DOI: 10.1534/genetics.106.065300

Temporal Estimates of Effective Population Size in Species With Overlapping Generations

Robin S. Waples*,[1] and Masashi Yokota*,†

*Northwest Fisheries Science Center, Seattle, Washington 98112 and †Tokyo University of Marine Science and Technology, Tokyo 108-8477, Japan

Manuscript received August 25, 2006
Accepted for publication October 18, 2006

ABSTRACT

The standard temporal method for estimating effective population size (N_e) assumes that generations are discrete, but it is routinely applied to species with overlapping generations. We evaluated bias in the estimates $\hat{N}_e$ caused by violation of this assumption, using simulated data for three model species: humans (type I survival), sparrow (type II), and barnacle (type III). We verify a previous proposal by Felsenstein that weighting individuals by reproductive value is the correct way to calculate parametric population allele frequencies, in which case the rate of change in age-structured populations conforms to that predicted by discrete-generation models. When the standard temporal method is applied to age-structured species, typical sampling regimes (sampling only newborns or adults; randomly sampling the entire population) do not yield properly weighted allele frequencies and result in biased $\hat{N}_e$. The direction and magnitude of the bias are shown to depend on the sampling method and the species' life history. Results for populations that grow (or decline) at a constant rate paralleled those for populations of constant size. If sufficient demographic data are available and certain sampling restrictions are met, the Jorde–Ryman modification of the temporal method can be applied to any species with overlapping generations. Alternatively, spacing the temporal samples many generations apart maximizes the drift signal compared to sampling biases associated with age structure.

BECAUSE effective population size (N_e) is an important parameter in evolutionary biology but is notoriously difficult to measure in natural populations, considerable interest has focused on genetic methods for estimating N_e (reviewed by Beaumont 2003; Leberg 2005; Wang 2005). By far the most widely used genetic approach for estimating contemporary N_e is the temporal method (Krimbas and Tsakas 1971; Nei and Tajima 1981), so called because it depends on estimates of allele frequency taken from a population at two or more points in time. In addition to other simplifying assumptions, the standard temporal method assumes that generations are discrete, whereas many species are age structured and hence have generations that overlap. Two variations of the temporal method can account for effects of age structure, at least in some circumstances. Waples (1990) developed a modified temporal method for species with life histories like Pacific salmon (semelparous with variable age at maturity), but this model is not intended for use with iteroparous species. A more general model for age-structured, iteroparous species was developed by Jorde and Ryman (1995), who showed that the magnitude of allele-frequency change is determined not only by effective size and the sampling interval, but also by age-specific survival and birth rates. They derived an adjustment to the standard model to account for age-structure effects, and subsequent evaluations (*e.g.*, Jorde and Ryman 1996; Palm *et al.* 2003) documented the biases in estimates of effective size that can result from application of the standard temporal method without accounting for overlapping generations. However, the Jorde–Ryman method requires detailed demographic information and the ability to age individuals or group them into single-cohort samples and perhaps for these reasons has not been widely applied (but see Turner *et al.* 2002 and Palm *et al.* 2003 for examples). In spite of its obvious limitations, the discrete-generation temporal method has been and continues to be widely applied to species with overlapping generations (*e.g.*, Scribner *et al.* 1997; Johnson *et al.* 2004; Hoffman *et al.* 2004; Kaeuffer *et al.* 2004; Poulsen *et al.* 2006). *A priori*, we expect bias in the resulting estimate of N_e when an otherwise unbiased method is applied to situations that violate assumptions of the model. With respect to application of the standard temporal method to species with overlapping generations, it is possible that the biases could be small if the elapsed time between samples is long enough that the drift signal strongly dominates sampling considerations (as suggested by Jorde and Ryman 1995 and assumed, for example, by Miller and Kapuscinski 1997 and Hauser *et al.* 2002). However, under what specific circumstances this might be true cannot be

[1]*Corresponding author:* Northwest Fisheries Science Center, 2725 Montlake Blvd. E., Seattle, WA 98112. E-mail: robin.waples@noaa.gov

determined without a quantitative analysis, nor can the magnitude and direction of biases associated with samples more closely spaced in time.

In this article we evaluate performance of the standard temporal method when it is applied to iteroparous species with overlapping generations. Central to our evaluations is establishing a point of reference for evaluating the true N_e of such a population. For this we draw on two important contributions of previous authors. First, Hill (1972) showed that discrete-generation models for N_e can be modified to apply to organisms with overlapping generations, provided that the population is of constant size (N) and demographically stable. This means, for example, that over t generations of genetic drift, a population with overlapping generations should experience the same amount of allele-frequency change as a population with discrete generations and the same effective size per generation. But to test this, one must be able to measure (or estimate) the population allele frequency at a given point in time. This is straightforward if generations are discrete, but how can one measure the parametric allele frequency of a population when generations overlap? Felsenstein (1971) proposed that the correct way to calculate a population allele frequency in this case is to weight each individual by its reproductive value. We first show that this method correctly predicts the rate of allele-frequency change in simulated, iteroparous populations of constant size. Next, we evaluate bias of estimates of N_e using the standard temporal method and how they vary as a function of the species' life history and various commonly used sampling strategies.

Populations that change in size are of particular interest to evolutionary biologists and conservation biologists. Changing population size does not present a problem for the standard, discrete-generation temporal model (if population size varies the method estimates the harmonic mean N_e over the time between samples), but most models for N_e in species with overlapping generations depend on the assumption of constant population size. However, Felsenstein (1971) derived an expression for the variance-effective size in species with overlapping generations that grow (or decline) at a constant rate, in which case age structure remains constant. To complete our evaluations, we used Felsenstein's model to establish the benchmark "true" effective size in populations that deterministically change in size and evaluated performance of the standard temporal method under these conditions.

TABLE 1

Notation used

t	Time units (generally in years)
N_1	No. of age-1 individuals (newborns) produced in one time interval
N_x	No. of individuals of age x in the population
$\sum N_x$	Total no. of individuals in the population
j	Age at first maturity
q	Age at senescence
N	Population size, expressed as the total census size ($\sum N_x$), no. of reproductive adults ($N_{adult} = \sum N_{x;q>x\geq j}$), or no. of newborns ($N_1$)
T	Generation length (mean age of parents)
b_x	Mean no. of offspring produced in one time interval by individuals of age x
l_x	Fraction of a cohort that survives to age x
s_x	Fraction of individuals of age x that survive to age $x + 1$
d_x	Fraction of individuals of age x that die before reaching age $x + 1$; $d_x = 1 - s_x$
v_x	Reproductive value for individuals of age x
α_x	Fraction of a cohort that reaches age x and then dies; $\alpha_x = l_x d_x$
λ	Population growth rate per time period
k	No. of gametes contributed to the next generation by an individual over its lifetime
$\bar{k}_x$	Mean lifetime k for individuals that die between ages x and $x + 1$
σ^2_x	Variance in lifetime k of individuals that die between ages x and $x + 1$
$\bar{k}$	Mean lifetime k for all individuals breeding within a generation
V_k	Variance in lifetime k among all individuals within a generation
N_e	Effective population size for a generation
$\hat{N}_e$	An estimate of effective size based on genetic data for two samples taken 1 or more time units apart
$N^*_{e(t)}$	Instantaneous effective size that describes the actual rate of change in the population at a point t in time
S	No. of individuals per time interval sampled for genetic analysis
L	Time interval between samples
g	Elapsed no. of generations between samples; $g = L/T$
F	Standardized variance of allele frequencies
$P_{x(t)}$	Allele frequency in age class x at time unit t
$\overline{P}_{S(t)}$	Population allele frequency at time t, weighting each individual equally
$\overline{P}_{W(t)}$	Population allele frequency at time t, weighting each individual by its reproductive value
C	A constant that depends on life-history parameters of the population, used to estimate N_e in the temporal method of Jorde and Ryman (1995)

METHODS

Definition of N_e when generations overlap: *Constant N:* The models developed by Felsenstein (1971) and Hill (1972, 1979) both have features that are useful for our analyses. Notation is consistent with that used by Felsenstein (1971) and is summarized in Table 1. The species is a monoecious diploid with the possibility of selfing. Demographic parameters are fixed; exactly N_1 individuals are born in each time period, so the population will eventually reach a stable age distribution. Time units, indexed by t, are in years except as noted. The fraction of newborns that survive to age x

is $l_x = N_x/N_1$, where N_x is the number in age class x. During a single time interval, individuals of age x produce an average of b_x offspring that survive to the beginning of age class 1, so the probability that a newborn has a parent of age x is $l_x b_x$. The cohort size of newborns is $N_1 = \sum b_x N_x$. In this model, $\sum l_x b_x = 1$, generation length is given by $T = \sum x l_x b_x$, and reproductive value (Fisher 1958; Felsenstein 1971) can be calculated as $v_x = (1/l_x)\sum_{i\geq x} l_i b_i$.

