



## Extinction in Fragmented Habitats Predicted from Stochastic Birth–death Processes with Density Dependence

TORMOD VAALAND BURKEY\*

*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, U.S.A.*

*(Received on 20 April 1998, Accepted in revised form on 5 May 1999)*

Habitat loss, the reduction of the habitat area available, is known to greatly reduce resident species' expected time to extinction. This process is widely recognized, if not adequately understood or quantified except in very simple models. However, it is not well understood how the time to extinction will change if the remaining habitat is distributed across a set of smaller, isolated patches, instead of being left in one single, continuous tract. The effect of habitat fragmentation on population persistence under demographic stochasticity has not been resolved. Specifically, it is not known whether a single large population will persist longer than an aggregate set of several smaller populations (with the same total size). Analytical studies of birth–death processes typically report the mean time to extinction for a single population as a function of the maximum population size, but omit higher moments. To estimate the overall persistence time, or the probability of extinction as a function of time, for a set of small populations, the entire distribution of extinction times must be known for a single population of each size. Knowing all the moments of the distribution of extinction times is not adequate, unless one can reconstruct the distribution from them. Here I analyse stochastic birth–death processes with linear density dependence in per capita birth and death rates, and obtain analytical expressions and numerical solutions for the distribution of extinction times in both subdivided and continuous populations. This is a single-species model that deals with demographic stochasticity only, and assumes independence of extinction events in different patches. These assumptions are relaxed elsewhere. Habitat fragmentation, even without any loss of overall area, has a great and detrimental effect on the persistence time of populations across all temporal and spatial scales. The effect is similar across spatial scales, but shifted in time—larger populations take longer to go extinct but the extinction risk relative to that of a smaller or more fragmented population is the same across spatial scales for the available habitat.

© 1999 Academic Press

### 1. Introduction

Diamond (1975) and Wilson & Willis (1975) first suggested that a single large reserve was a better

reserve strategy than a set of small reserves of the same total area. The SLOSS debate (Single Large Or Several Small reserves) that ensued was distorted by a confusion of goals (see Burkey, 1989). Several authors (Simberloff & Abele, 1976, 1982; Abele & Connor, 1979; Gilpin & Diamond, 1980; Higgs & Usher, 1980; Higgs, 1981; Järvinen, 1982; Margules *et al.*, 1982; Quinn & Harrison,

\* Present address: United Nations Development Programme (UNDP), P.O. Box 1041, Port Moresby, Papua New Guinea. E-mail: [tvburkey@alumni.princeton.edu](mailto:tvburkey@alumni.princeton.edu).

1988) held that one could increase the number of species (initially) present in the reserve system by strategically establishing several small reserves with slightly different species compositions. Others (Wilson & Willis, 1975; Diamond, 1975, 1976; Terborgh, 1976; Fahrig & Merriam, 1985) held that the risk of extinction was lower in a single large reserve than in a set of small reserves (of the same total area), though this has not been adequately demonstrated (but see Wright & Hubbell, 1983; Forney & Gilpin, 1989; Burkey, 1989, 1996). Minimizing per species extinction probabilities is the most appropriate goal for conservation purposes (Diamond, 1976; Terborgh, 1976; Whitcomb *et al.*, 1976). Some trivial arguments for why the risk of extinction may be lower in a single large reserve can be brought to bear (e.g. each of the small areas may not be large enough to protect a critical watershed, or to cover the home range of large, widely roaming animals), but the answer is by no means obvious in less contrived scenarios. We still cannot say with confidence whether a population of 400 individuals in a continuous habitat area of size  $A$  (large enough that the population does not go rapidly and deterministically to extinction) will persist longer or shorter than two populations of 200 individuals each in two isolates of size  $A/2$  each. While we may suspect that the answer may depend on the species in question and the spatial scale of the habitat areas, the spatial correlation of environmental variation and disturbances, and perhaps other factors as well, we cannot say what level of environmental fluctuations or what degree of spatial autocorrelation of such disturbances would change the answer in one direction or the other.

Fluctuations in population size due to random events of births and deaths—demographic stochasticity—is widely recognized as an important cause of extinction in finite populations. Finite populations are usually modelled as stochastic birth–death processes and analysed as Markov chains (e.g. MacArthur & Wilson, 1967; Goel & Richter-Dyn, 1974; Goodman, 1987; Mangel & Tier, 1993). Even in simple cases without environmental stochasticity, and just demographic stochasticity, it is not clear what the effect of habitat fragmentation on extinction risk is (assuming the overall area remains constant).

Density dependence is critical to the study of relative extinction rates in fragmented and unfragmented landscapes under demographic stochasticity (Burkey, 1989). Without density dependence, the fates of all individuals are independent regardless of whether or not they are isolated in different habitat fragments, and fragmentation becomes irrelevant to the extinction process [see e.g. the models by Bailey (1964, p. 95) or Järvinen (1982)]. Much work has been done on stochastic birth–death processes (e.g. Feller, 1939; Bailey, 1964; MacArthur & Wilson, 1967; Goel & Richter-Dyn, 1974; Pielou, 1977), but explicit solutions have not been found for birth–death processes with biologically realistic density dependence.

We do not know how extinction times for single populations are distributed, much less what the aggregate distribution of extinction times for a set of two or more populations may be. Typically, authors have resorted to finding the moments of the time to extinction for a single population of a given size, and further analysis has been restricted to the first and second moments. Unfortunately, assessing the overall persistence time for a set of populations requires the entire distribution of extinction times for single populations, not just individual moments. To estimate the persistence of a set of isolated, finite, reserves under demographic stochasticity, we need either the probability of extinction as a function of time for a population in a reserve of that size or the complete distribution of times to extinction. The aggregate distribution of overall extinction times for a set of  $n$  populations is the maximum of  $n$  realizations drawn from the distribution of extinction times in a single population of the relevant size. If the distribution of extinction times have a very long tail, we might expect that a large set of small populations, which have many shots at “hitting” this long tail could persist longer than a single large population.

