

Estimation of the variance effective population size in age structured populations



Fredrik Olsson*, Ola Hössjer

Department of Mathematics, Div. of Mathematical Statistics, Stockholm University, Stockholm, Sweden

ARTICLE INFO

Article history:

Received 7 January 2015

Available online 19 February 2015

Keywords:

Variance effective population size
Temporal method
Effective number of independent alleles
Overlapping generations
Confidence interval

ABSTRACT

The variance effective population size for age structured populations is generally hard to estimate and the temporal method often gives biased estimates. Here, we give an explicit expression for a correction factor which, combined with estimates from the temporal method, yield approximately unbiased estimates. The calculation of the correction factor requires knowledge of the age specific offspring distribution and survival probabilities as well as possible correlation between survival and reproductive success. In order to relax these requirements, we show that only first order moments of these distributions need to be known if the time between samples is large, or individuals from all age classes which reproduce are sampled. A very explicit approximate expression for the asymptotic coefficient of standard deviation of the estimator is derived, and it can be used to construct confidence intervals and optimal ways of weighting information from different markers. The asymptotic coefficient of standard deviation can also be used to design studies and we show that in order to maximize the precision for a given sample size, individuals from older age classes should be sampled since their expected variance of allele frequency change is higher and easier to estimate. However, for populations with fluctuating age class sizes, the accuracy of the method is reduced when samples are taken from older age classes with high demographic variation. We also present a method for simultaneous estimation of the variance effective and census population size.

© 2015 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In population genetics, it is necessary to be able to summarize the population with some key concepts. By imposing different assumptions, analysis of complex population systems can be manageable and comparisons of various populations as well as population scenarios can be made. However, the degree of realism of the assumptions varies, and for real populations they might be violated to some extent. Hence, it is important to study the effect of the assumptions for the concept under study, and, for significant effects, the model may need to be refined to obtain more realistic and accurate results.

In this paper, we study methods for estimating one such concept known as the effective population size (N_e). It was presented by Wright (1931) as a method for approximating populations by an idealized reference population with respect to some criterion. Many closely related effective sizes depending on the criterion under consideration have been developed and studied, such as

the variance, inbreeding, coalescence and eigenvalue effective size (Crow and Denniston, 1988; Wang and Caballero, 1999; Kaj et al., 2001; Waples, 2002; Ewens, 2004; Sagitov and Jagers, 2005; Sjödin et al., 2005; Charlesworth, 2009; Hössjer, 2011). Here, we focus on the variance effective population size (N_{eV}), defined as the number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift as the population under consideration (Crow, 1954). Expressions for N_{eV} have previously been derived for age structured populations with fixed population size (Felsenstein, 1971; Hill, 1972, 1979; Waples et al., 2011), for populations with fluctuating population size (Engen et al., 2005; Olsson et al., 2013), for determining the influence of e.g. mating structure and variation in fecundity (Nunney, 1991, 1993, 1996). These models are important in order to study e.g. how harvest strategies affect the population (Hard et al., 2006).

There exist a number of methods for estimating N_e in real populations (Waples, 1989b; Wang and Whitlock, 2003; Luikart et al., 2010). For N_{eV} , the temporal method (Krimbas and Tsakas, 1971) is commonly used. It is based on the idea that if genetic drift is the only cause of allele frequency change, N_{eV} can be estimated from temporal allele frequency shifts. Different methods to estimate the standardized variance of allele frequency change have been

* Corresponding author.

E-mail address: fredriko@math.su.se (F. Olsson).

presented (Nei and Tajima, 1981; Pollak, 1983; Jorde and Ryman, 2007), and comparisons between them can be found in Waples (1989b) and Waples and Yokota (2007).

A crucial assumption for all the above mentioned methods is that the population has discrete generations. However, for populations with overlapping generations, the genetic drift is no longer the only source of allele frequency change (Jorde and Ryman, 1995). Effective population sizes calculated based on allele frequency measurements close in time will be affected by the age structure of the population. Jorde and Ryman (1995) presented a modification of the temporal method, using a correction factor calculated from the demographic parameters of the population. By combining estimates of the allele frequency change in one age class and assumptions about distributions and parameters in the demographic population model, they derived a correction factor which, when applied to the estimate from the temporal method, yield approximately unbiased estimates of the expected long term standardized variance of allele frequency change.

Another way to reduce the bias due to the age structure when estimating N_{ev} is to increase the time between measurements (Jorde and Ryman, 1995; Waples and Yokota, 2007; Olsson et al., 2013).

In this paper we extend the estimator of Jorde and Ryman (1995), and their correction factor, to models with overlapping generations in several ways. We present a method for weighting samples from different age classes to form a population sample. The assumption in Jorde and Ryman (1995), that the number of births of individuals are Poisson distributed, is extended to situations with overdispersion. We also allow for correlated survival and reproductive success. With this model we can study how the correction factor depends on different parameters, such as the time between samples and the age classes that are present in the sample.

Another important criterion of an estimator, apart from low bias, is a small variance. Here, we give an approximate formula for the asymptotic coefficient of standard deviation of our proposed estimator of N_{ev} . Since the standard deviation of the estimator is typically much larger than the bias, this formula is useful when planning a study in order to determine how many individuals that should be sampled from the different age classes and also how many polymorphic loci we need genetic data from as well as optimal ways of weighting information from them. It is a very explicit function of key parameters of the model, such as sample size, sampling interval and the effective number of independent alleles, and therefore it is of independent interest. Using the approximate formula for the asymptotic coefficient of standard deviation, we also present approximate confidence intervals for the estimator of N_{ev} .

If individuals are sampled from all age classes and weighted by their reproductive numbers (Fisher, 1958), the correction factor equals one. Hence, if we use the reproductive numbers as weights we do not need to specify the variances of reproductive success for the different age classes, or possible correlation between reproductive success and survival. The same is true when the time interval between samples becomes large.

The paper is structured as follows: In Section 2 we describe the demographic model that is used throughout the paper. Next, in Section 3 we describe the estimation procedure of N_{ev} for populations with discrete generations and we generalize the estimators of Jorde and Ryman (1995, 2007) in Section 4. By means of simulation we illustrate the performance of our proposed estimator in various situations. In Section 5 we present a method for simultaneous estimation of the variance effective population size and the census population size and in Section 6 we present an optimal way of weighting loci. A discussion is found in Section 7, derivations and some examples are gathered in the appendices and a summary of the most important notation is collected in Table 1.

Table 1

List of notation used in the paper.

Notation	Definition
N_{ev}	Variance effective population size
N_t	Total number of individuals at time t
\mathbf{N}_t	Age composition at time t
\mathbf{g}	Expected projection matrix
b_j	Mean number of offspring for an individual in age class j
l_j	Probability that an individual survives to age class j
s_j	Probability that an individual in age class j survives to age class $j + 1$
ρ_j	Correlation between survival and number of offspring for an individual in age class j
λ	Multiplicative growth rate and largest eigenvalue of \mathbf{g}
\mathbf{u}	Vector with approximate equilibrium age distribution for \mathbf{N}_t
\mathbf{v}	Vector with reproductive values
\mathbf{w}	Vector of age class weights w_j normalized such that $\sum_{j=0}^{J-1} w_j = 1$
T	Generation time i.e. mean age of parents of newborns
F	Standardized variance of allele frequency change
F_s	Estimator of F presented by Jorde and Ryman (2007)
F^w	Standardized variance of weighted allele frequency change
\hat{F}_{OH}^w	Estimator of F^w before bias correction
F_{OH}^{w*}	Approximately unbiased estimator of F^w
n_{1j}	Number of sampled individuals from age class j at the first time point
n_{2j}	Number of sampled individuals from age class j at the second time point
n_1	Effective number of sampled individuals at the first time point
n_2	Effective number of sampled individuals at the second time point
\bar{n}	Time averaged effective sample size
τ	Time between measurements when applying the temporal method for assessing N_{ev} . One time unit represents the age difference of two successive age classes
\mathbf{A}	Projection matrix for standardized covariances
λ_A	Largest eigenvalue of \mathbf{A}
C	Correction factor for estimated genetic drift based on population demographics
L	Number of loci
L_e	Effective number of independent alleles
κ_l	Inverse information at locus l for estimating N_{ev}

2. Population model

We will consider a population of monocious diploid individuals divided into J age classes, with two homologous chromosomes per individual. At time t the population consists of $N_t = \sum_{j=0}^{J-1} N_{tj}$ individuals and the age composition is contained in the column vector $\mathbf{N}_t = (N_{t0}, \dots, N_{t,J-1})'$, where N_{tj} is the number of individuals in age class j and $'$ denotes transposition. Let Y_{tjh} be the number of progeny in the next time step of chromosome h in age class j and let Y_{tjh} be independent and identically distributed random variables with expected value $E[Y_{tjh}] = b_j$ and variance $\text{Var}(Y_{tjh}) = \sigma_j^2$. The survival of this chromosome, I_{tjh} , is Bernoulli distributed with mean s_j and the correlation, $\text{Corr}(Y_{tjh}, I_{tjh})$, between survival and number of offspring is denoted ρ_j . Let $N_{t+1,j0}$ be the total number of progeny at time $t + 1$ of all individuals in age class j at time t . Then,

$$\begin{aligned}
 2N_{t+1,j+1} &= \sum_{h=1}^{2N_{tj}} I_{tjh}, \quad j = 0, \dots, J-2, \\
 2N_{t+1,j0} &= \sum_{h=1}^{2N_{tj}} Y_{tjh}, \quad j = 0, \dots, J-1, \\
 N_{t+1,0} &= \sum_{j=0}^{J-1} N_{t+1,j0},
 \end{aligned} \tag{1}$$

describe the time dynamics of the population.

Let $\mathbf{g} = (g_{ij})$ be a Leslie matrix (Leslie, 1945) with non-zero entries $g_{0j} = b_j, j = 0, \dots, J-1$ and $g_{j+1,j} = s_j, j = 0, \dots, J-2$. Perron–Frobenius Theorem ensures that \mathbf{g} has a unique largest positive eigenvalue, λ , which is the multiplicative growth rate of

the population. The right eigenvector, $\mathbf{u} = (u_0, \dots, u_{J-1})'$ corresponding to λ , normalized so that $\sum_{j=0}^{J-1} u_j = 1$, is the approximate equilibrium age distribution, and the left eigenvector $\mathbf{v} = (v_0, \dots, v_{J-1})$, normalized so that $\mathbf{v}\mathbf{u} = 1$, contains the reproductive values (Fisher, 1958). The expected time dynamics of the population can then be described by means of matrix recursions

$$E(\mathbf{N}_{t+1}|\mathbf{N}_t) = \mathbf{g}\mathbf{N}_t.$$

We will focus on two types of progeny distributions, either the number of progeny of individuals in age class j follows a Poisson distribution $Y_{tjh} \sim \text{Po}(b_j)$, so that

$$P(Y_{tjh} = y) = \exp(-y) \frac{b_j^y}{y!}, \tag{2}$$

and $\sigma_j^2 = b_j$; or a negative binomial distribution $Y_{tjh} \sim \text{NegBin}(m_j, q_j)$, with

$$P(Y_{tjh} = y) = \binom{m_j + y - 1}{y} (1 - q_j)^{m_j} q_j^y, \tag{3}$$

corresponding to

$$E(Y_{tjh}) = m_j(1 - q_j)/q_j = b_j, \\ \text{Var}(Y_{tjh}) = m_j(1 - q_j)/q_j^2 = \sigma_j^2 > b_j.$$

For joint survival and reproduction, we use a logistic regression model

$$P(I_{tjh} = 1|Y_{tjh} = y) = \frac{\exp(\beta_{0j} + \beta_{1j}y)}{1 + \exp(\beta_{0j} + \beta_{1j}y)} =: s_{jy},$$

so that

$$s_j = P(I_{tjh} = 1) = \sum_{y=0}^{\infty} s_{jy}P(Y_{tjh} = y),$$

$$b_j s_j + \sigma_j \sqrt{s_j(1 - s_j)} \rho_j = E(Y_{tjh} I_{tjh}) = \sum_{y=0}^{\infty} y s_{jy} P(Y_{tjh} = y),$$

gives a one-to-one correspondence between (s_j, ρ_j) and (β_{0j}, β_{1j}) , for $j = 0, 1, \dots, J - 2$, if the offspring distribution has already been specified. In particular, $\rho_j = 0$ corresponds to $\beta_{1j} = 0$, and then Y_{tjh} and I_{tjh} are independent, with

$$s_j = \frac{\exp(\beta_{0j})}{1 + \exp(\beta_{0j})}.$$

3. Estimation of N_{eV} for populations with discrete generations

For a population with non-overlapping generations, the variance effective population size N_{eV} is obtained by solving

$$1 - \left(1 - \frac{1}{2N_{eV}}\right)^\tau = F,$$

which yields

$$N_{eV} = \frac{1}{2} \left\{ 1 - \left[1 - \frac{E([p_{t+\tau} - p_t]^2 | p_t)}{p_t(1 - p_t)} \right]^{1/\tau} \right\}^{-1} \\ = \frac{1}{2} [1 - (1 - F)^{1/\tau}]^{-1}$$

as a function of the allele frequencies p_t and $p_{t+\tau}$ of a selectively neutral gene at times t and $t + \tau$, where

$$F = F_t = \frac{E([p_{t+\tau} - p_t]^2 | p_t)}{p_t(1 - p_t)} \tag{4}$$

is the standardized variance of the allele frequency change. With this definition, the standardized genetic drift F is the same as for a Wright–Fisher population of diploid size N_{eV} .