Felsenstein (1971) showed that N_e in species with overlapping generations can be calculated directly from the age-specific survival and fecundity parameters contained in a standard Leslie matrix. In his model, each year within each age class there is random (binomial) variation among individuals in reproductive success, random survival of individuals between age classes, and no correlation between mortality and fecundity. Under these conditions, and assuming constant population size, inbreeding and variance-effective sizes are the same and are given by

$$N_e \approx \frac{N_1 T}{1 + \sum l_x s_x d_x v_{x+1}^2} \tag{1}$$

(Felsenstein 1971, Equation 10), where d_x is the probability of death at the end of age x and $s_x = 1 - d_x$.

Hill (1972, 1979) considered variance N_e in a model similar to Felsenstein's except that Hill made no particular assumption about variation in reproductive success among individuals. For monoecious diploids with random selfing, Hill showed that effective size is given by

$$N_e \approx \frac{4N_1 T}{V_k + 2}, \tag{2}$$

where N_1 and T are as defined above and V_k is the lifetime variance in reproductive success (production of newborns) by the N_1 individuals making up a cohort of newborns. With random variation in (lifetime) reproductive success ($V_k \approx 2$), $N_e = N_1 T$, which is the number of individuals entering the population over a period of one generation. In general, however, age structure leads to $V_k > 2$ and $N_e < N_1 T$.

Equation 2 is more general than Equation 1 but the latter is more convenient to use with data from a typical life table. However, if one assumes that the distribution of reproductive success is Poisson within each age class, Equations 1 and 2 are comparable (Johnson 1977). Under this assumption, a reformulation of Equation 2 (see appendix a) provides a way of implementing Hill's model on the basis of population vital rates,

$$N_e \approx \frac{4N_1 T}{\sum l_x d_x \bar{k}_x (1 + \bar{k}_x) - 2}, \tag{3}$$

where $\bar{k}_x = \sum_{i=1,x} b_i$ is the average lifetime reproductive success of individuals that die between years x and $x + 1$. Effective size calculated from Equation 3 is referred to as $N_e(H)$.

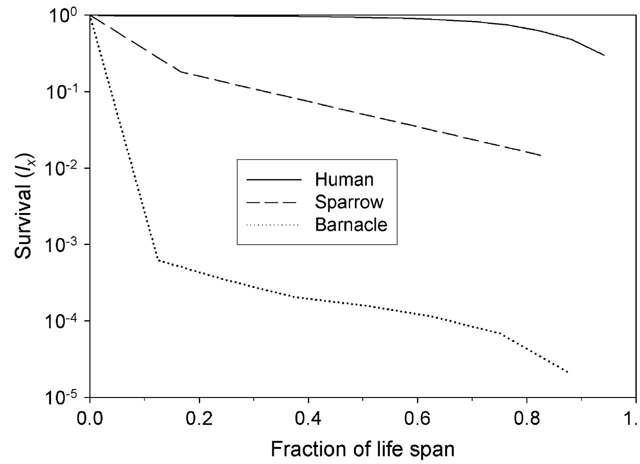

Figure 1.—Survivorship curves for the three model species. See Table 2 for data sources.

Changing N: Felsenstein (1971) also developed an expression for variance N_e in populations that are changing size at a constant rate λ per time period. λ is the dominant eigenvalue of the Leslie matrix and is the unique real solution to the discrete-time version of the Euler–Lotka equation $\sum l_x b_x \lambda^{-x} = 1$. When population size changes deterministically, analogs to the life-history parameters described above are $T = \sum x l_x b_x \lambda^{-1}$ and $v_x = (\lambda^{i-1}/l_x)\sum_{j\geq x} l_j b_j \lambda^{-j}$, and effective size is given by

$$N_e \approx \frac{N_1' T}{1 + \sum l_x s_x d_x v_{x+1}^2 \lambda^{-i} - \sum l_x b_x^2 \lambda^{-i}} \tag{4}$$

(Felsenstein 1971, Equation 24), where N_1' is the number of newborns in the next time interval. If $\lambda = 1$, $N_1' = N_1$ and Equation 4 is identical to Equation 1 except for the last term in the denominator, which is zero if the number of offspring per parent in one time interval is Poisson (Felsenstein 1971). Effective size calculated from Equation 4 is referred to as $N_e(F)$.

Model species: We chose three model species (Figure 1; Table 2), each representative of one of the three basic survival schedules. Humans are a classic type I survivorship species, with high survival well into adulthood followed by a period of rapidly increasing mortality. The white-crowned sparrow (*Zonotrichia leucophrys nuttalli*; Baker *et al.* 1981) has a modified type II survivorship curve, with a constant survival rate after an episode of high early mortality. The barnacle (*Balanus glandula*; Connell 1970) exhibits a classical type III survivorship curve with very high early mortality, even after we reduced fecundity and age-1 mortality by an order of magnitude from published values to make the simulations more tractable. We used the human demographic data analyzed by Felsenstein (1971), which are arranged into 5-year age classes. For the other two species, life-history parameters were modified slightly from published values to provide for a constant population size

TABLE 2

Life-history parameters for the model species, scaled to constant population size

	Human			Sparrow			Barnacle		
Age (x)	l_x	b_x	v_x	l_x	b_x	v_x	l_x	b_x	v_x
1	1.000	0	1.000	1.000	0	1.000	1.00000	0	1.0
2	0.979	0	1.022	0.180	2.546	5.556	0.00062	359.2	1612.9
3	0.978	0.017	1.023	0.095	2.754	5.679	0.00034	679.4	2279.4
4	0.975	0.290	1.009	0.051	2.921	5.518	0.00020	898.1	2666.6
5	0.972	0.344	0.720	0.027	3.130	4.899	0.00016	991.8	2267.2
6	0.968	0.215	0.378	0.014	3.339	3.339	0.00011	991.8	1771.4
7	0.964	0.114	0.164	—	—	—	0.00007	991.8	1299.3
8	0.957	0.045	0.050	—	—	—	0.00002	991.8	991.8
9	0.946	0.006	0.006	—	—	—	—	—	—
10	0.928	0	0	—	—	—	—	—	—
Generation (T)		5.26			3.0			4.0	

Sources: humans, Felsenstein (1971), based on data for United States white females taken from U. S. Department of Health, Education, and Welfare (1969, Vol. II, Part A, Table 5-2); white-crowned sparrow, modified from Baker *et al.* (1981); barnacle, modified from Connell (1970). Ages are 5-year units in humans and 1 year in the other two species.

and integer generation lengths in units of years ($T = 3$ years for the sparrow and 4 years for the barnacle). The published data for all three species are for females only, so for the purposes of this exercise we assumed that the same values apply to the entire population.

Computer simulations: We used computer simulations to model drift variance in allele frequency in populations with demographic parameters characteristic of the three life-history types. We considered both stable and growing populations.

Constant population size: For each species we considered a "small" and a "large" population size, indexed by N_1, the number of newborns produced each year ["newborns" were enumerated as the number of live births (humans), clutch size (sparrow), and plankton just after hatching (barnacle)]. The small and large N_1 values were: human, 10^2, 10^3; sparrow, 10^3, 10^4; and barnacle, 10^5, 10^6 (Table 3). These population sizes translated into N_e values of order 10^2 for the small population sizes and of order 10^3 for the large population

TABLE 3

Fixed numbers in each age class and effective population size for constant populations of the three model species

Age	Human		Sparrow		Barnacle	
1	100	1,000	1,000	10,000	100,000	1,000,000
2	98	979	180	1,800	62	620
3	98	978	95	954	34	341
4	97	975	51	506	20	205
5	97	972	27	268	16	160
6	97	968	14	142	11	115
7	96	964	—	—	7	69
8	96	957	—	—	2	21
9	95	946	—	—	—	—
10	93	927	—	—	—	—
$>10^a$	628	6,283	—	—	—	—
$\sum N_x$	1,502	15,020	1,367	13,670	10^5	10^6
N_{adult}	676	6,760	367	3,670	152	1,521
N_e (F)	511.8	5,118.1	370.2	3,702.2	120.0	1,199.6
N_e (H)	511.8	5,118.1	370.3	3,702.8	119.9	1,199.2
N_e (JR)	511.8	5,118.1	370.1	3,702.1	119.8	1,199.1
N_e (F)/$\sum N_x$	0.34	0.34	0.27	0.27	0.0012	0.0012
N_e (F)/N_1	5.12	5.12	0.37	0.37	0.0012	0.0012
N_e (F)/N_{adult}	0.76	0.76	1.01	1.01	0.79	0.79

$N_{\text{adult}} = \sum N_{x;q>x\geq j}$; N_e (F), Equation 4 from Felsenstein (1971); N_e (H), Equation 3 modified from Hill (1972); N_e (JR), equilibrium value calculated by iterating Equations 2–8 in Jorde and Ryman (1995).