The mean time to extinction, which is commonly reported from analytical extinction models, is a notoriously bad metric of the extinction process (Quinn & Hastings, 1987) and the overall persistence for sets of identical populations cannot be calculated from the mean. The median time to extinction would be a lot more useful, since from it we can calculate the probability of

extinction for a set of independent, identical populations at that particular time. Even all the moments of the distribution of extinction times are not adequate for the task of comparing extinction times for a set of small populations with the corresponding single large population—unless a way can be found to reconstruct the distribution from the moments. Closed-form solutions for the moment-generating functions have not been found and the distributions are not easily reconstructed. The most useful metric for comparing persistence times for fragmented and unfragmented systems is probably the probability of extinction as a function of time for populations of given sizes. Under the assumption that isolated subpopulations are independent, the probability of overall extinction for several identically distributed populations at any given time is simply the  $n$ -th power of the probability that a single such population goes extinct by that time.

Here I explore the effects of habitat fragmentation on extinction probabilities under demographic stochasticity, under the assumption that the persistence times of isolated subpopulations are independent random variables with the same distribution. To isolate the effect of habitat subdivision from that of habitat loss *per se*, I will compare the extinction proneness of a single large population in a large area with a set of

the behaviour of such a system. This is a single-species model that deals with demographic stochasticity only, and assumes independence of extinction events in different patches. I show exact solutions of this process, which for moderate to large populations rely on numerical solution of the critical equations for the populations. Multi-species systems with predator-prey interactions, linked subpopulations, and environmental as well as demographic stochasticity are modeled elsewhere (Burkey, 1996).

**Model**

Extinction under demographic stochasticity with density dependence can be simulated (Burkey, 1989, 1996), and a stochastic birth-death process can be explored numerically as a branching process on a computer. Here I introduce an exact solution of a density-dependent stochastic birth-death process.

Consider the system

$$\frac{d\mathbf{P}}{dt} = \mathbf{M}\mathbf{P}$$

where  $\mathbf{P}$  is an  $N_{max} \times 1$  vector with elements  $P_i$  representing the probability that a population is at size  $i \in [1, N_{max}]$  at time  $t$ , and

$$\mathbf{M} = \begin{bmatrix} -d_1 - b_1 & d_2 & 0 & \dots & 0 \\ b_1 & -d_2 - b_2 & d_3 & 0 & \vdots \\ 0 & b_2 & \ddots & & 0 \\ \vdots & & & & d_{N_{max}} \\ 0 & 0 & \dots & b_{N_{max}-1} & -d_{N_{max}} - b_{N_{max}} \end{bmatrix}$$

smaller populations isolated in smaller areas that add up to the same area overall as the single large area. I will assume the initial population size, maximum population size and carrying capacity of each habitat area scales linearly with its size (area). I present an explicit model of a stochastic birth-death process with linear density dependence in per capita birth and death rates to study

where  $d_i$  and  $b_i$  are death and birth rates ( $\neq$  zero), respectively, at population size  $i$ . Since  $\mathbf{P}$  does not contain the probability that the population size is zero,  $P_0(t)$ , the decay of this system over time represents the probability of extinction. All the eigenvalues have negative real parts. From Gerschgorin's Theorem (Råde & Westergren, 1990), we know that the eigenvalues are all within

the union of a circle with radius  $b_1$  around  $-d_1 - b_1$  and the circle with radius  $r = \max_{i=2}^{N_{max}}(d_i + b_i)$  around  $-r$ . It is easy to show that zero is not an eigenvalue.

Let  $\lambda_i$  and  $\varepsilon_i$  be the  $i$ -th eigenvalue and eigenvector of  $\mathbf{M}$ , respectively. We have  $\mathbf{P} = c_1 \varepsilon_1 e^{\lambda_1 t} + c_2 \varepsilon_2 e^{\lambda_2 t} + \dots + c_{N_{max}} \varepsilon_{N_{max}} e^{\lambda_{N_{max}} t}$ .

Set

$$Q_i = c_i \sum_{j=1}^{N_{max}} \varepsilon_{ij},$$

where  $\varepsilon_{ij}$  is the  $j$ -th element in the  $i$ -th eigenvector. Then the probability that the population is of a size between one and  $N_{max}$  at time  $t$ , is

$$P_T(t) = P_1(t) + P_2(t) + \dots + P_{N_{max}}(t) \\ = \sum_{i=1}^{N_{max}} Q_i e^{\lambda_i t}.$$

The cumulative distribution function for the time to extinction is  $F(t) = 1 - P_T(t)$ . The probability density function for the time to extinction  $f(t) = \sum_{i=1}^{N_{max}} -Q_i \lambda_i e^{\lambda_i t}$ .

Consider a given habitat area, defined by its carrying capacity,  $K$ , and its maximum population size,  $N_{max}$ .  $K$  is defined as the size at which the population is expected to just replace itself, and  $N_{max}$  is a population ceiling that cannot be exceeded. Then  $d_K = b_K$  ( $b_{N_{max}} = 0$ ). Assume that  $K$  and  $N_{max}$  are proportional to the area of the available habitat. If we divide the area into  $n$  isolated pieces without any loss of area, we get  $n$  patches each of carrying capacity  $K/n$  and maximum population size  $N_{max}/n$ . In a set of  $n$  small populations,  $d_i$ ,  $b_i$ ,  $\lambda_i$ ,  $\varepsilon_i$ ,  $c_i$  and  $Q_i$  must all be functions of the size of the habitat patch—which depends on  $n$ . Denote the distribution of extinction times for a single such patch as  $f_n(t)$ , and its cumulative distribution function as  $F_n(t)$ . Then the cumulative distribution function for the overall extinction time of the set of  $n$  independent small populations is  $F(t) = F_n(t)^n$ , and the probability density function is

$$f(t) = n F_n(t)^{n-1} f_n(t) \\ = n \left[ 1 - \sum_{i=1}^{N_{max}} Q_i e^{\lambda_i t} \right]^{n-1} \left[ \sum_{i=1}^{N_{max}} -Q_i \lambda_i e^{\lambda_i t} \right].$$

We can find eigenvalues and eigenvectors of  $\mathbf{M}$  and solve for the  $c$ 's, given initial conditions and functions for  $d_i$  and  $b_i$ .

