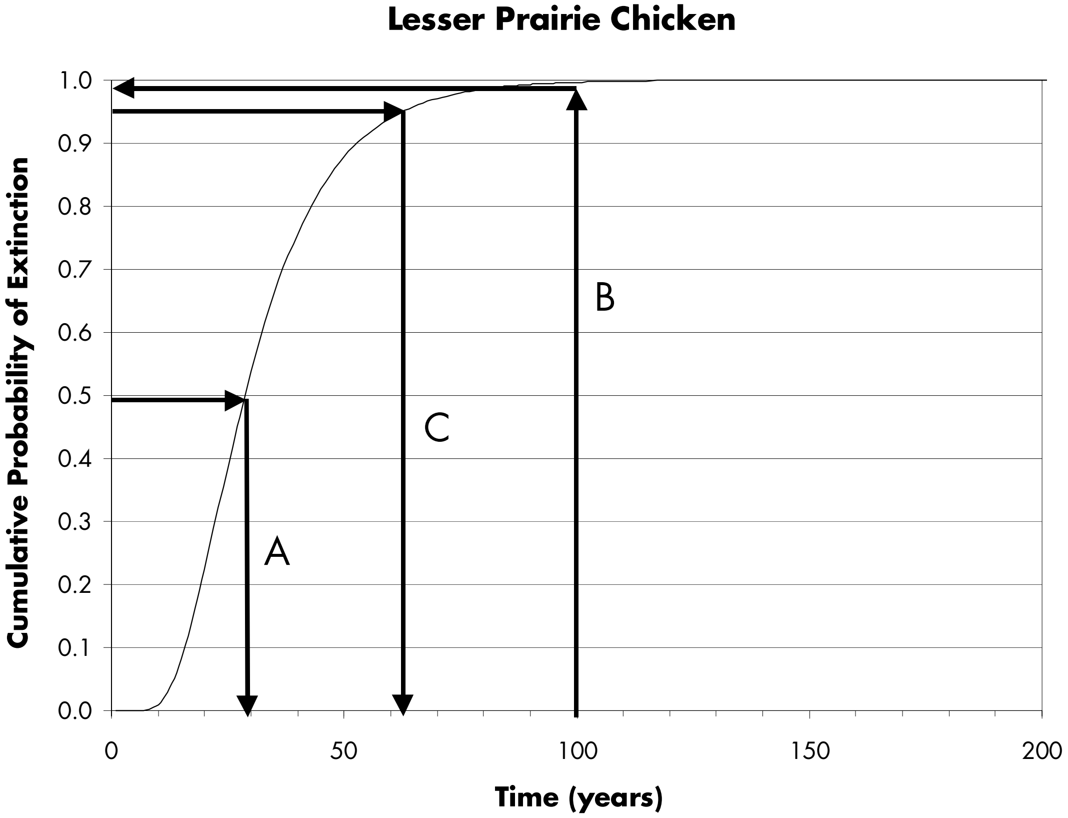
appropriate threshold, based upon biological, political, and economic considerations, we can define a population above the threshold to be viable, and can calculate both the mean time to attain the threshold given that it is reached and the probability that the population has fallen below the threshold by a specified time in the future (Box 3.3).

Because it is relatively easy to calculate, much theoretical work has focused on the mean time until an extinction threshold is reached. But we must be careful here, because the mean time to extinction will almost always *overestimate* the time it takes for most doomed realizations to reach a threshold. This fact traces back to the skewness in population abundance that develops in a stochastic environment (Fig. 3.2B). The large fraction of realizations that hover at low abundance are likely to dip below the threshold at relatively short times, while the few realizations that grow rapidly at first will likely take a very long time before they experience the long string of bad years necessary to carry them below the threshold. These later trajectories have a disproportionate effect on the mean time to extinction. For this reason, the mean time to extinction is a potentially misleading metric for PVA.

A better measure of the time required for most populations to attain the threshold is the median time to extinction, which is one of several useful measures that can easily be obtained from the so-called “cumulative distribution function,” or CDF, of the conditional time to extinction (Box 3.3). In effect, the conditional extinction time CDF asks the question: “if the extinction threshold is going to be reached eventually, what is the probability that a population starting at a specified initial size will have already hit the threshold at a certain time in the future?” Thus the conditional extinction time CDF considers only those population realizations that will eventually fall below the threshold; this will include all possible realizations if μ is less than or equal to zero, but only a subset of realiza-

FIGURE 3.6

The cumulative distribution function of extinction time for the lesser prairie chicken estimated from the data in Fig. 3.1E. Arrows indicate how the CDF can be used to calculate: A) the median time to extinction given that extinction occurs (note that the median, 29 years, is less than the mean time to extinction of 32.1 years); B) the probability of extinction by 100 years; and C) the number of years at which there is only a 5% chance of population persistence.



tions if μ is greater than zero (see Box 3.3). ***When the estimated value of μ is negative (so that eventual extinction is certain), the conditional extinction time CDF is the single most useful viability measure one can compute.*** From the CDF, one can read the median time to extinction as the time at which the probability of extinction first reaches a value of 0.5 (Fig. 3.6); for the reason given above, the median time to extinction is typically shorter than the mean time to extinction. The time to any other "event", such as the probability of extinction first exceeding 5% (or to put it in a more positive light, the probability of population persistence first falling below 95%), can also be easily read off the CDF. The CDF also clearly shows the probability of population persistence until any given future time horizon, which itself may be dictated by management considerations. ***Even if the estimated value of μ is positive (so that only a subset of the possible realizations will ever hit the extinction threshold), calculating the conditional extinction time CDF is still valuable, because it can be used in combination with the***

(continued on page 22)

BOX 3.3 (Key): Methods of Analysis
Measures of Viability Based on the Parameters μ and σ^2

Once the procedure outlined in Box 3.2 has been performed to estimate the parameters μ and σ^2, a suite of measures describing the population's growth and risk of extinction can be calculated (Dennis et al. 1991). In this box, we show how to calculate four of these measures that are either the most commonly used or the most informative. To make these calculations easier, we give protocols to perform them using Microsoft Excel (readers interested in the original mathematical formulae should consult Dennis et al. 1991). An Excel spreadsheet that calculates these measures and others given by Dennis et al. (1991) can be obtained by contacting W.F. Morris.

1) The Average Population Growth Rate, $\bar{\lambda}$

The average population growth rate $\bar{\lambda}$ is simply the base of natural logarithms, e, raised to the power $\mu + ½ \sigma^2$. To calculate $\bar{\lambda}$, its continuous-time analog, and their approximate 95% confidence intervals using Excel, follow these steps:

1) In cells **E3** and **E4** of an Excel worksheet, enter the duration of the counts in years (i.e. the year of the last count minus the year of the first count) and the total number of inter-census intervals (the number of counts minus one), respectively. Note that these numbers will not be equal if censuses were not performed in some years; for example, for Knowlton's cactus (Box 3.2), these numbers are 11 and 10.

2) In cells **E5** and **E8**, enter the estimated values of μ and σ^2, respectively.

3) Enter the following formulae in the appropriate cells, which will yield the measures indicated:

Measure	Cell	Excel Formula
Continuous rate of increase, $\bar{r}$:	E12	=E5 + E8/2
Lower 95% confidence limit for $\bar{r}$:	E13	=E12 + NORMSINV(0.025) * SQRT(E8*((1/E3)+(E8/(2*(E4-1)))))
Upper 95% confidence limit for $\bar{r}$:	E14	=E12 - NORMSINV(0.025) * SQRT(E8*((1/E3)+(E8/(2*(E4-1)))))
Average finite rate of increase, $\bar{\lambda}$:	E15	=EXP(E12)
Approximate lower 95% confidence limit for $\bar{\lambda}$:	E16	=EXP(E13)
Approximate upper 95% confidence limit for $\bar{\lambda}$:	E17	=EXP(E14)

For Knowlton's cactus (Box 3.6, Fig. 3.1A), the estimated average population growth rate $\bar{\lambda}$ is 1.04 with a 95% confidence of 0.96 to 1.14. Thus while the best estimate indicates the average of the possible population realizations will grow, the data do not allow us to rule out the possibility of a decline.

2) The Probability of the Population Reaching a Lower Extinction Threshold

If μ is negative, then all population realizations will decline eventually, and the probability is 1 that any threshold lower than the current population size will eventually be reached

(continued on page 20)

BOX 3.3 *(continued)*

(i.e. extinction is certain to occur). However, even if μ is positive, an extinction threshold may nevertheless be reached owing to the chance occurrence of a sequence of bad years (see Box 3.1). To calculate the probability of ultimate extinction, make the following additions to worksheet you began above:

1) In cell **E11**, enter the formula **=(E4-1)*E8/E4**, which will calculate a slightly different estimate of σ^2 to be used in subsequent calculations.

2) In cells **E19** and **E20**, enter the current population size and the extinction threshold, respectively.

3) Calculate the probability that the extinction threshold is eventually reached by entering the formula **=IF(E5<0,1,(E20/E19)^(2*E5/E11))** in cell **E22**.

For Knowlton's cactus, the estimate of m is positive (Box 3.2), and using a current population size of 323 (the size of the last count taken) and an extinction threshold of 10 plants, the estimated probability of ultimate extinction is only 0.000014. Thus the available data suggest that the risk of extinction faced by the only known population of this cactus is not very great (provided that poaching can be prevented). In contrast, the lesser prairie chicken population in Fig. 3.1E (for which the estimated μ is negative; see Box 3.2) has an ultimate extinction probability of 1.

3) The Mean Time to Extinction

The mean time to extinction given that the extinction threshold is reached (which is *not* certain to occur if μ is greater than zero; see Section 2 in this Box) is simply the difference between the natural logs of the initial population size and the extinction threshold, divided by the absolute value of μ. It is easily calculated by adding the following elements to the Excel worksheet begun above:

1) Enter the formula **=LN(E19/E20)** in cell **E21**. Cell **E21** will now contain the difference between the natural logarithms of the initial and threshold population sizes.

2) Enter the formula **=E21/ABS(E5)** in cell **E23**, which will then contain the mean time to extinction. Incidentally, entering the formulae **=IF(E23+NORMSINV(0.025)*SQRT((E21^2)*E$11/((E5^4)*E3))<0,0,(E23+NORMSINV(0.025)*SQRT((E21^2)*E11/((E5^4)*E3))))** and **=E23-NORMSINV(0.025)*SQRT((E21^2)*E11/((E5^4)*E3))** in cells **E24** and **E25** will yield the lower and upper 95% confidence limits for the mean time to extinction, respectively.

