A Generalized Approach for Estimating Effective Population Size From Temporal Changes in Allele Frequency

Robin S. Waples

Northwest and Alaska Fisheries Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington 98112

ABSTRACT

The temporal method for estimating effective population size (N_e) from the standardized variance in allele frequency change (F) is presented in a generalized form. Whereas previous treatments of this method have adopted rather limiting assumptions, the present analysis shows that the temporal method is generally applicable to a wide variety of organisms. Use of a revised model of gene sampling permits a more generalized interpretation of N_e than that used by some other authors studying this method. It is shown that two sampling plans (individuals for genetic analysis taken before or after reproduction) whose differences have been stressed by previous authors can be treated in a uniform way. Computer simulations using a wide variety of initial conditions show that different formulas for computing $\hat{F}$ have much less effect on $\hat{N}_e$ than do sample size (S), number of generations between samples (t), or the number of loci studied (L). Simulation results also indicate that (1) bias of $\hat{F}$ is small unless alleles with very low frequency are used; (2) precision is typically increased by about the same amount with a doubling of S, t, or L; (3) confidence intervals for $\hat{N}_e$ computed using a χ^2 approximation are accurate and unbiased under most conditions; (4) the temporal method is best suited for use with organisms having high juvenile mortality and, perhaps, a limited effective population size.

POPULATION geneticists have been very success-ful in describing the theoretical behavior of genes in terms of a few key parameters. Obtaining reliable estimates of these same parameters in natural populations, however, has often proved more difficult. Of these parameters, effective population size (N_e) is arguably both the most important and the most difficult to evaluate directly. For this reason, several authors have explored the possibility of estimating N_e indirectly by measuring temporal changes in allele frequency. The logic for this approach is that the drift variance of allele frequency between generations is $P(1 - P)/(2N_e)$, where P is the population frequency in the initial generation. The effects of initial allele frequency can be compensated for by using some variation of Wright's standardized variance (F), and this approach has formed the basis for several efforts to relate N_e to observed changes in allele frequencies.

If P_t is the population frequency in generation t, parametric F takes the form $F = (P - P_t)^2/[P(1 - P)]$. $E(P - P_t)^2 = P(1 - P)\{1 - [1 - 1/(2N_e)]^t\}$, so $E(F) \approx 1 - [1 - 1/(2N_e)]^t$ and $\hat{N}_e$ is approximately $t/(2F)$ if t is not large. However, since one generally has access to sample (not population) allele frequencies, F must be estimated by $\hat{F}$, which is also affected by random error in drawing the samples and by the method used to estimate $P(1 - P)$. In the initial use of the temporal method to estimate effective population size, KRIMBAS and TSAKAS (1971) computed mean $\hat{F}$ using a number of alleles and subtracted the quantity $[1/(2S_0) + 1/(2S_t)]$ to account for sampling error. PAMILO and VARVIO-AHO (1980) argued that this correction is not adequate, noted the large variance of $\hat{F}$, and concluded that the temporal method is not very promising for estimating N_e. NEI and TAJIMA (1981) pointed out that some of the difficulties with the previous analyses were due to incorrect assumptions about the scheme of gene sampling used. They identified two different sampling plans (samples for genetic analysis taken before or after reproduction) and also suggested a better method for estimating F. POLLAK (1983) extended the analysis to samples taken at more than two points in time and suggested another way to compute $\hat{F}$. Some of POLLAK's claims regarding properties of the various estimators $\hat{F}$ were disputed by TAJIMA and NEI (1984).

The approaches described above have not been entirely satisfactory, in part because of the intrinsically large variance associated with estimates of N_e, but also because of rather restrictive assumptions in the models used and the fragmented treatment of the two methods of sampling. The objectives of the present paper are fourfold: (1) to define the model in such a way that $\hat{N}_e$ can be interpreted in a more generalized fashion than was possible with previously used models; (2) to show that the different sampling plans can be treated in a uniform way, thus demonstrating the general applicability of the temporal method to a wide

range of organisms; (3) to examine the distribution of the estimates of N_e and evaluate properties of confidence intervals (CIs) for $\hat{N}_e$; and (4) to identify situations in which the temporal method can (and cannot) reasonably be expected to provide important information about effective population size.

THE MODEL

Authors previously studying the temporal method for estimating N_e (e.g., NEI and TAJIMA 1981; POLLAK 1983) considered a diploid, random mating population of size N, from which samples for genetic analysis (S_0, S_t individuals) were drawn at generations 0 and t, yielding sample allele frequencies x and y, respectively. Generations were assumed to be discrete, and selection, migration, and mutation were presumed to be unimportant. These assumptions will also be adopted in the present model, although we shall see that population size is not a factor if sampling is before reproduction. The two sampling plans (Figure 1) also correspond to those described by NEI and TAJIMA (1981). In plan I, individuals are taken after reproduction or are replaced before reproduction occurs; this sampling plan would apply to human populations or others that can be sampled nondestructively, or to species that can be sampled after age of reproduction (e.g., Pacific salmon, *Oncorhynchus* spp., collected after spawning). Under sampling plan II, individuals are taken before reproduction and not replaced. This type of sampling would apply to a wide range of organisms, particularly those with high fecundity that are sampled as juveniles.

In the model proposed by NEI and TAJIMA (1981) and followed by POLLAK (1983) (hereafter the N-T model), the point of reference for measuring allele frequency change was the initial frequency in a finite population, meaning that the individuals taken for genetic analysis and the breeding population in generation zero were hypergeometric samples from the N total individuals. Not only does this complicate the sampling process, it also mandates a rather restrictive definition of effective population size: N_e represents the actual number of breeding individuals, which are assumed to have a binomial distribution of progeny number.

These difficulties can be resolved by defining P to be the allele frequency in the gamete pool preceding generation 0. Sampling in generation 0 is now binomial, meaning that the $2N_e$ genes representing the effective population size need not correspond to any particular number of individuals. Furthermore, defining the allele frequency change ($x - y$) in terms of the initial gamete pool simplifies the equations for the variance of allele frequencies, $V(x - y) = V(x) + V(y) - 2\text{Cov}(x, y)$, where $\text{Cov}(x, y)$ is the covariance of x and y.

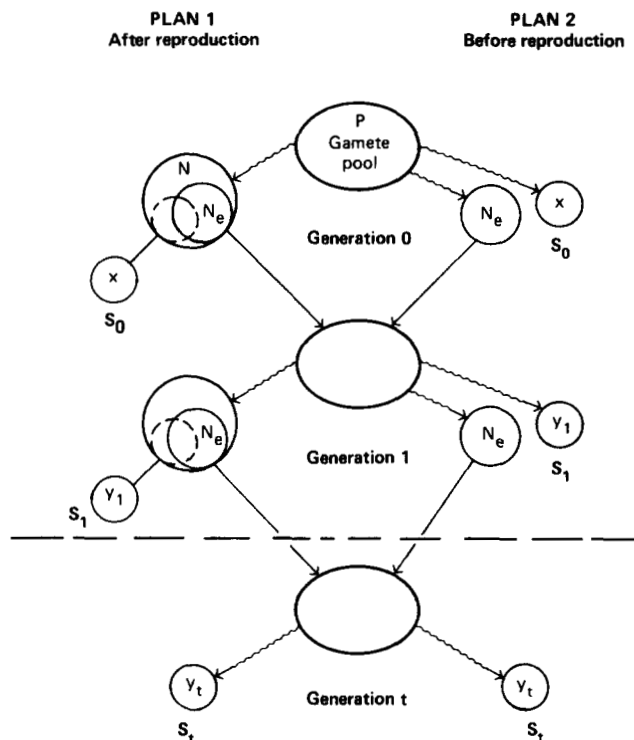FIGURE 1.—Two sampling plans considered in the analysis. In both plans, P is frequency of an allele in gamete pool preceding generation 0, x and y_t are allele frequencies in samples (of S_0 and S_t individuals) for genetic analysis taken at generations 0 and t, respectively, N is total population size at time of the initial sample, and N_e is variance effective population size. *Plan I:* sample S_0 is taken after reproduction, so it may contain some of $2N_e$ genes representing effective population size. Sample allele frequencies x and y_t are positively correlated with respect to P because samples S_0 and S_t are derived from same population (size N) at generation 0. *Plan II:* sample is taken before reproduction and not replaced, so the samples S_0 and N_e are mutually exclusive and can be considered to be independent binomial draws from initial gamete pool. Total population size is not a factor, and x and y_t are uncorrelated.

We begin by finding expressions for $V(x) = E(x - P)^2$ and $V(y) = E(y - P)^2$. For sampling plan II, the $2S_0$ genes sampled at time $t = 0$ are binomially drawn from the initial gamete pool, so

$$V(x) = E(x - P)^2 = \frac{P(1 - P)}{2S_0}. \qquad (1)$$

In plan I, the genes in sample S_0 are drawn from a population of finite size, which itself is binomially drawn from the initial gamete pool. As pointed out by NEI and TAJIMA (1981), this two-step procedure is equivalent to a single binomial sample, so $V(x)$ for sampling plan I is also given by (1).