### Numerical Solution by Iteration

In a discrete-time model, let  $d_N = d(N)N \Delta t$  and  $b_N = b(N)N \Delta t$  be the probabilities of a death and a birth, respectively, in  $\Delta t$  if the population size is  $N$ . Let the patch-specific carrying capacity,  $K$ , be the population size at which a population isolated in a given area has expected number of births equal to the expected number of deaths, i.e.  $d(K) = b(K)$ . Consequently,  $d$  and  $b$  are functions of  $K$ , or analogously, the area of the habitat patch. Making the time step,  $\Delta t$ , so small that the probability of multiple events in  $\Delta t$  is negligible, we have

$$P_{i,K}(t) = P_{i,K}(t - \Delta t)(1 - d_{i,K} - b_{i,K}) \\ + P_{i+1,K}(t - \Delta t)d_{i,K} + P_{i-1,K}(t - \Delta t)b_{i,K}.$$

The probability of extinction by time  $t$  is

$$P_{e,K}(t) = P_{e,K}(t - \Delta t) + P_{1,K}(t - \Delta t)d_{1,K},$$

where  $d_{1,K} = d_K(1) \Delta t$  is the probability of a death if the population size is one. The cumulative distribution of extinction times is

$$F_K(t) = P_{e,K}(t)$$

and the distribution of extinction times is

$$f_K(t) = P_{e,K}(t) - P_{e,K}(t - \Delta t).$$

The cumulative distribution of extinction times for a set of  $n$  independent small populations, each of size  $K/n$  is

$$F_{n,K/n}(t) = F_{1,K/n}(t)^n = P_{e,K/n}(t)^n,$$

and the distribution of extinction times for a set of  $n$  small populations is

$$f_{n,K/n}(t) = \frac{dF_{1,K/n}(t)^n}{dt} \\ = n F_{1,K/n}(t)^{n-1} \frac{dF_{1,K/n}(t)}{dt} \\ = n F_{1,K/n}(t)^{n-1} f_{1,K/n}(t).$$

Choosing  $d(N)$  and  $b(N)$  and an initial distribution of population sizes, the system can be solved by iteration. In the following I will assume that  $P_K(0) = 1$ , i.e. all populations start at their carrying capacity.  $d(N)$  and  $b(N)$  change with available habitat area such that  $d(0)$  and  $b(0)$  are constant irrespective of area, and  $d(K)|_{K=K'} = b(K)|_{K=K'}$  where  $K'$  is the area-specific carrying capacity.

For instance, let  $d(N)$  and  $b(N)$  be linear in  $N$ ,  $d$  increasing with  $N$  and  $b$  decreasing with  $N$ . Make  $b(N_{max}) = 0$  so that the population size never exceeds  $N_{max}$ . Hence,  $d(i)$  and  $b(i)$  in a patch of  $K = K'/n$  must have slopes  $n$  times the slopes of the comparable rates in a patch of  $K = K'$ . This is the way in which density dependence is incorporated in all the calculations presented in this paper.

## Results

There is extremely close agreement, at all time-scales, between the exact solution of the density-dependent birth–death process and the approximation based on the dominant eigenvalue,  $1 - e^{-\lambda_1 t}$ . It is well known that the dominant eigenvalue of such systems has the strongest effect on the solution and tends to dominate it after many iterations (e.g. Gantmacher, 1959; Fuller, 1962; Edelstein-Keshet, 1988). In the present model, the approximation tends to be quite good from the start. In many cases, the approximation is so close that the two cannot be told apart when showed together [Fig. 1(a)]. Displaying the difference between the probability of extinction as a function of time in the two models is therefore often necessary to show the two together (Fig. 1).