For Knowlton's cactus, the mean time required for the census counts to decline from 323 to 10 plants is 114 years. But remember, this calculation only applies to those realizations of population growth that eventually reach the extinction threshold; our calculation of the probability of ultimate extinction in Section 2 of this Box indicates that the estimated mean time to extinction will apply to only a small fraction of all possible realizations (about one in 100,000). For this reason, we suggest that *the mean time to extinction should only be used as a measure of extinction risk when the estimate of* μ *is negative* (because only then does it apply to every population realization). One such example is the lesser prairie chicken (Fig. 3.1E). The mean time to an extinction threshold of 10 individuals for this sharply declining population is only 32.1 years, indicating not only that ultimate extinction is a certainty (see Section 2 above), but that the extinction threshold is likely to be reached quite soon.

Due to skewness in population size (Fig. 3.3), the confidence interval for the mean time to extinction is usually large; for the lesser prairie chicken, it ranges from 0 to 77 years.

BOX 3.3 *(continued)*

4) The Cumulative Distribution Function (CDF) for the Conditional Time to Extinction

Next, we can extend our Excel worksheet to calculate the conditional extinction time CDF given that the extinction threshold will be attained. The extinction time CDF gives the probability, *considering only those realizations of population change that ultimately fall below the extinction threshold*, that the threshold has already been reached at a given time. Hence as with the mean time to extinction, the extinction time CDF applies to all realizations if $\mu<0$, but to only a subset of realizations if $\mu>0$. To calculate it, we use the standard normal cumulative distribution function NORMSDIST provided by Excel:

1) Fill column **B** downward from cell **B31** with a series of times at which you wish to compute the CDF. For most purposes, every 5 years from 5 to 1000 years is adequate (the sequence "Edit-Fill-Series" from the pull-down menu will allow you to accomplish this easily).

2) In cell **D31**, enter the following formula:
=NORMSDIST((-E21+ABS(E5)*$B31)SQRT($E$11*$B31))+EXP(2*E21*ABS (E5)/ E11)*NORMSDIST((-E21-ABS(E5)*$B31)/SQRT($E$11*$B31)).

Now select cell **D31**, drag downward to the row corresponding with the last entry in Column B that you created in step 2, and then type Ctrl-D; column **D** will now be filled with the values of the CDF that correspond to the times in Column B. You can treat these 2 columns as a table in which you can look up values of the CDF at different times, or you can use Excel to create a graph of the CDF versus time (for example, Fig. 3.6 shows a graph of the CDF for the lesser prairie chicken, and indicates how several measures of population viability (including the median time to extinction) can be read from the graph).

5) Using the Extinction Time CDF When μ is Positive

When μ is positive, the extinction time CDF must be interpreted with caution, *because it does not apply to all population realizations* (only to those that will eventually reach the extinction threshold). For example, the median time to extinction from the CDF (not shown) for the Knowlton's cactus population in Fig. 3.1A is 105 years (using a current population size of 323 and an extinction threshold of 10 plants). This does NOT mean that half of all realizations will have reached the extinction threshold by 105 years, but instead that *half of the realizations that will eventually hit the threshold (which represent only about 1 in 100,000 possible realizations) will have done so by 105 years*. Given the positive value of μ, the underlying population model predicts that the remaining 99,999 of 100,000 realizations will NEVER hit the extinction threshold. Nevertheless, the conditional extinction time CDF is still valuable even when μ is positive, for the following reason. We can calculate the *total probability* that the population has gone extinct by a given future time horizon, *accounting for ALL possible realizations*, if we calculate both the probability that the extinction threshold is reached eventually (see Section 2 of this Box) and the conditional extinction time CDF. The total probability that extinction occurs by, say, 100 years is the probability that extinction will occur eventually multiplied by the *conditional* probability that extinction will have occurred by 100 years *given that it will occur eventually*, which is precisely what the conditional extinction time CDF tells us. For Knowlton's cactus, the value of the CDF at 100 years is 0.455, so the total probability of reaching an extinction threshold of 10 plants by 100 years is 0.000014 (see Section 2 above) multiplied by 0.455, or 0.0000064, a rather small number. By performing a calculation such as this, we could compare the relative viabilities of two populations, one with postive and one with negative μ, whereas it would be inappropriate to compare directly the CDFs of the two populations.

probability of ultimate extinction to compute the likelihood that extinction will have occurred by a given future time horizon (see Box 3.3).

Uses of the Extinction Time Cumulative Distribution Function in Site-based and Ecoregional Planning

Because the conditional extinction time CDF encapsulates so much useful information about population viability, we now give three examples that show how the CDF can be used to inform decisions about the viability of individual element occurrences (EOs), or about which of several EOs should receive the highest priority for acquisition or management.

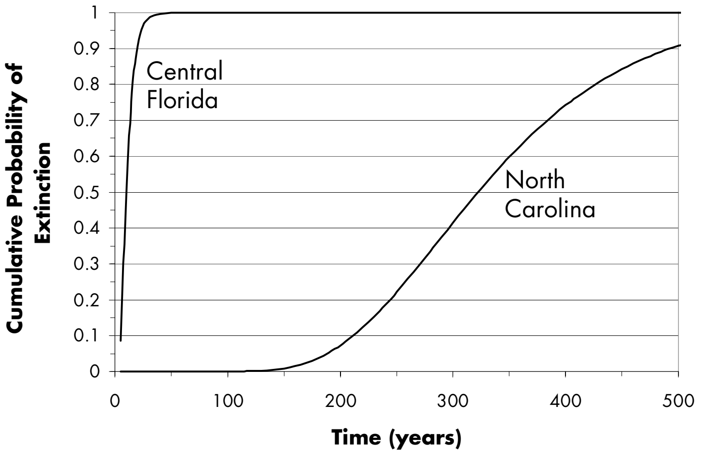
Perhaps the most valuable use of the CDF is to make comparisons between the relative viabilities of 2 or more EOs. Ideally, we would have a series of counts from each EO. For example, Fig. 3.1B & C show the number of adult birds during the breeding season in populations of the federally-listed red-cockaded woodpecker in central Florida and in North Carolina. Applying the methods outlined in Boxes 3.2 and 3.3 yields the CDFs in Fig. 3.7A. Both because it has a more negative estimate for μ (-0.083 vs -0.011) and a smaller initial size, the Florida population has a much greater probability of extinction at any future time than does the North Carolina population.

Often we will not have independent census data from each EO about which we must make conservation decisions. However, if we have a single count of the number of individuals of a particular species in one EO, we can use count data from multiple censuses of the same species at a second location to make a ***provisional*** viability assessment for the first EO when no other data are available. For example, Dennis et al. (1991) calculated the extinction time CDF for the population of grizzly bears (*Ursus arctos*) in the Greater Yellowstone ecosystem, using values of μ and σ^2 estimated from aerial counts of the number of adult females over 27 years, a starting population of 47 females (the number estimated in 1988, the last census available to them), and an "extinction" threshold of 1 female (Fig 3.7B). A second isolated population of grizzly bears occupying the Selkirk Mountains of southern British Columbia consists of about 20 adults, or roughly 10 adult females. If we have no information about the Selkirk Mountains population other than its current size, we may as well use the CDF for the Yellowstone population to give us a *relative* sense of the viability of the Selkirks population. In so doing, we are assuming that the environments (including the magnitude of inter-annual variation) and the human impacts at the two locations are similar, an assumption which could be evaluated using additional information on habitat quality, climatic variation, and land-use patterns. Accepting these assumptions, and using the CDF from the Yellowstone population, the Selkirks population of 10 females would have an 31-fold greater probability of extinction at 100 years (Fig 3.7B; for an extension of this analysis to multiple sites, see Box 5.2). For species of particular concern, it may be possible to improve upon this approach by compiling count data from multiple locations. We could then estimate *average* values for the parameters μ and σ^2 to provide ballpark assess-

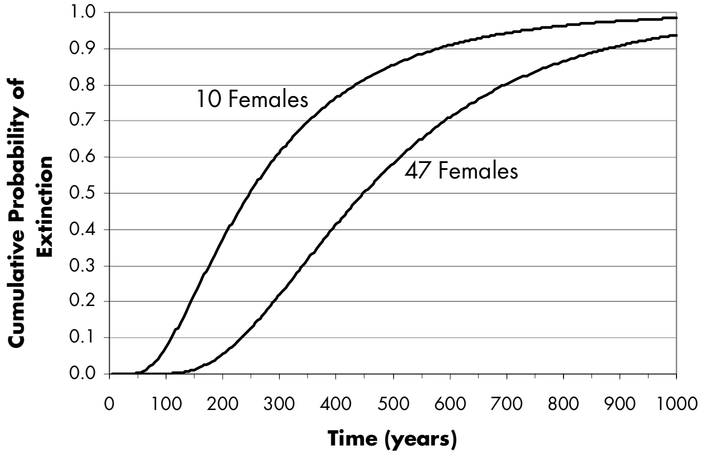
FIGURE 3.7

How to use the extinction time CDF in site-based and ecoregional planning. A) Comparing the CDF's for the two red-cockaded woodpecker populations in Fig. 3.1B, C (for both curves, initial population size equaled the last available count and the extinction threshold was 10 birds). B) CDF's for the Yellowstone grizzly bear (Fig. 3.1D) assuming initial population sizes of 10 or 47 females and an extinction threshold of 1 female; C) The effect of the variance parameter σ^2 on the CDF, using the data for the Yellowstone grizzly bear with the observed variance (σ^2=0.9) or one half the observed variance (σ^2=0.45)

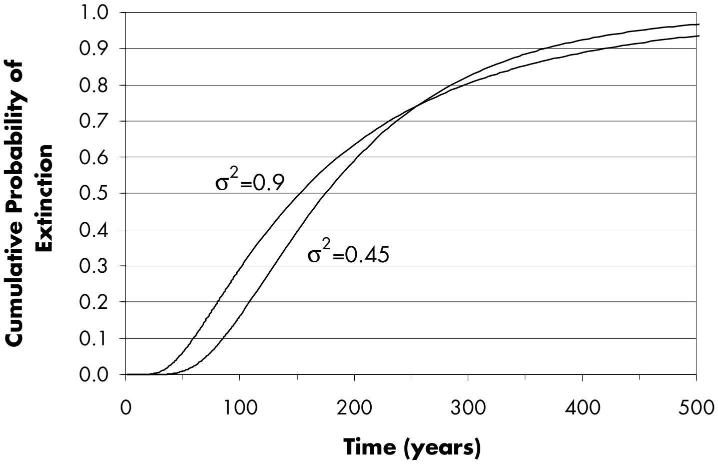
A) Red-cockaded Woodpecker



B) Yellowstone Grizzly Bear



C) Yellowstone Grizzly Bear



ments of viability for EOs with only a single census, or choose the location with the most similar environment for comparison.