$V(y)$ is more complicated to evaluate because it also involves genetic drift between generations. If we let P_t be the allele frequency in the gamete pool from which generation t is drawn, then the variance of P_t with respect to P (i.e., the variance due to t generations

of genetic drift) is

$$V(P_t|P) = E(P_t - P)^2$$
$$= P(1 - P)[1 - \left(1 - \frac{1}{2N_e}\right)^t]. \quad (2)$$

To get $V(y|P) = E(y - P)^2$, first note that the variance of y with respect to P_t (i.e., the sample variance of y) is

$$V(y|P_t) = E(y - P_t)^2 = \frac{P_t(1 - P_t)}{2S_t}. \quad (3)$$

We now take advantage of the basic property of conditional probability (i.e., that the variance is equal to the mean of the conditional variance plus the variance of the conditional mean (RAO 1973, p. 97)):

$$V(y|P) = E[V(y|P_t)] + V[E(y|P_t)]$$
$$= E\left[\frac{P_t(1 - P_t)}{2S_t}\right] + V(P_t|P) \quad (4)$$
$$= P(1 - P)[1 - \left[\left(1 - \frac{1}{2N_e}\right)^t \left(1 - \frac{1}{2S_t}\right)\right].$$

As was the case for (1), expression (4) is the same for both sampling plans. This point has been missed by authors using the N-T model, who have provided separate expressions for $V(y)$ for the two plans.

The variance of $(x - y)$ can be expressed for both plans I and II as

$$V(x - y) = P(1 - P)\left[\frac{1}{2S_0} + 1 - \right.$$
$$\left(1 - \frac{1}{2N_e}\right)^t \left(1 - \frac{1}{2S_t}\right)\right] - 2\text{Cov}(x, y). \quad (5)$$

Examination of (5) indicates that the only difference between the two plans is the covariance term. If sampling is before reproduction (plan II), the initial sample and the $2N_e$ gametes representing the effective population size (and, hence, all future generations) can be considered to be independent binomial samples from the pool of gametes preceding generation 0, so that $\text{Cov}(x, y) = 0$. Under sampling plan I, the allele frequencies in samples S_0 and S_t are positively correlated with respect to P because they are both derived from the same population (size N) at generation 0. If these frequencies are expressed as follows (letting P' be the population allele frequency in generation 0):

$$x = P + (P' - P) + (x - P') = P' + (x - P')$$
$$y = P + (P' - P) + (y - P') = P' + (y - P'),$$

it can be seen that the covariance of x with y is simply the covariance of P' with itself. That is, for plan I sampling,

$$\text{Cov}(x, y) = V(P') = E(P' - P)^2 = \frac{P(1 - P)}{2N}. \quad (6)$$

In this context, it is important to note that N is the size of the population subject to sampling in generation 0; if juveniles as well as adults can appear in the sample for genetic analysis, N will be larger than if only adults are sampled. After generation 0, N is not a parameter of interest unless samples are taken in more than one subsequent generation.

Substitution of (6) in (5) yields

Plan I:

$$V(x - y) = P(1 - P)\left[\frac{1}{2S_0} + 1\right.$$
$$\left. - \left(1 - \frac{1}{2N_e}\right)^t \left(1 - \frac{1}{2S_t}\right) - \frac{1}{N}\right], \quad (7a)$$

whereas if sampling is before reproduction, $\text{Cov}(x, y) = 0$ and (5) reduces to

Plan II:

$$V(x - y) = P(1 - P)\left[\frac{1}{2S_0} + 1\right.$$
$$\left. - \left(1 - \frac{1}{2N_e}\right)^t \left(1 - \frac{1}{2S_t}\right)\right]. \quad (7b)$$

Calculating $\hat{F}$: In the original formulation of the temporal method, KRIMBAS and TSAKAS (1971) used the following expression to calculate $\hat{F}$ for a single locus:

$$\hat{F}_a = \frac{1}{K} \sum_{i=1}^{K} \frac{(x_i - y_i)^2}{x_i(1 - x_i)},$$

with K being the number of segregating alleles. However, this method has a drawback: $\hat{F}_a$ is infinitely large if an allele is found at time t but not at time 0 (i.e., if any $x_i = 0$). NEI and TAJIMA (1981) proposed an alternative estimator, $\hat{F}_c$:

$$\hat{F}_c = \frac{1}{K} \sum_{i=1}^{K} \frac{(x_i - y_i)^2}{(x_i + y_i)/2 - x_i y_i}, \quad (8)$$

which avoids this problem. Another method for calculating $\hat{F}$ was suggested by POLLAK (1983). His measure, $\hat{F}_k$, is

$$\hat{F}_k = \frac{1}{K - 1} \sum_{i=1}^{K} \frac{(x_i - y_i)^2}{(x_i + y_i)/2}. \quad (9)$$

Often, data for multiple gene loci with a varying number of alleles (K_j alleles at the jth locus) are available. In this case, weighted means of the single locus $\hat{F}$ values are computed as mean $\hat{F}_c = \sum K_j \hat{F}_{cj} / \sum K_j$ and mean $\hat{F}_k = \sum (K_j - 1)\hat{F}_{kj} / \sum (K_j - 1)$, where, for example, $\hat{F}_{cj} = \hat{F}_c$ for the jth locus computed using (8) (TAJIMA and NEI 1984).

As $\hat{F}$ is a ratio, it is difficult to give its expectation exactly, but an approximation for $E(\hat{F}_c)$ is:

$$E(\hat{F}_c) \approx \frac{E(x - y)^2}{E[(x_i + y_i)/2 - x_i y_i]} = \frac{V(x - y)}{P(1 - P) - \text{Cov}(x, y)}.$$

This agrees with the result given by NEI and TAJIMA (1981), except that they omitted the covariance term. In the current model, $Cov(x, y) = 0$ for plan II but not for plan I. In the latter case, $E(\hat{F}_c) = V(x - y)/\{P(1 - P)[1 - 1/(2N)]\}$, so taking $Cov(x, y)$ into consideration amounts to increasing $\hat{F}$ by approximately the factor $1/(2N)$. This is a very small adjustment (smaller than the difference between $\hat{F}_c$ and $\hat{F}_k$) that has little effect on estimates of N_e. Therefore, $Cov(x, y)$ may safely be ignored in deriving the expectation of $\hat{F}_c$, and

$$E(\hat{F}_c) \approx \frac{V(x - y)}{P(1 - P)}, \qquad (10)$$

whether sampling is before or after reproduction.

For plan II, substituting (5) into (10) and canceling the term $P(1 - P)$ yields

$$E(\hat{F}_c) \approx \frac{1}{2S_0} + 1 - \left(1 - \frac{1}{2N_e}\right)^t \left(1 - \frac{1}{2S_t}\right),$$

which, if $t/(2N_e)$ is small, is well approximated by

$$E(\hat{F}_c) \approx \frac{1}{2S_0} + \frac{1}{2S_t} + \frac{t}{2N_e}.$$

This latter expression suggests the estimator

Plan II:

$$\hat{N}_e = \frac{t}{2[\hat{F}_c - 1/(2S_0) - 1/(2S_t)]}, \qquad (11)$$

which is identical to the formula given by NEI and TAJIMA (1981; Equation 18) for the special case of $N \gg N_e$. Therefore, this estimator can be used in a more general sense. In fact, it is clear from (7b) that $V(x - y)$ and $E(\hat{F})$ are independent of total population size if sampling is before reproduction.

$E(\hat{F}_c)$ for plan I can be obtained in a similar fashion, the only difference being the covariance term:

$$E(\hat{F}_c) \approx \frac{1}{2S_0} + \frac{1}{2S_t} + \frac{t}{2N_e} - \frac{1}{N}.$$

Therefore, N_e can be estimated by the following:

Plan I:

$$\hat{N}_e = \frac{t}{2[\hat{F}_c - 1/(2S_0) - 1/(2S_t) + 1/N]}. \qquad (12)$$

Equation 12 differs from that used by NEI and TAJIMA (1981) and POLLAK (1983) (see APPENDIX for discussion). Note, however, that for the special case of $N = N_e$, (12) can be written as

$$\hat{N}_e = \frac{t - 2}{2[\hat{F}_c - 1/(2S_0) - 1/(2S_t)]}, \qquad (13)$$

which is identical to equation (16) in NEI and TAJIMA (1981). Note also that for $N = \infty$, (12) reduces to (11),

as it should (plans I and II are equivalent with infinite population size).

Often N will not be known, but a rough estimate of $r = N/N_e$ can be made. In this case, a useful formula is

$$\hat{N}_e = \frac{rt - 2}{2r[\hat{F}_c - 1/(2S_0) - 1/(2S_t)]}. \qquad (14)$$

Various values of r can be tried to generate a range of possible estimates $\hat{N}_e$.