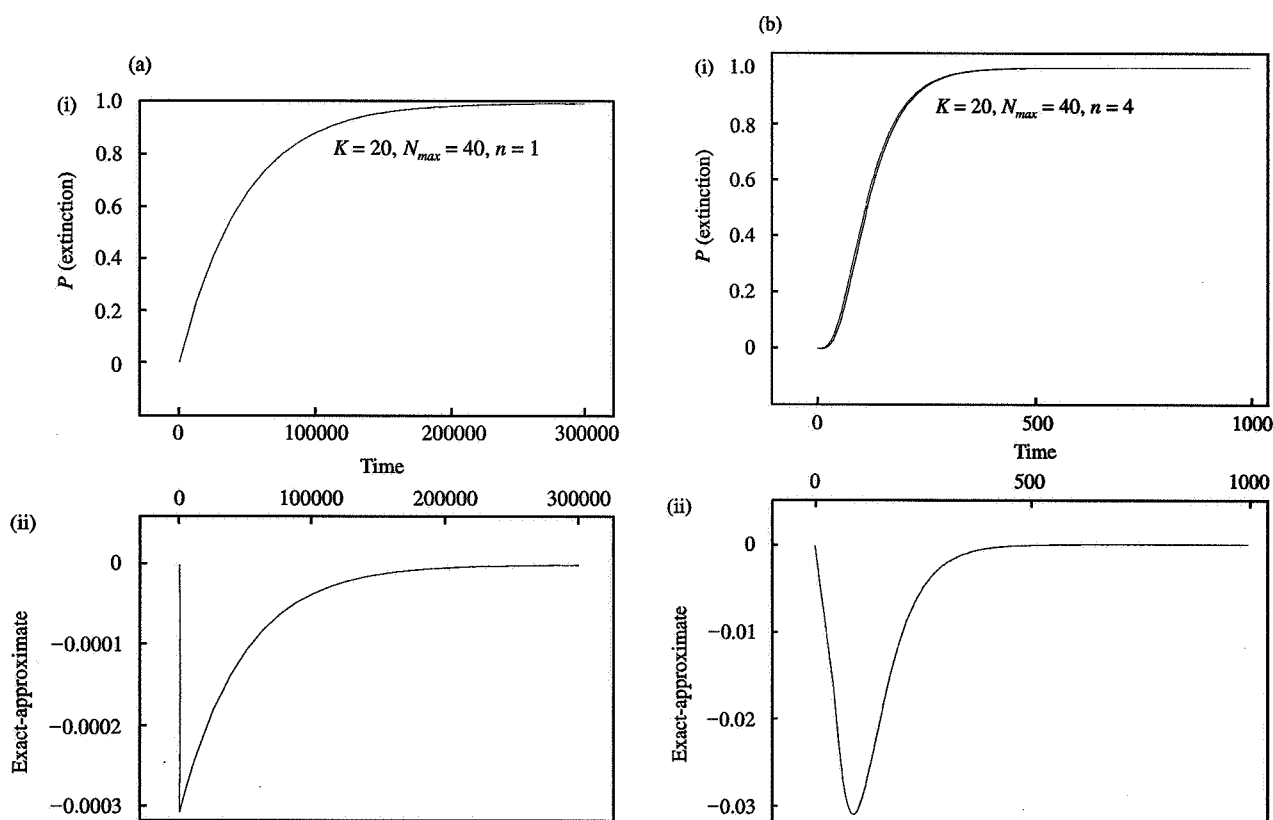


FIG. 1. The cumulative distribution of extinction times (or the probability of extinction as a function of time), for the stochastic birth–death process with linearly density dependent demographic rates—exact and approximate solution. (i) approximation and exact solution together, (ii) the difference between the exact and the approximate solution.  $K = 20$ ,  $N_{max} = 40$  in an unfragmented area (a), or the same area divided into four independent fragments (b);  $d_0 = 0.2$ ,  $b_0 = 0.5$ . The close proximity between the exact solution and the approximation is due to the fact that the dominant eigenvalue is huge compared to the subsequent eigenvalues.

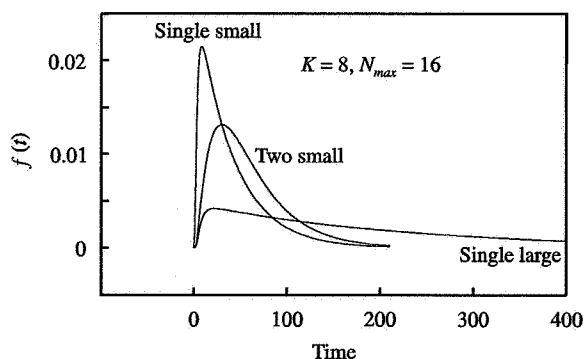


FIG. 2. The distribution of extinction times for a single large population in an area with  $K = 8$ ,  $N_{max} = 16$ , a single small population of  $K = 4$ ,  $N_{max} = 8$ , and two such small populations with a combined size equal to that of the single large population. The single large population persists much longer than both a single small population and a set of two small populations which together have the same amount of available habitat as the single large population.  $d_0 = 0.2$  and  $b_0 = 0.5$ . The distributions resemble gamma distributions with  $\alpha > 1$ , and the shape for a pair of populations or larger single populations resemble gamma distributions with higher values of  $\beta$ .

Small populations go extinct sooner than large populations (this is well known from other studies: MacArthur and Wilson, 1967; Goel and Richter-Dyn, 1974; Ludwig, 1974; Burkey, 1989, 1995, 1997). A fragmented population goes extinct sooner than an unfragmented population of the same overall size (Fig. 2). The distribution of extinction times is approximately gamma distributed with  $\alpha > 1$  at all spatial scales, and the shape for a pair of populations or larger single populations changing as by increasing  $\beta$  (the gamma distribution can be written as  $f(t) = t^{\alpha-1}e^{-t/\beta} / \beta^\alpha \Gamma(\alpha)$ , where  $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t} dt$  and  $\alpha\beta$  is the mean while  $\alpha\beta^2$  is the variance). A single large population has a distribution of extinction times that resembles a gamma distribution with larger  $\beta$  value than both a set of small populations and a single small population (Figs 2 and 3). The maximum likelihood estimate for  $\beta$  increases exponentially with the magnitude of  $K$  for both single populations and pairs of populations. On a smaller scale,  $\alpha$  decreases with  $K$ — $\alpha$  approaches unity asymptotically but never reaches this value, where the gamma distribution would have switched to a negative exponential distribution. The combined result of these two trends is that the mean of the distribution,  $\mu = \alpha\beta$ , in-

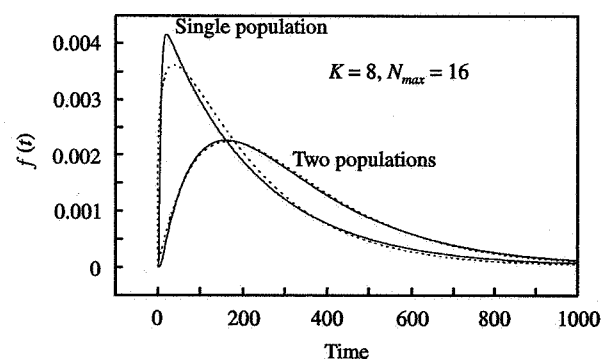


FIG. 3. The distribution of extinction times for a single population in an area with  $K = 8$ ,  $N_{max} = 16$ , and the aggregate distribution of extinction times for two such populations, showing the exact numerical solution and the best fitting gamma distribution for each. Demographic parameters as in Fig. 2. The best fitting gamma distributions, as estimated by simulated annealing, have parameter estimates  $\alpha = 1.22$ ,  $\beta = 258.74$  and  $\alpha = 2.04$ ,  $\beta = 670.35$ , for the single population and the set of two populations, respectively. Numerical solution (—); best-fit gamma distribution ( $\cdots$ ).

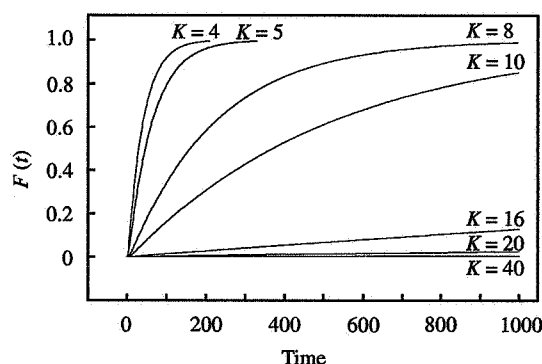


FIG. 4. The probability of extinction as function of time for habitat areas of different sizes (i.e. the accumulative distribution of extinction times).  $d_0 = 0.2$  and  $b_0 = 0.5$ . The probability of extinction for  $n$  reserves of a given size, assuming independence, is the  $n$ -th power of the probability of extinction for an area of that size.

creases exponentially with  $K$ . The  $\alpha$  values for a pair of populations are larger than for a single population, making the distribution more rounded and dragged out.