As a final use of the CDF, we point out that even in the absence of any count data for a species of critical concern, knowledge of how the CDF is affected by its underlying parameters can help us to make *qualitative* assessments of relative viability, especially if we can use natural history information to make inferences about the local environment of an EO or about the life history of the species in question. For example, we will frequently be able to make an educated guess that one EO's environment is likely to be more variable than another's in ways that will affect population growth. Similarly, some species will have life history features (e.g. long-lived adults) that buffer their populations against year-to-year environmental variation. If the effects of environmental variation on the population growth rate are less for one species or EO than another, then its σ^2 value will be smaller. Such differences in σ^2 influence the CDF even when its other determinants (μ and the starting and threshold population sizes) are fixed (Fig. 3.7C). Thus we can state that, all else being equal, the greater the environmentally-driven fluctuations in population growth rate the greater will be the risk of extinction at early time horizons, a qualitative statement that nonetheless provides some useful guidance.

Assumptions in Using the Method of Dennis et al.

As with any quantitative model of a complex biological process, PVA using count data relies upon simplifying assumptions. In Box 3.4, we list the most important assumptions we are making when we apply the method of Dennis et al. to a series of counts and then estimate measures of population viability. The fact that these assumptions are explicit is an advantage of a quantitative approach to evaluating viability, relative to an approach based upon general natural history knowledge or intuition. By evaluating whether the assumptions are met, we can determine whether our analysis is likely to give unreliable estimates of population viability, but more importantly, we can often determine whether violations of the assumptions are likely to render our estimates (e.g. of time to extinction) optimistic or pessimistic. By "optimistic" and "pessimistic", we mean, for example, that the true time to extinction is likely to be shorter or longer than the estimated value, respectively. If we know that the estimated time to extinction for an EO is short but pessimistic, we should be more cautious in assigning a low viability ranking, while a long but optimistic estimate should not inspire complacency.

We now give a few brief examples illustrating how, by evaluating the assumptions in Box 3.4, we can make more informed viability assessments. One life history feature that may cause Assumption 1 (Box 3.4) to be violated is dormant or diapausing stages in the life cycle, such as seeds in a seed bank or diapausing eggs or larval stages of insects and freshwater crustaceans. Because they are difficult to census accurately, these stages are typically ignored in population counts, but as a result the counts may not represent a constant fraction of the total population. For example, when the number of above-ground

BOX 3.4 (Key): Important Assumptions
Assumptions of PVA Using the Method of Dennis *et al.*

1) The data represent exhaustive counts of individuals in the population, estimates of total population size, or counts of a subset of the individuals (e.g. adult females, individuals in quadrats that sample a constant proportion of the area occupied by the population, etc.) comprising a fraction of the entire population that does not change over time.

2) The year-to-year variation in the counts reflects the true magnitude of environmentally-driven variation. We must have performed censuses in a sufficient number of years to accurately assess the year-to-year variation in the population growth rate. Furthermore, the variation in the counts must not be due primarily to observation error in estimating population size each year.

3) Inter-annual environmentally-driven variation is not extreme. In particular, we assume there are no large-magnitude fluctuations caused by extreme catastrophes or unusually good years.

4) The population growth rate is not affected by density, and thus does not change as the population increases or decreases.

individuals in a plant population is zero, total population size is not necessarily zero, as some individuals may remain in the seed bank. If the subpopulation in the seed bank is more buffered from environmentally-driven fluctuations than is the above-ground population (as is likely to be the case in environments that favor the evolution of dormant life stages in the first place), then extinction times estimated from the above-ground population alone may underestimate the true value for the entire population, and thus provide a (potentially highly) pessimistic measure of population viability. For organisms such as desert annual plants in which a large and persistent fraction of the population is likely to go uncensused, the method of Dennis et al. is probably not an appropriate way to estimate extinction risk.

In Assumption 2 (Box 3.4), observation error is the failure to count accurately the true number of individuals in a population at any one time. Observation error can be caused by a host of factors, such as complex background vegetation that makes some individuals difficult to detect, multiple counts of the same individual for mobile organisms (or by different members of the census team), incorrect species identification, or sampling variation introduced when a partial census (e.g. quadrat or transect sampling) is use to infer total population size or when indirect measures of abundance (e.g. scat, tracks or hair snags) are used. Such errors will lead to a pessimistic measure of viability over the short term, because they will cause the estimated value of σ^2 to be an overestimate of the true environmentally-driven component of variation in the counts, and a higher σ^2 predicts a greater likelihood of extinction over short times (see Fig. 3.7C). Repeated sampling of the same area (see Chapter 6 of this handbook) and “ground-truthing” indirect measures of abundance are two ways to estimate the magnitude of observation errors. We must also be aware of the fact that short sequences of counts will tend to misrepresent the true environmental component of variability, because they will tend not to include extreme values.

One violation of the assumptions that will cause viability estimates to be optimistic is the existence of intermittent catastrophes (Assumption 3), such as rare ice storms, droughts, severe fires, etc., which introduce the possibility of sudden declines in abundance not accounted for in our estimate of σ^2. More detailed methods have been developed to include catastrophes in estimates of time to extinction (see the methods of Mangel and Tier 1993 and Ludwig 1996, which also allow density dependence; see below). However, with most short-term count data, we will lack sufficient information to estimate the frequency and severity of rare catastrophes, information that more detailed methods require if they are to provide more accurate assessments of extinction time. Thus in practice, we may need to be content with the statement that if catastrophes do indeed occur, our assessments of extinction risk based upon short-term census data will likely underestimate the true risk. If catastrophes do occur but with similar intensity and frequency across multiple EOs, we can still use the method of Dennis et al. to assess *relative* viability. Of course the converse, failure to account for rare good years, will have a pessimistic effect on the estimated extinction risk.

The ways in which density dependence (i.e., the tendency for population growth rate to change as density changes; see Assumption 4) may alter our estimates of extinction risk are more complex. A decline in the population growth rate as density increases will tend to keep the population at or below a carrying capacity. Unlike the predictions of the geometric growth model upon which the viability measures of Dennis et al. are based, such regulated populations cannot grow indefinitely, and the probability of ultimate extinction is always 1 (although the time to extinction may be extremely long). On the other hand, declining populations may receive a boost as density decreases and resources become more abundant; because estimates of μ based upon counts taken during the decline do not account for this effect, they may result in pessimistic estimates of extinction risk. Finally, the opposite effect may occur if a decline in density leads to difficulties in mate finding or predator defense and a consequent *reduction* in population growth rate. The downward spiral created by these so-called "Allee effects" results in extinction risks that become greater and greater as the population declines, and causes estimates of extinction risk made by ignoring these effects to be overly optimistic.

As with catastrophes, including density dependence in viability assessments will generally require more data, but there are ways to do so. Statistical methods developed by Pollard et al. (1987) and by Dennis and Taper (1994) allow one to test whether the counts show any evidence of density dependence (but see Shenk et al. 1998). If the population growth rate does depend upon density, density-dependent versions of the geometric growth equation 3.1 can be fit to the count data, using nonlinear rather than linear regression techniques (Middleton and Nisbet 1997). However, it is an inescapable fact that, because density-dependent models require us to estimate at least three parameters (for example, the carrying capacity in addition to μ and σ^2), they will require more censuses to achieve a similar degree of estimation accuracy. In addition, for most density-dependent population growth models, there are no simple mathematical formulae for extinction probability or time to extinction, and we must rely upon computer simulations to calculate them (for example, see Ludwig 1996 and Middleton and Nisbet 1997). One exception that has received a great deal of attention from theoretical population biologists is a model in which the population grows exponentially up to a ceiling, above which it cannot go. In this case, mathematical formulae for the mean time to extinction have been derived by several authors (Box 3.5). Examining the relationship between time to extinction and the "height" of the population ceiling provides a way to ask how the maximum population size a particular EO will support should influence its rank (Box 3.5).

BOX 3.5 (Optional): Theoretical Underpinnings Density-Dependent Models

When a population's growth rate declines with increasing density, the population will not continue to grow exponentially, but will typically approach a carrying capacity, usually denoted K. A mathematically simple way to approximate this effect is to allow the population to grow exponentially until it reaches K, when further population growth ceases. This model of "exponential growth up to a ceiling" allows us to ask how the mean time to extinction of a population currently at the ceiling increases as the ceiling itself increases. Because the "height" of the ceiling will be determined by the amount of habitat available to the population, this is another way of asking how the spatial extent of an EO will influence population viability. Approximate formulae for relationship between the mean time to extinction and the height of the population ceiling have been derived by several authors, including Lande (1993), Foley (1994), and Middleton et al. (1995).

Here we give a brief overview of the results of Lande (1993). When the parameter μ is positive, the mean of the normal distribution of the log of population size will increase over time (Fig. 3.4); that is, most population realizations will grow. Nevertheless, some realizations will fall below the extinction threshold, and this outcome will be more likely if σ^2 is large. In fact, the magnitude of the environmentally-driven variation in the population growth rate, as embodied by the parameter σ^2, will have a pervasive effect on how the extinction time depends on the height of the population ceiling, K. Lande showed that if μ is positive and the ceiling is sufficiently high, the mean time to extinction will be approximately proportional to $K^{2\mu/\sigma^2}$. The ratio of μ to σ^2 thus determines the shape of the relationship between the mean time to extinction and the population ceiling (Fig. 3.8). If σ^2 is small relative to μ (i.e. less than 2μ), the extinction time increases faster than linearly as the ceiling is increased; increasing σ^2 until it equals and then surpasses 2μ will cause the scaling of extinction time vs height of the ceiling to become linear and then less-than-linear. Thus if environmental variability is high, we may face a law of diminishing returns; by increasing the ceiling (for example by increasing the amount of habitat available to the population), we may gain a disproportionately small increase in the

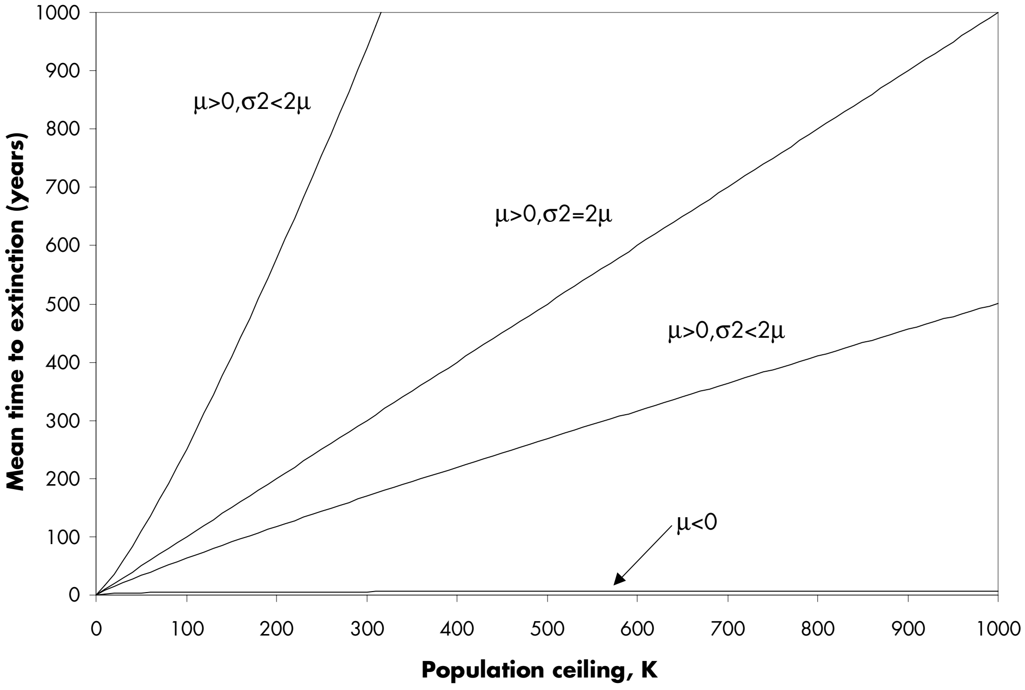
(continued on page 28)