Although the above formulae were derived for $\hat{F}_c$, they can be used for estimates of N_e based on $\hat{F}_k$ as well (POLLAK 1983; TAJIMA and NEI 1984). If sample size varies among loci, S_0 and S_t in (11)–(14) can be replaced by their harmonic means weighted by the number of independent alleles $(K - 1)$ per locus.

Computer simulations: There has been some disagreement in the literature regarding means and variances of $\hat{F}_c$ and $\hat{F}_k$ (POLLAK 1983; TAJIMA and NEI 1984). To address this issue and to evaluate the properties of $\hat{F}_c$ and $\hat{F}_k$ as estimators of N_e, a series of computer simulations was done based on the model described above. In each replicate simulation, $2N_e$ genes representing the effective population size in generation 0 were first chosen binomially from the initial gamete pool (allele frequency P). The initial sample (S_0 individuals) for plan II was also chosen binomially (and independently) from this gamete pool. For sampling plan I, the $2N_e$ genes representing the effective population size were first incremented to $2N$ genes by additional binomial sampling. The sample S_0 was then chosen without replacement from the $2N$ genes representing the total population size. In both plans, this process was repeated for generations $t = 1$–10, yielding samples $S_1 \cdots S_{10}$. Five thousand replicates were performed for each parameter set P, N_e, N, and $S = S_0 = S_t$. At each generation in each replicate, $\hat{F}_c$ and $\hat{F}_k$ were computed (equivalent to single locus $\hat{F}$ values). Mean $\hat{F}_c$ and mean $\hat{F}_k$ were then computed over all 5000 replicates for each value of t, these mean values being used to estimate N_e using (11) and (12). Simulations were also done using three alleles at a locus (initial frequency $= P_i$ for the ith allele), in which case sampling was multinomial. In some simulations using alleles at low frequency, sample allele frequencies in generations 0 and t were both 0. $\hat{F}$ was not computed for replicates in which this occurred. This corresponds to procedures that would normally be followed in sampling from real populations (i.e., unobserved alleles would not be included in the analysis).

RESULTS

Plan I vs. plan II: In the simulations involving sampling after reproduction, parametric N was used in (12) to estimate N_e. In practice, N will generally not
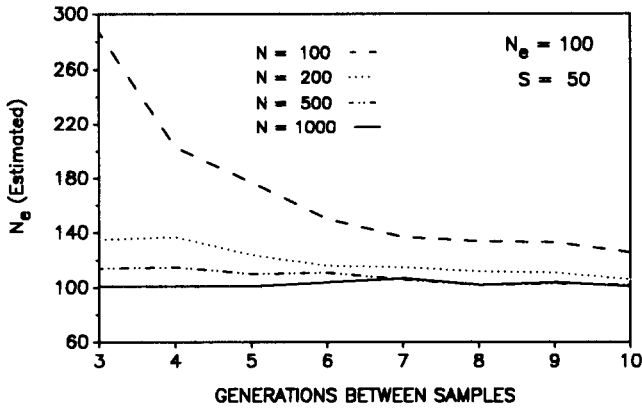
FIGURE 2.—Estimates of N_e for plan I sampling for various actual values of total population size (N) when (11) was used to compute $\hat{N}_e$ [equivalent to setting $N = \infty$ in (12)]. Bias was relatively small unless true value of $r = N/N_e$ was less than 2. Results shown are for mean $\hat{F}_k$ computed from 5000 replicate simulations of a single locus with two alleles. Initial allele frequency was 0.5.

be known exactly and must be estimated. One approach is to assume that N is large in comparison to N_e ($r = N/N_e \approx \infty$), in which case there is little difference between plans I and II and (12) is well approximated by (11). Figure 2 illustrates two points about the potential bias resulting from using (11) to estimate N_e if sampling is according to plan I. First, bias is fairly small unless $N < 2N_e$ and is negligible if $N \geq 5N_e$. Although the true ratio $r = N/N_e$ will not be known exactly in most cases, often it will be clear whether N is less than $2N_e$. The second point is that the magnitude of bias is inversely related to time between samples [because the proportional contribution of the term $1/N$ to the denominator of (12) decreases with time]. For the simulation shown in Figure 2, the bias in using (11) was only about 50% or less even for the extreme case of $N = N_e$, provided that 6 or more generations separated the two samples. For $t < 3$, however, estimates of N_e for plan I are very sensitive to the accuracy with which N is estimated, so caution must be used in these cases. Because (11) appears in general to be a fairly robust estimator of N_e for sampling after reproduction, the remaining analyses will focus on plan II sampling, for which N is not a factor.

Comparison of $\hat{F}_c$ and $\hat{F}_k$: For a locus with two alleles and equal sample sizes S_0 and S_t, $\hat{F}_c$ and $\hat{F}_k$ can be expressed as follows:

$$\hat{F}_c = \frac{(x - y)^2}{(x + y)/2 - xy};$$

$$\hat{F}_k = \frac{(x - y)^2}{(x + y)/2 - [(x + y)/2]^2}.$$

Except in the trivial case where both x and $y = 0$ or 1, $[(x + y)/2]^2$ is always larger than $[xy]$, so $\hat{F}_k$ is larger than $\hat{F}_c$ because it has a smaller denominator. This result can be seen in the simulations (Table 1). If there are more than two alleles, which estimator is smaller

depends on the allele frequencies in the samples. In simulations using three alleles, mean $\hat{F}_c$ was generally slightly smaller than mean $\hat{F}_k$ unless the initial frequency for the least common allele was less than about 0.02 (Table 1).

Because $\hat{F}$ appears as a positive term in the denominator of (11) and (12), smaller estimates of F lead to larger estimates of N_e. In the simulations, overall accuracy of mean $\hat{F}_c$ and mean $\hat{F}_k$ as estimators of N_e was very good, although both tended to overestimate N_e slightly, with $\hat{F}_k$ frequently being more accurate because its larger value led to a lower estimate of N_e (Table 1). However, mean $\hat{F}_c$ provided slightly more accurate estimates for loci with three alleles and extreme allele frequencies.

The good overall agreement of $\hat{N}_e$ with the actual value used in the simulations indicates that the formulae for $E(\hat{F})$ used to obtain (11) and (12) are very good approximations. These approximations are not quite so good if the distribution of allele frequencies at a locus is very uneven. In simulations using such frequencies, mean $\hat{F}_c$ and mean $\hat{F}_k$ were somewhat lower than given by the approximation for $E(\hat{F})$, leading to estimates of N_e that were too high. However, as shown in Figure 3 for $\hat{F}_k$, this bias was relatively minor unless alleles with very low frequency were involved. For $K = 2$, $N_e = 100$, $S = 50$, mean $\hat{F}_k$ overestimated N_e by no more than about 20% through generation 10 unless initial allele frequency was greater than 0.95. If diallelic loci with higher allele frequencies than 0.95 are used, the downward bias in $\hat{F}$ can be substantial and increases with t. For loci with three alleles, estimates of N_e appear to be even less sensitive to the effects of allele frequency. Even with initial frequency of one allele as small as 0.01, $\hat{N}_e$ was no more than 40% above true N_e after 10 generations (Figure 3). Considering the difficulty in obtaining even approximate estimates of N_e by other methods, this bias does not seem unduly large. Effects of allele frequency were somewhat less in simulations using larger values of N_e or S.

TAJIMA and NEI (1984) disputed POLLAK's (1983) claim that the variance of $\hat{F}_k[V(\hat{F}_k)]$ is approximately equal to $V(\hat{F}_c)$ for diallelic loci, and results of the simulations support the former authors: $V(\hat{F}_c)$ was slightly smaller than $V(\hat{F}_k)$ for all simulations with $K = 2$ (Table 1). For $K = 3$, both authors agreed that $V(\hat{F}_c)$ should be smaller if frequencies of the various alleles are fairly even; otherwise, $V(\hat{F}_k)$ should be smaller. Simulation results generally support this conclusion (Table 1). Apparently, however, it is not whether the allele frequencies are uneven, but whether the frequency of any single allele is small, that determines the relative values of $V(\hat{F}_c)$ and $V(\hat{F}_k)$. For $K = 3$, $\hat{F}_c$ had a smaller variance in simulations with $N_e = 100$, $S = 50$ and initial frequencies (0.5, 0.4, 0.1) and (0.7,