The inclusion of greater habitat area available to a population vastly enhances its ability to persist through time (Fig. 4).

Fragmented systems go extinct much sooner than continuous systems of the same size. Figure 5 compares the distribution of extinction

times and the cumulative distribution times in single large, two small and a single small (half the size of the large) reserve at two spatial scales. The large reserve in Fig. 5(a) is identical to the small reserve in Fig. 5(b). Note the increase in the critical temporal scale, and the similarity of the fragmentation effect across spatial scales. Apart from an initial relatively flat section, the cumulative distribution of extinction closely resembles the function  $1 - e^{-ct}$  (i.e. 1 minus a negative exponential distribution with  $c$  a positive constant). This follows from the close approximation of the exact solution by the model based on the dominant eigenvalue only. The flat section near  $t = 0$  follows from the fact that a stochastic birth-death process can only go extinct after  $N(0)$  time steps (albeit with  $\Delta t$  small), and that the number of paths to extinction is restricted at small time-scales. The number of combinations of birth and death events that lead to extinction by a given time increases dramatically at longer

time-scales. The probability of extinction for a set of small populations remains flat longer since it constitutes a higher power of a distribution with small values at small time-scales. Larger systems take longer to go extinct, but the comparison of a single large population with a single small (half the size of the large population), and a pair of two such populations, is consistent across spatial scales. In all solutions shown here, the model is initialized so that the maximum attainable population size is twice the carrying capacity,  $K$  (the population size at which the expected number of births per time interval equals the expected number of deaths).

The uniform effect of habitat loss and fragmentation across spatial scales becomes evident on a logarithmic time-scale (Fig. 6). Figure 6 shows extinction probabilities for single populations (dotted lines) and pairs of populations (solid lines) of a given size, and the persistence of a single large population can be compared with

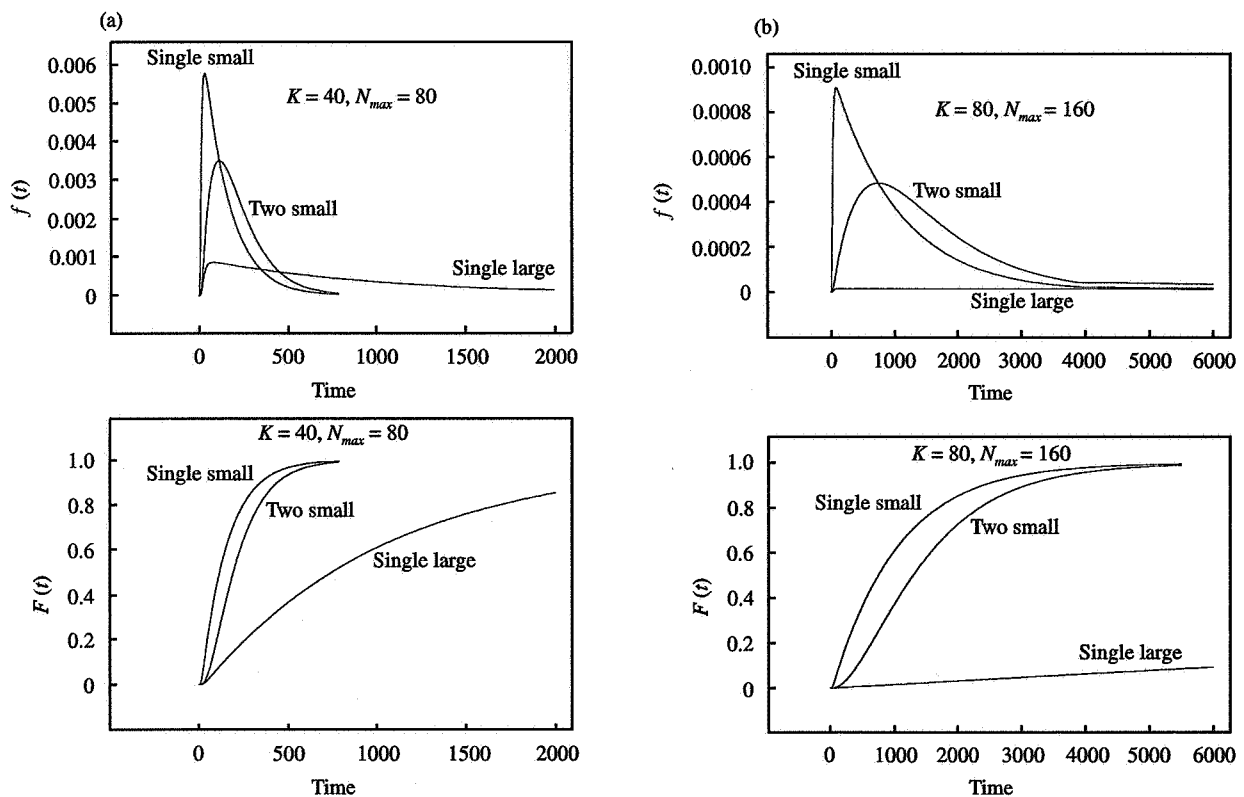


FIG. 5. Distribution of extinction times and cumulative distribution of extinction times for a single large population, a single small population half the size of the large population, and two such small populations with a combined size equal to that of the single large population.  $d_0 = 0.3$  and  $b_0 = 0.5$ . (a)  $K = 40$ ,  $N_{max} = 80$  in the large population, (b)  $K = 80$ ,  $N_{max} = 160$  in the large population. A single large population in (a) is identical to the single small population in (b).