[a] Proportion of human females in age classes >10 based on estimate by Felsenstein (1971).

sizes (Table 3), on the basis of application of Equations 1 and 3 to the life-history data in Table 2. Given a fixed value of N_1, the stable age distribution (and hence the number in each age class) is given by the right eigenvector of the Leslie matrix, which we calculated using the power method described by Caswell (2001). For the simulations, we rounded the number of individuals in each age class to the nearest integer, yielding (for each population size in each species) a vector of N_x values representing the (fixed) number of individuals in age class x.

For each fixed demographic trajectory, we modeled the stochastic process of genetic drift by randomly drawing genes to represent birth of newborns and survival from one age class to the next. At each time t, we calculated the frequency of a "gamete pool" of infinite size as the weighted mean of the allele frequencies in each age class of reproductive adults: $P_{\mathrm{gamete}(t+1)} = \sum_{x=1}^{\max\,\mathrm{age}} 2b_x N_{x(t)} P_{x(t)} / 2N_{1(t+1)}$. In generating this gamete pool, all individuals within an age class contribute equally, but the total contribution differs among age classes on the basis of age-specific survival and birth rates. Therefore, the process for reproduction simulated an array of Wright–Fisher subpopulations with the possibility of selfing, stratified by age. [Whether a species is monoecious or dioecious, and whether or not selfing is allowed, has a negligible effect on N_e (Crow and Denniston 1988; Caballero 1994); therefore, this aspect of the model should not have appreciably influenced the results.] Next, we calculated the allele frequency in the N_1 newborns in time unit $t + 1$ ($P_{1(t+1)}$) by drawing $2N_1$ genes binomially from this gamete pool. Finally, for each age class $x > 1$, allele frequency at time $t + 1$ was calculated by sampling hypergeometrically (without replacement) from the $2N_{x-1}$ genes representing the frequency in age class $x - 1$ at time t. This entire random process of sampling genes due to births and deaths was then repeated to generate an age-structured vector of allele frequencies for the next time unit. Simulated populations were initialized using the stable age distribution and with the same allele frequency in all age classes at time 0 [$P_{x(0)} = 0.5$, $x = 1, 2, \ldots$]. Each replicate simulation was run for 100 time units. Each run of 100 time units was taken to represent the trajectory of one gene locus, and the entire procedure was repeated for each additional locus (5000 loci total).

At each time unit, population allele frequencies were computed parametrically (by exhaustive population census) and estimated by sampling subsets of individuals. Parametric frequencies were computed two ways: (1) the standard method, which involves counting alleles in all individuals in the population and leads to an unweighted frequency $\overline{P}_S$, and (2) the weighted method, in which each individual is weighted by its reproductive value ($\overline{P}_W = \sum_x v_x N_x P_x / \sum_x v_x N_x$; Felsenstein 1971).

Sampling of individuals was simulated by random draw of the required number of genes from specified age classes, and sample allele frequencies were computed by counting alleles (hence the sample frequencies were unweighted). We considered three different sample sizes ($S_1 = S_2 = 25$, 50, and 100) and three different sampling regimes: (a) all individuals are at equal risk of being sampled regardless of their age, (b) only reproductive adults (those in age classes with $b_x > 0$) are subject to sampling, and (c) only newborns are subject to sampling (so sampling is from a single cohort). In all cases, genes were counted without replacement, so no individual could be sampled more than once in any time period. However, after the sampling (or enumeration) was completed, all genes were returned to the population, so sampling did not affect the future demographic or genetic trajectory of the population.

After initializing all age classes at the same frequency ($P_0 = 0.5$), a period of time is needed before the population reaches a dynamic equilibrium with respect to the amount of random allele-frequency variation among years and among age classes within years. In agreement with results reported by previous authors (Jorde and Ryman 1995; Waples 2002), we found that this dynamic equilibrium was reached quickly (within a few generations) under the conditions we modeled (data not shown). Therefore, we allowed each replicate to "warm up" for 20 time periods (periods 0–19) before collecting data.

Beginning in period 20, parametric or estimated population allele frequencies were compared at several different time intervals (L): 1 time unit and 1, 5, or 10 generations (T, $5T$, and $10T$ time units, respectively). For each time interval, we used the method of Nei and Tajima (1981) to estimate F, the standardized variance of allele frequencies at two points in time. For diallelic loci such as those considered here, this measure is calculated as

$$\hat{F}_c = \frac{1}{a} \sum_{i=1}^{a} \frac{(P_{i1} - P_{i2})^2}{(P_{i1} + P_{i2})/2 - P_{i1}P_{i2}}, \qquad (5)$$

where a is the number of loci and P_{i1} and P_{i2} are gene frequencies at the ith locus in the first and second samples, respectively. We also considered a closely related measure ($\hat{F}_k$) proposed by Pollak (1983). As previous authors have shown that the two measures generally lead to comparable results, we focused on $\hat{F}_c$ because it has a smaller variance for diallelic loci (Tajima and Nei 1984; Waples 1989). However, we found that in some circumstances $\hat{N}_e$ differed substantially depending on which estimator of F was used, and these cases are noted below.

For each time interval, we computed an overall $\hat{F}$ as the mean across the 5000 replicate loci. After advancing the initial time period by one generation (*e.g.*, first sample occurs in time period $20 + T$), this process was repeated for each interval of L years, generating a vector of mean $\hat{F}$ values for a number of intervals of the same duration equally spaced one generation apart across the last 80 years of each replicate. For some analyses, we also averaged across the elements of these vectors to arrive at an overall mean $\hat{F}$ characteristic of a particular sampling interval of L years.

To evaluate accuracy of the demographic formula for N_e, we compared observed rates of allele frequency change in simulated populations with the magnitude of change expected assuming that true N_e was as given in Table 3. The expected variance of parametric population allele frequency at time period t among replicate populations or gene loci is given by

$$\mathrm{Var}(P_t \mid P_0) = E(P_0 - P_t)^2 = P_0(1 - P_0)[1 - (1 - 1/(2N_e))^g], \quad (6)$$

where P_0 = initial allele frequency (0.5) and $g = t/T$ = elapsed time in generations. For the simulated populations, we calculated an observed variance in allele frequency at time period t as the variance in $\overline{P}_{W(t)}$ among 5000 replicate gene loci.

Estimating N_e with the temporal method: In the temporal method, effective size is estimated by relating the observed amount of allele frequency change to that expected under pure drift. The standard temporal method (Krimbas and Tsakas 1971; Nei and Tajima 1981; Pollak 1983; Waples 1989) is commonly referred to as the moment method because it relies on an estimate of F. Under a pure drift model, the expected value of $\hat{F}$ depends on N_e, the number of generations between samples (g), and the sampling regime. Here we consider nonlethal sampling of the population (with replacement, as in collecting a biopsy for DNA analysis). This corresponds to sampling plan I described by Nei and Tajima (1981) and Waples (1989). Under these conditions,

$$E(\hat{F}) \approx \frac{g}{2N_e} + \frac{1}{2S_1} + \frac{1}{2S_2} - \frac{1}{N}$$

and

$$\hat{N}_e = \frac{g}{2(\hat{F} - 1/(2S_1) - 1/(2S_2) + 1/N)},$$

where S_1 and S_2 are the sample sizes at times 1 and 2, N is the number of individuals at risk of being sampled at the time of the first sample, and $g = L/T$ = elapsed time in generations. Because sample sizes S_1 and S_2 were the same in our analyses of constant population size, the above equation simplifies to

$$\hat{N}_e = \frac{g}{2(\hat{F} - 1/S + 1/N)}. \quad (7)$$

The term $1/N$ accounts for the covariance of allele frequencies at times 1 and 2 that arises in plan I sampling because some individuals in the initial sample can also contribute genes to future generations. The appropriate value of N depends on the sampling scheme. If sampling is random with respect to the entire population, the estimator of N_e becomes

$$\hat{N}_e = \frac{g}{2\left(\hat{F} - 1/S + 1/\sum N_x\right)} \quad (8)$$

(sampling from entire population). For an exhaustive population census, $S = \sum N_x$, leading to

$$\hat{N}_e = \frac{g}{2\hat{F}} = \frac{L}{2T\hat{F}} \quad (9)$$

(complete population census). If only reproductive adults are sampled, then the number subject to sampling is the total in age classes equal to or older than the age at first maturity (j) and younger than the age of senescence, q (at which age $b_x = 0$ again):

$$\hat{N}_e = \frac{g}{2\left(\hat{F} - 1/S + 1/\sum_{q>x\geq j} N_x\right)} \quad (10)$$

(sampling mature adults). Finally, if each sample is taken from a single cohort of N_1 newborns, the appropriate estimator is

$$\hat{N}_e = \frac{g}{2(\hat{F} - 1/S + 1/N_1)} \quad (11)$$

(sampling newborns). Assuming that survival among age classes is random, Equation 11 also applies to any random sample of a single cohort; it is necessary to substitute for N_1 only the total number remaining in the cohort at the time of sampling.