TABLE 1

Means, standard deviations (s), and estimates of N_e for $\hat{F}_c$ and $\hat{F}_k$

| | | | | | | Mean $\hat{F}$ | | $s_{\hat{F}}$ | | $(K\text{-}1)s_{\hat{F}}^2/\bar{F}^2$ | | $\hat{N}_e$ | |
| | | | | | | | | | | | | | |
N_e	S	Allele frequency			t	$\hat{F}_c$	$\hat{F}_k$	$\hat{F}_c$	$\hat{F}_k$	$\hat{F}_c$	$\hat{F}_k$	$\hat{F}_c$	$\hat{F}_k$
$K = 2$													
100	50	0.500	0.500		3	0.0348	0.0357	0.0473	0.0496	1.84	1.93	101	96
					10	0.0689	0.0726	0.0946	0.1049	1.89	2.09	102	95
100	50	0.950	0.050		3	0.0338	0.0346	0.0413	0.0429	1.49	1.54	109	103
					10	0.0574	0.0593	0.0617	0.0656	1.15	1.23	134	127
500	50	0.950	0.050		3	0.0226	0.0230	0.0294	0.0303	1.70	1.74	578	509
					10	0.0290	0.0296	0.0368	0.0382	1.61	1.66	554	522
500	100	0.950	0.050		3	0.0131	0.0132	0.0176	0.0179	1.82	1.84	489	471
					10	0.0193	0.0196	0.0253	0.0256	1.72	1.76	538	523
$K = 3$													
100	50	0.333	0.333	0.333	3	0.0339	0.0349	0.0327	0.0345	1.86	1.96	108	101
					10	0.0681	0.0713	0.0643	0.0697	1.78	1.90	104	97
100	50	0.700	0.200	0.100	3	0.0351	0.0359	0.0340	0.0343	178	1.82	100	94
					10	0.0668	0.0692	0.0620	0.0632	1.72	1.66	107	102
100	50	0.500	0.450	0.050	3	0.0343	0.0347	0.0347	0.0333	2.04	1.84	105	102
					10	0.0648	0.0657	0.0610	0.0580	1.88	1.56	112	109
100	50	0.500	0.490	0.010	3	0.0323	0.0321	0.0323	0.0281	2.00	1.54	122	125
					10	0.0577	0.0552	0.0600	0.0532	2.16	1.86	133	142
100	50	0.900	0.090	0.010	3	0.0318	0.0315	0.0301	0.0265	1.78	1.42	127	130
					10	0.0565	0.0536	0.0535	0.0471	1.80	1.54	137	149

Results are from 5000 replicate simulations (equivalent to sampling 5000 loci) with indicated initial allele frequencies and $S = S_0 = S_t =$ sample size. t is the number of generations between samples; K is the number of alleles per locus.

0.2, 0.1), while $V(\hat{F}_k)$ was smaller in all simulations with any single frequency less than 0.07 (Table 1).

All of the above results for plan II sampling also held for plan I sampling [when parametric N was used in (12) to estimate N_e], except that the upward bias in $\hat{N}_e$ for diallelic loci was not as consistently observed. In some simulations, $\hat{F}_k$ led to estimates that were slightly too low, and those based on $\hat{F}_c$ were more accurate.

These differences in the properties of $\hat{F}_c$ and $\hat{F}_k$ having been noted, it must be pointed out that they do not result in major differences in the estimates of N_e. In no simulations did $\hat{N}_e$ differ by more than a few percent for mean $\hat{F}_c$ and mean $\hat{F}_k$. It appears that $V(\hat{F}_c)$ is always less than $V(\hat{F}_k)$ if $K = 2$; however, the difference is not large, and even smaller differences were found in simulations with $K = 3$. In practice, the choice of which estimator of F to use has a relatively small effect on $\hat{N}_e$. This point is illustrated in Figure 4, which shows the distribution of $\hat{N}_e$ values based on $\hat{F}_c$ and $\hat{F}_k$ from a pair of simulations using very different initial conditions. To simplify presentation of the rest of the analyses, therefore, results are given for $\hat{F}_k$ only.

Distribution of $\hat{N}_e$: Although the simulation results indicate that mean $\hat{F}$ computed over 5000 replicates provides an accurate estimate of N_e, $\hat{F}$ for natural populations must be based on far fewer loci. There-

fore, the usefulness of the temporal method also depends heavily on precision of $\hat{N}_e$. To evaluate precision, it might seem reasonable to compute the variance of $\hat{N}_e$, but there are two good reasons for not using this quantity. First, the distribution of $\hat{F}$ (on which $\hat{N}_e$ is based) is skewed, and Figures 4–7 show that the distribution of $\hat{N}_e$ also is far from normal unless a very large number of alleles is used. Second, from (11) it is clear that $\hat{N}_e$ is infinitely large if $\hat{F}$ is exactly equal to $[1/(2S_0) + 1/(2S_t)]$, while $\hat{N}_e$ is negative if $\hat{F}$ is less than this quantity. In either case, computation of the variance of $\hat{N}_e$ is problematical.

A different approach is based on the observation (LEWONTIN and KRAKAUER 1973) that $n\hat{F}/E(\hat{F})$ [$n = \sum (K_j - 1) =$ total number of independent alleles surveyed] is distributed approximately as chi square with n degrees of freedom. NEI and TAJIMA (1981) reported a good fit for $\hat{F}_c$, but they examined plan I sampling only (and only for $K = 2$ and $N = N_e$) and suggested that the approximation should be somewhat poorer if sampling is before reproduction. MUELLER et al. (1985) pointed out that many questions regarding the distribution of $\hat{F}$ remain to be answered.

To examine this topic more thoroughly, the distribution of $\hat{N}_e$ based on a variable number (L) of loci was found by sampling with replacement from the single locus $\hat{F}$ values generated in the simulations. Two thousand such samples (each yielding a mean $\hat{F}_k$
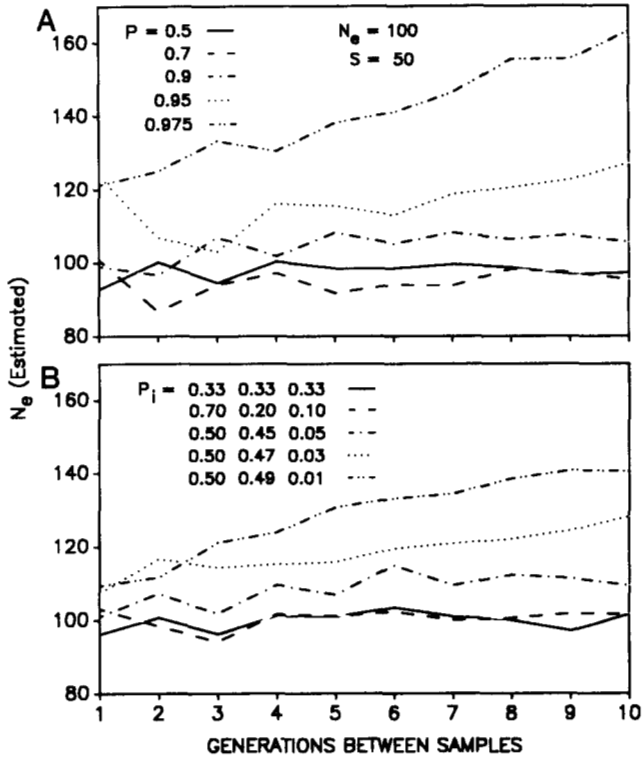
FIGURE 3.—Effects of allele frequency on estimates of N_e. Results shown are for simulations with two alleles (A) or three alleles (B) at each locus and initial frequencies as indicated. Sampling was according to plan II; mean $\hat{F}_k$, based on 5000 replicate simulations of a single locus, was used to compute $\hat{N}_e$ using (11).

FIGURE 4.—Comparison of $\hat{F}_c$ and $\hat{F}_k$ as estimators of N_e. Distributions shown are for $\hat{N}_e$ values based on 2000 mean $\hat{F}_c$ and mean $\hat{F}_k$ values sampled with replacement from single locus $\hat{F}$ values generated in the simulations. Results shown are for simulations with two alleles (A) or 3 alleles (B) at each locus and mean $\hat{F}$ values based on data for 5 loci (A) or 20 loci (B). Other initial parameters were as indicated. Sampling was according to plan II; $\hat{N}_e$ was computed using (11).

value) were drawn for each parameter set. In calculating the expected distribution of $\hat{N}_e$, the value of $\hat{F}$ corresponding to a given value of $\hat{N}_e$ was first computed from the relationship $E(\hat{F}) \approx 1/(2S_0) + 1/(2S_t) + t/(2N_e)$. Next, the χ^2 value was computed as $\chi^2 = n\hat{F}/E(\hat{F})$ using $E(\hat{F})$ = mean $\hat{F}$ observed in the simulations, and the corresponding P value was found from a table of the χ^2 distribution. For example, consider the simulation depicted in Figure 4B ($N_e = 500$, $S_0 = S_t = 100$, 3 alleles per locus [$P_i = 0.9, 0.07, 0.03$]; mean $\hat{F}_k = 0.0198$ at $t = 10$ in this simulation), and assume we want the percentage of $\hat{N}_e$ values expected in the range 400–600. This is equivalent to determining the probability that $\hat{F}$ lies between $\hat{F} = 1/200 + 1/200 + 10/800 = 0.0225$ and $\hat{F} = 1/200 + 1/200 + 10/1200 = 0.01833$. For estimates based on $L = 20$ loci ($n = 40$ independent alleles), the χ^2 values for the desired N_e values are $\chi^2 = 40(0.0225)/0.0198 = 45.45$ and $\chi^2 = 40(0.01833)/0.0198 = 37.04$, respectively. From a table of the χ^2 distribution with $n = 40$ degrees of freedom, the probability that $\chi^2 > 37.04 \approx 0.61$ and the probability that $\chi^2 > 45.45 \approx 0.26$, so $\hat{N}_e$ is expected to be in the range 400–600 with probability $0.61 - 0.26 = 0.35$. In the simulation, the actual proportion was 0.37.