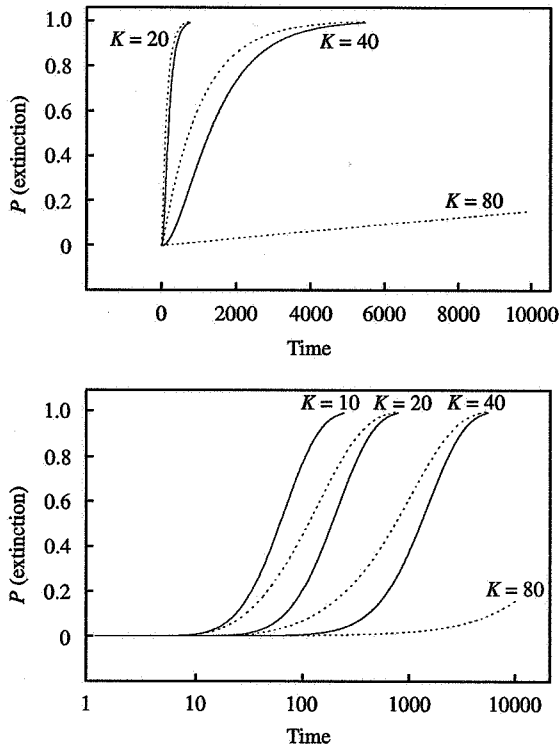


FIG. 6. The probability of extinction as a function of time for single populations and pairs of populations in habitat areas of different sizes. Compare single populations of size  $K$  with a set of two populations each of size  $K/2$ . Demographic rates are identical to those in Fig. 5. The effect of fragmentation is recognized across spatial scales, at different temporal scales. Single population ( $\cdots$ ); two populations (—).

that of two independent populations of the same total size. The median time to extinction is the time at which a line through 0.5 on the ordinate axis intersects each curve. Single, continuous populations persist *much* longer than the comparable fragmented systems.

Figure 7 makes more explicit the comparison of a continuous population with populations fragmented to varying degrees. Subdividing a habitat area into increasingly smaller pieces (without any loss of overall area, as before) causes more rapid extinction of the focal species. The more fragmented a system is, the shorter its persistence time. This effect is strong across differential spatial scales—note the logarithmic time-scale of the comparison.

The decline in persistence time continues with increasing fragmentation across all degrees of

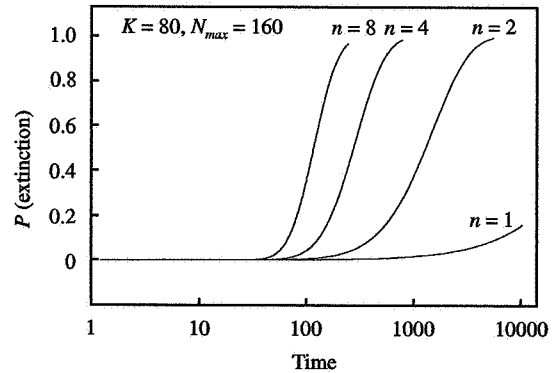


FIG. 7. The effect of habitat subdivision on the probability of extinction as a function of time for an area of size  $K = 80$ ,  $N_{max} = 160$  split into one, two, four or eight sub-populations. Demographic rates are identical to those in Fig. 5.

fragmentation. The median time to extinction (the time until the probability of extinction reaches 50%) increases exponentially with  $K$  (Fig. 8). This is also true for pairs of populations of the same size, as well as increasingly fragmented reserve systems. The rate of increase in persistence time with available habitat is so large that a less fragmented system always persists longer than a more fragmented system of the same total size. Scaled logarithmically, the difference between median extinction times for more fragmented systems and less fragmented systems of the same size is identical at all spatial scales (see Fig. 8). Compare, for instance, the persistence time for a single reserve with that of two reserves each half the size and four reserves each a quarter the size of the continuous reserve. These results are robust to a variety of different kinds of density dependence that can be incorporated into the model [different shapes of  $d(n)$  and  $b(n)$ ]. Different shapes and values for  $d(n)$  and  $b(n)$  merely alter the critical time-scale of comparisons.

### Discussion

In this stochastic birth–death model with density dependence in demographic rates, populations in fragmented landscapes go extinct much sooner than unfragmented population in continuous landscapes of the same overall size. Median extinction times increase exponentially with available area, but the effect of fragmentation is



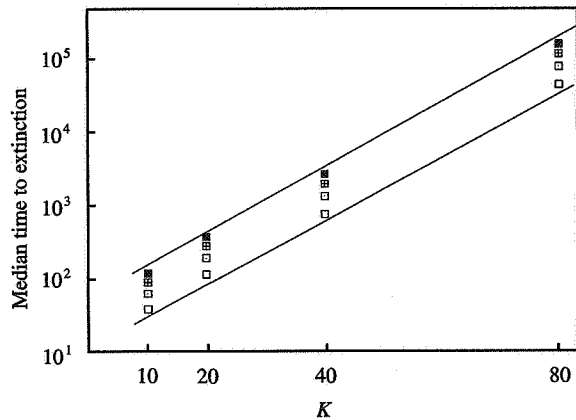


FIG. 8. Median time to extinction (the time until the probability of extinction reaches 0.5) as a function of  $K$ , for single populations, for a pair of populations, four populations and eight populations, each of size  $K$ . The maximum population size is two times  $K$ . Demographic rates as in Fig. 5. The linear relationship and the constant distance between median extinction times for one population and two populations indicates that the effect of fragmentation is identical across spatial scales—only the relevant temporal scale changes. Compare for instance the medium time to extinction for a single reserve of size 40 with two of size 20, or four of size 10. In this figure, we can also see the combined effect of habitat loss and subdivision of the remaining habitat. Time can be thought of in years, but the actual timing of extinction will be affected (reduced) by many other factors which the model does not consider, and will vary from species to species. Number of population: one ( $\square$ ); two ( $\square$ ); four ( $\boxplus$ ); eight ( $\blacksquare$ ).

the same across spatial scales. Applied to the design of nature reserves for the conservation of endangered species, these results indicate that nature reserve systems should be as large as possible and as unfragmented as possible. Recall, however, that the present model incorporates only demographic stochasticity and ignores environmental stochasticity. The effect of adding environmental stochasticity and dispersal between subpopulations is studied elsewhere (Burkey, 1989, 1996).

While density dependence is critical to the understanding of fragmentation effects in populations subject to demographic stochasticity (see Järvinen, 1982; Burkey, 1989, 1996), the results shown here are robust to changes in the shape of the birth and death rates as a function of density. Even death rates that decline with density yield results similar to those displayed here. The main requirement is to limit the population size of a given habitat patch in a way that is related to its

area. The actual time to extinction obviously depends on the difference between birth and death rates at different densities, but the qualitative effects of fragmentation are robust to changes in parameter values and the shape of the density dependence. Many other mechanisms and factors will affect the temporal scale of extinction, reducing extinction times relative to those calculated for these models that ignore them.