Jorde and Ryman (1995) modified the standard temporal method to account for overlapping generations. Their modified estimator is (in current notation and assuming plan I sampling)

$$\hat{N}_e = \frac{C}{2T(\hat{F} - 1/(S) + 1/N_1)}, \quad (12)$$

where C is a constant that depends on life-history features of the population. Jorde and Ryman's model assumes that population size and demographic parameters are constant and is based on a comparison of consecutive cohorts, which can be achieved by sampling single cohorts or sampling mixed cohorts and sorting individuals by age to reconstruct single-cohort samples. We sampled randomly from the N_1 newborns in successive years and used Equation 12 to estimate N_e using the Jorde–Ryman method. The recurrent formulas (10–13) and (23) in Jorde and Ryman (1995) were used to calculate C. This iterative process rapidly converged on a constant C value for each life-history scenario we considered.

Changing population size: We also modeled "human" and "sparrow" populations that grew deterministically, increasing each time period by the multiplicative factor λ. We used the λ values estimated from the original published data (1.037 for humans and 1.063 for the sparrow). In addition, we considered a human population that declined at the rate $\lambda = 0.994$ per time period. These populations were initiated and allowed to run for

20 time periods (periods 0–19) at constant size (determined by low N_1) and demographic parameters given in Table 2. At time period 20, population growth was generated by scaling the birth rates to provide the desired λ, while the l_x values remained unchanged. At time period 20 the age structure was also adjusted to conform to the stable age distribution for a growing (or declining) population. In each successive time period, the vector of N_x values was obtained by applying the birth and death rates (Table 6) to the current population: the newborn cohort was generated according to $N_{1(t+1)} = \sum_{x=1}^{\max\text{age}} b_x N_{x(t)}$, and random mortality as described above determined which genes survived from one age class to the next. We generated 100 time periods of population sizes using real numbers for the N_x but rounded these to integers in each time period before conducting the simulations. Under the modeled growth rates, the human population (excluding senescent individuals) increased from 826 to >31,000, and the sparrow population size climbed from 1326 to $>5 \times 10^5$ (Table 6). In the human decline scenario, population size dropped from 5118 to 2807.

With changing population size, effective size also changes every time interval. At a given time t, Equation 4 gives a per-generation effective size ($N_e(F_t)$) that characterizes the expected amount of gene frequency drift that occurs between times t and $t + T$ (Felsenstein 1971). This definition creates an analytical difficulty here. Because there is a different $N_e(F_t)$ for every time period, genetic change in the population is described by an array of $N_e(F_t)$ values that apply to partially overlapping periods of one generation each, and it is not immediately clear how to determine the appropriate N_e value for any given point in time. In the present context, we need to evaluate genetic change over one or more individual time units that are fractions of a generation. Therefore, we need a way to calculate $E(N_{e(t)})$, which is the effective size that describes genetic change between times t and $t + 1$. As shown in appendix b, $E(N_{e(t)})$ is just the instantaneous N_e for the point midway between times t and $t + 1$,

$$E(N_{e(t)}) = N^*_{e(t+0.5)} = N_e(F_{t-T/2+0.5}) = \text{geomean}[N^*_{e(t)}, N^*_{e(t+1)}], \quad (13)$$

where $N^*_{e(t)}$ is an instantaneous effective size that describes the actual rate of change in the population at time t. $E(N_{e(t)})$ from Equation 13 can be compared with $\hat{N}_e$ estimated from samples taken in time periods t and $t + 1$. For longer time intervals, let $E(N_{e(t,t+x)})$ represent the effective size that determines the rate of change in the population between time periods t and $t + x$. Then $E(N_{e(t,t+x)})$ can be calculated as the harmonic mean of the $E(N_{e(t)})$ values for time periods t through $t + x - 1$.

When effective size changes over time, the expected variance of parametric population allele frequency at time period t among replicate populations or gene loci is

$$\text{Var}(P_t \mid P_0) = E(P_0 - P_t)^2 = P_0(1 - P_0)[1 - \prod_{i=0,t-1}\{1 - 1/(2N_{e(i)})\}],$$

where $N_{e(i)}$ is the effective size for the time period i to $i + 1$. We used Equation 13 to calculate $E(N_{e(i)})$ for each time period.

To estimate N_e, we used the procedure described above to calculate $\hat{F}$ for various time intervals and sampling regimes. With changing population size, however, it is necessary to adjust the terms for N to correspond to the correct time period. Therefore, in Equations 8, 10, and 11 we used the values for $\sum N_x$, $\sum N_{x(q>x\geq j)}$, and N_1, respectively, that applied to the time period of the first sample. The sample sizes of 25, 50, and 100 were fixed, but when complete population enumeration was done to calculate parametric population allele frequencies in growing populations, the sample size at time 2 was larger than that at time 1: $S_2 > S_1 = \sum N_x$. Therefore, Equation 9 was modified as follows:

$$\hat{N}_e = \frac{g}{2(\hat{F} - 1/(2S_1) - 1/(2S_2) + 1/\sum N_x)} = \frac{g}{2(\hat{F} - 1/(2S_2) + 1/(2\sum N_x))} \quad (14)$$

(complete population census; growing population).

RESULTS

Constant population size: For the three model species, we compared the demographic calculations of N_e (based on Felsenstein 1971) with values calculated by two other methods (Equation 3, modified from Hill 1972, and Jorde and Ryman 1995). The three methods produced essentially identical values of N_e under both low and high N_1 (Table 3). For simplicity, in what follows we use only Equation 4 to calculate N_e from demographic data.

Comparison of expected and observed rates of change in allele frequency: As shown in Figure 2, the observed variance in weighted population allele frequencies $\overline{P}_{W(t)}$ very closely tracked the expected variance in all three model species, indicating that the putative N_e values in Table 3 accurately describe the variance effective size of these populations. For the remainder of this section, therefore, we refer to the values in Table 3 as the "true" N_e.

Genetic estimates of N_e: Figure 3 compares true N_e values with estimates of N_e based on the standard temporal method, using complete population census to calculate weighted allele frequencies ($\overline{P}_{W(t)}$). For both small and large cohort sizes in all three species, agreement of expected and true N_e was very good. Under every scenario, $\hat{N}_e$ for a given time interval fluctuated randomly around the true value. Figure 3 shows results for low N_1 for samples taken 1 year apart; similar results

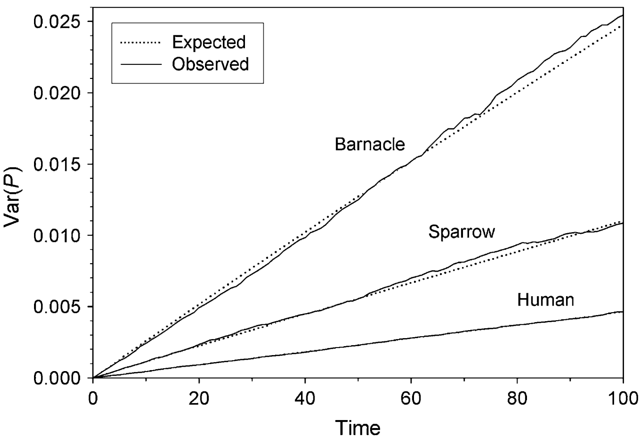

FIGURE 2.—Comparison of observed and expected Var(P_t) for the three model species under constant population size and low N_1. Comparable results were found for high N_1 (data not shown).