One way to estimate N_{eV} is by the temporal method where we assume that n_1 and n_2 individuals are sampled at time t and $t + \tau$ respectively. Previously, two different sampling plans have been identified for populations with discrete generations (Nei and Tajima, 1981; Waples, 1989a). Under plan I individuals are sampled without replacement after reproduction or sampled before reproduction and subsequently returned to the population. Under plan II, individuals are sampled destructively before they reproduce.

From the sampled individuals, the allele frequencies at a number of loci are used to estimate the standardized variance F of allele frequency change. There exist many estimators F (e.g. Nei and Tajima, 1981; Pollak, 1983). In particular

$$F_s = \frac{\sum_{a=1}^A (\hat{p}_{t+\tau,a} - \hat{p}_{t,a})^2}{\sum_{a=1}^A \frac{\hat{p}_{t+\tau,a} + \hat{p}_{t,a}}{2} \left(1 - \frac{\hat{p}_{t+\tau,a} + \hat{p}_{t,a}}{2}\right)}, \tag{5}$$

is the estimator presented by Jorde and Ryman (2007) for one locus. Here A is the number of alleles and $\hat{p}_{t,a}$ is an estimator of $p_{t,a}$, the allele frequency of allele a at time t , and τ is the time between samples. With this estimator, alleles with allele frequency close to 0.5 will be weighted higher than rarer alleles. Correcting F_s for sampling bias, Jorde and Ryman (2007) found an approximately unbiased estimator F'_s of the standardized allele frequency change F_s . The variance effective population size can then be estimated by

$$\hat{N}_{eV} = \frac{1}{2} [1 - (1 - F'_s)^{1/\tau}]^{-1} \approx \frac{\tau}{2F'_s},$$

where the approximation is accurate when $\tau N_{eV} \gg 1$.

4. Estimation of N_{eV} for populations with overlapping generations

In order to extend the temporal method to populations with overlapping generations, we first need to clarify the sampling mechanism as well as how the allele frequencies in the population are defined.

4.1. Sampling

For populations with overlapping generations the interpretation of the two sampling schemes I and II is unclear (Jorde and Ryman, 1995). Here, we will only consider situations where sampling is non-destructive but we distinguish between sampling with or without replacement.

One advantage with plan II is that in order to compute the estimate of F we need no knowledge of the total population size (Waples, 1989a). This also holds if we sample non-destructively with replacement. However, this sampling scheme might be unpractical since we have to return each individual to the population before the next one is drawn. Our other scheme, sampling non-destructively without replacement, corresponds to plan I above.

4.2. Allele frequency

Following the approach in Olsson et al. (2013), we let

$$p_{tal}^w = \sum_{j=0}^{J-1} w_j p_{tjal} \tag{6}$$

be a weighted average over all age classes of p_{tjal} , the frequency of allele a in age class j at time t at locus l , and $\mathbf{w} = (w_0, \dots, w_{J-1})$ are

non-negative weights such that $\sum_{j=0}^{J-1} w_j = 1$. When estimating these quantities in the population, we let \hat{p}_{tjal} be the estimate of p_{tjal} based on n_{1j} sampled individuals of age class j , and

$$\hat{p}_{tal}^w = \sum_{j=0}^{J-1} w_j \hat{p}_{tjal}$$

the corresponding estimate of the age averaged allele frequency. However, if $n_{1j} = 0$ we let $\hat{p}_{tjal} = 0$ for all a and l and we also require that $w_j = 0$. The sampling variance for a particular locus l and allele a at time t equals

$$\text{Var}(\hat{p}_{tal}^w | \{p_{tjal}^w\}_{j=0}^{J-1}) = \sum_{j=0}^{J-1} \frac{D_j w_j^2 p_{tjal}^w (1 - p_{tjal}^w)}{2n_{1j}}, \quad (7)$$

where $D_j = (2N_{tj} - 2n_{1j}) / (2N_{tj} - 1)$ when sampling without replacement, and $D_j = 1$ when sampling with replacement.

4.3. Temporal method for an age structured population

It is well known that it is intractable to quantify genetic drift for structured populations (Ewens, 1982; Hössjer and Ryman, 2014). For a fixed allele a at locus l we could treat $p_t^w = p_{tal}^w$ as an allele frequency of a homogeneous population and quantify genetic drift using F in (4). However, for a structured population, due to the conditional expectations in (4), this formula will in general depend on p_t^w . It is more convenient to use

$$F^w = \frac{E((p_{t+\tau}^w - p_t^w)^2)}{E(p_t^w [1 - p_t^w])} \quad (8)$$

where the conditional expectation in (4) is replaced by an expectation with respect to generation 0. In order to compute an approximation of (8), we will use a multilocus variant

$$\hat{F}_{OH}^w = \frac{\sum_{l=1}^L \sum_{a=1}^{A_l} (\hat{p}_{t+\tau,al}^w - \hat{p}_{tal}^w)^2}{\sum_{l=1}^L \sum_{a=1}^{A_l} \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2} \left(1 - \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2}\right)} \quad (9)$$

of (5) as our estimator of the standardized variance of the allele frequency change, where L is the number of loci and A_l the number of alleles at locus l . We show in Appendix A that F_s' in the previous section can be extended to an approximately unbiased estimator

$$F_{OH}^{w*} = \frac{\hat{F}_{OH}^w [1 - 1/(4\bar{n})] - 1/\bar{n}}{(1 + \hat{F}_{OH}^w/4)[1 - 1/(2n_2)]} \quad (10)$$

of the standardized variance F^w of allele frequency change between time t and $t + \tau$ for weighting scheme w , where

$$\begin{aligned} 1/n_1 &= \sum_{j=0}^{J-1} \frac{w_{1j}^2}{n_{1j}}, \\ 1/n_2 &= \sum_{j=0}^{J-1} \frac{w_{t+\tau,j}^2}{n_{2j}}, \end{aligned} \quad (11)$$

if genes are sampled with replacement, with n_{2j} the number of individuals sampled in age class j at time $t + \tau$, and

$$\begin{aligned} 1/n_1 &= \sum_{j=0}^{J-1} \frac{w_{1j}^2}{n_{1j}} \frac{2N_{tj} - 2n_{1j}}{2N_{tj} - 1} \approx \sum_{j=0}^{J-1} \frac{w_{1j}^2}{n_{1j}} \left(\frac{u_j - n_{1j}/N_t}{u_j - 1/[2N_t]} \right), \\ 1/n_2 &= \sum_{j=0}^{J-1} \frac{w_{t+\tau,j}^2}{n_{2j}} \frac{2N_{t+\tau,j} - 2n_{2j}}{2N_{t+\tau,j} - 1} \\ &\approx \sum_{j=0}^{J-1} \frac{w_{t+\tau,j}^2}{n_{2j}} \left(\frac{u_j - n_{2j}/N_{t+\tau}}{u_j - 1/[2N_{t+\tau}]} \right), \end{aligned} \quad (12)$$

when genes are sampled without replacement. The numbers n_1 and n_2 can be interpreted as the effective sample sizes at time t and $t + \tau$ and \bar{n} is their harmonic mean,

$$\frac{1}{\bar{n}} = \frac{1}{2n_1} + \frac{1}{2n_2}, \quad (13)$$

interpreted as a time averaged effective sample size.

4.4. Generalization of the Jorde–Ryman correction factor

For most choices of w , the temporal allele frequency changes will be affected by the age structure. Since we are generally interested in the long time average rate at which the standardized variance of allele frequency change increases, F_{OH}^{w*} will typically not describe the quantity of interest. To overcome this problem, Jorde and Ryman (1995) derived a correction factor C based on population demographics. They showed that this factor can be applied to samples close in time from one age class in order to predict the long time average rate. Jorde (2012) improved the method to allow for time intervals of arbitrary length between the samples. Here, we generalize their results to accommodate samples from multiple age classes, general weighting schemes, overdispersion ($b_i < \sigma_i^2$) and correlated reproduction and survival ($\rho_i \neq 0$).

In Appendix B, we show that the estimator (10) has expected value

$$E(F_{OH}^{w*}) \approx F^w.$$

However, the standardized variance of allele frequency change, F^w , depends on the weighing scheme w , and does not describe the long term genetic drift for age structured populations in general. Whereas the average rate at which the variance of allele frequency change increases per time step during a time interval of length τ is $1 - (1 - F^w)^{1/\tau}$ when using weighting scheme w , the long term rate at which the variance of allele frequency change increases per time step is $1 - \lambda_A$, where λ_A is derived in Appendix B. It is the largest eigenvalue of a certain matrix A , and it can be calculated from the assumptions of the demographic model. By dividing these two rates, we generalize the expression for the constant C in Jorde and Ryman (1995) to

$$C = \frac{1 - (1 - F^w)^{1/\tau}}{1 - \lambda_A}. \quad (14)$$

If we put $\tau = 1$, $w = (1, 0, \dots, 0)$ and let $N_{tj} \rightarrow \infty$ for all t and $j = 0, \dots, J - 1$, in (14) we obtain the correction factor of Jorde and Ryman (1995). In Appendix C we give an example of how to calculate the correction factor, assuming constant age class sizes and Poisson reproduction, and another example in Appendix D with non-constant age class sizes and overdispersion.

As a next step, we define the (diploid) variance effective population size per generation implicitly through

$$\lambda_A = \left(1 - \frac{1}{2N_{ev}}\right)^{1/T}, \quad (15)$$

where

$$T = \sum_{j=0}^{J-1} (j+1) l_j b_j \lambda^{-j-1},$$

is the mean age of parents of a newborn. The corresponding estimator of N_{ev} is obtained by combining (14) and (15) with F_{OH}^{w*} instead of F^w and solve for N_{ev} . This yields

$$\hat{N}_{ev} = \frac{1}{2} \left\{ 1 - \left[1 - \frac{1 - (1 - F_{OH}^{w*})^{1/\tau}}{C} \right]^T \right\}^{-1} \approx \frac{C\tau}{2TF_{OH}^{w*}}. \quad (16)$$

Omitting T in (16), we end up instead with an expression for an estimator of a version of N_{ev} that quantifies drift per time step.

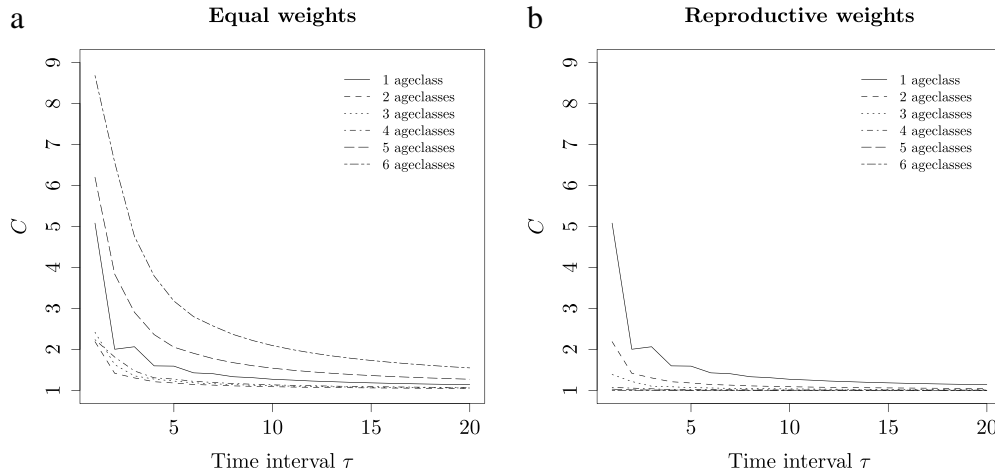


Fig. 1. The correction factor (14) as a function of the time interval τ between samples. Demographic parameters are according to the sparrow data of Table 2. The lines correspond to different choices of weight vectors where the i :th line consists of the first i age classes having equal weights, $w_j = 1/i, j = 0, \dots, i - 1$ in subplot (a), and weights proportional to their reproductive values $w_j = u_j v_j / \sum_{k=0}^{i-1} u_k v_k, j = 0, \dots, i - 1$ in subplot (b).