BOX 3.5 *(continued)*

mean time to extinction. *The important conservation message is that highly variable populations will continue to be prone to rapid extinction caused by a chance sequence of bad years even when the maximum population size becomes large*. Thus even if we are unable to estimate K accurately, obtaining estimates of μ and σ^2 using the method of Dennis et al. (Box 3.2) may still give us generally useful information about how increasing the available habitat will buffer a population from extinction. Finally, if μ is negative (that is, the population declines over the long run), then the mean extinction time scales as the natural log of the population ceiling, which also increases in a less-than-linear fashion as the ceiling increases (Fig. 3.8). Useful approximations for the distribution of extinction times (not just the mean extinction time) as a function of the height of the ceiling were derived by Middleton et al. (1995).

FIGURE 3.8

The scaling of the mean time to extinction versus the height of the population ceiling in a model of "exponential growth up to a ceiling" (after Lande 1993)



Summary and Conclusions: Prospects for Using Count Data to Perform Population Viability Analyses

In summary, the procedure we have described above provides a straightforward method to obtain quantitative measures of population viability using data from a series of population censuses. The principal advantage of the method is its simplicity, in terms of both its data requirements and the ease of calculating viability measures. Other than simply recording the presence or absence of a species, population censuses are likely to be the simplest way that field biologists can collect data. However, to apply the methods described above, censuses must be conducted over a reasonable number of years. *We recommend that ten censuses should be viewed as an absolute minimum when applying this method, and more would be preferable* (see Chapter 7). But when an adequate number of censuses are available, both the linear regression and the calculation of viability measures such as the extinction time CDF can be performed in short order using software available on almost any desktop computer.

Users of count-based PVA need to be aware of its limitations when the underlying assumptions (Box 3.4) are violated. Because of these limitations, the method we have presented here is not a panacea for making conservation decisions in a world of sparse data, but neither are these limitations a fatal flaw that renders the method unusable. We reiterate that because the assumptions are explicit, the exercise of calculating the extinction time CDF (for example) is still useful, because we will often be able to gauge in which direction our estimate is in error (something which cannot always be said about viability measures obtained in other ways, such as through expert opinion). Moreover, if we know that a particular factor (such as density dependence in the population growth rate) has been omitted in a consistent way across multiple PVAs, they may still provide us with useful guidance about how risks are likely to differ among species or EOs. In essence, we argue that count-based PVA is best viewed as a tool that provides us with *relative* measures of the "health" of two or more populations, measures that decision-making agendas such as ecoregional planning require us to make. That is, while we would not put much credence in a particular numerical value of a viability measure (for example, the estimated median time to extinction for the lesser prairie chicken population of 22 years (Fig. 3.6)), we can be more comfortable accepting that median extinction time estimates of, for example, 100 and 500 years for two EOs warn us of potentially real differences in the viability of those populations. Assessing the reliability of viability measures when the underlying assumptions of the method are violated is currently an area of active research in population biology. But at the present time, we believe it would be a shame not to make use both of the available data (e.g. Fig. 2.1) and of the simple tools at hand to bring some quantitative rigor to the process of determining the relative ranks of multiple occurrences of rare and unique species.

CHAPTER FOUR
Projection Matrix Models

While simple counts of the number of individuals in a population are more likely to be available (See Chapter 2, Fig. 2.1), for a few species of particular concern we may have access to detailed demographic information describing the rates of survival, reproduction, and growth of individuals that differ in size, age or other characteristics. Viability analyses that use such individual-based data most often rely upon population projection matrices. Matrix models predict long-term population growth rates, transient population dynamics, and probabilities of extinction over time. A basic requirement for constructing a projection matrix is that we must be able to classify individuals unambiguously into classes based on their age, size or life-history stage. We must then quantify the survival, growth, and reproduction of numerous individuals in each class over a minimum of 2 to 3 years, preferably longer. These data are used to calculate so-called "vital rates", or the class-specific annual rates of survival, growth and fecundity, as well as the degree of year-to-year variation in those rates. In turn, we use the vital rates to compute the entries in the projection matrix, which we then use to predict the likely future trajectory of the population and its risk of extinction.

Projection matrix models have three advantages over simpler models that track only the total number of individuals in a population, such as those we reviewed in Chapter 3. (The tradeoff that balances these advantages is that we require far more data to construct a projection matrix than we do to perform a count-based PVA.) First, projection matrices make it possible to assess the influence that the vital rates of particular classes have on the growth of the population as a whole. For example, it is possible in a matrix analysis to assess how much juvenile mortality (as distinct from the mortality of other life stages) drives population decline. In an analysis based on total population counts, we would not be able to discern whether juvenile mortality plays a role in the decline—we would simply know that the population is declining. Yet if juvenile mortality is a key factor causing the decline, management efforts aimed at improving juvenile survival should be more effective than efforts aimed at other vital rates or life stages.

A second advantage to a matrix model is that it may more accurately predict future population trends for a long-lived species that has undergone either recent changes in one or more vital rates (e.g. due to a novel human impact, or a recently-imposed management plan) or a perturbation in the population structure (i.e. the distribution of individuals among classes). For example, if the fecundity of adults suddenly drops, or if a large number of juveniles unexpectedly migrates from the population, counts of adults may show little change for several years, despite the fact that the adult population will certainly decline once the reduced juvenile cohort matures. Put another way, count data for long-lived species can reflect past conditions more than current ones. If recent changes in vital rates or population structure have occurred, more accurate predictions are possible with a matrix approach.

A third advantage of projection matrix models is that they are particularly well suited to evaluating management alternatives, provided demographic data from contrasting situations exist. For example, an experiment in which different management techniques are applied randomly to different plots, or data from sites with different management histories, would facilitate the construction of contrasting matrices for each management scenario. These matrices would serve to integrate the various effects of management on multiple vital rates into an overall effect on population growth. Moreover, manipulation of the parameters in matrix models allows one to perform computer "experiments" to predict the likely effects of proposed management practices. Importantly, these models can be informative about the magnitude of threats and the potential for improvements, and allow determination of how much effort will be necessary to rescue a declining population.

Like count-based models, projection matrix models can also be used to assess the relative viability of different EOs. The difficulty is that in order to do this, demographic data must be collected from several sites. Obviously, only very rarely will it be possible to undertake this type of analysis, but at the end of this chapter we present one example of this comparative approach using a TNC data set for the rare plant *Aletes humilis*.

In this chapter, we describe the construction and analysis of projection matrix models, but we do not cover the topic comprehensively (thorough reviews are given by Caswell (1989) and Tuljapurkar and Caswell (1997)). There are a large number of complexities and nuances to this type of PVA; hence users would be well-advised to consult with a population modeller to go beyond the simpler uses of matrix models covered here. However, it is important to have an understanding of the principal issues in model construction and interpretation even if collaborating with a mathematical population biologist.

Constructing and using a projection matrix model

The most laborious and time-intensive step in matrix-based modeling is the collection of demographic data on known individuals over a number of years. The details of how to gather such data are well described in a variety of references (e.g. Southwood 1978, Krebs 1989), and we do not review them here. Instead, we describe in Box 4.1 the steps one must follow to use raw data on individuals to produce a projection matrix and to use the matrix to predict future population sizes. The basic steps are:

STEP 1: Determine what feature of individuals (age, size, or life stage) best predicts differences in vital rates. Then divide the population into classes based upon the feature you have chosen.

STEP 2: Use demographic data on known individuals to estimate the vital rates for each class, and use them to construct a population matrix, according to the rules specified in Box 4.1.

STEP 3: Construct a population vector by specifying the initial number of individuals in each class in the population. A population vector is a list of the number of individuals in each class; the sum of the elements in the vector equals the total population size.

STEP 4: Use the matrix and the population vector to project the population forward in time, thus predicting the future size of the population, the long-term population growth rate, λ, and the risk of future extinction. This step involves simple rules of linear algebra (Box 4.1, Step 4).

While these simple steps apply to any matrix-based PVA, there are a host of additional complexities that may be added to the models, and many additional analyses that could be performed. The most essential complexity to be included in a PVA, provided the data are available, is year-to-year variation in vital rates. We discuss this complexity in detail in the following section of this chapter. One additional analysis that is usually desirable is to perform "experiments" by looking at effects on λ and on extinction risk of altering matrix entries to reflect changes in the demographic rates of particular life stages resulting from differing forms of management. Often this is the most instructive step in using matrix models. After we review how to include stochastic variation into matrix models, we give an example of a matrix-based PVA using TNC data, illustrating how projection matrices can be used to explore management options.

(continued on page 36)

BOX 4.1 (Key): Methods of Analysis
Constructing and Using a Matrix Model: A Step-by-Step Guide

STEP 1. Decide if the population is best classified by age, size, or life stage.

To put together a model we must first divide the population into different classes. We can do this by AGE (e.g., all individuals a year old, all individuals 2 years old, etc.), SIZE (e.g., all individuals <10 cm long, all individuals 10-20 cm long, etc.), or STAGE (e.g., all seeds, all seedlings, all saplings, all adults). The choice of which type of classification to use depends on the life history of the organism and on the available data.

Age-structured populations: For many species the age of the individual is the best descriptor of how the vital rates react to the environment. For example, it could be that 1-yr olds survive worst, and 3-yr olds best. An age-structured matrix is often the method of choice when our data are taken from known-age individuals in a cohort analysis. Determinate growers (e.g. mammals and birds) are often modeled in this way, or with a modified formulation which adds a final adult stage. This simplification is only appropriate for species in which there is little variation in adult survival and fecundity rates as individuals age. Obviously, an age-based classification is restricted to species that are easy to age in the field.