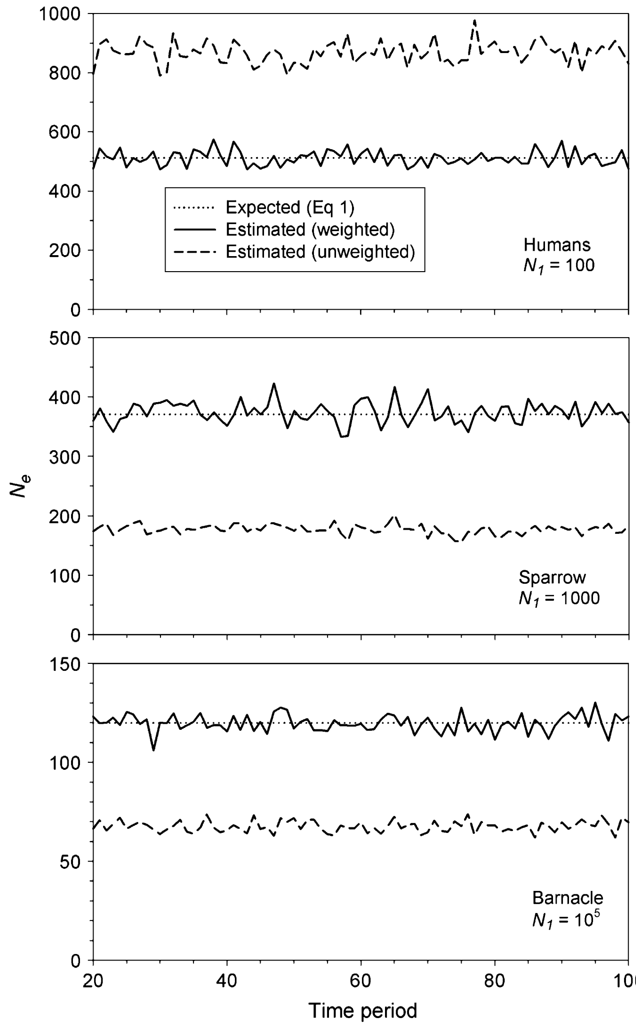

FIGURE 3.—Comparison of estimated N_e ($\hat{N}_e$) and expected N_e [$E(N_e)$] for the low N_1 simulations of the three model species under constant population size. Expected N_e uses Equation 4 and the life-table data in Table 2. Estimated N_e was calculated using the standard temporal method (Equation 9) on the basis of samples taken 1 year apart; $\hat{F}_c$ was computed by complete population enumeration, using either unweighted allele frequencies or frequencies weighted by reproductive value.

were found for high N_1 and longer time between samples ($L = 1$, 5, or $10T$; Table 4). At the extreme, deviations of $\hat{N}_e$ from true N_e were <3% (sampling interval of $10T$ in humans and barnacle).

We also estimated N_e using the temporal method based on a complete population census using unweighted allele frequencies ($\overline{P}_{S(t)}$); this produced estimates of N_e that were substantially biased in all cases: ~50% too high for humans and ~50% too low for the other two species (Figure 3; Table 4). That is, complete enumeration of all individuals alive at a given time does not provide a reliable way of calculating parametric population allele frequencies in species with overlapping generations.

The array of different sampling regimes produced a range of outcomes in the three species (Table 4). Two general patterns can be identified. First, small samples ($S = 25$) produced estimates of N_e that had large biases under many scenarios, even when $\hat{F}$ was averaged across 5000 replicate gene loci. Sample sizes of 50 or 100 produced N_e estimates that were less variable and more accurate.

Second, under most sampling regimes in most species, $\hat{N}_e$ was severely biased when F was estimated over short time intervals but approached true N_e as the sampling interval increased. For example, Figure 4 plots $\hat{N}_e/E(N_e)$ as a function of time between samples for random samples from either newborns or the entire population. In all six scenarios, the longest sampling interval ($10T$) produced the most accurate estimate of N_e. In five of the six cases the estimates for shorter time intervals were biased more strongly downward; however, sampling from the entire population for humans led to overestimates of N_e, and the bias was high for $L = 1$ generation (Figure 4). Figure 4 shows results for conditions under which the temporal method has the most power: large sample size ($S = 100$) and low N_1 (hence relatively small N_e). The same basic pattern (more accurate estimates of N_e over longer time periods) is seen, albeit sometimes less clearly, in the scenarios with smaller sample sizes and larger N_1 (Table 4). Sampling only from adult age classes produced results qualitatively similar to those for sampling from the entire population under all scenarios (Table 4).

Finally, we evaluated estimates of N_e from samples (rather than complete population census) of individuals weighted by reproductive value. These estimates were consistently biased downward (data not shown), a result that we determined arises because weighting the samples increases the sampling variance of allele frequency beyond that expected for the nominal sample size S. Although it is possible, for a given set of individual

TABLE 4

Effective population size estimated for the three model species using the standard temporal method

Sampling regime	S	Sampling interval		Human		Sparrow		Barnacle	
			N_1:	100	1,000	1,000	10,000	100,000	1,000,000
			N_e (F):	511.8	5,118.1	370.2	3,702.2	120.0	1,199.6
Complete census									
Unweighted P	$\sum N_x$	1		874	8,761	175	1,755	67	673
		1T		893	8,868	289	2,867	96	955
		5T		568	5,708	353	3,535	116	1,126
		10T		547	5,536	365	3,631	121	1,142
Weighted P	$\sum N_x$	1		511	5,116	370	3,709	120	1,203
		T		510	5,094	371	3,683	120	1,195
		5T		519	5,173	370	3,718	122	1,182
		10T		522	5,271	372	3,726	123	1,174
Random sampling									
All age classes	25	1		Inf	Inf	1,826	Inf	116	Inf
		1T		Inf	Inf	572	Inf	122	Inf
		5T		626	Inf	446	Inf	129	1,801
		10T		606	26,676	414	8,228	131	1,435
	50	1		Inf	Inf	226	Inf	74	Inf
		1T		1,320	Inf	320	Inf	102	1,631
		5T		645	12,865	367	4,679	120	1,270
		10T		580	5,509	377	5,004	125	1,316
	100	1		1,139	Inf	184	5,838	73	838
		1T		1,076	5,268	285	2,179	100	992
		5T		577	5,881	358	3,618	118	1,159
		10T		563	4,940	371	3,650	123	1,162
Adults	25	1		Inf	Inf	222	Inf	60	Inf
		1T		82,131	Inf	377	Inf	106	Inf
		5T		711	Inf	401	76,308	122	1,706
		10T		607	Inf	404	9,308	127	1,415
	50	1		2,579	Inf	135	Inf	51	Inf
		1T		846	Inf	278	Inf	100	1,620
		5T		552	11,641	359	6,463	119	1,223
		10T		554	8,161	367	5,222	124	1,154
	100	1		538	Inf	131	3,863	49	566
		1T		893	29,797	251	5,103	97	896
		5T		563	6,041	348	3,820	117	1,150
		10T		555	5,880	365	3,847	122	1,178
Newborns	25	1		10	1,811	118	Inf	141	Inf
		1T		64	2,991	240	Inf	123	Inf
		5T		195	5,655	352	156,798	127	1,591
		10T		280	5,260	390	9,502	131	1,519
	50	1		9	119	82	Inf	80	3,838
		1T		60	922	201	Inf	101	1,437
		5T		184	1,754	329	5,138	119	1,169
		10T		264	2,834	351	4,591	125	1,199
	100	1		9	93	75	1,109	69	892
		1T		59	618	189	3,323	96	997
		5T		181	1,884	321	3,489	118	1,137
		10T		265	2,790	350	3,571	123	1,153

Allele frequencies were computed on the basis of either a complete population census (using weighted or unweighted allele frequencies) or unweighted allele frequencies in random samples from the population as a whole (all age classes), only adults, or only the cohort of newborns. Samples of size S individuals were taken at the indicated time intervals. Results are based on mean $\hat{F}_c$ values calculated over 5000 replicate gene loci. Inf, a point estimate of infinity.

weights, to account for this increase in variance to arrive at an effective sample size, in natural populations the information necessary to do so will rarely be available. Consequently, attempting to weight samples by reproductive value is unlikely to be a practical solution and we did not pursue this idea further.

The Jorde–Ryman method was designed for application to data for two successive cohorts (equivalent to

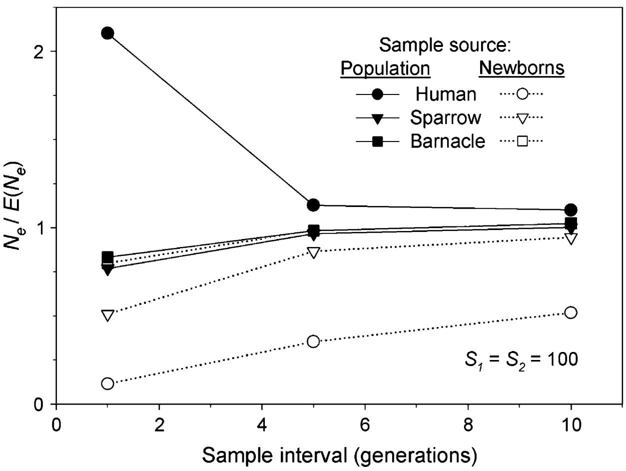

Figure 4.—The ratio $\hat{N}_e/E(N_e)$ for the three model species under constant population size and low N_1. $E(N_e)$ was computed using Equation 4 and the life-table data in Table 2. $\hat{N}_e$ was calculated using $\hat{F}_c$ in the standard temporal method, using random samples either from the population as a whole (Equation 8; solid symbols) or only from the cohort of newborns (Equation 11; open symbols). Sample sizes at times 1 and 2 were fixed at $S_1 = S_2 = 100$ but the interval between samples varied from $L = 1$ to 10 generations.

sampling newborns in two consecutive years in our model), and under that scenario $\hat{N}_e$ for the Jorde–Ryman method asymptotically approached the true N_e as sample size increased (Table 5). In their published example (which was for a species with type I survivorship), Jorde and Ryman found that with $S = 30$, $\hat{N}_e$ based on $\hat{F}_c$ was ~10% too high. With the human life history, we found comparable results for $S = 25$ and low N_1 (8% upward bias in $\hat{N}_e$; Table 5), but this bias was considerably less for larger samples. For the sparrow and barnacle life histories, the Jorde–Ryman method also produced accurate estimates for large sample size but the upward bias was more extreme (up to 100%) for $S =$ 25. A more pronounced upward bias was also found for high N_1 simulations, especially for small sample sizes (Table 5). When we used the Jorde–Ryman method with Pollak's measure $\hat{F}_k$, we found the same general pattern ($\hat{N}_e$ asymptotically accurate for large samples) except that small samples led to underestimates (rather than overestimates) of N_e (data not shown).