4.5. Reproductive weights

It has previously been shown that weighting individuals according to their reproductive value when calculating the effective population size for populations with overlapping generations eliminates the fluctuations due to the age structure (Felsenstein, 1971; Waples and Yokota, 2007). With the reproductive weighting scheme, the rate at which the variance of allele frequency change increases per time step will equal the long term rate $1 - \lambda_A$ (Olsson et al., 2013), and in Appendix E, we show that indeed, the correction factor $C = 1$ regardless of the demographic model.

With these weights we only need to specify first order moments of the distributions in the demographic model, whereas the variance of the offspring distribution, as well as the correlation between survival and reproduction, need not be known. However, it is required that individuals are sampled from all age classes with positive reproductive values.

4.6. Increasing time between samples

Another method to reduce the fluctuations in the estimate of the standardized variance of allele frequency change due to the age structure, is to increase the time between the samples (Jorde and Ryman, 1995; Waples and Yokota, 2007; Olsson et al., 2013). In Appendix F we show that, regardless of weighting scheme, when the time τ between the samples increases, the correction factor C approaches 1. However, the time needed until C is sufficiently close to 1, varies a lot depending on the weighting scheme.

In Fig. 1 the correction factor is shown as a function of time and the number of age classes present in the sample. As a model species, we use a population of sparrows with demographic parameters according to Table 2, given by Baker et al. (1981). Age classes are added from the youngest to create different weight vectors. In Fig. 1(a), equal weights are given to the age classes with positive weights, and in Fig. 1(b), the weights are proportional to their reproductive values. We see in Fig. 1 that for the equal weighting scheme, sampling from more age classes does not necessarily result in lower values of C .

4.7. Asymptotic variance and distribution of the estimator

In real populations, it is not sufficient for the estimator to have low bias. Since we usually only have one estimate, the variability of the estimator is of great importance in order to obtain reliable

Table 2

Life table data for sparrows, where b_j is the mean number of progeny for an individual in age class j , $l_j = \prod_{i=0}^{j-1} s_i$ is the probability for an individual to survive to age class j and s_j the probability that an individual in age class j survives to age class $j + 1$. Each age class represents 1 year.

Age class	Sparrow	
	l_j	b_j
0	1	0
1	0.167	3.018
2	0.083	3.202
3	0.048	3.416
4	0.012	3.602
5	0.006	3.842

estimates. In Appendix G, we derive the following approximate formula

$$\frac{\text{AsVar}(\hat{N}_{eV})}{N_{eV}^2} \approx \frac{\text{AsVar}(F_{OH}^{w*})}{(F^w)^2} \approx \frac{2}{L_e} \left(1 + \frac{2N_{eV}T}{C\tau\bar{n}} \right)^2, \quad (17)$$

for the asymptotic coefficient of variance of the estimator of the effective population size, where

$$L_e = \frac{\left[\sum_{l=1}^L \sum_{a=1}^{A_l} p_{tal}^w (1 - p_{tal}^w) \right]^2}{\sum_{l=1}^L \kappa_l \left[\sum_{a=1}^{A_l} p_{tal}^w (1 - p_{tal}^w) \right]^2}, \quad (18)$$

can be interpreted as the effective number of independent alleles and

$$\kappa_l = \frac{\sum_{a=1}^{A_l} (p_{tal}^w)^2 (1 - p_{tal}^w)^2 + \sum_{a \neq b} \sum_{a=1}^{A_l} (p_{tal}^w)^2 (p_{tbl}^w)^2}{\left[\sum_{a=1}^{A_l} p_{tal}^w (1 - p_{tal}^w) \right]^2}, \quad (19)$$

is a number between 0 and 1 that quantifies how correlated genetic drift is among the A_l alleles at locus l , with $\kappa_l = 1$ at biallelic loci. If we have L loci with the same number A of alleles and equal allele frequency $1/A$, the effective number of independent alleles equals $(A - 1)L$, as shown in Appendix H.

It is clear from (17) that precision increases when the product $C\tau\bar{n}$ increases. We also see that when either the time between

samples or the effective sample size tend to infinity, the asymptotic variance of the estimator equals $2/L_e$ and hence only depends on the effective number of independent alleles.

It has previously been shown that estimators of F by e.g. [Krimbas and Tsakas \(1971\)](#) and [Nei and Tajima \(1981\)](#) can be normalized to follow approximately a χ^2 -distribution ([Lewontin and Krakauer, 1973](#); [Waples, 1989a](#)). [Jorde and Ryman \(1995\)](#) argued that this approximation also holds for their estimator. For our estimator (10), it follows from the derivations in [Appendix G](#) that

$$\frac{L(F_{OH}^{w*} + \frac{1}{\bar{n}})}{F + \frac{1}{\bar{n}}} \approx \chi^2(L) \tag{20}$$

if all loci are biallelic with the same allele frequencies. In this case, a confidence interval with confidence level $1 - \alpha$ is given by

$$I_{F^w, \alpha} = \left(\frac{L}{\chi^2_{(\alpha/2)}(L)} \left[F_{OH}^{w*} + \frac{1}{\bar{n}} \right] - \frac{1}{\bar{n}}, \frac{L}{\chi^2_{(1-\alpha/2)}(L)} \left[F_{OH}^{w*} + \frac{1}{\bar{n}} \right] - \frac{1}{\bar{n}} \right), \tag{21}$$

where $\chi^2_{(\alpha/2)}(L)$ is the $(1-\alpha/2)$ -quantile of a $\chi^2(L)$ distribution. For the general case with multiple alleles or unequal allele frequencies there is no simple explicit distribution. However, if multiple loci are used, we could construct an approximate confidence interval

$$I_{F^w, \alpha} = \left(\frac{F_{OH}^{w*} + \frac{1}{\bar{n}}}{1 + \lambda_{\alpha/2} \sqrt{\frac{2}{L_e}}} - \frac{1}{\bar{n}}, \frac{F_{OH}^{w*} + \frac{1}{\bar{n}}}{1 - \lambda_{\alpha/2} \sqrt{\frac{2}{L_e}}} - \frac{1}{\bar{n}} \right) \tag{22}$$

using normal approximation where $\lambda_{\alpha/2}$ is the $(1 - \alpha/2)$ -quantile of a standard normal distribution. By transforming the limits in (21) and (22) using (16) we obtain an approximate confidence interval for N_{eV} where the negative lower bound in (22) can be interpreted as an infinite upper bound in the corresponding confidence interval for N_{eV} .

4.8. Simulations

In order to verify the performance of the method and to justify the approximations that have been made, we have performed a number of simulations. In all simulations we use the demographic parameters of [Table 2](#) and simulate a population of $N = 10\,000$ individuals with L unlinked biallelic loci using initial allele frequency 0.5 for all loci. We sample $n = \sum_{j=0}^{J-1} n_{1j} = \sum_{j=0}^{J-1} n_{2j}$ individuals with replacement at two consecutive time steps and estimate N_{eV} using (16). The simulations have been made according to methods defined in [Olsson and Hössjer \(2014\)](#), which allow us to simulate multiple alleles for a given population and to eliminate the need for a burn in period.

The values of L and n as well as the sample and weighting scheme have been varied. For each combination we repeated the simulation $M = 10\,000$ times and let $\hat{N}_{eV, i}$ and $F_{OH, i}^{w*}$ be the estimates of $N_{eV, i}$ and F_i^w in simulation i , with $\hat{N}_{eV, (1)} \leq \dots \leq \hat{N}_{eV, (M)}$ and $F_{OH, (1)}^{w*} \leq \dots \leq F_{OH, (M)}^{w*}$ the corresponding ordered values. We let

$$\widehat{\text{RelMedBias}}(\hat{N}_{eV}) = \frac{\hat{N}_{eV, (0.5M)} - N_{eV}}{N_{eV}} \tag{23}$$

be an estimate of the relative median bias for \hat{N}_{eV} and

$$\widehat{\text{StDev}}(\hat{N}_{eV}) = \frac{\hat{N}_{eV, (0.75M)} - \hat{N}_{eV, (0.25M)}}{N_{eV} [\Phi(0.75) - \Phi(0.25)]} \tag{24}$$

be a robust estimate of the coefficient of standard deviation, where Φ is the standard normal cumulative distribution function. The corresponding functions for F_{OH}^{w*} are defined analogously.

In the first simulation we let $L = 1000$, the age class sizes N_{tj} equal $N_t u_j$ rounded to the nearest integers and sample $n = 500$ individuals from one age class at each time point. We see in [Fig. 2\(a\)](#) that the estimated relative median bias (23) is small compared to the estimated coefficient of standard deviation (24). We also see that the precision increases with the age of the sampled individuals. This is expected since older age classes consist of fewer individuals and the standardized variance of allele frequency change is expected to be larger and hence, easier to estimate.

In [Fig. 2\(b\)](#) we let the initial age class size distribution vary according to its equilibrium distribution. We see that the relative bias still is small, however for the later age classes, the coefficient of standard deviation deviates from the expected value calculated as the square root of (17). By introducing demographic variability, the coefficient of standard deviation will increase since the actual value of (8) will vary between the simulations. This is not accounted for when deriving the asymptotic variance (17) of F_{OH}^{w*} in (10). However, repeating the simulation for one choice of initial age class composition, we see in [Fig. 2\(c\)](#) that the estimated coefficient of standard deviation is close to the theoretical value. On the other hand, if individuals are sampled from older age classes, where the demographic variance is larger, we introduce a bias.

In [Fig. 2\(d\)](#), we fix $L = 1000$ and $n = 500$ but vary the sampling scheme and the weight vector. In the left part, we use equal weights as in [Fig. 1\(a\)](#), but let the number of sampled individuals be proportional to the stable age distribution. In the right part, we use the reproductive weighting schemes of [Fig. 1\(b\)](#), and the number of sampled individuals are proportional to the reproductive weights for the various age classes.

In [Fig. 3\(a\)](#), we let $L = 1000$ and sample individuals from age class 2 with n varied from 20 to 750 and in [Fig. 3\(b\)](#) we also sample individuals from age class 2 but fix $n = 500$ and let L range between 100 and 10 000. Since the demographic variance is small for age class 2, it is only for large values of L that the estimated coefficient of standard deviation differs noticeably from the square root of (17). We also see that, if we consider the average bias from multiple simulated populations, the standard deviation of \hat{N}_{eV} is much larger than the bias, so the standard deviation can be used as performance measure.

5. Simultaneous estimation of population census size and effective population size

Usually, not only the effective population size but also the population census size is of interest, but both are in general difficult to estimate. However, using knowledge of the ratio N_e/N , one could save both time and money by inferring one parameter from the other, if the ratio is relatively stable ([Luikart et al., 2010](#)). In our model, the assumptions needed to calculate the correction factor (14) can be used to calculate the ratio of the effective population size and census population size, N_e/N ([Felsenstein, 1971](#); [Hill, 1972, 1979](#); [Engen et al., 2005](#); [Waples et al., 2011](#); [Olsson et al., 2013](#)) which is constant over time. By combining these assumptions with genetic data from the temporal method we propose a method for simultaneous estimation of the effective population size and the census population size.

The estimator (16) depends on the census population size through the correction factor (14). If we sample without replacement, the census population size also affects F_{OH}^{w*} through the effective sample size (12), and in any case the estimation procedure turns into an iterative process. In [Appendix I](#) we provide an algorithm for computing estimates (\hat{N}_{eV}, \hat{N}) of the variance effective and census sizes jointly.