As noted above, trials for which $\hat{F} \leq [1/(2S_0) + 1/(2S_t)]$ lead to unusual results ($\hat{N}_e = \infty$ or negative), but
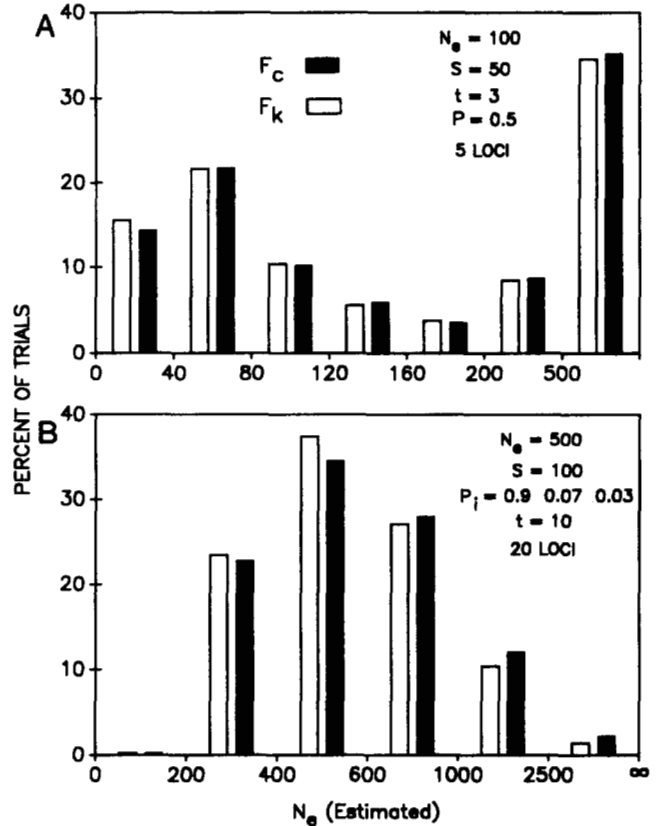
they can be interpreted in a straightforward manner. The quantity $[1/(2S_0) + 1/(2S_t)]$ accounts for the expected "spurious" contribution to $\hat{F}$ that results from taking a finite sample of individuals for analysis. If $\hat{F} \leq [1/(2S_0) + 1/(2S_t)]$, all of the temporal differences in allele frequency can be explained by sampling error without invoking genetic drift at all (i.e., there is no evidence that N_e is finite). In this case, the only feasible estimate of N_e is infinity, as noted by LAURIE-AHLBERG and WEIR (1979) and HILL (1981) in a similar context using linkage disequilibrium data. A related phenomenon (a negative estimate for a parameter normally constrained to be non-negative) can occur with estimates of genetic distance of F_{ST} corrected for sampling errors (NEI 1978; NEI and CHESSER 1983; WEIR and COCKERHAM 1984). In Figures 4–7, estimates of N_e corresponding to $\hat{F} \leq [1/(2S_0) + 1/(2S_t)]$ have been plotted in the category that includes $N_e = \infty$.

It turns out that the χ^2 approximation is actually quite good for either sampling plan for a broad range of initial conditions (Figures 5–6). Substantial departures from the χ^2 distribution were found only for
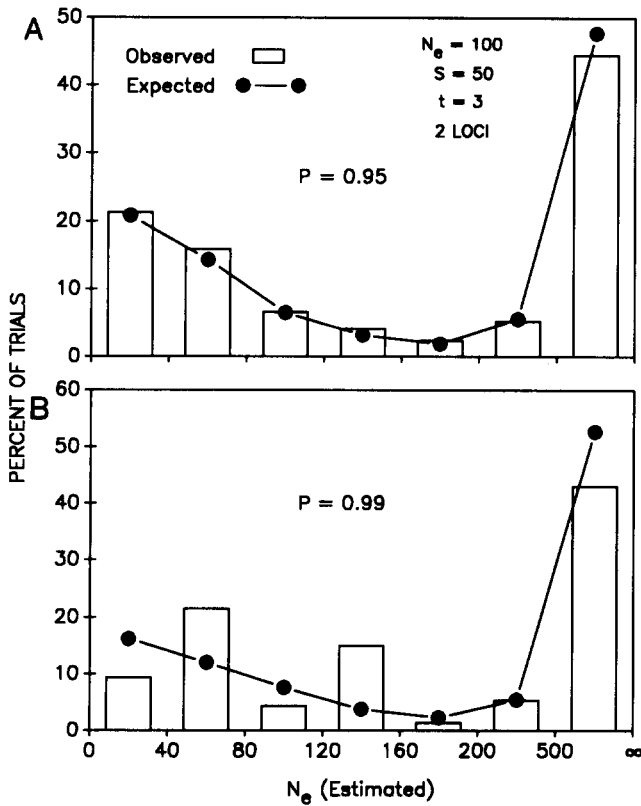
FIGURE 5.—Comparison of the χ^2 distribution ("expected") with observed distribution of $\hat{N}_e$ computed using 2000 two-locus mean $\hat{F}_k$ values sampled with replacement from single locus $\hat{F}$ values generated in the simulations. Each locus had two alleles. Initial parameters were as indicated. Sampling was according to plan II; $\hat{N}_e$ was computed using (11).

simulations with low frequencies of one or more alleles. For example, for $K = 2$, $N_e = 100$, $S = 50$, and $t = 3$, the observed distribution of $\hat{N}_e$ based on data for 2 loci was very close to that expected even with initial allele frequency as high as 0.95 (Fig. 5A). For extreme allele frequencies ($P = 0.99$; Figure 5B), the observed distribution of $\hat{N}_e$ was narrower than the chi square, particularly as t increased or more loci were used to compute mean $\hat{F}$. For simulations with $K = 3$, the fit also was quite good unless some alleles were at very low frequency. With $N_e = 100$, $S = 50$, $t = 10$, and $\hat{N}_e$ based on data for 5 loci, very close agreement was found for $P_i = (0.7, 0.2, 0.1)$, but the distribution of $\hat{N}_e$ was noticeably narrower than the chi square for $P_i = (0.96, 0.02, 0.02)$ (Figure 6, A and B).

As noted by NEI and TAJIMA (1981), the quantity $(K - 1)s_F^2/\bar{F}^2$ provides an indication of how close the distribution of $n\hat{F}/E(\hat{F})$ is to the χ^2, a value of exactly 2 expected for a perfect fit. Results in Tables 1 and 2 indicate that observed values for both $\hat{F}_c$ and $\hat{F}_k$ were generally slightly lower than 2, indicating less dispersion (a narrower distribution) than the χ^2. $(K - 1)s_F^2/\bar{F}^2$ was smaller for $\hat{F}_c$ in all simulations using two alleles, but was smaller for $\hat{F}_k$ with $K = 3$ except
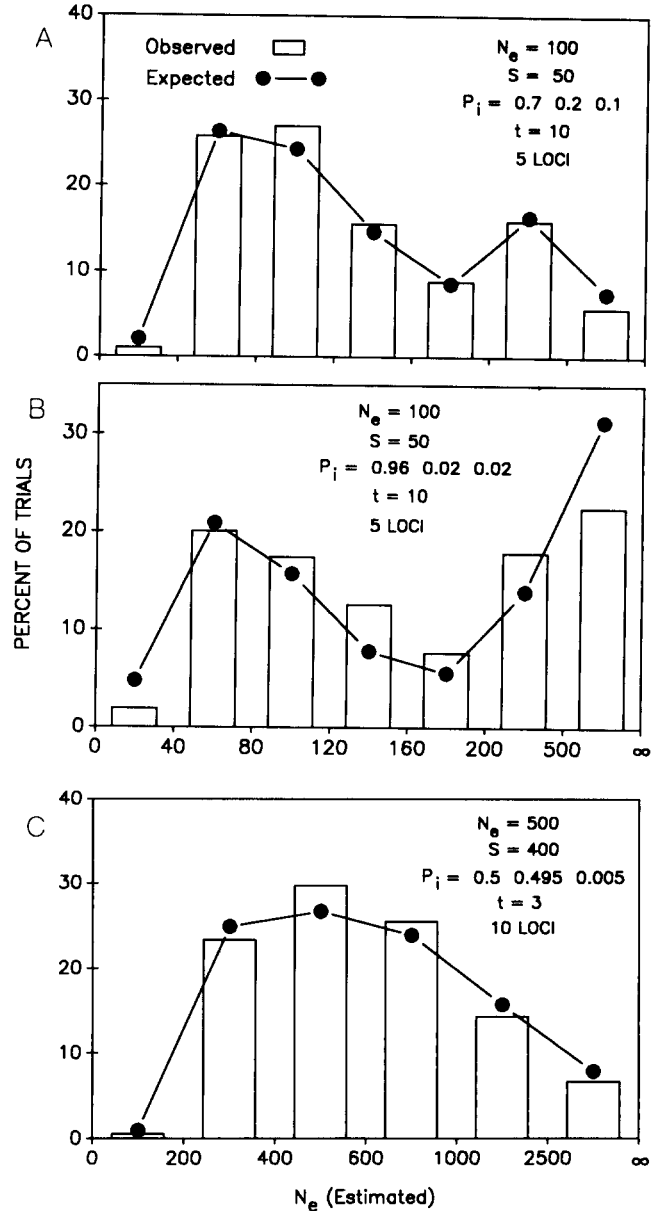