N_e/N ratio: As pointed out by previous authors (*e.g.*, Nunney 1993; Frankham 1995), the N_e/N ratio is sensitive to the choice of which age classes to include in calculating N. Results in Table 3 illustrate this point. For the barnacle, N_e/N is 1.2×10^{-3} if population size is taken to be all individuals alive at a given time ($N = \sum N_x$) but is 0.79 (almost three orders of magnitude higher) if only mature adults are counted ($N = N_{\text{adult}}$). [Note that to make our simulations more tractable we reduced fecundity and increased age-1 survival of the barnacle by an order of magnitude; the difference between N_e/N ratios would be even more extreme using the published data]. The effects are not as dramatic in the other species, but even for humans the ratio N_e/N is over twice as high if N is taken to be the number of reproductive adults rather than the total population size (0.76 *vs.* 0.34; Table 3).

Changing population size: *Comparison of expected and observed rates of change in allele frequency:* For all three scenarios (human growth and decline and sparrow growth; see Table 6), the observed variance in weighted population allele frequencies $\overline{P}_{W(t)}$ closely tracked the expected variance (data not shown, but results are comparable to those shown in Figure 2), indicating that Equation 13 accurately predicts the effective size as it changes over time.

Genetic estimates of N_e: Agreement between $\hat{N}_e$ and $E(N_e)$ for populations that changed in size was excellent for all three scenarios considered. Median $\hat{N}_e/E(N_e)$ ratios were almost exactly 1.0 for all time periods when weighted population allele frequencies were used to calculate $\hat{F}$ (Table 7). Other sampling regimes produced results that roughly paralleled those for populations of constant size. In humans, a population census with unweighted frequencies led to overestimates of N_e for short time periods, with lower bias over longer time periods. Conversely, sampling newborns led to gross underestimates of N_e for short time periods (Table 7). In the sparrow, a complete population census using

TABLE 5

Effective population size estimated for constant populations of the three model species using the method of Jorde and Ryman (1995)

Sampling regime	S		Human ($C = 56.80$)		Sparrow ($C = 4.947$)		Barnacle ($C = 1.793$)	
		N_1:	100	1,000	1,000	10,000	100,000	1,000,000
		N_e (F):	511.8	5,118.1	370.2	3,702.2	120.0	1,199.6
Newborns	25		553	102,835	583	Inf	251	Inf
	50		523	6,778	405	Inf	143	6,848
	100		515	5,286	372	5,488	124	1,591

Estimates are based on unweighted allele frequency differences between samples taken in two consecutive time periods; other conditions are as described in Table 4. Parameter C is calculated by iterating Equations 10–13 in Jorde and Ryman (1995).

TABLE 6

Life-history parameters for the model species in populations that change in size

	Human growth			Human decline			Sparrow growth		
Age (x)	l_x	b_x	v_x	l_x	b_x	v_x	l_x	b_x	v_x
1	1.000	0.000	1.000	1	0	1	1.000	0.000	1.000
2	0.979	0.000	1.059	0.979	0	1.016	0.180	3.050	5.905
3	0.978	0.021	1.100	0.978	0.017	1.010	0.095	3.300	6.090
4	0.975	0.352	1.123	0.975	0.281	0.990	0.051	3.500	5.987
5	0.972	0.416	0.816	0.972	0.333	0.705	0.027	3.750	5.404
6	0.968	0.260	0.432	0.968	0.208	0.370	0.014	4.000	3.763
7	0.964	0.138	0.189	0.964	0.110	0.160			
8	0.957	0.054	0.059	0.957	0.043	0.049			
9	0.946	0.007	0.007	0.946	0.006	0.006			
10	0.928	0	0.000	0.928	0	0			
		$T = 5.2105$			$T = 5.2695$			$T = 2.919$	
		$\lambda = 1.037$			$\lambda = 0.994$			$\lambda = 1.063$	

unweighted allele frequencies and random sampling from newborns both produced severe underestimates of N_e for short time intervals, but $\hat{N}_e$ approached $E(N_e)$ for longer time intervals. In contrast, sampling from the entire population or only adults led to substantial overestimates for short time periods. These patterns are all consistent with those found for constant population size (Tables 4 and 7).

TABLE 7

Ratio of estimated ($\hat{N}_e$) to expected [$E(N_e)$] effective size for the three model species in populations that change in size

Sampling regime	L	Human growth	Human decline	Sparrow growth
Census				
Unweighted	1	1.22	1.80	0.44
	T	1.42	1.79	0.75
	5T	1.03	1.12	0.94
	10T	1.00	1.08	0.97
Weighted	1	1.00	1.00	1.01
	T	1.01	1.00	1.01
	5T	1.01	1.01	1.01
	10T	1.01	1.01	1.02
Sampling[a]				
All age classes	1	Inf	Inf	3.19
	T	5.28	1.83	1.28
	5T	1.12	1.21	1.02
	10T	1.06	1.15	1.00
Adults	1	Inf	Inf	Inf
	T	Inf	Inf	Inf
	5T	1.17	1.27	2.56
	10T	1.02	1.11	1.36
Newborns	1	0.12	0.11	0.23
	T	Inf	Inf	0.66
	5T	1.25	1.04	0.91
	10T	1.06	1.07	0.97

Values shown are median $\hat{N}_e/E(N_e)$ ratios for samples taken L time units apart; other conditions are as in Table 4.

[a] Sample size = 100; unweighted allele frequencies.

DISCUSSION

Felsenstein (1971) proposed, but did not prove, that weighting each individual by its reproductive value is the correct way to calculate parametric population allele frequencies in species with overlapping generations. Our results demonstrate that this is indeed the case, regardless of the species' life history and regardless of whether the population is fixed or changing in size at a constant rate. In contrast, measuring the magnitude of population change using unweighted population allele frequencies led to large departures from expected values in all scenarios considered, even when the entire population was sampled.

These results indicate that caution is needed in interpreting standard-model temporal estimates of effective population size in species with overlapping generations. The resulting estimates will be biased unless the entire population is sampled and individual genotypes are weighted by their reproductive value. Unfortunately, this is not a practical solution to the problem of estimating N_e. Rarely is it feasible to sample an entire population, and, even if this were possible (*e.g.*, as might be the case in small captive populations), the age-specific life history data needed to compute individual reproductive values typically are not available. As discussed above, taking samples and weighting allele frequencies by reproductive value also are generally not feasible. One solution to this problem is to use the Jorde–Ryman method (Jorde and Ryman 1995), which we found generally performed well when it was applied to situations for which it was designed: analysis of samples from consecutive cohorts. This method can be used with any age-structured species, but in practice it has not been widely applied, perhaps because it requires both detailed demographic information to compute the correction factor C and samples from single cohorts in successive years. Most applications of the temporal method for species with overlapping generations continue to use the standard discrete-generation model, so it is important to consider the nature and magnitude of the bias that is likely to result from common sampling schemes. These results are complex, but some patterns are apparent.

First, bias in $\hat{N}_e$ is largest for short time intervals and in many cases largely disappears after 5–10 generations between samples (Figure 4; Table 4). This pattern, which was also seen in analyses of growing populations (Table 7), can be understood by noting that the magnitude of $\hat{F}$ (and hence $\hat{N}_e$) is determined by two major components: a component related to drift (the signal) and one related to sampling a finite number of individuals (the noise). The bias arises because the adjustment for sampling derived for the discrete-generation model does not capture all of the complexities associated with sampling from age-structured populations (Jorde and Ryman 1995). The magnitude of this sampling bias is fixed by the sample size and the nature of the sampling in relation to the species' life history; it does not change with the amount of time between samples. In contrast, a longer sampling interval allows more episodes of genetic drift to influence $\hat{F}$, thus increasing the signal-to-noise ratio in the data.