It has been argued (e.g. Goodman, 1987) that the exponential increase in mean extinction times with increasing habitat area in this kind of model makes extinction in all but the smallest reserves unlikely to the extent that demographic stochasticity is relatively unimportant in causing the extinction of populations. However, the detrimental effect of fragmentation in such models is strong at all spatial scales, and the pattern of this effect scales log-linearly with area (Fig. 8). Since the persistence times of real populations will be reduced by many factors ignored in these models—e.g. predation, disease, genetic effects, Allee effects, environmental variation, variation in sex and age structure—I would argue instead that the persistence of many fragmented populations will be reduced to time-scales that must be considered in conservation planning and management. For instance, Mangel & Tier (1993) show that when moderate environmental disturbances are included in the MacArthur—Wilson model, the extremely long extinction times for modest values of  $K$  disappear. Such disturbances will also tend to bring finite populations repeatedly into the range of population sizes where the risk of extinction from random events is substantial (see Burkey, 1996).

The actual time to extinction will be hugely variable from species to species, but the effect of fragmentation on their viability may be relatively invariable across species—albeit at different temporal scales. Even if the estimated time to extinction is long relative to the time-scale at which policy makers and managers usually operate, I see no reason to make decisions about habitat management that, although innocuous in the short term, are bound to be detrimental in the long run. Furthermore, interactions between species and the potential for cascade effects following the loss of particularly extinction prone species makes it important to make policy decisions

based on the large, area sensitive species that exist at low densities and are likely to be negatively affected by fragmentation on short to moderate time scales. Since the model ignores sex and age structure among other things, the carrying capacities and maximum population sizes used here to define the size of a habitat patch should be thought of in terms of the number of reproductive females in the population.

Extinction times of both single populations and fragmented populations are approximately gamma distributed. The gamma distribution of extinction times emerges in the following manner: in the discrete model with small  $\Delta t$ , only one event can occur in any given unit of time so the earliest time at which the population can go extinct is after  $N(0)$  time steps. There is only one way in which this can occur (a death in every consecutive time step). After that the number of different ways in which the population can go extinct by subsequent units of time increases rapidly as the number of different permutations of births and deaths leading to the required number of net deaths. This causes the increase phase of the gamma distribution. The schedule of per capita birth and death rates remain constant, however, so the distribution of probabilities of population sizes [leaving out  $P(N = 0 | t)$ ] eventually becomes fixed—analogueous to the way an age structure eventually stabilizes under a fixed schedule of mortality and reproduction. The height of this distribution of course is not constant, as a greater proportion of the trajectories are absorbed at  $N = 0$ . Extinction now occurs as a constant proportion,  $d(1)$ , of a steadily declining number,  $P(N = 1 | t)$ , dies—yielding the negative exponential, or geometric, distribution that is the right-hand tail of the observed distribution of extinction times.

In general, the gamma distribution arises as the waiting time for  $n$  independent processes decaying as a negative exponential. In the present case, it arises as the waiting time for  $N(0)$  different lineages to go extinct. Since the present process has density-dependent birth and death rates, the extinction of the different lineages is not entirely independent. Hence, the discrepancy between the best fitting gamma distribution obtained by a crude Metropolis algorithm for simulated annealing (Kirkpatrick *et al.*, 1983) may be ascribed

to this dependence (or to the crudeness of the Metropolis algorithm). In a set of independent populations the interdependence between lineages is reduced (since some of them are in different patches), and the best fitting gamma distribution has a correspondingly improved fit.

Since the distribution of extinction times has a long (infinite) tail, one might think that a set of small reserves having several opportunities to “hit” this long tail, might have a substantial chance of having at least one subpopulation persisting longer than the population in the comparable single large reserve. However, the decline in persistence time with decreasing habitat area is so rapid in these models that this does not happen—the tails of the distribution of a small reserve is so thin that the thicker tail of the distribution for the larger reserve overwhelms the benefit of having “back-up” populations.

The present analysis apparently disagrees with earlier simulation results (Gabriel & Bürger, 1992) that the distribution of extinction times for single populations in models of this kind are distributed approximately as the negative exponential or the geometric (in the discrete case). This discrepancy is explained by the fact that these authors displayed their results in histogram form, pooling across a relatively large interval on the time axis. Thus, the increase phase of the gamma distribution is obscured in bars which also include the peak of the distribution (although in some cases it is hinted at by slightly lower than expected numbers at the left tail of the distribution). Initiating the system near the quasi-stationary distribution of population sizes (probabilities of being at a given population size) would also yield an approximately geometric distribution of extinction times.

The calculation of extinction probabilities for sets of small reserves assumes that the fates of populations in isolated reserves are independent of one another. This assumption can be violated in two ways. Either the reserves are not completely isolated from each other so that some dispersal between subpopulations is possible. This “rescue effect” (Brown & Kodric-Brown, 1977) enhances the persistence of metapopulations under some conditions (Burkey, 1989, 1996; but see Simberloff *et al.*, 1992; Hess, 1994; Burkey, 1997). Alternatively, the fates of

subpopulations may be linked through large spatial scale environmental variation, "catastrophes" or "environmental stochasticity". This spatial correlation leads to a reduction of persistence times relative to those estimated under the assumption of independence. The assumptions of independence and the absence of environmental stochasticity are relaxed elsewhere (Burkey, 1989, 1996).

To the mechanisms discussed here should be added an assortment of edge effects, population genetic, pragmatic and economic effects, epidemiological effects, population dynamics and autecological effects—most of which, I believe, favor unfragmented systems. It could be argued simply that habitat fragmentation is the anthropogenic process of making the world more fragmented than it once was, and that most animals and plants were adapted to the previous state and their populations were doing fine before this process got as far as it has in our human-dominated world.

I thank Steve Pacala for insightful advice and encouragement on this paper. I am indebted to Ben Bolker and George Hurtt for help developing the code for the metropolizer used to find the maximum likelihood parameters for the Gamma distributions, for helpful and insightful discussion throughout, and for comments on earlier drafts of this manuscript. I thank Steve Hubbell, Simon Levin, Thomas Hansen and two anonymous referees for helpful comments on previous drafts.