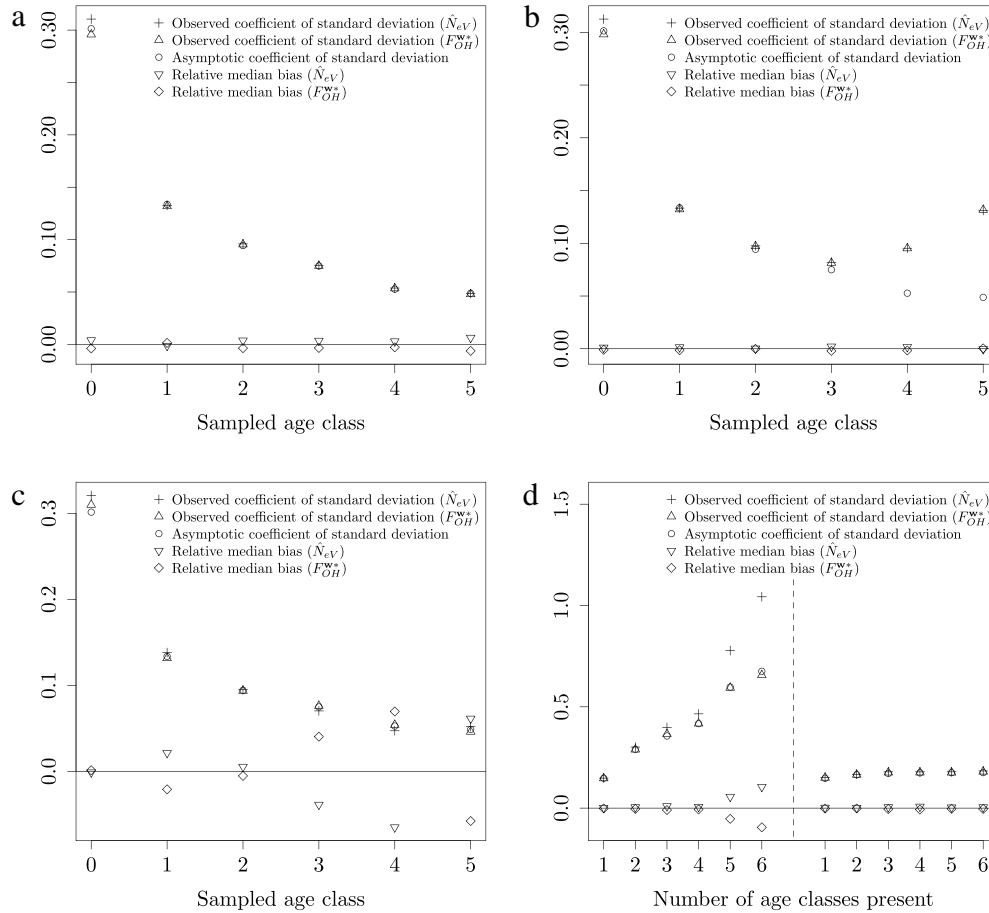


Fig. 2. Observed and asymptotic coefficients of standard deviation of \hat{N}_{eV} and F_{OH}^{w*} according to (24) and (17) and the observed relative median bias (23) as a function of which age class that is sampled (a)–(c) and of the sampling and weighting scheme (d). Age class sizes are constant in (a) but random in the other subplots. In (c), the age class sizes are the same in all $M = 10000$ repetitions. In (d), age classes are sampled according to a truncated stable age (left) or reproductive (right) distribution, using equal (left) or reproductive (right) weights that are truncated to 0 above the highest sampled age class. Demographic parameters are according to Table 2, $L = 1000$ loci are generated and $n = 500$ individuals are sampled for each of the two consecutive time points.

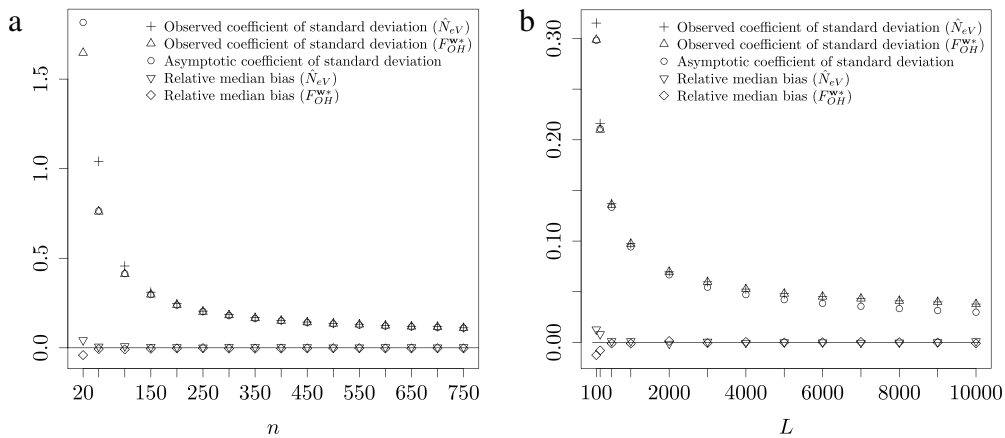


Fig. 3. Observed and asymptotic coefficients of standard deviation of \hat{N}_{eV} and F_{OH}^{w*} according to (24) and (17) and the observed relative median bias (23) as a function of sample size (a) and number of loci (b). In (a) the number of loci $L = 1000$ is fixed and the sample size $n = 500$ at each time point is fixed in (b). All individuals are sampled from age class 2 at two consecutive time points, and demographic parameters are according to Table 2.

To illustrate the performance of the method we simulated 10000 data sets by methods in Olsson and Hössjer (2014) and demographic parameters according to Table 2. We sampled 500 individuals both with and without replacement from age class 0 at two consecutive time points. In Fig. 4(a), estimates of the census population size are shown for both sampling schemes. We see that, regardless of sampling scheme, both box plots are centered around

the true value. However, the variance of the estimates is larger when individuals are sampled without replacement. Estimates of the variance effective population size are shown in Fig. 4(b). When estimating N_{eV} , we assume that the census size in the two boxes to the left is unknown, whereas we assume it is known for the estimates in the two boxes to the right. We see that, when sampling with replacement, we obtain a similar distribution of

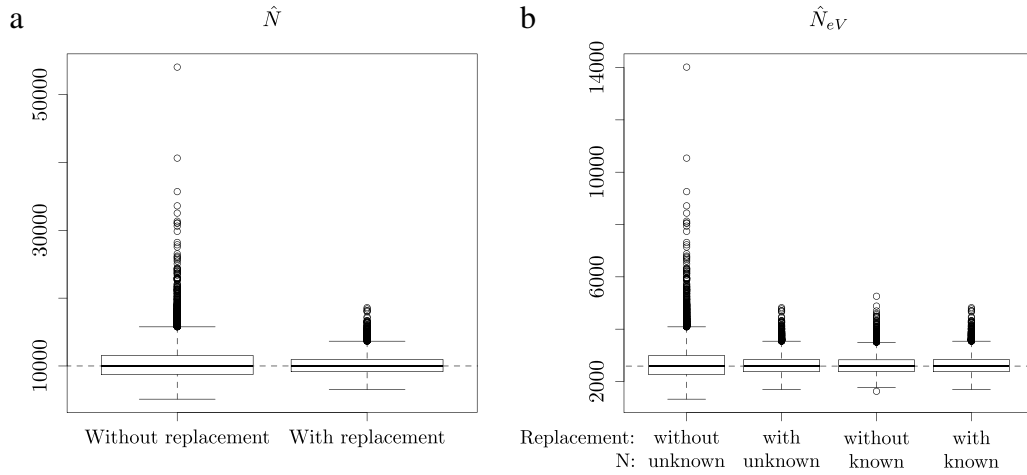


Fig. 4. Box plots of 10 000 repeated simulations based on the demographic parameters of Table 2. The census population size and the variance effective population size are estimated simultaneously according to (1.1)–(1.2). Estimates of the census population size N for sampling without and with replacement are shown in (a). The two boxes in (b) to the left show the corresponding estimates of the variance effective population size N_{eV} , whereas the two boxes to the right illustrate the distribution of the estimates of N_{eV} if we assume that N is known. The dashed lines correspond to the true values $N = 10\,000$ and $N_{eV} = 2589$. Estimated coverage is 95.1% for calculated 95% confidence intervals according to (21) and (22) when sampling is conducted with replacement, regardless if N is known or not. For sampling conducted without replacement the estimated coverage is 94.9% if N is known and 78% if N is unknown. Demographic parameters are according to Table 2, $L = 1000$ loci are generated and $n = 500$ individuals are sampled from age class 0 at two consecutive time points.

the estimates of N_{eV} regardless of whether we assume that N is known or not, and the corresponding 95% confidence intervals calculated by (21) and (22) all have an estimated coverage of 95.1%. On the other hand, for sampling without replacement, the lack of knowledge of N increases the variance of the estimates of N_{eV} and the corresponding 95% confidence intervals estimated coverage is only 78%. The confidence intervals become too narrow since the uncertainty of the estimates of N is not accounted for, and refined methods for more accurate confidence intervals is an interesting problem for future research. However, if N is known, the confidence intervals for sampling without replacement perform well and we obtain an estimated 94.9% coverage of the corresponding 95% confidence intervals.

6. Optimal weighting of loci

When using the estimator (9)–(10), alleles with frequency close to 0.5 will be given the highest weight as in Jorde and Ryman (2007). This way of weighting alleles might not be optimal. For example, if we first consider L biallelic loci with minor allele frequency 0.01 the effective number of independent alleles equals L . By adding a single biallelic locus with allele frequency 0.5 the effective number of independent alleles is reduced since this locus will have a much higher weight than the other loci. This serves as a motivation to extend (9) to

$$\hat{F}_{OH}^w = \frac{\sum_{l=1}^L \omega_l \sum_{a=1}^{A_l} (\hat{p}_{t+\tau,al}^w - \hat{p}_{tal}^w)^2}{\sum_{l=1}^L \omega_l \sum_{a=1}^{A_l} \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2} \left(1 - \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2}\right)} \quad (25)$$

for an appropriate sequence of locus specific weights ω_l . It can be seen, by similar calculations as in Appendix G, that (17) still holds for the generalized estimator of N_{eV} based on (25), with

$$L_e = \frac{\left[\sum_{l=1}^L \omega_l \sum_{a=1}^{A_l} p_{tal}^w (1 - p_{tal}^w) \right]^2}{\sum_{l=1}^L \kappa_l \left[\omega_l \sum_{a=1}^{A_l} p_{tal}^w (1 - p_{tal}^w) \right]^2} \leq \sum_{l=1}^L \frac{1}{\kappa_l}, \quad (26)$$

with the maximum attained when $\omega_l \propto \left[\kappa_l \sum_{a=1}^{A_l} p_{tal}^w (1 - p_{tal}^w) \right]^{-1}$, by the Cauchy–Schwarz inequality. This optimal weighting scheme leads to $L_e = L$ when all loci are biallelic. On the other hand, these weights have to be estimated from data, and thus introduces additional variance to the estimator of N_{eV} . We also see from (26) that $1/\kappa_l$ quantifies how much information each locus l provides in terms of estimating N_{eV} .

7. Discussion

Jorde and Ryman (1995) presented an unbiased estimator of N_{eV} for populations with overlapping generations. Their main idea was to sample individuals from one of the age classes and correct the estimate for overlapping generations using a correction factor. This method has since then been modified (Jorde, 2012) and used to estimate N_{eV} from real data (Jorde and Ryman, 1996; Charlier et al., 2012).

In this paper we have generalized the estimator presented in Jorde and Ryman (1995) in several ways. In the original version, the correction factor was calculated using an iterative process. Here, we derive a direct way to calculate it through formulas (14) and (16). We include the actual age class sizes when deriving the correction factor which gives a more accurate constant, especially for populations with few individuals in the sampled age classes. An optimal way of weighting multiple loci is given and we also derive a general formula for the coefficient of asymptotic standard deviation, with accompanying approximate confidence intervals. This approximate and very explicit formula implies that in order to reduce the standard deviation for a specified population one can do one or more of the following:

- Increase the effective number of independent alleles L_e ,
- increase the effective sample size \tilde{n} ,
- increase the time τ between the samples,
- sample from age classes that maximize C .

Even though the standard deviation of the estimator decreases as the time τ between samples increases, one has to be careful while increasing τ by one or a few time steps, since the correction factor also depends on τ . This means that the product $C\tau$ can initially get smaller when τ increases, which will result in a higher standard deviation of \hat{N}_{eV} .