The second noteworthy pattern is that the direction of bias in $\hat{N}_e$ based on $\hat{F}_c$ differs markedly between the human and the barnacle: random sampling from the population (or a complete population census with unweighted allele frequencies) leads to an overestimate of N_e in humans but an underestimate in the barnacle. In the barnacle, the vast majority of individuals alive at any time are newborns, so a random sample will primarily reflect frequencies in the newborns, which are derived only from the fraction of the population that reproduced the previous year. The result is a sampling variance greater than would be expected from a random sample from the generation as a whole, with the consequence that $\hat{F}$ is higher than expected and N_e is underestimated. Humans are much more evenly distributed across age classes, including a substantial contingent of older individuals with low reproductive value that represent a sort of genetic inertia in the population. If their frequencies are weighted equally with younger individuals that have higher reproductive value, the result will be an underestimate of the rate of genetic change and an overestimate of N_e. Sampling only human newborns, however, leads to an underestimate of N_e, as it does in the barnacle, because the progeny are drawn from a relatively small fraction of the population that reproduces in any given time period. Use of $\hat{F}_k$ exacerbates the downward bias in $\hat{N}_e$ for the barnacle (data not shown), because with diallelic loci $\hat{F}_k$ is always larger than $\hat{F}_c$ (Waples 1989), resulting in a lower $\hat{N}_e$. Use of $\hat{F}_k$ with humans led to quite variable results, with $\hat{N}_e$ being biased either upward or downward depending on the sampling regime, sample size, and elapsed time between samples (data not shown).

Results for the sparrow are somewhat intermediate but closer to the barnacle than to the human and consistent between the constant size and growth scenarios (Figures 3 and 4; Tables 4 and 7). Presumably this reflects high mortality in the sparrow between ages 1 and 2, which causes a large fraction of the population to have low reproductive value, as in the barnacle. Thus, although the sparrow displays a classical type II survival curve after age 1, the episode of high juvenile mortality has important consequences for estimation of effective population size.

Finally, an interesting result is that, at least for small N_1, the magnitude of bias in $\hat{N}_e$ was larger for the human than for the other two species (Table 4, Figure 4). This might mean that species with type I survivorship are particularly prone to bias with use of the standard temporal method. However, the generality of this result requires further evaluation, as other factors such as longer generation time could also be involved.

Precision: Our analyses have focused on bias, but precision is also an important consideration for any genetic method of estimating N_e. Although it is beyond the scope of this article to evaluate precision in any comprehensive way, some general observations can be made on the basis of previously published results (Nei and Tajima 1981; Waples 1989; Jorde and Ryman 1995). First, the same proportional increase in sample size, elapsed time between samples, and number of independent alleles used in the estimate of $\hat{N}_e$ all have approximately the same effect on precision. Our results show that in some cases it is possible to get largely unbiased estimates of N_e from samples taken only a fraction of a generation apart. In general, however, such estimates would have very low precision unless N_e were very small and samples of individuals and loci very large. As the number of alleles used to compute $\hat{F}$ increases, the mean contributions from sampling and drift both stabilize around their expected values and precision increases. With existing technology it is quite feasible to collect data for 10–20 highly polymorphic microsatellite loci in many natural populations, which can produce precise estimates of N_e under certain conditions.

Second, the temporal method can be used most effectively to study populations with small N_e, because the signal from drift is large relative to sampling error. With any genetic method it can be difficult to distinguish a large population from a very large one. For example, under many scenarios with high N_1 the point estimates of N_e were infinity, even when $\hat{F}$ was averaged across 5000 gene loci (Tables 4 and 5). A similar effect was seen in growing populations, which reached a large size before the end of the simulations (Table 7).

Finally, results obtained by Jorde and Ryman (1995) suggest that the coefficient of variation of $\hat{F}$ when generations overlap is comparable to that for the discrete-generation model, so parametric confidence intervals for $\hat{N}_e$ for the standard temporal method (Waples 1989) should also be applicable to iteroparous species.

Likelihood-based methods for estimating N_e: In recent years a number of likelihood-based methods (Williamson and Slatkin 1999; Anderson *et al.* 2000; Wang 2001; Berthier *et al.* 2002) have been proposed for estimating N_e. In general, results for these methods are roughly comparable to the standard temporal method for moderate allele frequencies but are less biased and more precise when many low-frequency alleles are considered. These methods are all computationally demanding and were not evaluated here. However, all depend on estimates of allele-frequency change in samples taken at two or more points in time. Therefore, they should all be affected in the same general way by the various age-structure biases considered in this article. This conjecture should be tested empirically. The empirical Bayesian method proposed by Tallmon *et al.* (2004) should be able to deal with overlapping generations by modifying life-history parameters of the modeled species.

Fluctuating population size: Our results confirmed the basic elements of Felsenstein's (1971) model for populations that change in size at a constant rate, regardless of whether they are increasing or decreasing. However, many populations fluctuate around a "mean" population size. Waples (2005) evaluated performance of the temporal method for semelparous, age-structured species following the Pacific salmon model, but comparable analyses have not been done for iteroparous species. The model recently developed by Engen *et al.* (2005) for defining N_e in populations with overlapping generations that fluctuate in size could form the basis for such an evaluation in the future.

Conclusions: Results discussed above lead to the following conclusions regarding application of the temporal method to iteroparous, age-structured species:

The Jorde–Ryman method (Jorde and Ryman 1995) is perhaps the best option if appropriate samples and demographic data are available. Ideally, samples from a number of consecutive cohorts can be analyzed, so that an overall mean $\hat{F}$ (and associated $\hat{N}_e$) can be calculated that accounts for small variations over time in population size and demographic parameters (Jorde and Ryman 1996; Palm *et al.* 2003). If single-cohort samples cannot be collected directly, it might be possible to reconstruct them by ageing individuals. Jorde and Ryman (1995) suggested that if it is not possible to reconstruct *consecutive* cohorts, the correction factor C might be modified to reflect a different timing of the samples. Some uncertainty in demographic parameters might not seriously affect the results, depending on the species' life history (Jorde and Ryman 1995). Biases associated with small samples can be minimized by using at least 50 individuals in each cohort.

If the available data do not allow use of the Jorde–Ryman model, sampling biases can be minimized (and precision enhanced) by taking samples spaced far apart in time (at least three to five generations, preferably more). This in fact has been the assumption adopted by some authors who have used the temporal method with age-structured species, and our results show that it can be a reasonable one in at least some cases. Recent improvements in the ability to extract DNA from historical material can provide opportunities for retrospective analyses that span large numbers of generations (Hauser *et al.* 2002; Johnson *et al.* 2004; Poulsen *et al.* 2006).

In some cases (*e.g.*, for species with type I survivorship), estimates of N_e using the standard temporal method for closely spaced, random samples might not be severely biased, but precision of such estimates is unlikely to be satisfactory for most applications. Therefore, when generations overlap and samples are closely spaced in time, standard temporal estimates of N_e should be interpreted with extreme caution. This is particularly true because in these cases the choice of which estimator of F to use can have a profound effect on $\hat{N}_e$. The minor differences between $\hat{F}_c$ and $\hat{F}_k$ have less impact on $\hat{N}_e$ when a longer time elapses between samples and the drift signal is stronger. As a precaution, those employing the temporal method should compute $\hat{N}_e$ using both $\hat{F}_c$ and $\hat{F}_k$ and use considerable caution in interpreting scenarios under which the estimate of effective size depends heavily on the estimator of F.

Large samples of individuals are important, not only for precision but also to help minimize bias arising from various violations of implicit sampling assumptions.

Understanding as much as possible about the species' life history is important to design a sampling strategy to minimize potential biases.

We thank Per Erik Jorde, Nils Ryman, and two anonymous reviewers for insightful comments on an earlier draft of the manuscript, and we are grateful to Shuichi Kitada and Toshihide Kitakado for stimulating discussions.

LITERATURE CITED

Anderson, E. C., E. G. Williamson and E. A. Thompson, 2000 Monte Carlo evaluation of the likelihood for N_e from temporally spaced samples. Genetics **156:** 2109–2118.

Baker, M. C., L. R. Mewaldt and R. M. Stewart, 1981 Demography of white-crowned sparrows (*Zonotrichia leucophrys nuttalli*). Ecology **62:** 636–644.

Beaumont, M. A., 2003 Conservation genetics, pp. 751–792 in *Handbook of Statistical Genetics*, Ed. 2, edited by D. J. Balding, M. Bishop and C. Cannings. Wiley, Chicester, UK.

Berthier, P., M. A. Beaumont, J.-M. Cornuet and G. Luikart, 2002 Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach. Genetics **160:** 741–751.

Caballero, A., 1994 Developments in the prediction of effective population size. Heredity **73:** 657–679.

Caswell, H., 2001 *Matrix Population Models*, Ed. 2. Sinauer, Sunderland, MA.

Connell, J. H., 1970 A predator-prey system in the marine intertidal region. I. *Balanus glandula* and several predatory species of *Thais*. Ecol. Monogr. **40:** 49–78.

Crow, J. F., and C. Denniston, 1988 Inbreeding and variance effective population numbers. Evolution **42:** 482–495.