Size-structured populations: If the exact age is not as important as the physical size of an individual in determining its vital rates, then a size-based model should be used. Examples include plants, for which size often influences how many seeds an individual can produce, and many fish species, in which size influences survival. Size may also be used when age is unknown, because size can be measured easily

BOX 4.1 *(continued)*

in the field. Typically, the divisions between size categories are determined by looking for critical sizes at which vital rates change abruptly, or by assuring that the number of individuals in the data set that will fall into each size class will be adequate to estimate vital rates (see below).

Stage-structured populations: Organisms in which the life stage of an individual (such as simply being a newborn, a juvenile, or a reproductive adult, regardless of age or size within these stages) has the greatest influence on vital rates should be modeled as a stage-structured population. Stage-structured models can be the easiest to use because recognizing distinct stages may be more practical than determining age or size.

Next we must determine the number of classes to use. This decision must strike a balance between two opposing goals. First, each class should be as homogeneous as possible—that is, all individuals within a class should be quite similar in terms of their vital rates, while differing from individuals in other classes. Second, the number of individuals in each class must be large enough to estimate vital rates accurately. Using too few classes will cause us to lump in a single class individuals that possess different demographic rates, whereas using too many classes means that some will contain too few individuals for accurate estimation of vital rates. See Vandemeer (1979) for one approach to balancing these two goals.

STEP 2. Construct a projection matrix.

We are now ready to transform raw data into class-specific vital rates, and then put these into a matrix format. All projection matrices, regardless of the type of classification used, have the form:

$$\begin{array}{cc} & \text{FROM CLASS} \\ & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \text{TO CLASS} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \left[\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \end{array}$$

Each matrix entry (a_{11} - a_{44}) has two subscripts, the first indicating its row and the second indicating its column in the matrix. Each entry corresponds to the annual rate of the transition FROM the class indicated by the column subscript TO the class indicated by the row subscript. Thus a_{32} is the annual rate of transition from class 2 to class 3. Fecundities are placed in the first row of the matrix because they represent contributions from adult classes to the newborn class (which by convention is class 1), and in this case correspond to entries a_{12}, a_{13}, and a_{14}. All other entries in the matrix represent class transitions of individuals other than newborns (for example a_{23}, a_{33}, and a_{43} give the fraction of class 3 individuals that shrink to class 2, remain the same size, and grow to class 4 in one year, respectively).

The details of taking raw data and calculating the probabilities that go into the matrix are discussed using two examples.

(continued on page 34)

BOX 4.1 *(continued)*

Example 1: A stage-structured matrix. We will use as our first example a modified age-based matrix for the American Oystercatcher, a threatened shorebird that breeds along the eastern and gulf coasts of the U.S. (from Davis and Groom, unpublished manuscript). American Oystercatchers can be divided into three stages: juveniles, subadults, and adults. Data on the fates of banded oystercatchers and the success of known nests have been gathered by a variety of researchers for more than a decade. American oystercatchers have very low fecundity due to high rates of nest loss to predators and storm events. Each adult oystercatcher nesting on South Core Banks, NC, produces an average of only 0.054 offspring that survive to the following year (Davis and Simons 1997). Juvenile survivorship is also low, estimated to be 66.5% per year. However, adult survivorship is high, and birds may live as long as 20 years. Annual adult and subadult survivorships are estimated to be 95% and 72.4%, respectively.

We have arranged these data on fecundity and survival rates of individuals in each stage into a matrix, **M** (see equation 4.1 below). Fecundity values are placed on the first row, and survival rates are placed in the other two rows so as to correctly represent transitions from one stage to another (e.g., survival from the subadult stage to the adult stage (0.724) goes in column 2 (FROM subadults) of row 3 (TO adults):

Equation 4.1

$$
\begin{array}{ll|ccc|}
 & & & \text{FROM:} & \\
 & & \text{juveniles} & \text{subadults} & \text{adults} \\
\hline
 & \text{juveniles} & 0 & 0 & 0.054 \\
\text{TO:} & \text{subadults} & 0.665 & 0 & 0 \\
 & \text{adults} & 0 & 0.724 & 0.95 \\
\hline
\end{array} = M
$$

Example 2: A size-structured matrix. Nantel et al. (1996) performed a PVA to determine the harvest potential of American Ginseng (*Panax quinquefolium*). Ginseng is a perennial plant that is difficult to age, but individuals vary in their survival and fecundity according to the number of leaves. Hence, plants are divided into 4 size classes, those with 1, 2, 3, and 4 leaves, and into seed and seedling classes.

In the case of a size-structured population, many more transitions among classes are possible than in an age-structured or even a stage-structured population, because individuals may grow one or more size classes in a single year, and may even drop from a larger to a smaller class (many plants, for example, may shrink between years). To measure the rates at which these numerous transitions occur in the population, it is best to perform a census of marked individuals in all size categories (as was done with *Aletes humilis*; see Tables 4.1 and 4.2). We do not have the original data that Nantel et al. used to construct their matrices, but we could imagine they followed 100 marked individuals in each size class (including seedlings, which are called the 0 size class) and obtained the following results:

		SIZE CLASS WHEN MARKED:				
		0	**1**	**2**	**3**	**4**
SIZE NEXT YEAR:	**0**	0	0	0	0	0
	1	16	25	0	0	0
	2	2	59	52	6	0
	3	0	0	40	85	19
	4	0	0	3	5	74

BOX 4.1 *(continued)*

Note that the numbers in each column do not sum to the initial number of marked individuals (due to mortality), that most individuals with 1 leaf grow if they survive, and that shrinkage occurs only rarely (6 individuals with 3 leaves shrank to having just two, and 19 individuals with 4 leaves shrank to having just 3). Fecundity values represent the production by aboveground plants of seeds in the soil (represented by size class "S"). Estimating fecundities is more complex than estimating other matrix elements; we discuss the estimation of fecundities in more detail in Box 4.2. In this example, counting the average number of seeds that each marked individual produces during the year is the most straightforward way of calculating fecundity (although these numbers must be corrected for the death of seeds between the time they are produced and the time of the next census; see Box 4.2). The estimated fecundities for ginseng are 0, 0, 1.35, 13.24, and 18.50 surviving seeds per year for classes 0 to 4, respectively. Nantel et al. estimated that 9% of seeds survive and emerge as seedlings in the following year. Using the fecundity estimates in the first row of the matrix, and dividing each number in the table above by 100 (the initial number of marked ginseng plants in each size class) to get the corresponding matrix entries, yields the matrix:

$$\begin{array}{c|cccccc|}
 & \mathbf{S} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\
\mathbf{S} & 0 & 0 & 0 & 1.35 & 13.24 & 18.50 \\
\mathbf{0} & .09 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{1} & 0 & .16 & .25 & 0 & 0 & 0 \\
\mathbf{2} & 0 & .02 & .59 & .52 & .06 & 0 \\
\mathbf{3} & 0 & 0 & 0 & .4 & .85 & .19 \\
\mathbf{4} & 0 & 0 & 0 & .03 & .05 & .74
\end{array}$$

Step 3. Construct an initial population vector. The initial population vector is simply a column of the numbers of individuals currently in each class. For example, if the oystercatcher population in Example 1 above has 40 juveniles, 50 subadults, and 60 adults in year t, the population is described by the vector:

$$\mathbf{n}_t = \begin{bmatrix} 40 \\ 50 \\ 60 \end{bmatrix}$$

The total population size is obtained by adding together the numbers in the vector (so this population has a total of 150 individuals).

Step 4. Project the population by multiplying the projection matrix by the population vector. Projecting the number of individuals in each class one year from now given a current population vector consists of multiplying the projection matrix by the population vector to generate a new population vector. Multiplication is achieved as follows. For each row of the matrix, take the product of the first entry of that row and the first entry of the vector, add that to the product of the second entry of that row and the second entry of the vector, and repeat for all

(continued on page 36)

BOX 4.1 *(continued)*

other entries of the row and vector. The result is the first entry in the new population vector. Repeat this process for all other rows of the matrix. For example, to project the oystercatcher population, we multiply the matrix **M** (see equation 4.1 above) by the population vector n_t:

$$\begin{array}{ccccccc} \mathbf{M} & \text{x} & \mathbf{n_t} & = & \mathbf{n_{t+1}} & & \\ \begin{bmatrix} 0 & 0 & .054 \\ .665 & 0 & 0 \\ 0 & .724 & .95 \end{bmatrix} & \text{x} & \begin{bmatrix} 40 \\ 50 \\ 60 \end{bmatrix} & = & \begin{bmatrix} (0\text{x}40)+(0\text{x}50)+(.054\text{x}60) \\ (.665\text{x}40)+(0\text{x}50)+(0\text{x}60) \\ (0\text{x}40)+(.724\text{x}50)+(.95\text{x}60) \end{bmatrix} & = & \begin{bmatrix} 3.2 \\ 26.6 \\ 93.2 \end{bmatrix} \end{array}$$

Note that the total population decreased from 150 individuals in year t (the sum of the entries in n_t) to 123 individuals in year t+1. Thus we use information on vital rates embodied in the projection matrix M, and census figures on the numbers of individuals at time t, n_t, to give us a projection of the number of individuals we will find in each class at time t+1, or n_{t+1}. To project farther into the future, we simply continue to multiply the matrix by the newest population vector, as shown below for one more iteration:

$$\begin{array}{ccccccc} \mathbf{M} & \text{x} & \mathbf{n_t} & = & \mathbf{n_{t+2}} & & \\ \begin{bmatrix} 0 & 0 & .054 \\ .665 & 0 & 0 \\ 0 & .724 & .95 \end{bmatrix} & \text{x} & \begin{bmatrix} 3.2 \\ 26.6 \\ 93.2 \end{bmatrix} & = & \begin{bmatrix} (0\text{x}3.2)+(0\text{x}26.6)+(.054\text{x}93.2) \\ (.665\text{x}3.2)+(0\text{x}26.6)+(0\text{x}93.2) \\ (0\text{x}3.2)+(.724\text{x}26.6)+(.95\text{x}93.2) \end{bmatrix} & = & \begin{bmatrix} 5 \\ 2.1 \\ 107.8 \end{bmatrix} \end{array}$$

Now the population has declined further, to approximately 115 in year t+2. Ultimately, the population will take on a constant rate of population growth or shrinkage (or in rare cases, reach a constant size), and this rate will correspond to the so-called "dominant eigenvalue" of the matrix, λ ("lambda"; for methods to calculate λ, see Caswell 1989). Values of λ larger than 1 indicate population growth, while values smaller than one indicate population decline. This constant population growth rate will be achieved only if all the rates in the matrix remain constant (i.e., if the model is deterministic). The realized population growth rate that results when matrix elements vary is most easily determined by computer simulation, which may be performed using commercially available PVA software (RAMAS, Alex, Vortex, etc.), or by writing simple programs in mathematical software such as MATLAB or Mathematica.