Usually, in real populations, neither the effective population size, nor the census population size is known. However, both the effective population size and the census population size are generally of interest. Here, we present a method for simultaneous estimation of both sizes. This method relies on the ratio N_{ev}/N of the effective and census sizes being known when the life table is, so that an estimator of N_{ev} naturally gives rise to an estimator of N . A number of other one- and two-sample demographic and genetic estimators of N exist, as reviewed by Luikart et al. (2010). If such data is available it should of course be used for estimating N . Since most of these estimators of N do not require a life table, they could be compared with ours in order to test whether the life table is correctly specified or not.

The temporal method has two main drawbacks. First, in order to calculate the correction factor C , a number of assumptions have to be made. In this paper we have generalized the original assumption of Poisson distributed number of offspring to allow for overdispersion. We also include the possibility of correlation between survival and reproductive success. Such detailed information about a studied population is often not known. But if individuals are sampled from all age classes with positive reproductive numbers, we show that, when using reproductive weights, the correction factor equals one. Since assumptions about variance of reproductive success and its correlation with survival are only needed in order to calculate the correction factor, in this case, such assumptions can be relaxed. However, to determine the reproductive weights, we still need to know the expected number of births per age class and the age specific survival probabilities.

Even if accurate life-table information can be difficult to assess for many species, we still believe it is important to use it for species where such data can be estimated, see for instance Jorde and Ryman (1996), Turner et al. (1999) and Serbezov et al. (2012). But it is also possible to extend our method to include uncertainty of C . Write $C = C(\theta)$ as a function of the set θ of all relevant demographic parameters. Then, if $\hat{\theta}$ is the estimate of θ , we can estimate C as $\hat{C} = C(\hat{\theta})$. Given that the demographic uncertainty is specified in terms of a distribution of $\hat{\theta}$, we may use parametric bootstrap to estimate $\text{Var}(\hat{C})$. It then follows from (16) that a second term $\text{AsVar}(\hat{C})/C^2$ should be added in (17) to the asymptotic coefficient of variance of \hat{N}_{ev} .

Second, the temporal method requires genetic data from at least two time points. In order to avoid this, Wang (2009) and Wang et al. (2010) use information from one sample of multilocus genotypes. The idea is to estimate identity-by-descent-sharing between sampled individuals first, to assign either their sibship or parental relations, and then to use this for estimating N_e . One may also utilize that genetic drift causes linkage disequilibrium between pairs of loci in order to estimate N_e of an age structured populations from one single sample, see Hill (1981), Waples et al. (2014) and references therein. We believe these methods are of great value, and a good complement to the temporal method even when data from several time points are available.

The framework in this paper is for monocious populations. Since animal species are diecious, our method should be extended to account for this. Effective size formulas for age-structured two sex models have been derived by Johnson (1977), Emigh and Pollak (1979), Engen et al. (2005) and Pollak (2011). Engen et al. (2007) showed how effective size estimates for males and females are combined into one when birth and survival rates of each sex are age independent. This could be used as a first approximation to combine two separate estimators (16) for males and females into one. A more elaborate approach is to extend the estimator of this paper to diecious models. The theory for calculating C in Appendix B relies on haploid results of Hössjer et al. (2014). We have recently developed a corresponding theory for diploid and diecious species in Hössjer et al. (2015), and it would be of great interest to apply it in order to derive a diecious extension of the Jorde–Ryman estimator of N_{ev} .

Acknowledgments

We are grateful for the comments of two anonymous reviewers that considerably improved the paper. Financial support from the Swedish Research Council, contract No. 621-2008-4946 and 621-2013-4633, and the Gustafsson Foundation for Research in Natural Sciences and Medicine to Ola Hössjer is acknowledged.

Appendix A. Derivation of (10)

Let E denote expectation with respect to $t = 0$. We will motivate that F_{OH}^{w*} in (10) is an approximately unbiased estimator of F^w in (8), which is assumed to be independent of a and l . Starting with the unadjusted estimator \hat{F}_{OH}^w it follows from (9) that

$$E(\hat{F}_{OH}^w) \approx \frac{\sum_{l=1}^L \sum_{a=1}^{A_l} E([\hat{p}_{t+\tau,al}^w - \hat{p}_{tal}^w]^2)}{\sum_{l=1}^L \sum_{a=1}^{A_l} E\left(\frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2} \left[1 - \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2}\right]\right)} \tag{A.1}$$

when L is large. By conditioning on all p_{tal}^w and $p_{t+\tau,al}^w$, and since $E(\hat{p}_{tal}^w) = p_{tal}^w$, $E(\hat{p}_{t+\tau,al}^w) = p_{t+\tau,al}^w$, the nominator in (A.1) can be written as

$$\begin{aligned} & \sum_{l=1}^L \sum_{a=1}^{A_l} E([\hat{p}_{t+\tau,al}^w - \hat{p}_{tal}^w]^2) \\ &= \sum_{l=1}^L \sum_{a=1}^{A_l} E[E([\hat{p}_{t+\tau,al}^w - \hat{p}_{tal}^w]^2 | p_{tal}^w, p_{t+\tau,al}^w)] \\ &= \sum_{l=1}^L \sum_{a=1}^{A_l} \{E[(p_{t+\tau,al}^w - p_{tal}^w)^2] \\ & \quad + E[\text{Var}(\hat{p}_{t+\tau,al}^w | p_{t+\tau,al}^w)] + E[\text{Var}(\hat{p}_{tal}^w | p_{tal}^w)]\}, \end{aligned} \tag{A.2}$$

and the denominator, by the same type of argument, as

$$\begin{aligned} & \sum_{l=1}^L \sum_{a=1}^{A_l} E\left(\frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2} \left[1 - \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2}\right]\right) \\ &= \sum_{l=1}^L \sum_{a=1}^{A_l} E\left[E\left(\frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2} \right. \right. \\ & \quad \times \left. \left. \left[1 - \frac{\hat{p}_{tal}^w + \hat{p}_{t+\tau,al}^w}{2}\right] \middle| p_{tal}^w, p_{t+\tau,al}^w \right)\right] \\ &= \sum_{l=1}^L \sum_{a=1}^{A_l} \left\{ E\left(\frac{p_{tal}^w + p_{t+\tau,al}^w}{2} \left[1 - \frac{p_{tal}^w + p_{t+\tau,al}^w}{2}\right]\right) \right. \\ & \quad \left. - (E[\text{Var}(\hat{p}_{t+\tau,al}^w | p_{t+\tau,al}^w)] + E[\text{Var}(\hat{p}_{tal}^w | p_{tal}^w)])/4 \right\}, \end{aligned} \tag{A.3}$$

where $\text{Var}(\hat{p}_{tal}^w | p_{tal}^w)$ is the sampling variance which approximates (7) and depends on the sampling scheme. For binomial sampling, i.e. sampling with replacement

$$\begin{aligned} \text{Var}(\hat{p}_{tal}^w | \{p_{tj,al}^w\}_{j=0}^{J-1}) &= \sum_{j=0}^{J-1} w_{tj}^2 \frac{p_{tj,al}^w (1 - p_{tj,al}^w)}{2n_{1j}} \\ &\approx \frac{p_{tal}^w (1 - p_{tal}^w)}{2n_1} \\ &\approx \text{Var}(\hat{p}_{tal}^w | p_{tal}^w), \end{aligned} \tag{A.4}$$

because of (11), and analogously

$$\begin{aligned} \text{Var}(\hat{p}_{t+\tau,al}^{\mathbf{w}} | \{p_{t+\tau,al}^{\mathbf{w}}\}_{j=0}^{J-1}) \\ &= \sum_{j=0}^{J-1} w_{t+\tau,aj}^2 \frac{p_{t+\tau,aj}(1-p_{t+\tau,aj})}{2n_2} \\ &\approx \frac{p_{tal}^{\mathbf{w}}(1-p_{tal}^{\mathbf{w}})}{2n_2} - \frac{p_{tal}^{\mathbf{w}}(1-p_{tal}^{\mathbf{w}}) - p_{t+\tau,al}^{\mathbf{w}}(1-p_{t+\tau,al}^{\mathbf{w}})}{2n_2} \\ &\approx \text{Var}(\hat{p}_{t+\tau,al}^{\mathbf{w}} | p_{t+\tau,al}^{\mathbf{w}}). \end{aligned} \quad (\text{A.5})$$

The right hand sides of (A.4) and (A.5) are still valid for sampling without replacement, if we redefine n_1 and n_2 as in (12).

Putting (A.2), (A.4) and (A.5) together, we find that $E(p_{tal}^{\mathbf{w}}[1 - p_{tal}^{\mathbf{w}}] - p_{t+\tau,al}^{\mathbf{w}}[1 - p_{t+\tau,al}^{\mathbf{w}}]) = E(p_{tal}^{\mathbf{w}}[1 - p_{tal}^{\mathbf{w}}])F^{\mathbf{w}}$, where $E(p_{t+\tau,al}^{\mathbf{w}} | p_{tal}^{\mathbf{w}}) \approx p_{tal}^{\mathbf{w}}$ is used, $F^{\mathbf{w}}$ is defined in (8) and assumed to be independent of a, l . Then we have that

$$\begin{aligned} \sum_{l=1}^L \sum_{a=1}^{A_l} E([\hat{p}_{t+\tau,al}^{\mathbf{w}} - \hat{p}_{tal}^{\mathbf{w}}]^2) \\ &= [F^{\mathbf{w}} + 1/(2n_1) + 1/(2n_2) - F^{\mathbf{w}}/(2n_2)] \\ &\quad \times \sum_{l=1}^L \sum_{a=1}^{A_l} E[p_{tal}^{\mathbf{w}}(1 - p_{tal}^{\mathbf{w}})]. \end{aligned} \quad (\text{A.6})$$

In the same way, (A.3) can be written as

$$\begin{aligned} \sum_{l=1}^L \sum_{a=1}^{A_l} E\left(\frac{\hat{p}_{tal}^{\mathbf{w}} + \hat{p}_{t+\tau,al}^{\mathbf{w}}}{2} \left[1 - \frac{\hat{p}_{tal}^{\mathbf{w}} + \hat{p}_{t+\tau,al}^{\mathbf{w}}}{2}\right]\right) \\ &= [1 - F^{\mathbf{w}}/4 - 1/(8n_1) - 1/(8n_2)(1 - F^{\mathbf{w}})] \\ &\quad \times \sum_{l=1}^L \sum_{a=1}^{A_l} E[p_{tal}^{\mathbf{w}}(1 - p_{tal}^{\mathbf{w}})], \end{aligned} \quad (\text{A.7})$$

so that the double sums in (A.6) and (A.7) cancel when taking the ratio of these two expressions, and

$$E(\hat{F}_{OH}^{\mathbf{w}}) \approx \frac{F^{\mathbf{w}} + 1/(2n_1) + 1/(2n_2) - F^{\mathbf{w}}/(2n_2)}{1 - F^{\mathbf{w}}/4 - 1/(8n_1) - 1/(8n_2)(1 - F^{\mathbf{w}})}. \quad (\text{A.8})$$

Solving for $F^{\mathbf{w}}$ in (A.8) and using the definition of \tilde{n} , we find that (10) is an approximately unbiased estimator of the standardized variance of allele frequency change between time t and $t + \tau$ for weighting scheme \mathbf{w} . \square

Appendix B. How to calculate (14)

Let $p_{ij} = p_{ijal}$ and $p_t^{\mathbf{w}} = p_{tal}^{\mathbf{w}}$ be age-specific and age-averaged frequencies of a particular allele a at some locus l . It is shown in the previous subsection, that the expected value of $F_{OH}^{\mathbf{w}}$ is well approximated by $F^{\mathbf{w}}$ in (8). By using the definition of $p_t^{\mathbf{w}}$ in (6) and expanding the expected values of the numerator and denominator of (8), we can rewrite $F^{\mathbf{w}}$ as

$$F^{\mathbf{w}} = \frac{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j (f_{t+\tau,ij} + f_{t,ij} - \frac{2\text{Cov}(p_{t+\tau,i}, p_{t,j})}{p(1-p)})}{1 - \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j f_{t,ij}}, \quad (\text{B.1})$$

where E and Cov denote expectation and covariance with respect to $t = 0$, and

$$f_{ij} = \text{Cov}(p_{ti}, p_{tj}) / (p(1-p)), \quad (\text{B.2})$$

is a standardized covariance at time t between age classes i and j , assuming that all age classes j have the same allele frequency, $p = p_{0j}$, at time 0.