Engen, S., R. Lande and B.-E. Saether, 2005 Effective size of a fluctuating age-structured population. Genetics **170:** 941–954.

Felsenstein, J., 1971 Inbreeding and variance effective numbers in populations with overlapping generations. Genetics **68:** 581–597.

Fisher, R. A., 1958 *The Genetical Theory of Natural Selection*, Ed. 2. Dover, New York.

Frankham, R., 1995 Effective population size/adult population size ratios in wildlife—a review. Genet. Res. **66:** 95–107.

Hauser, L., G. J. Adcock, P. J. Smith, J. H. B. Ramirez and G. R. Carvalho, 2002 Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). Proc. Natl. Acad. Sci. USA **99:** 11742–11747.

Hill, W. G., 1972 Effective size of populations with overlapping generations. Theor. Popul. Biol. **3:** 278–289.

Hill, W. G., 1979 A note on effective population size with overlapping generations. Genetics **92:** 317–322.

Hoffman, E. A., F. W. Schueler and M. S. Blouin, 2004 Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. Evolution **58:** 2536–2545.

Johnson, D. L., 1977 Inbreeding in populations with overlapping generations. Genetics **87:** 581–591.

Johnson, J. A., M. R. Bellinger, J. E. Toepfer and P. Dunn, 2004 Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. Mol. Ecol. **13:** 2617–2630.

Jorde, P. E., and N. Ryman, 1995 Temporal allele frequency change and estimation of effective size in populations with overlapping generations. Genetics **139:** 1077–1090.

Jorde, P. E., and N. Ryman, 1996 Demographic genetics of brown trout (*Salmo trutta*) and estimation of effective population size from temporal change of allele frequency. Genetics **143:** 1369–1381.

Kaeuffer, R., D. Pontier, S. Devillard and N. Perrin, 2004 Effective size of two feral domestic cat populations (*Felis catus* L.): effect of the mating system. Mol. Ecol. **13:** 483–490.

Krimbas, C. B., and S. Tsakas, 1971 The genetics of *Dacus oleae*. V. Changes of esterase polymorphism in a natural population following insecticide control—Selection or drift? Evolution **25:** 454–460.

Leberg, P., 2005 Genetic approaches for estimating the effective size of populations. J. Wildl. Manage. **69:** 1385–1399.

Miller, L. M., and A. R. Kapuscinski, 1997 Historical analysis of genetic variation reveals low effective population size in a Northern pike (*Esox lucius*) population. Genetics **147:** 1249–1258.

Nei, M., and F. Tajima, 1981 Genetic drift and estimation of effective population size. Genetics **98:** 625–640.

Nunney, L., 1993 The influence of mating system and overlapping generations on effective population size. Evolution **47:** 1329–1341.

Palm, S., L. Laikre, P. E. Jorde and N. Ryman, 2003 Effective population size and temporal genetic change in stream resident brown trout (*Salmo trutta*, L.). Conserv. Genet. **4:** 249–264.

Pollak, E., 1983 A new method for estimating the effective population size from allele frequency changes. Genetics **104:** 531–548.

Poulsen, N. A., E. E. Nielsen, M. H. Schierup, V. Loeschcke and P. Grønkjær, 2006 Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). Mol. Ecol. **15:** 321–331.

Scribner, K. T., J. W. Arntzen and T. Burke, 1997 Effective number of breeding adults in *Bufo bufo* estimated from age-specific variation at minisatellite loci. Mol. Ecol. **6:** 701–712.

Tajima, F., and M. Nei, 1984 Note on genetic drift and estimation of effective population size. Genetics **106:** 569–574.

Tallmon, D. A., G. Luikart and M. A. Beaumont, 2004 Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. Genetics **167:** 977–988.

Turner, T. F., P. Wares and J. R. Gold, 2002 Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). Genetics **162:** 1329–1339.

U. S. Department of Health, Education, and Welfare, 1969 *Vital Statistics of the United States, 1967. Volume II: Mortality*. U. S. Government Printing Office, Washington, D.C.

Wang, J., 2001 A pseudo-likelihood method for estimating effective population size from temporally spaced samples. Genet. Res. **78:** 243–257.

Wang, J., 2005 Estimation of effective population sizes from data on genetic markers. Philos. Trans. R. Soc. Ser. B **360:** 1395–1409.

Waples, R. S., 1989 A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics **121:** 379–391.

Waples, R. S., 1990 Conservation genetics of Pacific salmon. III. Estimating effective population size. J. Hered. **81:** 277–289.

Waples, R. S., 2002 Effective size of fluctuating salmon populations. Genetics **161:** 782–791.

Waples, R. S., 2005 Genetic estimates of contemporary effective population size: To what time periods do the estimates apply? Mol. Ecol. **14:** 3335–3352.

Waples, R. S., 2006 Seed banks, salmon, and sleeping genes: effective population size in semelparous, age-structured species with fluctuating abundance. Am. Nat. **167:** 118–135.

Williamson, E. G., and M. Slatkin, 1999 Using maximum likelihood to estimate population size from temporal changes in allele frequencies. Genetics **152:** 755–761.

Communicating editor: J. Wakeley

APPENDIX A

Consider a population with constant demographic parameters that produces N_1 newborns per time period. The cohort of newborns can be partitioned into x demographic classes of individuals on the basis of age at death. The number of individuals living to age x, reproducing, and then dying before the next time interval is $N_1 l_x d_x$. The average lifetime reproductive success of individuals that die at the end of year x is $\bar{k}_x = \sum_{i=1,x} b_i$. Population size is constant, so the mean reproductive success of the entire cohort of N_1 individuals is $\bar{k} = 2$, which is equal to the weighted mean for each age.

Now let σ_x^2 be the variance in lifetime reproductive success of individuals that die at the end of age x, and let $\alpha_x = l_x d_x$ be the fraction of the initial cohort that reaches age x and then dies. It can be shown (see online Appendix A in Waples 2006 for details, and see Hill 1972 for a related treatment based on continuous time intervals) that the overall variance in reproductive success for the entire cohort is

$$\begin{aligned} V_x &= \frac{\sum \alpha_x N_1 \sigma_x^2 + \sum \alpha_x N_1 \bar{k}_x^2}{N_1} - \bar{k}^2 \\ &= \sum \alpha_x \sigma_x^2 + \sum \alpha_x \bar{k}_x^2 - 4. \end{aligned} \tag{A1}$$

If the distribution of reproductive success is Poisson within each age class, so that $\sigma_x^2 = \bar{k}_x$, this simplifies to $V_x = \sum \alpha_x \bar{k}_x (1 + \bar{k}_x) - 4$ and (from Equation 2)

$$\begin{aligned} N_e &\approx \frac{4N_1 T}{\sum \alpha_x \bar{k}_x (1 + \bar{k}_x) - 2} \\ &= \frac{4N_1 T}{\sum l_x d_x \bar{k}_x (1 + \bar{k}_x) - 2}. \end{aligned} \tag{A2}$$

APPENDIX B

The objective is to calculate $E(N_{e(t)})$, which is the effective size that describes genetic change between times t and $t + 1$. This can be accomplished by treating N_e as a continuous function of time. Effective size is directly proportional to total reproductive value of the population, which increases at the rate λ per time unit (Felsenstein 1971). Therefore, N_e can be described by the exponential growth model

$$N_e(F_t) = N_e(F_0)\lambda^t, \tag{B1}$$

where $N_e(F_0)$ is the effective size in the generation immediately preceding population growth (506.2 in our human model after adjusting to stable age structure in a growing population). $N_e(F_t)$ describes the "average" rate of change over the generation that starts at time t and ends at time $t+T$; however, since the population is growing geometrically, the actual effective size must be smaller early in the generation and larger later in the generation. Let us define an instantaneous effective size, denoted by $N^*_{e(t)}$, that describes the actual rate of change in the population at a point t in time. N^*_e increases exponentially over the interval t to $t+T$, and it can be shown that $N^*_e = N_e(F_t)$ at the midpoint of the interval. That is, $N^*_{e(t+T/2)} = N_e(F_t)$, which is also equal to the geometric mean of the N^*_e values over the period of the generation. It follows that the instantaneous N_e for any time period t can be calculated as

$$N^*_{e(t)} = N_e(F_{t-T/2}). \tag{B2}$$

That is, instantaneous N_e at time t is just $N_e(F)$ from one-half generation earlier.

We want to find $E(N_{e(t)})$, which is analogous to $N_e(F_t)$ except that it describes genetic change over the next time period rather than the next generation. The above logic indicates that $E(N_{e(t)})$ is just the instantaneous N_e for the point midway between times t and $t+1$:

$$\begin{aligned} E(N_{e(t)}) &= N^*_{e(t+0.5)} = N_e(F_{t-T/2+0.5}) \\ &= \text{geomean}[N^*_{e(t)}, N^*_{e(t+1)}]. \end{aligned} \tag{B3}$$