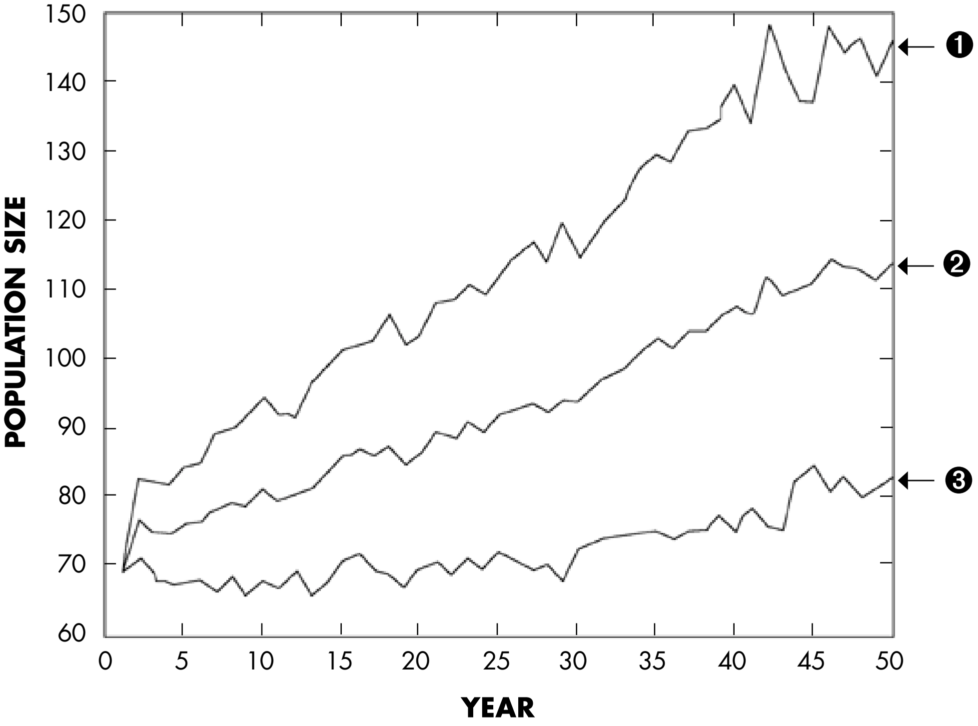
Greater detail on the technical aspects of matrix model formulation and analysis can be found in Caswell (1989).

Stochastic matrix models: incorporating variability and uncertainty

Matrix population models can be either deterministic (all vital rates are constants) or stochastic (vital rates vary over time). The long-term population growth rate in a matrix model is analogous to λ in the simple population growth model (see equation 3.1, Fig. 3.2). If environmental variation is low, or if we have no information regarding its magnitude, a deterministic matrix model is often used; in this case, the long-term population growth rate will be a constant. If

FIGURE 4.1

Trends in the Cap Rock population of the cushion plant *Aletes humilis* (Apiaceae) projected for 50 years using a stochastic projection matrix model. Line ❷ is the average of 100 independent runs of the simulation, and lines ❶ and ❸ are the 95% confidence limits on the population size over this 50 year interval.



environmental variation is incorporated into the model, then the annual change in population size will vary over some range of values, just as it does in the analysis of count data in Chapter 3. Projecting population growth in a stochastic matrix model involves the same rules of linear algebra as in Box 4.1; the only difference is that the matrix itself will vary from year to year. In practice, stochastic matrix projections are usually performed with the aid of a computer. A number of PVA software packages (RAMAS, ALEX, Vortex, etc.) have been developed for this purpose (for a review of these packages, see Lindenmayer et al. 1995). The most commonly used way to present the results of stochastic simulations is to display both the average and the 95% confidence limits for a series of population realizations over some time interval of interest, say the next 20, 50 or 100 years (for example, see Fig 4.1; for methodology see Heyde & Cohen, 1985; Alvarez-Buylla & Slatkin, 1991). In this way, population size projections can be compared with new data from ongoing population censuses; deviations between actual and predicted trajectories would then suggest that changes in vital rates or population structure have occurred, or that there are errors in the model that need to be corrected.

In addition to projecting future population size, we can also use stochastic matrix models to quantify extinction risk. For a deterministic matrix

BOX 4.2 (Key): Methods of Analysis Calculating Fecundity Values

Special care must be taken in calculating fecundity values because matrices are calculated and multiplied on an annual cycle. Thus, fecundity values that only account for the number of offspring produced per female per year will overestimate the true fecundity values because they fail to account for the mortality incurred either by reproductive females or the offspring themselves during each annual cycle. The timing of a census with respect to breeding will have different consequences for the calculation of fecundity. If we imagine that a census just occurred immediately after breeding (a post-breeding census), then to accurately predict the number of offspring that each female in our census will produce before the next census, we need to account for the mortality of those females before the next breeding event. It is likely that immediately prior to the next breeding event, most of the females that are fated to perish over the coming year will have already died. Hence when calculating F_i (the fecundity of class i females) for a post-breeding census, we discount the birth rate b_i by female survival P_i:

Equation 4.2 $F_i = P_i\, b_i$

If instead the census occurs immediately before breeding (a pre-breeding census), we will need to account for the mortality of the offspring in our estimates, because certainly not all the young born immediately after the census will still be alive nearly a year later at our next census. The realized number of offspring per female in a population censused prior to breeding is:

Equation 4.3 $F_i = l_1\, b_i$

where l_1 is the first-year survival rate of offspring.

For many species, censuses will be conducted neither immediately before nor immediately after breeding, in which case both the mortality of offspring from the time of birth until the next census and the mortality of females prior to the next breeding event must be accounted for:

Equation 4.4 $F_i = l_p\, P_i^{1-p}\, b_i$

where p is the amount of time (measured in a fraction of a year) between the time of birth and the next census and l_p is the proportion of offspring that survive for a time interval of length p.

These details are commonly ignored in the construction of matrix models, but must be adhered to in order to calculate fecundity correctly. See Caswell (1989) for more detail.

model, only three outcomes are possible (stasis, growth to infinity or decline to extinction), just as in the simple geometric growth model of equation 3.1 (see Fig. 3.2). If the population is declining deterministically, it is a simple matter to project the population until the number of individuals falls below the threshold, thus determining the predicted time to extinction. For models that incorporate variation in vital rates, extinction is a stochastic event, and its probability will be related both to the average value of λ and to its variance. Just as in the simpler count models, when λ is more variable the risk of extinction tends to rise, even in populations whose average growth rate is greater than 1.

Most commonly, variability is added to a projection matrix in one of two ways: by constructing separate matrices that represent different environmental conditions (which we will call the "multiple matrices" approach) or by constructing a matrix whose entries are random variables (the "variable entries" approach). Below we briefly describe both approaches, and provide references to studies that have used each approach. Two studies that incorporated elements of both approaches addressed the viability and management of populations of endangered plants (Maschinski et al. 1997, Gross et al. 1998).

The multiple matrices approach involves constructing a separate matrix for each of several years of data, or from multiple situations, and then analyzing population trends based on either random draws or specified sequences of these matrices. For example, the extreme weather associated with El Niño events may be reflected in lower-than-average values for both survivorship and fecundity rates. Because of these concerted effects, we might best model this kind of variation by using two kinds of matrices, one for "normal" years, and another for El Niño years. Then, to build projections of the potential fate of the population, we could use the normal matrix as a default, and substitute the El Niño matrix at randomly chosen intervals that correspond to the typical frequency of El Niño events. Because the exact sequence of matrices used will differ among runs, it is necessary to repeat these simulations many times (at least 1000 runs is recommended, see Harris et al. 1987), and summarize the results of these runs in terms of the probable population trajectories, extinction risk, and population growth rate. This approach is also useful for modeling populations subject to periodic events such as controlled burns (Gross et al. 1998) or years of high rainfall that occur at regular intervals (Beissinger 1995); in this case, the sequence of matrices would be fixed, not random.

One limitation of the multiple matrices approach is that it captures only the specific relationships among vital rates that were observed in the years in which the demographic data were collected. For instance, in the example above, the specific combinations of values for survivorship and fecundity measured for El Niño and non-El Niño years will be the only combinations ever used to represent variation in environmental conditions. While the multiple matrices approach has the advantage that it does not stray beyond the observed data, it has the disadvantage that it thereby restricts (perhaps unrealistically) the possible combinations of vital rates that the stochastic model can generate. Using the multiple matrices approach is most appropriate when simulating the effect of discrete events (e.g., fires, hurricanes, floods, logging) on the population of interest, because such events typically result in large, strongly correlated changes in vital rates.

The variable entries approach involves drawing each matrix element at random during each time step of the simulation. Unlike the multiple matrices approach in which a limited number of matrices are employed, the variable entries approach involves the assembly of a unique matrix at each time step, and thus explores a much wider range of combinations of vital rates than does the multiple matrices approach. As above, it is necessary to perform the simulation many times to account for variation among runs. PVAs employing the variable entries approach have been performed for desert tortoises (Doak et al. 1994) and northern spotted owls (Lamberson et al. 1992).

There are two ways that matrix entries can be drawn. First, each entry can be chosen at random from a list of discrete values, most likely those that were actually observed in the data set (note that individual entries are being drawn, not entire matrices as in the multiple matrices approach). Second, entries can be drawn from a continuous range of possible values specified by probability distributions whose means and variances are estimated from the demographic data. When using the second approach, we must first decide which of many probability distributions (e.g., uniform, normal, log normal, beta, gamma) to use. When insufficient data exist to clearly favor one probability distribution over others, one can explore several distributions in parallel models whose results are examined for shared predictions about extinction risk.

There is another major question that must be answered when building a stochastic matrix model: should we incorporate correlations in vital rates into the model? There are two types of correlation to consider: correlations among the different entries in the matrix, and correlations between the values of a single matrix entry in subsequent years. The first type of correlation arises because different vital rates will often change in a concerted fashion. For example, the fecundity values of *all* classes may increase in good years. Similarly, years of high survivorship may also be years of high fecundity, and thus survivorship values and fecundity values should be positively correlated. Furthermore, management efforts directed at a particular life stage may show correlated effects in multiple vital rates associated with that stage. Correlations among vital rates will be automatically built in to the multiple matrices approach, but must be specified in the variable entries approach.

Correlation among the values of a single matrix entry can arise if the environment is correlated among years (e.g. good years may tend to be followed by good years). Environments may be periodic, having exceptional years (good or bad) occurring on a roughly predictable schedule (e.g., 10-year snowshoe hare cycles that affect food availability for predators such as lynx and owls). The decisions on whether and how to incorporate correlation structure into a model can have a large influence on predictions (Groom and Pascual 1998; for examples of how to include correlations among different matrix elements, see Doak et al. 1994 and Gross et al. 1998).