Let $\mathbf{p}_t = (p_{t0}, \dots, p_{t,J-1})'$ be a column vector of the generic, conditional on demographics, allele frequencies in all age classes

at time t . It follows from Hössjer et al. (2014) that the expected time dynamics of the allele frequencies is given by

$$E[\mathbf{p}_{t+\tau} | \mathbf{p}_t] = \mathbf{B}_{t+\tau-1} \dots \mathbf{B}_{t+1} \mathbf{p}_t \quad (\text{B.3})$$

where

$$\mathbf{B}_t = (B_{tik})_{i,k=0}^{J-1} = \begin{pmatrix} Q_{t0} & Q_{t1} & \dots & Q_{t,J-2} & Q_{t,J-1} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \quad (\text{B.4})$$

is a backward migration matrix between time points $t + 1$ and t , and

$$Q_{tj} = N_{tj} b_j / N_{t+1,0}$$

is the probability that the parent of an offspring belongs to age class j . If population size changes slowly over the time interval $[t, t + \tau]$ we have approximately $\mathbf{B}_t = \dots = \mathbf{B}_{t+\tau} = \mathbf{B}$, and

$$\begin{aligned} \text{Cov}(p_{t+\tau,i}, p_{tj}) &= \text{Cov}(E[p_{t+\tau,i} | p_{tj}], p_{tj}) \\ &= \text{Cov}([\mathbf{B}^{\tau} \mathbf{p}]_i, p_{tj}) = p(1-p) \sum_{k=0}^{J-1} (\mathbf{B}^{\tau})_{ik} f_{tkj}, \end{aligned}$$

and combining this with (B.1), we have

$$E(F_{OH}^{\mathbf{w}*}) \approx F^{\mathbf{w}} = \frac{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j (f_{t+\tau,ij} + f_{t,ij} - 2 \sum_k (\mathbf{B}^{\tau})_{ik} f_{tkj})}{1 - \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j f_{t,ij}}. \quad (\text{B.5})$$

Let

$$\mathbf{f}_t = \text{vec} \left((f_{tj})_{i,j=0}^{J-1} \right)$$

be a column vector of length J^2 that stacks all standardized covariances. From theory developed in Hössjer et al. (2014) and Hössjer (2014), it follows that

$$\mathbf{f}_{t+1} = \mathbf{A}_t \mathbf{f}_t + (\mathbf{I} - \mathbf{A}_t) \mathbf{1}, \quad (\text{B.6})$$

where $\mathbf{1}$ is a column vector of J^2 ones, \mathbf{I} an identity matrix of order J^2 ,

$$\mathbf{A}_t = (A_{t,ij,kl})_{0 \leq i,j \leq J-1, 0 \leq k,l \leq J-1}$$

a square matrix of order J^2 with elements

$$A_{t,ij,kl} = \left(1 - \frac{1}{2N_{t+1,i}}\right)^{\{i=j\}} \left(\frac{1 - \pi_{tj,k}}{1 - \frac{1}{2N_{ik}}}\right)^{\{k=l\}} B_{tik} B_{tjl}, \quad (\text{B.7})$$

and $\pi_{tj,k}$ is the coalescence probability that two individuals in age classes i and j at time $t + 1$ that have both of their parents at time t in age class k , have the same parent.

Assuming that demographics changes slowly, so that $\mathbf{A}_t \approx \mathbf{A}$, we have from (B.6) for large t that

$$\mathbf{f}_t = \mathbf{1} - \mathbf{A}^t \mathbf{1} \approx \mathbf{1} - \lambda_A^t \mathbf{r} \mathbf{1} = \mathbf{1} - \lambda_A^t \mathbf{r}, \quad (\text{B.8})$$

where λ_A is the largest eigenvalue of \mathbf{A} , $\mathbf{1}$ and \mathbf{r} are the left and right eigenvectors, corresponding to λ_A , normalized such that $\mathbf{1} \mathbf{1} = \mathbf{r} \mathbf{1} = 1$. By inserting (B.8) into (B.5) and dividing by λ_A^t in the numerator and denominator, we have that

$$E(F_{OH}^{\mathbf{w}*}) \approx F^{\mathbf{w}} \approx \frac{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j \left[2 \sum_k (\mathbf{B}^{\tau})_{ik} r_{kj} - (1 + \lambda_A^{\tau}) r_{ij} \right]}{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j r_{ij}}. \quad (\text{B.9})$$

Finally, we compute C by inserting the right hand side of (B.9) into (14).

For a demography that changes cyclically with period c time steps so that $\mathbf{A}_{t+c} = \mathbf{A}_t$ for all t , we let $\mathbf{A} = (\mathbf{A}_1 \cdot \dots \cdot \mathbf{A}_c)^{1/c}$ so that λ_A and \mathbf{r} in (B.9) is the largest eigenvalue and corresponding right eigenvector of this matrix. In order to take fluctuating demographics into account, we have to replace \mathbf{A} by its expected value with respect to such fluctuations. \square

Appendix C. Constant age class sizes and Poisson reproduction

This is the scenario treated by Jorde and Ryman (1995), with $N_{ij} = N_j$, $\pi_{tijk} = \pi_{ijk}$ and $\mathbf{A}_t = \mathbf{A}$. Since the subpopulation sizes are constant, we drop index t for all quantities.

We only need to specify the coalescence probabilities for age class triples i, j, k such that $B_{ik}B_{jk} \neq 0$. This includes three possibilities; a common parent in age class k of two newborns, with

$$\pi_{00k} = \frac{1}{2N_k}, \tag{C.1}$$

that two adults in age class $k \in \{1, \dots, J - 1\}$ originate from the same individual (in age class $k - 1$) at the previous time point, which is not possible, so that

$$\pi_{kk,k-1} = 0, \tag{C.2}$$

and an adult in age class k being the parent of a newborn at the previous time point,

$$\pi_{0,k+1,k} = \pi_{k+1,0,k} = \frac{1}{2N_k}. \tag{C.3}$$

Inserting (C.1)–(C.3) into (B.7), one can show that recursion (B.6) for the vector \mathbf{f}_t of the standardized allele frequency covariances is the same as in Jorde and Ryman (1995). \square

Appendix D. Non-constant age class sizes and overdispersion

We will extend the previous example by dropping the assumptions of constant subpopulation sizes, Poisson distributed reproduction numbers and independence between survival and reproduction.

In order to achieve this, we need to generalize (C.1)–(C.3). Starting with (C.2), this is straightforward, since

$$\pi_{t,kk,k-1} = 0. \tag{D.1}$$

Generalizing (C.1) is a well known problem of coalescence theory, see for instance Durrett (2008) and references therein. It follows from the exchangeability of the terms in the middle equation of (1) that

$$\begin{aligned} \pi_{t,00,k} &= E \left(\frac{\sum_{h=1}^{2N_{tk}} \binom{Y_{tkh}}{2}}{\binom{2N_{t+1,k0}}{2}} \middle| N_{tk}, N_{t+1,k0} \right) \\ &= \frac{N_{tk} E(Y_{tk1}[Y_{tk1} - 1] | N_{tk}, N_{t+1,k0})}{N_{t+1,k0}(2N_{t+1,k0} - 1)} \\ &= \frac{N_{tk}}{N_{t+1,k0}(2N_{t+1,k0} - 1)} \left[(N_{t+1,k0}/N_{tk})^2 - N_{t+1,k0}/N_{tk} \right. \\ &\quad \left. + \text{Var}(Y_{tk1} | N_{tk}, N_{t+1,k0}) \right], \end{aligned} \tag{D.2}$$

since

$$E(Y_{tk1} | N_{tk}, N_{t+1,k0}) = \frac{N_{t+1,k0}}{N_{tk}}. \tag{D.3}$$

When the number of offspring has a Poisson distribution, before conditioning on the total number of offspring in the age class, the distribution after such conditioning is binomial, so that

$$\text{Var}(Y_{tk1} | N_{tk}, N_{t+1,k0}) = N_{t+1,k0} \cdot \frac{1}{N_{tk}} \left(1 - \frac{1}{2N_{tk}} \right), \tag{D.4}$$

which inserted into (D.2) yields

$$\pi_{t,00,k} = \frac{1}{2N_{tk}}.$$

From this we see that (D.2) generalizes (C.1) to varying population sizes and non-Poissonian offspring distributions. For instance, for negative binomial offspring distribution, we use instead (3), and find that

$$\begin{aligned} Y_{tk1} | N_{t+1,k0} &\sim \text{Bin}(2N_{t+1,k0}, P_{tk1}), \\ P_{tk1} | N_{tk} &\sim \text{Beta}(m_k, (2N_{tk} - 1)m_k). \end{aligned}$$

Hence the variance term in (D.2) can be computed as

$$\begin{aligned} \text{Var}(Y_{tk1} | N_{tk}, N_{t+1,k0}) &= E \left[\text{Var}(Y_{tk1} | N_{tk}, N_{t+1,k0}, P_{tk1}) \right] \\ &\quad + \text{Var} \left(E[Y_{tk1} | N_{tk}, N_{t+1,k0}, P_{tk1}] \right) \\ &= E \left(2N_{t+1,k0} P_{tk1} [1 - P_{tk1}] | N_{tk} \right) + \text{Var}(2N_{t+1,k0} P_{tk1}) \\ &= N_{t+1,k0} \cdot \frac{1}{N_{tk}} \left(1 - \frac{1}{2N_{tk}} \right) + 2N_{t+1,k0} (2N_{t+1,k0} - 1) \text{Var}(P_{tk1}) \\ &= N_{t+1,k0} \cdot \frac{1}{N_{tk}} \left(1 - \frac{1}{2N_{tk}} \right) \cdot \left(1 + \frac{2N_{t+1,k0} - 1}{2N_{tk}m_k + 1} \right). \end{aligned} \tag{D.5}$$

We notice that (D.4) is a special case of (D.5) that corresponding to $m_k = \infty$. Inserting (D.5) into (D.2), we get a general expression for $\pi_{t,00,k}$ when population sizes vary and the number of offspring has a negative binomial distribution.

Finally, we need to generalize (C.3). By a similar argument as in (D.2), one can show that

$$\begin{aligned} \pi_{t,0,k+1,k} &= \pi_{t,k+1,0,k} \\ &= E \left(\frac{\sum_{h=1}^{2N_{tk}} Y_{tkh} I_{tkh}}{4N_{t+1,k0} N_{t+1,k+1}} \middle| N_{tk}, N_{t+1,k0}, N_{t+1,k+1} \right) \\ &= \frac{N_{tk} E(Y_{tk1} I_{tk1} | N_{tk}, N_{t+1,k0}, N_{t+1,k+1})}{2N_{t+1,k0} N_{t+1,k+1}}. \end{aligned} \tag{D.6}$$

In particular, if survival and reproduction are independent ($\rho_k = 0$), it follows from (D.3) and

$$E(I_{tk1} | N_{tk}, N_{t+1,k+1}) = \frac{N_{t+1,k+1}}{N_{tk}} \tag{D.7}$$

that

$$\pi_{t,0,k+1,k} = \pi_{t,0,k+1,k} = \frac{1}{2N_{tk}},$$

in agreement with (C.3). More generally, we compute the expected value in the numerator of (D.6) as

$$\begin{aligned} E(Y_{tk1} I_{tk1} | N_{tk}, N_{t+1,k0}, N_{t+1,k+1}) \\ = \frac{N_{t+1,k+1}}{N_{tk}} \sum_{y=0}^{2N_{t+1,k0}} y P(Y_{tk1} = y | I_{tk1} = 1, N_{t+1,k0}, N_{t+1,k+1}), \end{aligned}$$

making use of (D.7) in the last step. The sum refers to the expected value of Y_{tk1} given $I_{tk1} = 1$, $N_{t+1,k0}$ and $N_{t+1,k+1}$, which could be approximated by a Monte Carlo simulation. \square

Appendix E. Correction factor for reproductive weights

First, let

$$w_j = \gamma_j = \frac{v_j N_{tj}}{\sum_{i=0}^{J-1} v_i N_{ti}} \approx \frac{v_j u_j}{\sum_{i=0}^{J-1} v_i u_i} = v_j u_j,$$

if the population is close to demographic equilibrium, $N_{tj} \approx N_t u_j$, so that an individual in age class j in the sample is weighted according to its age class reproductive value γ_j . With this choice of weights and letting

$$\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_{J-1}) \quad (\text{E.1})$$

be the vector of reproductive values it can be shown generally that $\boldsymbol{\gamma}$ is the equilibrium distribution of the backward migration matrix \mathbf{B} (Nagylaki, 1980; Hössjer and Ryman, 2014). Therefore

$$\boldsymbol{\gamma} \mathbf{B}^\tau = \boldsymbol{\gamma}, \quad (\text{E.2})$$

and it follows from (B.9) that

$$F^\gamma = \frac{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \gamma_i \gamma_j (1 - \lambda_A^\tau) r_{ij}}{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \gamma_i \gamma_j r_{ij}} = 1 - \lambda_A^\tau.$$