#### REFERENCES

- ABELE, L. G. & CONNOR, E. F. (1979). Application of island biogeography theory to refuge design: making the right decision for the wrong reason. In: *Proc. Conf. on Sci. Res. in the Nat. Parks* (LINN, R. M., ed.) pp. 89–94.
- BAILEY, N. T. J. (1964). *The Elements of Stochastic Processes—with Applications to the Natural Sciences*. New York: John Wiley & Sons, Inc.
- BROWN, J. H. & KODRIC-BROWN, A. (1977). Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* **58**, 445–449.
- BURKEY, T. V. (1989). Extinction in nature reserves: the effect of fragmentation and the importance of migration between reserve fragments. *Oikos* **55**, 75–81.
- BURKEY, T. V. (1995). Extinction rates in archipelagoes: implications for populations in fragmented habitats. *Conserv. Biol.* **9**, 527–541.
- BURKEY, T. V. (1996). *Extinction in Fragmented Landscapes: Demographic Mechanisms and Predator–Prey Interactions*. Ann Arbor: UMI Dissertation Services.
- BURKEY, T. V. (1997). Metapopulation extinction in fragmented landscapes: using bacteria and protozoa communities as model ecosystems. *Am. Nat.* **150**, 568–591.
- DIAMOND, J. M. (1975). The island dilemma: lessons of modern biogeography studies for the design of natural reserves. *Biol. Conserv.* **7**, 129–145.
- DIAMOND, J. M. (1976). Island biogeography and conservation: strategy and limitations. *Science* **193**, 1027–1029.
- EDELSTEIN-KESHET, L. (1988). *Mathematical Models in Biology*. New York: Random House.
- FAHRIG, L. & MERRIAM, G. (1985). Habitat patch connectivity and population survival. *Ecology* **66**, 1762–1768.
- FELLER, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums dasein in Wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheoretica* **5**, 11–40.
- FORNEY, K. A. & GILPIN, M. E. (1989). Spatial structure and population extinction: a study with *Drosophila* flies. *Conserv. Biol.* **3**, 45–51.
- FULLER, L. E. (1962). *Basic Matrix Theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- GABRIEL, W. & BÜRGER, R. (1992). Survival of small populations under demographic stochasticity. *Theor. Pop. Biol.* **41**, 44–71.
- GANTMACHER, F. R. (1959). *Applications of the Theory of Matrices*. New York: Interscience Publishers, Inc.
- GILPIN, M. E. & DIAMOND, J. M. (1980). Subdivision of nature reserves and the maintenance of species diversity. *Nature* **285**, 568–569.
- GOEL, N. S. & RICHTER-DYN, N. (1974). *Stochastic Models in Biology*. New York: Academic Press.
- GOODMAN, D. (1987). The demography of chance extinction. In: *Viable Populations for Conservation* (Soule, M. E., ed.), pp. 11–34. Cambridge: Cambridge University Press.
- HESS, G. R. (1994). Conservation corridors and contagious disease: a cautionary note. *Conserv. Biol.* **8**, 256–262.
- HIGGS, A. J. (1981). Island biogeography theory and nature reserve design. *J. Biogeography* **8**, 117–124.
- HIGGS, A. J. & USHER, M. B. (1980). Should nature reserves be large or small? *Nature* **285**, 568–569.
- JÄRVINEN, O. (1982). Conservation of endangered plant populations: single large or several small reserves? *Oikos* **38**, 301–307.
- KIRKPATRICK, S. GELATT, C. D. & VECCHI, M. P. (1983). Optimization by simulated annealing. *Science* **220**, 671–680.
- LUDWIG, D. (1974). *Stochastic Population Theories. Lecture Notes in Biomathematics*, Vol. 3. Berlin: Springer-Verlag.
- MACARTHUR, R. H. & WILSON, E. O. (1967). *The Theory of Island Biogeography*. Princeton: Princeton University Press.
- MANGEL, M. & TIER, C. (1993). Dynamics of metapopulations with demographic stochasticity and environmental catastrophes. *Theor. Pop. Biol.* **44**, 1–31.
- MANGEL, M. & TIER, C. (1994). Four facts every conservation biologist should know about persistence. *Ecology* **75**, 607–614.
- MARGULES, C., HIGGS, A. J. & RAFF, R. W. (1982). Modern biogeographic theory: are there any lessons for nature reserve design? *Biol. Conserv.* **24**, 115–128.
- PIELOU, E. C. (1977). *Mathematical Ecology*. New York: John Wiley & Sons.
- QUINN, J. F. & HARRISON, S. P. (1988). Effects of habitat fragmentation and isolation of species richness: evidence from biogeographic patterns. *Oecologia* **75**, 132–140.

- QUINN, J. F. & HASTINGS, A. (1987). Extinction in subdivided habitats. *Conserv. Biol.* **1**, 198–208.
- RÅDE, L. & WESTERGREN, B. (1990). *Beta  $\beta$  Mathematics Handbook*, 2nd edn. Chartwell-Bratt.
- SIMBERLOFF, D. S. & ABELE, L. G. (1976). Island biogeography theory: theory and conservation practice. *Science* **191**, 285–286.
- SIMBERLOFF, D. S. & ABELE, L. G. (1982). Refuge design and island biogeographic theory: effects of fragmentation. *Am. Nat.* **120**, 41–50.
- SIMBERLOFF, D. S., FARR, J. A., COX, J. & MEHLMAN, D. W. (1992). Movement corridors: conservation bargains or poor investments? *Conserv. Biol.* **6**, 493–504.
- TERBORGH, J. (1976). Island biogeography and conservation: strategy and limitations. *Science* **193**, 1029–1030.
- WHITCOMB, R. F., LYNCH, J. F., OPLER, P. A. & ROBBINS, C. S. (1976). Island biogeography and conservation: strategy and limitations. *Science* **193**, 1030–1032.
- WILSON, E. O. & WILLIS, E. O. (1975). Applied biogeography. In: *Ecology and Evolution of Communities* (Cody, M. L. & Diamond, J. M., eds), pp. 522–534. Harvard: Harvard University Press.
- WRIGHT, J. P. & HUBBELL, S. P. (1983). Stochastic extinction and reserve size: a focal species approach. *Oikos* **41**, 466–476.