Before moving on to an example of a stochastic matrix model, we warn readers that ***extreme caution should be used in interpreting the results of stochastic matrix models that are based on only a few years of data***. With only 3 or 4 years of demographic data, it is unlikely that the true range of environmental conditions will have been observed, or an anomalous year will be included and given more weight than its actual prevalence in longer time series warrants. Such small samples of data also may show spurious correlations among vital rates (e.g., fecundity and subadult survival varying together by chance when in reality they are unrelated, or vice versa). Further, it is extremely difficult to identify a best-fit probability distribution for a particular vital rate when using small samples of data. It is preferable to work with at least 6-10 years of data, as we discussed in Chapter 3, and to be cautious in accepting results obtained with fewer years of data.

Population status of Aletes humilis*: an example based on real data*

Aletes humilis (Apiaceae) is a rare, endemic plant found in Larimer and Boulder counties, Colorado. The plant is an herbaceous perennial that most commonly inhabits crevices of north-facing granite cliffs, but is also found in some other habitats (Schulz & Carpenter, unpublished manuscript). The plant forms cushions 5-20 cm in diameter and rarely reaches more than 10 cm in height. All populations of the plant are relatively small (50 to a few thousand individuals), and patchily distributed (usually isolated from neighboring populations by several km), yet they are composed mostly of reproductive adults.

The Nature Conservancy, concerned that this rare plant could slip into extinction, created two small preserves for *A. humilis* at Phantom Canyon and Cap Rock in Larimer County. The populations at Phantom Canyon and Cap Rock are small (roughly 1100 and 920 individuals respectively), but appear to be relatively stable. At Phantom Canyon (elevation 6100-6600 ft), an area of granitic outcrops typical of the species' habitat, the populations appear to be constrained by their ability to disperse to empty "safe sites," cracks in the boulders with sufficient soil. Once a plant colonizes a fissure in the rock, it typically grows rapidly to fill the entire opening, and then remains at roughly this size until it dies. Thus, it is possible that in these habitats, populations could be enhanced by distributing seeds to crevices. At Cap Rock (elevation 7500 ft), however, the plant is found beneath ponderosa pine canopy, where its recruitment may be inhibited by the thick layers of pine duff. Gopher disturbance may open sites for *A. humilis* recruitment, although it can also harm established plants (Schulz & Carpenter, in review).

Demographic data on *A. humilis* were collected on over 300 tagged individuals for 7 years at several sites within the Phantom Canyon Preserve and on 76 tagged plants for 5 years in the Cap Rock Preserve (Schultz and Carpenter, unpublished data). Tagged plants were measured each year (to obtain the two-dimensional area of each plant) and any fruiting was noted, as well as any degeneration or death. Few seedlings were noted at either site, but those that appeared were followed in a similar manner.

Constructing a population projection model from the demographic data collected by TNC involves the steps outlined in Box 4.1. Based on the available data, a size-based model is most appropriate. We do not know plant ages, and few discernable life stages are apparent for the plant (although a very simple stage-based model with seeds, seedlings and adults would be possible to construct). Division of the population into size classes makes sense only if the

TABLE 4.1

Sample data from the Phantom Canyon population of *Aletes humilis* (Schulz and Carpenter, unpublished data). Plants were tagged and measured, and re-censused yearly.

	YEAR						
TAG NUMBER	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**
29	188.69	153.94	176.72	132.73	132.73	95.03	dead
51	9.26	9.26	14.19	17.73	19.66	23.76	12.57
56	3.14	21.65	33.18	38.48	28.27	44.18	7.07
114	50.27	50.27	50.27	50.27	70.88	78.54	50.27
223	194.83	153.94	182.65	132.73	63.62	50.27	dead
224	2.41	4.91	7.07	9.62	11.04	dead	

probability of surviving or reproducing differs according to size. Examination of the data on the tagged individuals revealed a few differences among the probabilities of survival and fruiting among individuals of different sizes. At Phantom Canyon, survivorship was lowest among the smallest individuals (2-4 cm^2 in area), low in somewhat larger individuals (4-50 cm^2), and highest in the largest individuals (>300 cm^2 in area). Fruiting was rare in individuals smaller than 50 cm^2, moderate in individuals 50-100 cm^2, and highest in individuals >100 cm^2. At Cap Rock, fewer obvious differences were apparent among individuals in different size classes, particularly among the larger-sized individuals. Nevertheless, in order to compare analyses between these sites, we decided to divide both populations into the same 4 size classes (2 - 4 cm^2; >4 - 50 cm^2; >50 - 100 cm^2; >100 cm^2). Clearly, this step in the process involves a great deal of judgement. It is always prudent to look at the effects of choosing different size categories to see if those decisions greatly influence the predictions of the model. In this case, we explored several different scenarios, but all produced similar results.

The vital rates and transitions among these 4 size classes can be calculated easily from the demographic data. As an example, Table 4.1 presents data from Phantom Canyon on the size of six individuals in each of the 7 years of the census. Plants can survive or die in any given year-to-year transition. If they survive, they can either grow, shrink or stay the same size. By following individual plants, one can trace their history of growth and survival. For example, individual #114 had a size of 50.27 cm^2 in the first year, and thus was categorized as a >4-50 cm^2 plant. Between 1989 and 1992 it survived and stayed in the same size class every year. Then in 1993 it grew to 70.88 cm^2, and thus made a transition into the next size class. Individuals #29 and #223 both shrank from the >100 cm^2 category to the >50-100 cm^2 category, and then died at the end of the sampling period. By tabulating the fates of all individuals

TABLE 4.2

Population matrix for the Phantom Canyon population of *Aletes humilis* for the year 1994-1995.

$$
\begin{array}{l|ccccc|}
 & \text{Seedlings} & \text{2-4 cm}^2 & \text{>4-50 cm}^2 & \text{>50-100 cm}^2 & \text{>100cm}^2 \\
\text{Seedlings} & 0 & 0 & 0.060 & 0.046 & 0.044 \\
\text{2-4 cm}^2 & 0.14 & 0.25 & 0.000 & 0.000 & 0.004 \\
\text{>4-50 cm}^2 & 0.85 & 0.345 & 0.681 & 0.145 & 0.006 \\
\text{>50-100 cm}^2 & 0 & 0.000 & 0.214 & 0.554 & 0.127 \\
\text{>100 cm}^2 & 0 & 0.000 & 0.048 & 0.123 & 0.809
\end{array}
$$

in each size class each year, we calculated the "transition probabilities" or the fraction of plants in each size class that shrink by one size class, remain in the same size class, or grow into the next largest size class over one year. We entered these values directly into the matrix, as described in Box 4.1.

The final decision that must be made to model this population is how to represent the production of new individuals (i.e. fecundity). There are no data available on the demography of seeds once they are dispersed from the parent plant. Most seeds land directly below the parent plant, and if they germinate, are found as seedlings only if the parent dies and the space the parent occupied is made available. Thus to represent newborn individuals, we decided to add to the model a fifth category, "seedlings". Parentage was assigned to the plant occupying the site before the seedling was found, and the size of that individual was used to assign to a particular size class the successful production of a seedling. Plants at Cap Rock produced many more seedlings than those at Phantom Canyon.

From the available data it was possible to produce individual matrices for each pair of years in the data set (e.g., one for the 1994-95 transition: Table 4.2) for both populations, resulting in six matrices for Phantom Canyon and four for Cap Rock.

In projecting these matrices forward to calculate λ and extinction probabilities, we wrote a MATLAB program that took the multiple matrices approach. At each time step, the program selected one of the individual matrices calculated for each year-to-year transition in the data set. Each matrix had an equal probability of being drawn in any time step. We tried using the variable entries method, but with so few years of data with which to fit probability distributions, we felt that our efforts could lead to large inaccuracies with this approach.

Our analyses make different predictions about the viability of the Cap Rock and Phantom Canyon populations. We found an expanding population at Cap Rock (Fig. 4.1), that shows no risk of extinction over then next 50 or even 100 years (the model predicted a 0% probability of extinction in 100 years, and only a 10% probability of falling below 200 individuals).

However, at Phantom Canyon the population appears to be in decline (Fig. 4.2). The model shows a 50% and 100% probability of falling below 45 or 200 individuals, respectively, over the next 100 years (Fig. 4.3).

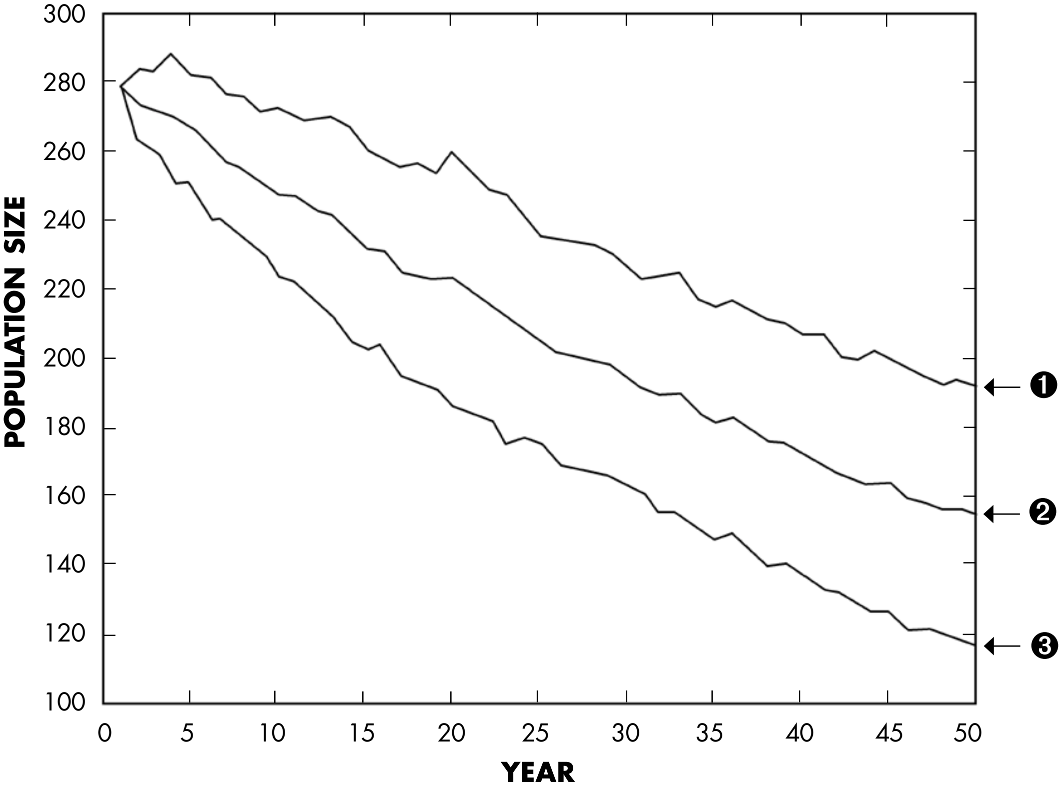
This example illustrates how population viability analyses based upon projection matrices could be used to compare the relative viability of 2 or more occurrences. The models suggest that the Cap Rock population is likely to persist if the estimated vital rates prevail, but that we may need to consider management to improve the viability of the Phantom Canyon population.