FIGURE 6.—As in Figure 5, but each locus had three alleles. Mean $\hat{F}_k$ was computed using data for the number of loci indicated in each graph.

for the case of completely even initial allele frequencies ($P_i = \frac{1}{3}, \frac{1}{3}, \frac{1}{3}$).

Confidence interval (CI) for N_e: Because the skewed distribution of $\hat{F}$ makes $V(\hat{F})$ and $V(\hat{N}_e)$ generally unsuitable for computing CIs for point estimates of N_e, NEI and TAJIMA (1981) suggested using instead the $\hat{N}_e$ values corresponding to $\hat{F}$ values for the 2.5% and 97.5% cumulative probabilities of the χ^2 distribution. CIs they provided indicate that they computed the 95% CI for $\hat{F}$ as

$$95\% \text{ CI for } \hat{F} = \left[\frac{\chi^2_{0.975[n]}\hat{F}}{n}, \frac{\chi^2_{0.025[n]}\hat{F}}{n} \right]. \quad (15)$$

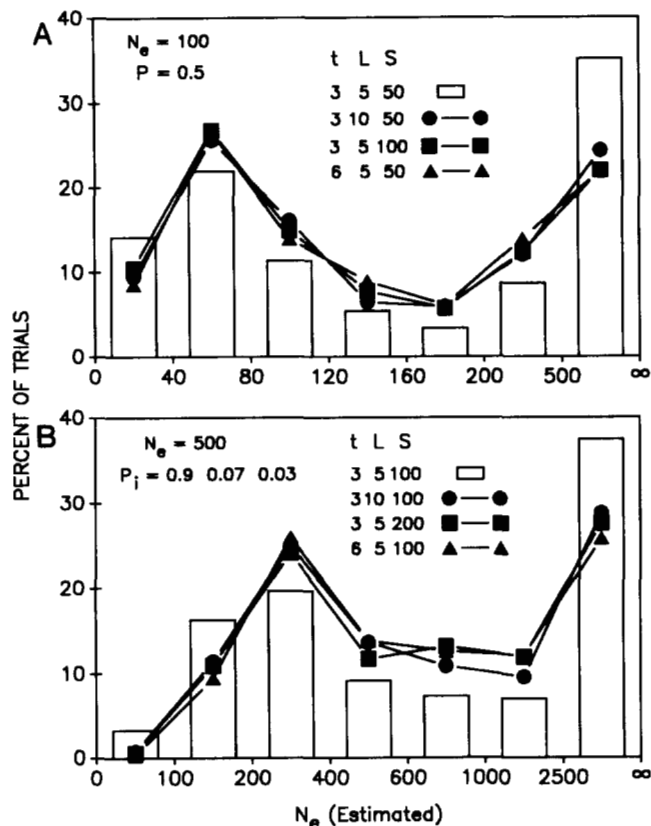This method, however, is only asymptotically correct

FIGURE 7.—Distribution of $\hat{N}_e$ as a function of sample size ($S = S_0 = S_t$), number of generations (t) between samples, and number of loci (L) used to compute $\hat{F}$. $\hat{N}_e$ was computed for 2000 mean $\hat{F}_k$ values sampled with replacement from single locus $\hat{F}$ values generated in the simulations. Simulations used either two (A) or three (B) alleles per locus; effective population size and initial allele frequencies were as indicated. Vertical bars show results for reference simulations; lines with symbols show distribution of $\hat{N}_e$ values as S, L, and t were independently doubled.

for large n. When $\hat{F}$ is based (as it must be) on a finite number of alleles, CIs computed using (15) will be biased upward and will include the true value of F less than 95% of the time. If we note that F is actually a variance of allele frequencies, the quantity $n\hat{F}/E(\hat{F})$ can be represented as ns^2/σ^2, and it is clear that the appropriate formula is that for the CI of a variance (e.g., SOKAL and ROHLF 1969, p. 153):

$$(1 - \alpha) \text{ CI for } \hat{F} = \left[\frac{n\hat{F}}{\chi^2_{\alpha/2[n]}}, \frac{n\hat{F}}{\chi^2_{1-\alpha/2[n]}} \right]. \quad (16)$$

These bounds for the estimate $\hat{F}$ can be used to calculate the CI for $\hat{N}_e$ using (11) or (12). That this approach generally leads to CIs with the desired properties (a proportion $1 - \alpha$ of the CIs containing the true value of N_e, with the remainder equally divided between CIs that are too high and too low) is demonstrated in Table 2. If anything, CIs computed using (16) *do* tend to be slightly conservative because the distribution of $\hat{F}_k$ is usually slightly narrower than the χ^2, but even this effect is small in most cases. Poor

results are obtained only for conditions where the estimate $\hat{N}_e$ is substantially biased, which may occur if alleles with frequencies very close to the boundaries (0,1) are used.

Example: Allele frequency data for the esterase A and B loci in *Dacus olea* presented by KRIMBAS and TSAKAS (1971) were reanalyzed by NEI and TAJIMA (1981), and some of their results are shown in Table 3. The correction terms for finite sampling [$1/(2S_0) + 1/(2S_t)$] were small, reflecting the large number of flies (281–474) examined. Sampling was apparently according to plan II, and an estimated $t = 4$ generations elapsed between the 1966 and 1967 samples. Because NEI and TAJIMA used (12) to compute $\hat{N}_e$, the estimate of N_e would be the same under the present model. However, the CIs for $\hat{N}_e$ reported by NEI and TAJIMA [apparently using (15)] are uniformly higher than those computed using (16).

One point regarding these data that has not been discussed is that they include a large number of alleles at low frequency. Based on allele frequencies for 1966–1967 given by KRIMBAS and TSAKAS (1971), of the 17 alleles for locus A, 13 have mean frequency less than 0.05, and 5 of these are at frequency less than 0.01. For locus B (13 total alleles), the numbers are 10 and 4, respectively. It is reasonable to ask whether estimates of N_e (or their CIs) can be relied upon under these circumstances. Although none of the simulations reported here used the large number of alleles found at these esterase loci, the one depicted in Figure 6C [$N_e = 500$, $S_0 = S_t = 400$, $t = 3$, $P_i = (0.5, 0.495, 0.005)$, $\hat{F}$ based on data for 10 loci ($n = 20$ independent alleles)] mimics the real situation in most other respects. Under these conditions, the distribution of $\hat{N}_e$ was close to the chi square distribution (Figure 6C), and other data for the same simulation (Table 2) suggest that the confidence intervals for $\hat{N}_e$ computed using (16) are, if anything, slightly conservative.