Hence, the correction factor

$$C = \frac{1 - (1 - (1 - \lambda_A^\tau))^{1/\tau}}{1 - \lambda_A} = 1$$

regardless of demographic model. \square

Appendix F. Correction factor in the limit of large time intervals

For an arbitrary vector of weights, \mathbf{w} , we insert (B.9) into (14), let $\tau \rightarrow \infty$ and find that

$$\lim_{\tau \rightarrow \infty} C = \lim_{\tau \rightarrow \infty} \frac{1 - \left\{ 1 - \frac{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} [2\gamma_i w_j - w_i w_j (1 + \lambda_A^\tau)] r_{ij}}{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j r_{ij}} \right\}^{1/\tau}}{1 - \lambda_A} \\ = \lim_{\tau \rightarrow \infty} \frac{1 - \left\{ 1 - \frac{2 \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \gamma_i w_j r_{ij}}{\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j r_{ij}} + (1 + \lambda_A^\tau) \right\}^{1/\tau}}{1 - \lambda_A}$$

since $\boldsymbol{\gamma}$ is the equilibrium distribution of \mathbf{B} (cf. (E.2)) and therefore $(\mathbf{B}^\tau)_{ik} \rightarrow \gamma_k$ as $\tau \rightarrow \infty$. Hence, if $\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \gamma_i w_j r_{ij} \approx \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j r_{ij}$, then the correction factor $C \rightarrow 1$ as $\tau \rightarrow \infty$. Since $\sum_{j=0}^{J-1} \gamma_j = \sum_{j=0}^{J-1} w_j = 1$ and $f_{tj} \approx 1 - \lambda_A^\tau r_{tj}$ according to (B.2), we thus need to show that

$$\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \gamma_i w_j f_{t,ij} - \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j f_{t,ij} = o(\lambda_A^t). \quad (\text{F.1})$$

We will motivate (F.1) for strong migration (Nagylaki, 1980), which is reasonable to assume for age-structured models.

By weighting the standardized covariances (B.2) and making use of (6), we have that

$$\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} w_i w_j f_{t,ij} = \frac{\text{Var}(p_t^w)}{p(1-p)},$$

and

$$\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \gamma_i w_j f_{t,ij} = \frac{\text{Cov}(p_t^\gamma, p_t^w)}{p(1-p)},$$

where

$$\text{Cov}(p_t^\gamma, p_t^w) = \text{Cov}(p_t^\gamma, p_t^\gamma + [p_t^w - p_t^\gamma]) \\ = \text{Var}(p_t^\gamma) + \text{Cov}(p_t^\gamma, [p_t^w - p_t^\gamma])$$

and

$$\text{Var}(p_t^w) = \text{Var}(p_t^w + p_t^\gamma - p_t^\gamma) \\ = \text{Var}(p_t^\gamma) + \text{Var}(p_t^w - p_t^\gamma) + 2\text{Cov}(p_t^\gamma, [p_t^w - p_t^\gamma]). \quad (\text{F.2})$$

Hence, in order to establish (F.1), we need to show that

$$\text{Var}(p_t^w - p_t^\gamma) = o(\lambda_A^t) \quad (\text{F.3})$$

and

$$\text{Cov}(p_t^\gamma, [p_t^w - p_t^\gamma]) = o(\lambda_A^t). \quad (\text{F.4})$$

Let $\mathbf{p}_t = (p_{t,0}, \dots, p_{t,J-1})'$ be the vector of allele frequencies in the different age classes at time t for a particular allele a and locus l , as in (B.3). Following the approach in Hössjer and Ryman (2014) we have the recursion

$$\mathbf{p}_t = \mathbf{B} \mathbf{p}_{t-1} + \boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\epsilon}_t = (\epsilon_{t,0}, \dots, \epsilon_{t,J-1})'$ is a random error vector with expected value

$$E(\boldsymbol{\epsilon}_t | \mathbf{p}_{t-1}) = \mathbf{0},$$

and covariance matrix

$$\text{Cov}(\boldsymbol{\epsilon}_t | \mathbf{p}_{t-1}) = \Omega(\mathbf{p}_{t-1}) = \mathcal{O}[p(1-p)(1-\lambda_A)\lambda_A^t].$$

Let

$$\mathbf{p}_t = p_t^\gamma \mathbf{1} + \mathbf{p}_t^0,$$

where \mathbf{p}_t^0 represents the deviation from equal allele frequencies in all age classes, and $\mathbf{1}$ is a column vector of J ones. Suppose that all age classes have the same allele frequency p at time $t = 0$, then

$$p_t^\gamma = p + \sum_{s=1}^t \epsilon_s$$

and

$$\mathbf{p}_t^w = \mathbf{w}(p_t^\gamma \mathbf{1} + \mathbf{p}_t^0) = p_t^\gamma + \mathbf{w} \sum_{s=1}^t (\mathbf{B}^0)^{t-s} \boldsymbol{\epsilon}_s^0,$$

where $\epsilon_s = \boldsymbol{\gamma} \boldsymbol{\epsilon}_s = \sum_{j=0}^{J-1} \gamma_j \epsilon_{s,j}$, $\boldsymbol{\epsilon}_s^0 = \boldsymbol{\epsilon}_s - \epsilon_s \mathbf{1}$ and $\mathbf{B}^0 = \mathbf{B} - \mathbf{1} \boldsymbol{\gamma}$. Let λ_2 denote the largest eigenvalue of \mathbf{B}^0 . It is the same as the second largest eigenvalue of \mathbf{B} , and quantifies how fast migration is. Then,

$$\text{Var}(p_t^w - p_t^\gamma) = \sum_{s=1}^t \text{Var}(\mathbf{w}(\mathbf{B}^0)^{t-s} \boldsymbol{\epsilon}_s^0) \\ = \mathcal{O} \left[\frac{p(1-p)(1-\lambda_A)\lambda_A^t}{1-\lambda_2^2/\lambda_A} \right]$$

since $\text{Var}(\mathbf{w}(\mathbf{B}^0)^{t-s} \boldsymbol{\epsilon}_s^0) \in \mathcal{O}(p(1-p)(1-\lambda_A)\lambda_A^s \lambda_2^{2(t-s)})$ and

$$\text{Cov}(p_t^\gamma, (p_t^w - p_t^\gamma)) = \sum_{s=1}^t \text{Cov}(\epsilon_s, \mathbf{w}(\mathbf{B}^0)^{t-s} \boldsymbol{\epsilon}_s^0) \\ = \mathcal{O} \left[\frac{p(1-p)(1-\lambda_A)\lambda_A^t}{1-\lambda_2/\lambda_A} \right]$$

since $\text{Cov}(\epsilon_s, \mathbf{w}(\mathbf{B}^0)^{t-s} \boldsymbol{\epsilon}_s^0) \in \mathcal{O}(p(1-p)(1-\lambda_A)\lambda_A^s \lambda_2^{t-s})$. Strong migration means that genetic drift is much slower than migration between age classes, i.e. $1 - \lambda_A \ll 1 - \lambda_2$, which implies $1 - \lambda_A \ll 1 - \lambda_2/\lambda_A$. Then (F.3)–(F.4) follow and hence also (F.1). \square

Appendix G. Asymptotic variance of \hat{N}_{eV}

We derive the asymptotic variance in two steps. First, consider a biallelic locus with allele frequency $p_t = p_t^w$ at time t . We assume that $F = F^w$ is the genetic drift in (8),

$$p_{t+\tau} \approx p_t + X_1 \sqrt{F p_t (1 - p_t)}, \tag{G.1}$$

is the allele frequency at time $t + \tau$,

$$\hat{p}_t \approx p_t + X_2 \sqrt{\frac{p_t (1 - p_t)}{2n_1}}, \tag{G.2}$$

is the estimated allele frequency at time t and

$$\hat{p}_{t+\tau} \approx p_{t+\tau} + X_3 \sqrt{\frac{p_t (1 - p_t)}{2n_2}}, \tag{G.3}$$

is the estimated allele frequency at time $t + \tau$, where X_1, X_2 and X_3 are independent standard normal distributed random variables, applying a normal approximation to the genetic drift from time t to $t + \tau$ and the sample variation of \hat{p}_t and $\hat{p}_{t+\tau}$. This approximation is accurate when F is small and n_1, n_2 are large.

In order to find an approximation of the sample variance of F_{OH}^{w*} , we can assume

$$\begin{aligned} F_{OH}^{w*} &\approx \frac{(\hat{p}_{t+\tau} - \hat{p}_t)^2}{\hat{p}_t (1 - \hat{p}_t)} - \frac{1}{\tilde{n}} \\ &\approx \frac{\left[X_1 \sqrt{F p_t (1 - p_t)} - X_2 \sqrt{\frac{p_t (1 - p_t)}{2n_1}} + X_3 \sqrt{\frac{p_t (1 - p_t)}{2n_2}} \right]^2}{p_t (1 - p_t) + (1 - 2p_t) X_2 \sqrt{\frac{p_t (1 - p_t)}{2n_1}}} - \frac{1}{\tilde{n}} \\ &= \frac{Y^2 \left(F + \frac{1}{\tilde{n}} \right)}{1 + \frac{1 - 2p_t}{p_t (1 - p_t)} X_2 \sqrt{\frac{p_t (1 - p_t)}{2n_1}}} - \frac{1}{\tilde{n}} \\ &\approx Y^2 \left(F + \frac{1}{\tilde{n}} \right) - \frac{1}{\tilde{n}}, \end{aligned} \tag{G.4}$$

where

$$Y = \frac{X_1 \sqrt{F} - X_2 \frac{1}{\sqrt{2n_1}} + X_3 \frac{1}{\sqrt{2n_2}}}{\sqrt{\frac{1}{2n_1} + F + \frac{1}{2n_2}}} \tag{G.5}$$

is a standard normal distributed random variable and \tilde{n} is defined in (13), keeping only terms of order $F + 1/\tilde{n}$ in the last step of (G.4). Since

$$\text{AsVar}(F_{OH}^{w*}) \approx \text{AsVar} \left(Y^2 \left[F + \frac{1}{\tilde{n}} \right] - \frac{1}{\tilde{n}} \right) = 2 \left(F + \frac{1}{\tilde{n}} \right)^2,$$

we may use (16)–(17) to conclude that

$$\frac{\text{AsVar}(\hat{N}_{eV})}{N_{eV}^2} \approx \frac{\text{AsVar}(F_{OH}^{w*})}{F^2} \approx 2 \left(1 + \frac{1}{F\tilde{n}} \right)^2 \approx 2 \left(1 + \frac{2N_{eV}T}{C\tau\tilde{n}} \right)^2.$$