"Experiments" with matrix models to determine best management practices

It is often instructive to conduct computer "experiments" to ask what effect particular changes in management may have on population growth rate or extinction risk. We can use the model to ask "how will population viability change in response to the expected change in a vital rate?" For example, if we planted seeds of *Aletes humilis* in empty crevices, with an expectation of doubling seedling recruitment, we could double that value in the matrix and determine how much the probability of extinction would decrease. Such experiments can serve as an

FIGURE 4.2

Population trajectory for the Phantom Canyon population of *Aletes humilis*. The projection based on field data shows a nearly 50% decline over 50 years. Line ❷ is the average trajectory taken over 100 simulations, and lines ❶ and ❸ are the 95% confidence limits.

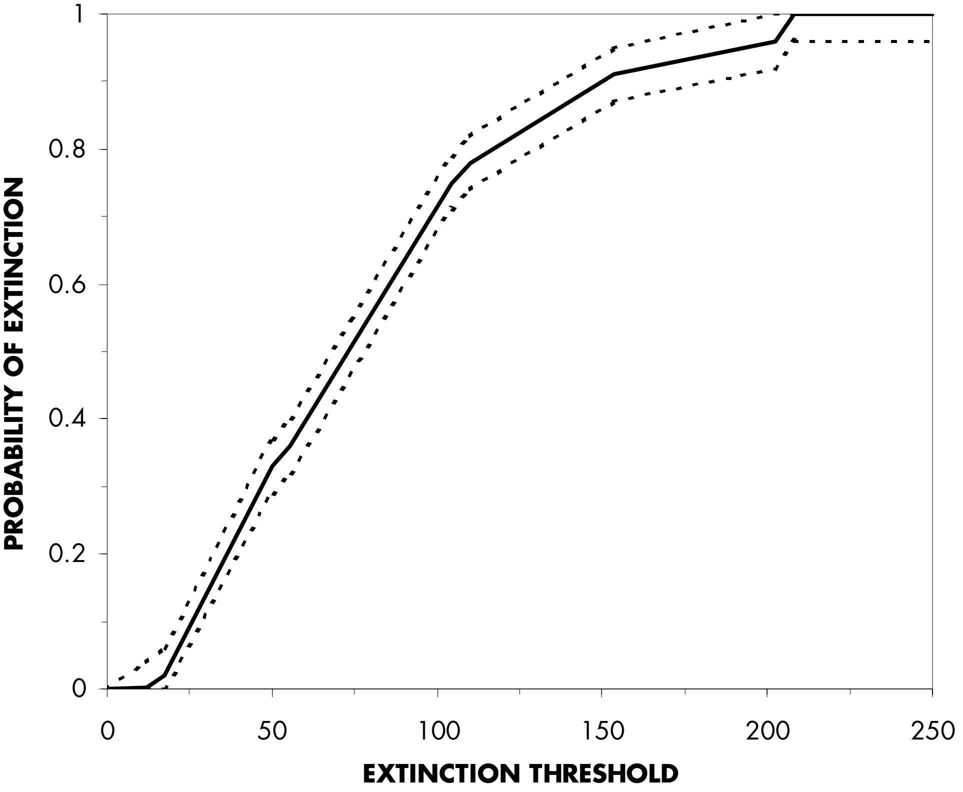


indicator of the potential utility of specific management strategies. We can also ask the reverse question: "in order to have an extinction risk below 5% over 100 years (or an extinction risk of 100% over 5 years for an exotic we wish to control), how large must a given matrix element or elements be?" For example, we could determine how much greater seedling recruitment would have to be to obtain a positive growth rate for the Phantom Canyon population of *Aletes humilis*, and then ask whether such a rate is practically achievable. Because recruitment is limited by dispersal of seeds to sites not currently occupied by adults, we asked whether doubling seedling recruitment via seedling transplants performed at various time intervals would be sufficient to divert the population from its projected decline. After several trials, we learned that only if we planted seedlings every year would the population begin to increase, although a few of the simulation runs still decreased (see Fig. 4.4). After performing additional analyses to be sure of our result, and once sure that some intervention is warranted, we could then assess whether this type of intervention is feasible. This approach can be extremely helpful for guiding management decisions in many cases.

As we pointed out at the beginning of this

FIGURE 4.3

Extinction risk profile for the Phantom Canyon population of *Aletes humilis*. The probability of dropping below a threshold population size by 100 years is shown in the bold line, and the 95% confidence envelope is shown in the dotted lines.

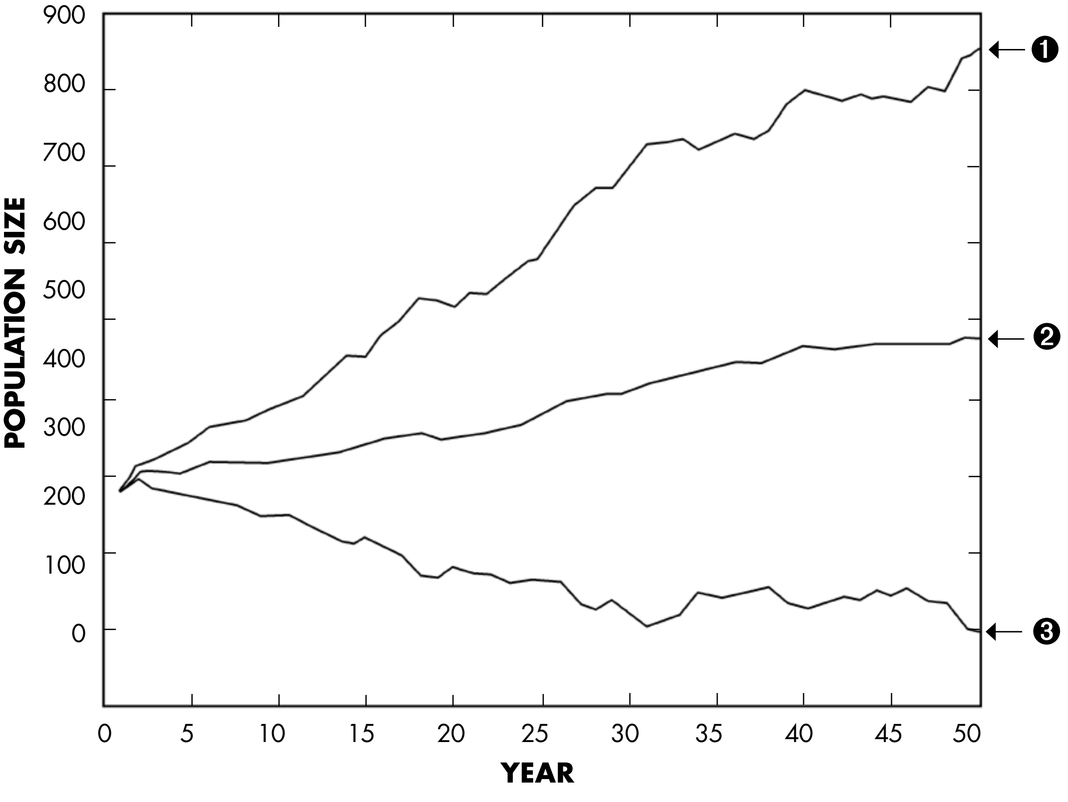


chapter, an informative use of matrix models is to identify which life stage or which vital rates have the greatest influence on overall population growth. We can explore this question through a series of computer "experiments" akin to the ones we just described. The idea is to systematically increase or decrease by a small amount each of the elements in the matrix, one at a time, and then examine the resulting rates of population growth (for an example, see Crouse et al. 1987). In this way we could ask, for example, whether it is more effective to manage populations of *A. humilis* by transplanting seedlings (i.e., enhancing fecundities) or reducing trampling of established plants (i.e., increasing survival of adults).

Finally, as our analysis of seedling transplantation for *A. humilis* indicates, one can use a matrix model to ask how frequently to apply a management regime. For example, if we have information on populations under two types of management, say burned and unburned plots, we could use a model to examine the effect of different burning cycles on population growth rates and thus determine the optimum burning cycle (see Gross et al. 1998).

FIGURE 4.4

A projection model "experiment" of doubling seedling recruitment via transplantation each year shows the average population increases in size over 50 years (although some replicates still decline - see lower confidence limit). Line ❷ is the average trajectory taken over 100 simulations, and lines ❶ and ❸ are the 95% confidence limits.



Prospects for the use of matrix-based PVAs in TNC decision-making

Often, structured models are seen as impractical for application in TNC planning efforts due to the amount of effort that must be put into gathering sufficiently detailed data for analysis. At a bare minimum, three years of effort must be dedicated to marking, recapturing, and measuring individuals in a population to even construct a stochastic matrix model. However, three years of data would be insufficient to characterize the true degree of variation in a population's vital rates unless they are highly invariant—put another way, the accuracy of these models will depend on how well a few years of data capture the range of variation the population experiences. For most populations many more years of data are required, especially when evaluating populations subject to long-term trends in vital rates or changes in management practices. In this sense, *projection matrix models are likely to require as many years of data as the count-based approaches, and of course much more effort is required to gather these more detailed data*. Because of their data requirements, matrix-based approaches will only rarely be of use in ranking EOs.

Yet, while it is certainly true that it will not be possible to perform projection matrix analyses for very many species or locations, there are cases in which this type of modeling will be more informative and useful than less data-intensive modeling efforts. Threatened and endangered species of particular prominence, species with large effects on ecosystems or on other species ("keystone" species), indicator species, umbrella species, or surrogates for a species of particular concern are all likely candidates for this type of PVA. When such species are long-lived, it will be particularly valuable to adopt a matrix-based approach to more accurately predict the population consequences of changes in vital rates and population structure. Results of matrix-based PVAs for taxa of special interest may help us to set goals for the size of particular reserves, or to judge the overall effect of a management practice. The analysis of matrix projection models can also help us to identify which life stages or vital rates most influence population growth rate, information that can be used to focus management and monitoring efforts. Similarly, by using the model to simulate different threats or by contrasting sites with different threats, it may be possible to evaluate which threats are most serious, and guide rescue efforts. Finally, because it is possible to test probable outcomes of different strategies by carrying out detailed "experiments" with the model, matrix-based approaches can serve an integral role in an adaptive context of reserve management.