DISCUSSION

The importance of the parameter N_e to many aspects of biology dictates that any method that can provide reasonable estimates of effective population size merits serious attention. The temporal method for analyzing allele frequency change is one such approach, but previous treatments have adopted a rather restrictive model that created doubts about the general usefulness of the method. For example, in the N-T model, it was necessary to assume that the effective population size consisted of N_e breeding individuals, with the distribution of progeny number per parent being binomial. This corresponds to an "ideal" population as defined by KIMURA and OHTA (1971) but is not likely to be an adequate description of most natural populations. The present treatment shows that

TABLE 2

Properties of $\hat{F}_k$ as an estimator of N_e

| | | | | | | | | | | Percent of 95% CI | | | | |
| | | | | | | | | 2 Loci | | | 20 Loci | | | |
N_e	S	Allele frequency			t	Mean $\hat{F}_k$	$\dfrac{(K-1)s_f^2}{\bar{F}^2}$	$\hat{N}_e$	L	C	H	L	C	H
$K = 2$														
100	50	0.800	0.200		3	0.0351	1.96	99	3	95	2	3	95	2
					10	0.0696	1.68	101	2	96	2	2	96	2
100	50	0.975	0.025		3	0.0313	1.26	133	0	98	2	0	98	2
					10	0.0507	1.16	163	1	98	1	0	94	6
500	100	0.990	0.010		3	0.0122	1.27	681	0	97	3	0	99	1
					10	0.0167	1.09	758	0	98	2	0	99	1
$K = 3$														
100	50	0.700	0.200	0.100	3	0.0359	1.82	94	2	95	3	2	95	3
					10	0.0692	1.66	102	1	97	2	1	97	2
100	200	0.700	0.200	0.100	3	0.0197	1.96	102	3	95	2	2	95	3
					10	0.0540	1.68	102	1	96	3	1	97	2
100	50	0.500	0.475	0.025	3	0.0331	1.62	114	1	97	2	1	97	2
					10	0.0590	1.66	128	1	97	2	0	92	8
500	100	0.333	0.333	0.333	3	0.0129	2.12	513	3	94	3	3	94	3
					10	0.0199	2.06	507	3	95	2	3	94	3
500	400	0.500	0.495	0.005	3	0.0053	1.70	533	1	97	2	1	96	3
					10	0.0108	1.82	605	2	96	2	0	93	7

Results are from 5000 replicate simulations (equivalent to sampling 5000 loci) with indicated initial allele frequencies and $S = S_0 = S_t =$ sample size. t is the number of generations between samples; K is the number of alleles per locus. CIs for $\hat{N}_e$ were based on CIs for mean $\hat{F}_k$ and were computed using (16). For each parameter set, 2000 mean $\hat{F}$ values were computed, each based on $\hat{F}$ values for 2 or 20 loci obtained by sampling with replacement from single locus $\hat{F}$ values in simulations. "C" indicates CI contained true value of N_e, "L" indicates CI was too low, and "H" indicates CI was too high.

TABLE 3

CI for $\hat{N}_e$ for *Dacus olea* computed using two methods

| | | | | | 95% CI for $\hat{N}_e$ | |
Locus	n	$\hat{F}_c$	$\left[\dfrac{1}{2S_0} + \dfrac{1}{2S_t}\right]$	$\hat{N}_e$	Equation 15	Equation 16
A	16	0.00609	0.00266	583	240, ∞	175, 2788
B	12	0.00474	0.00285	1056	314, ∞	198, ∞
$A + B$	28	0.00551	0.00274[a]	722	332, 7408	273, 2739

Based on allele frequency data for period 1966–1967 (KRIMBAS and TSAKAS 1971). $\hat{F}_c$ and CI for $\hat{N}_e$ based on (15) are from NEI and TAJIMA (1981). n is the number of independent alleles (one less than the total number of alleles) sampled at each locus. $\hat{N}_e$ was computed using (11) assuming plan II sampling and t = 4 generations elapsed between two samples.

[a] Harmonic mean for A and B.

this restrictive assumption regarding the breeding individuals is unnecessary; $\hat{N}_e$ computed from allele frequency data can be interpreted in a completely general way, corresponding to the variance effective number (CROW 1954; CROW and DENNISTON 1988).

Other authors using the temporal method have either (a) considered just one type of sampling scheme (SCHAFFER, YARDLEY and ANDERSON 1977; PAMILO and VARVIO-AHO 1980; WILSON 1980), or (b) examined only special cases of sampling plans I and II (NEI and TAJIMA 1981; POLLAK 1983). The latter authors considered plan II only for the case $N \gg N_e$, but it is clear from the model used here that $\hat{F}$ (and hence $\hat{N}_e$) is independent of N if sampling is before reproduc-

tion. Total population size must be considered if sampling is after reproduction (plan I), but the restriction that $N = N_e$ imposed by NEI and TAJIMA (1981) is unnecessary. Furthermore, because results for plans I and II converge rapidly if the ratio $r = N/N_e$ is larger than about 2, an adequate approximation to plan I sampling often will be possible even if N is not known exactly.

Robustness of assumptions: The present model thus demonstrates the general applicability of the temporal method to a wide variety of organisms. The usefulness of the method in practical terms will be discussed below, but first it is worthwhile to consider briefly the importance of some other assumptions

common to the current model and those previously proposed.

Constant N_e. If N_e changes over time, then $\hat{N}_e$ computed from (11) or (12) estimates the harmonic mean of the effective population sizes in the individual generations (NEI and TAJIMA 1981; POLLAK 1983).

Discrete generations. HILL (1979) showed that a discrete generation model yields robust estimates of N_e if generations overlap, provided that the population is demographically stable. If demographic parameters change over time, $\hat{F}$ may be biased upwards, leading to an estimate of N_e that is too small (POLLAK 1983).

Neutral alleles. Selection may cause $\hat{F}$ to be either higher or lower than the expectation under pure drift conditions. NICHOLAS and ROBERTSON (1976) and POLLAK (1983) showed that selection of constant intensity has a minor effect on $\hat{F}$ if t/N_e is small, as will typically be the case for the temporal method. MUELLER *et al.* (1985), however, showed that under symmetric variable selection [GILLESPIE's (1978) SAS-CFF model], $\hat{N}_e$ is biased downwards.

No mutation. The time scale feasible for most potential uses of the temporal method means that mutation can safely be ignored.

No migration. In a subdivided population, $\hat{N}_e$ may be strongly affected by migration between subpopulations. Under these conditions, it is also important to consider how the samples for genetic analysis are drawn; *i.e.*, are they taken randomly from the entire population, or from some smaller unit? See NEI and TAJIMA (1981) for a discussion of these points.

Practical considerations: A researcher faced with inevitable constraints on time and resources needs a strategy to maximize precision for a given effort. For this purpose, it is useful to examine an approximate expression for the variance of $\hat{N}_e$ (POLLAK 1983; Equations 28–29):

$$
\begin{aligned}
V(\hat{N}_e) &\approx \left[\frac{8N_e^4}{t^2 n}\right] [t/(2N_e) + 1/(2S_0) + 1/(2S_t)]^2 \\
&= \left[\frac{8N_e^4}{n}\right]\left[\frac{1}{4N_e^2} + \frac{1}{Nt\tilde{S}} + \frac{1}{t^2\tilde{S}^2}\right].
\end{aligned}
\tag{17}
$$

In (17) $\tilde{S}$ is the harmonic mean of S_0 and S_t, which are presumed to be the same for each locus. It is clear from (17) that increasing sample size can be expected to have the same effect on precision as the same proportional increase in time between samples. It can also be shown that if $t\tilde{S}/N_e > \sqrt{2}$, increasing the number of loci (or alleles) has a relatively greater effect on precision than does an increase in $\tilde{S}$ or t, while increasing sample size or time between samples is more effective if $t\tilde{S}/N_e < \sqrt{2}$. In practice, the effects of n are very similar to those of $\tilde{S}$ and t unless $t\tilde{S}/N_e$ is much larger or smaller than $\sqrt{2}$. For two sets of simulations with very different initial conditions [$t\tilde{S}/$

$N_e = 3 \times 50/100 = 1.5$ (Figure 7A), $= 3 \times 100/500 = 0.6$ (Figure 7B)], precision increased by about the same degree whether sample size, number of alleles surveyed, or number of generations between samples was doubled. Naturally, increasing all three items simultaneously is the best way to ensure greater precision.

Meaningful information about effective population size generally cannot be obtained unless some minimum experimental conditions are satisfied. First, data for a number of independent alleles are required. Use of uncommon alleles will increase precision of $\hat{N}_e$ without causing appreciable bias provided they are not at very low frequency. It appears that the primary factor leading to bias is not rare alleles *per se*, but the absence of alleles at intermediate frequency. Problems are most likely to occur for loci with a single allele at high frequency and the remaining alleles at low frequency, and even then bias may be small unless t is large. In cases where bias is a concern, some lumping of alleles may be desirable. Second, 50 individuals would appear to be a minimum sample size for reasonable precision unless a large number of loci can be surveyed. Because it will often be easier to increase sample size than either the number of polymorphic loci surveyed or the elapsed time of the experiment, those using the temporal method should consider using 100 or more individuals whenever possible.

Finally, a point that has not been fully appreciated is that the temporal method is not equally well suited to the analysis of populations of all sizes. POLLAK (1983) argued that indirect methods of estimating effective population size are necessary only if N_e is very large, but it is in just such cases that the temporal method is least reliable. Because the term $t/(2N_e)$ contributes proportionally less to $\hat{F}$ as N_e increases, the effects of genetic drift in large populations may be swamped by sampling error. As a result, often it will be difficult to distinguish a large population from an infinitely large one.

As pointed out by NEI and TAJIMA (1981), precision for the temporal method increases with the ratio S/N_e, which means that populations with small N_e are most effectively studied. The temporal method should thus be useful in the field of conservation biology, where inbreeding depression and loss of genetic variability due to small effective population size are a major concern. The total number of individuals can be fairly large even in populations with limited N_e, particularly for species with high fecundity and juvenile mortality. Examples include many plants, insects, and marine organisms with planktonic larvae. In such cases, it may be very difficult to obtain even an approximate idea of effective population size by other methods, and the temporal approach can be very useful in determining whether N_e is small relative to

N. The temporal method does not always yield a small $\hat{N}_e$ for populations with small effective size, as shown in Figures 4–7, but $\hat{N}_e$ is unlikely to be small if true N_e is large (Figures 4B, 6C and 7B), provided the number of individuals and loci samples are adequate. Therefore, although large estimates of $\hat{N}_e$ may be ambiguous, small $\hat{N}_e$ values can be a reliable indication that effective population size is indeed limited.

LITERATURE CITED

CROW, J. F., 1954 Breeding structure of populations. II. Effective population number, pp. 543–556, in *Statistics and Mathematics in Biology*, edited by O. KEMPTHORNE, T. A. BANCROFT, J. W. GOWEN and J. L. LUSH. Iowa State College Press, Ames, IA.

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

GILLESPIE, J. H., 1978 A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. Theor. Popul. Biol. **14**: 1–45.

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

HILL, W. G., 1981 Estimation of effective population size from data on linkage disequilibrium. Genet. Res. **38**: 209–216.

KIMURA, M., and T. OHTA, 1971 *Theoretical Aspects of Population Genetics.* Princeton University Press, Princeton, N.J.

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.

LAURIE-AHLBERG, C. C., and B. S. WEIR, 1979 Allozyme variation and linkage disequilibrium in some laboratory populations of *Drosophila melanogaster*. Genetics **92**: 1295–1314.

LEWONTIN, R. C., and J. KRAKAUER, 1973 Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. Genetics **74**: 175–195.

MUELLER, L. D., B. A. WILCOX, P. E. EHRLICH, D. G. HECKEL, and D. D. MURPHY, 1985 A direct assessment of the role of genetic drift in determining allele frequency variation in populations of *Euphydryas editha.* Genetics **110**: 495–511.

NEI, M., 1978 Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**: 583–590.

NEI, M., and R. K. CHESSER, 1983 Estimation of fixation indices and gene diversities. Ann. Hum. Genet. **47**: 253–259.

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

NICHOLAS, F. W., and A. ROBERTSON, 1976 The effect of selection on the standardized variance of gene frequency. Theor. Appl. Genet. **48**: 263–268.

PAMILO, P., and S. VARVIO-AHO, 1980 On the estimation of population size from allele frequency changes. Genetics **95**: 1055.

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

RAO, C. R., 1973 *Linear Statistical Inference and Its Applications*, Ed. 2. John Wiley & Sons, New York.

SCHAFFER, H. E., D. YARDLEY and W. W. ANDERSON, 1977 Drift or selection: a statistical test of gene frequency change over generations. Genetics **87**: 371–379.

SOKAL, R. R., and F. J. ROHLF, 1969 *Biometry.* W. H. Freeman, San Francisco.

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

WEIR, B. S., and C. C. COCKERHAM, 1984 Estimating F-statistics for the analysis of population structure. Evolution **38**: 1358–1370.

WILSON, S. R., 1980 Analyzing gene-frequency data when effective population size is finite. Genetics **95**: 489–502.

APPENDIX

The N-T model: The current model can be related to that used by NEI and TAJIMA (1981; the "N-T" model) and POLLAK (1983) in the following way. If $V(x)$ and $V(y)$ are defined with respect to the allele frequency (P') in the N individuals making up the total population in generation zero, then, for either sampling plan, the sample at time $t = 0$ is hypergeometric from a population of N individuals, so

$$V(x) = E(x - P')^2 = \frac{P'(1 - P')}{2S_0}\left[1 - \frac{(2S_0 - 1)}{(2N - 1)}\right]. \quad (A1)$$

NEI and TAJIMA (1981) showed that if $N_e = N$ and sampling is after reproduction (plan I),

$$V(y) = E(y - P')^2$$
$$= P'(1 - P')\left[1 - \left(1 - \frac{1}{2N}\right)^{t-1}\left(1 - \frac{1}{2S_t}\right)\right], \quad (A2)$$

while the corresponding equation for plan II is

$$V(y) = P'(1 - P')\left[1 - \left(1 - \frac{1}{2S_t}\right)\left(1 - \frac{1}{2N_e}\right)^{t-1}\right.$$
$$\left.\cdot\left(1 - \frac{1}{2N_e}\left[1 - \frac{2N_e - 1}{2N - 1}\right]\right)\right]. \quad (A3)$$

POLLAK (1983) pointed out that the restriction $N_e = N$ for plan I is not necessary, but he incorrectly identified (A2) as the proper expression for $V(y)$ if $N > N_e$. This ignores the hypergeometric sampling of the N_e breeding individuals in the initial generation. In fact, $V(y)$ is identical for the two sampling plans and is correctly given by (A3), which reduces to (A2) for plan I if $N_e = N$ [plan II is subject to the restriction that $N \geq (N_e + S_0)$].

Therefore, for the N-T model, the variance of $(x - y)$ is given by

$$V(x - y) = E(x - y)^2 = V(x) + V(y) - 2\text{Cov}(x, y)$$
$$= P'(1 - P')\left\{\frac{1}{2S_0}\left(1 - \frac{2S_0 - 1}{2N - 1}\right) + 1\right.$$
$$- \left(1 - \frac{1}{2S_t}\right)\left(1 - \frac{1}{2N_e}\right)^{t-1} \quad (A4)$$
$$\left.\cdot\left(1 - \frac{1}{2N_e}\left[1 - \frac{2N_e - 1}{2N - 1}\right]\right)\right\} - 2\text{Cov}(x, y)$$

for both plans I and II. As in the current model, $V(x)$ and $V(y)$ are identical for plans I and II, but in general $V(x - y)$ differs because $\text{Cov}(x, y)$ is not the same for the two sampling plans. The covariance terms differ from those in the current model because P', rather than P, is used as a point of

reference. In plan I of the N-T model, S_0 and S_t are independent with respect to P', so $\text{Cov}(x, y) = 0$. If sampling is according to plan II, the genes drawn for S_0 and N_e in the initial generation are negatively correlated, as, therefore, are x and y. NEI and TAJIMA and POLLAK examined plan II only for $N \gg N_e$, in which case $\text{Cov}(x, y)$ may be safely ignored. However, as shown below, this restriction is not necessary.

Relationship with present model: The N-T model and the model used here should be equivalent if N is infinitely large, because in that case $P' = P$ and it does not matter which is used as a point of reference. This can be verified by substituting $N = \infty$ in (A1) and (A3), which then reduce to (1) and (4), respectively. It is also possible to derive $V(x - y)$ for the present model from the corresponding expression (A4) for the N-T model. To do this, $P'(1 - P')$ in (A4) must be expressed in terms of $P(1 - P)$:

$$E[P'(1 - P')] = E(P') - E(P'^2) = P - [V(P') + P^2]$$

$$= P - P^2 - \frac{P(1 - P)}{2N} = P(1 - P)[1 - 1/(2N)].$$

Consider first plan I [$\text{Cov}(x, y) = 0$]. If $P(1-P)[1 - 1/(2N)]$ is substituted for $P'(1 - P')$ in (A4):

$$V(x - y) = P(1 - P)[1 - 1/(2N)] \times \left\{ \frac{1}{2S_0} \left(1 - \frac{2S_0 - 1}{2N - 1} \right) \right.$$

$$\left. + 1 - \left(1 - \frac{1}{2S_t} \right)\left(1 - \frac{1}{2N_e} \right)^{t-1}\left(1 - \frac{1}{2N_e}\left[1 - \frac{2N_e - 1}{2N - 1} \right] \right) \right\},$$

the result, after some manipulation, is (7). That is, the two ways of expressing $V(x - y)$ are equivalent.

This equivalence can be used to calculate $\text{Cov}(x, y)$ for plan II in the N-T model. In comparing $V(x, y)$ for plans I and II in the N-T model, we see that the difference is $-2\text{Cov}(x, y)$, while in the present model the difference is $P(1 - P)/N$. Therefore,

$$-2\text{Cov}(x, y) = \frac{P(1 - P)}{N} = \frac{P'(1 - P')}{N} \frac{(2N)}{2N - 1};$$

$$\text{Cov}(x, y) = -\frac{P'(1 - P')}{2N - 1}.$$

For plan II sampling in the N-T model, $\text{Cov}(x, y)$ is a decreasing function of N, while $V(x)$ and $V(y)$ [given by (A1) and (A4)] increase with N. The net result is that $V(x - y)$ is virtually independent of N for plan II in the N-T model; to a very good approximation, its value is given by

$$V(x - y) = P'(1 - P') \left[\frac{1}{2S_0} + 1 - \left(1 - \frac{1}{2N_e} \right)^{t}\left(1 - \frac{1}{2S_t} \right) \right],$$

regardless of the actual value of N.