As a next step, we generalize the derivation above and consider L loci, of which locus l has A_l alleles and aged averaged frequency $p_{tal} = p_{tal}^w$ of allele $a = 1, \dots, A_l$. Analogously with (G.1)–(G.3), we have that

$$\begin{aligned} (p_{t+\tau, 1l}, \dots, p_{t+\tau, Al}) &\approx (p_{t1l}, \dots, p_{tAl}) \\ &+ \left(X_{11l} \sqrt{F p_{t1l} (1 - p_{t1l})}, \dots, X_{1Al} \sqrt{F p_{tAl} (1 - p_{tAl})} \right), \\ (\hat{p}_{t1l}, \dots, \hat{p}_{tAl}) &\approx (p_{t1l}, \dots, p_{tAl}) \\ &+ \left(X_{21l} \sqrt{\frac{p_{t1l} (1 - p_{t1l})}{2n_1}}, \dots, X_{2Al} \sqrt{\frac{p_{tAl} (1 - p_{tAl})}{2n_1}} \right), \end{aligned}$$

and

$$\begin{aligned} (\hat{p}_{t+\tau, 1l}, \dots, \hat{p}_{t+\tau, Al}) &\approx (p_{t+\tau, 1l}, \dots, p_{t+\tau, Al}) \\ &+ \left(X_{31l} \sqrt{\frac{p_{t1l} (1 - p_{t1l})}{2n_2}}, \dots, X_{3Al} \sqrt{\frac{p_{tAl} (1 - p_{tAl})}{2n_2}} \right), \end{aligned}$$

where $A = A_l$ for simplicity of notation, and X_{i1l}, \dots, X_{iAl} are dependent standard normal random variables for $i = 1, 2, 3$ but independent for different i and l . Hence,

$$\begin{aligned} (\hat{p}_{t+\tau, 1l}, \dots, \hat{p}_{t+\tau, Al}) - (\hat{p}_{t1l}, \dots, \hat{p}_{tAl}) \\ \approx \sqrt{F + \frac{1}{\tilde{n}}} \left(\sqrt{p_{t1l} (1 - p_{t1l})} Y_{1l}, \dots, \sqrt{p_{tAl} (1 - p_{tAl})} Y_{Al} \right), \end{aligned}$$

where Y_{1l}, \dots, Y_{Al} are dependent standard normally distributed random variables, defined analogously as in (G.5), with Y_{al} a linear combination of X_{1al}, X_{2al} and X_{3al} . Then, summing over L loci in (9)–(10), and making a similar approximation as in (G.4), we find that

$$\begin{aligned} F_{OH}^{w*} &\approx \frac{\sum_{l=1}^L \sum_{a=1}^{A_l} (\hat{p}_{t+\tau, al} - \hat{p}_{tal})^2}{\sum_{l=1}^L \sum_{a=1}^{A_l} p_{tal} (1 - p_{tal})} - \frac{1}{\tilde{n}} \\ &\approx \left(F + \frac{1}{\tilde{n}} \right) \frac{\sum_{l=1}^L \sum_{a=1}^{A_l} Y_{al}^2 p_{tal} (1 - p_{tal})}{\sum_{l=1}^L \sum_{a=1}^{A_l} p_{tal} (1 - p_{tal})} - \frac{1}{\tilde{n}}. \end{aligned} \tag{G.6}$$

To calculate the variance of F_{OH}^{w*} we assume that each vector $(\hat{p}_{t+\tau, al} - \hat{p}_{tal})_{a=1}^{A_l}$ has the same covariance structure as a multinomial distribution i.e. that

$$\text{Var}(\hat{p}_{t+\tau, al} - \hat{p}_{tal}) = \left(F + \frac{1}{\tilde{n}} \right) p_{tal} (1 - p_{tal})$$

and

$$\text{Cov}(\hat{p}_{t+\tau, al} - \hat{p}_{tal}, \hat{p}_{t+\tau, bl} - \hat{p}_{tbl}) = - \left(F + \frac{1}{\tilde{n}} \right) p_{tal} p_{tbl}$$

for $a \neq b$. Hence,

$$\text{Var}(Y_{al}) = 1$$

and

$$\text{Corr}(Y_{al}, Y_{bl}) = - \sqrt{\frac{p_{tal} p_{tbl}}{(1 - p_{tal})(1 - p_{tbl})}} = \kappa_{abl}$$

for $a \neq b$. We can then decompose Y_{al} and Y_{bl} into independent standard random variables U, V and W by letting

$$Y_{al} = \sqrt{\kappa_{abl}} U + \sqrt{1 - \kappa_{abl}} V$$

and

$$Y_{bl} = \sqrt{\kappa_{abl}} U + \sqrt{1 - \kappa_{abl}} W$$

so that $\text{Var}(Y_{al}^2) = 2$ and

$$\begin{aligned} \text{Cov}(Y_{al}^2, Y_{bl}^2) &= \text{Cov} \left(\left[\sqrt{\kappa_{abl}} U + \sqrt{1 - \kappa_{abl}} V \right]^2, \left[\sqrt{\kappa_{abl}} U + \sqrt{1 - \kappa_{abl}} W \right]^2 \right) \\ &= \text{Cov} \left(\kappa_{abl} U^2 + (1 - \kappa_{abl}) V^2 + 2\sqrt{\kappa_{abl}(1 - \kappa_{abl})} UV, \right. \\ &\quad \left. \kappa_{abl} U^2 + (1 - \kappa_{abl}) W^2 + 2\sqrt{\kappa_{abl}(1 - \kappa_{abl})} UW \right) \\ &= \text{Var}(\kappa_{abl} U^2) \\ &= 2\kappa_{abl}^2 \end{aligned}$$

when $a \neq b$. Hence,

$$\begin{aligned} \text{Var} \left(\sum_{a=1}^{A_l} p_{tal}(1-p_{tal})Y_{al}^2 \right) &= \sum_{a=1}^A \sum_{b=1}^A p_{tal}(1-p_{tal})p_{tbi}(1-p_{tbi})\text{Cov}(Y_{al}^2, Y_{bl}^2) \\ &= 2 \sum_{a=1}^{A_l} p_{tal}^2(1-p_{tal})^2 + 2 \sum_{a \neq b} p_{tal}^2 p_{tbi}^2 \\ &= 2\kappa_l \left[\sum_{a=1}^{A_l} p_{tal}(1-p_{tal}) \right]^2 \end{aligned}$$

with κ_l as defined in (19). Letting L_e be the effective number of alleles in (18), taking the variance of the right hand side of (G.6) and dividing by $F^2 = (F^w)^2$ we obtain (17).

The confidence intervals (21)–(22) are also obtained from (G.6), looking at the pivotal statistic

$$\frac{F_{OH}^{w*} + \frac{1}{\bar{n}}}{F + \frac{1}{\bar{n}}},$$

which has distribution as in (20) when all loci are biallelic with the same allele frequencies and for the general case, its distribution is approximately $N(1, 2/L_e)$. \square

Appendix H. Effective number of independent alleles

Suppose that locus l has A_l alleles with equal allele frequency $1/A_l$ for $l = 1, \dots, L$. The effective number of independent alleles (18) can then be expressed as

$$\begin{aligned} L_e &= \frac{\left[\sum_{l=1}^L \sum_{a=1}^{A_l} \frac{1}{A_l} \left(1 - \frac{1}{A_l}\right) \right]^2}{\sum_{l=1}^L \kappa_l \left[\sum_{a=1}^{A_l} \frac{1}{A_l} \left(1 - \frac{1}{A_l}\right) \right]^2} \\ &= \frac{\left[\sum_{l=1}^L \left(1 - \frac{1}{A_l}\right) \right]^2}{\sum_{l=1}^L \kappa_l \left(1 - \frac{1}{A_l}\right)^2} \\ &= \frac{\left[\sum_{l=1}^L \left(\frac{A_l-1}{A_l}\right) \right]^2}{\sum_{l=1}^L \frac{A_l-1}{A_l^2}}, \end{aligned} \quad (\text{H.1})$$

since

$$\begin{aligned} \kappa_l &= \frac{\sum_{a=1}^{A_l} \frac{1}{A_l^2} \left(1 - \frac{1}{A_l}\right)^2 + \sum_{a \neq b} \frac{1}{A_l^2} \frac{1}{A_l^2}}{\left[\sum_{a=1}^{A_l} \frac{1}{A_l} \left(1 - \frac{1}{A_l}\right) \right]^2} \\ &= \frac{\frac{1}{A_l} \left(1 - \frac{1}{A_l}\right)^2 + A_l(A_l-1) \frac{1}{A_l^2}}{\left(1 - \frac{1}{A_l}\right)^2} \\ &= \frac{\frac{1}{A_l} \left(1 - \frac{1}{A_l}\right) + \frac{1}{A_l^2}}{\left(1 - \frac{1}{A_l}\right)} \\ &= \frac{1}{A_l - 1}. \end{aligned}$$

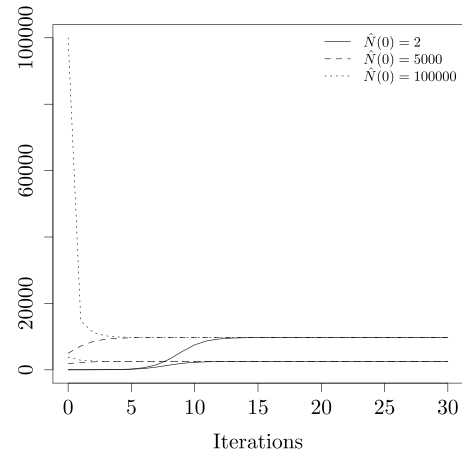


Fig. 5. Iterated estimates of the census (upper lines) and the variance effective population sizes (lower lines) for different initial estimates $\hat{N}(0)$ of N and one simulated data set. The true values are $N = 10\,000$ for the census size and $N_{ev} = 2589$ for the variance effective population size. Regardless of the starting value $\hat{N}(0)$, we obtain estimates $\hat{N} = \hat{N}(\infty) = 9715$ and $\hat{N}_{ev} = \hat{N}_{ev}(\infty) = 2519$. The population is simulated according to the demographic parameters of Table 2 with $L = 1000$ generated loci and $n = 500$ individuals sampled from age class 0 at two consecutive time points.

If all loci have the same number, A , of alleles then (H.1) simplifies to $L_e = (A - 1)L$. \square

Appendix I. Simultaneous estimation of effective and census population size

For simultaneous estimation of the effective population size and the population census size we make an initial guess $\hat{N}(0)$ of $N = N_t = N_{t+\tau}$. Then, starting with $i = 0$, we let

$$\hat{N}_{ev}(i) = \frac{1}{2} \left\{ 1 - \left[1 - \frac{1 - [1 - F_{OH}^{w*}(\hat{N}(i))]^{1/\tau}}{C(\hat{N}(i))} \right]^T \right\}^{-1}, \quad (\text{I.1})$$

be the estimate of the variance effective size in step i of the algorithm. Then we update our estimate of the census population size as

$$\hat{N}(i+1) = \hat{N}_{ev}(i) \frac{N}{N_{ev}}. \quad (\text{I.2})$$

After that, we let $i \leftarrow i + 1$, repeat (I.1)–(I.2) until both estimates converge, and let

$$(\hat{N}_{ev}, \hat{N}) = (\hat{N}_{ev}(\infty), \hat{N}(\infty)) \quad (\text{I.3})$$

be our final estimates.

To demonstrate this simultaneous estimation procedure, we simulated data by methods in Olsson and Hössjer (2014) and demographic parameters according to Table 2. In Fig. 5, we consider one simulated data set and illustrate the convergence rate of the estimates (I.1)–(I.2). We sampled 500 individuals without replacement from age class 0 at two consecutive time points. Three different starting guesses are used for $\hat{N}(0)$ and we see that regardless of which value we choose, all estimates converge to the same value.

References

- Baker, M., Mewaldt, L., Stewart, R., 1981. Demography of white-crowned sparrows (*Zonotrichia leucophrys nuttalli*). *Ecology* 62, 636–644.
Charlesworth, B., 2009. Effective population size and patterns of molecular evolution and variation. *Nature Rev. Genet.* 10, 195–205.

- Charlier, J., Laikre, L., Ryman, N., 2012. Genetic monitoring reveals temporal stability over 30 years in a small, lake-resident brown trout population. *Heredity* 109, 246–253.
- Crow, J.F., 1954. Breeding structure of populations. ii. Effective population number. *Stat. Math. Biol.* 543, 556.
- Crow, J., Denniston, C., 1988. Inbreeding and variance effective population numbers. *Evolution* 42, 482–495.
- Durrett, R., 2008. *Probability Models for DNA Sequence Evolution*. Springer.
- Emigh, T.H., Pollak, E., 1979. Fixation probabilities and effective population numbers in diploid populations with overlapping generations. *Theoret. Popul. Biol.* 15, 86–107.
- Engen, S., Lande, R., Saether, B., 2005. Effective size of a fluctuating age-structured population. *Genetics* 170, 941–954.
- Engen, S., Ringsby, T.H., Sæther, B.E., Lande, R., Jensen, H., Lillegård, M., Ellegren, H., 2007. Effective size of fluctuating populations with two sexes and overlapping generations. *Evolution* 61, 1873–1885.
- Ewens, W., 1982. On the concept of the effective population size. *Theoret. Popul. Biol.* 21, 373–378.
- Ewens, W., 2004. *Mathematical Population Genetics: I. Theoretical Introduction*. Vol. 27. Springer.
- Felsenstein, J., 1971. Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics* 68, 581–597.
- Fisher, R., 1958. *The Genetical Theory of Natural Selection*, second ed.. Dover Publications, Inc., New York.
- Hard, J.J., Mills, L.S., Peek, J.M., 2006. Genetic implications of reduced survival of male red deer cervus elaphus under harvest. *Wildl. Biol.* 12, 427–441.
- Hill, W., 1972. Effective size of populations with overlapping generations. *Theoret. Popul. Biol.* 3, 278–289.
- Hill, W., 1979. A note on effective population size with overlapping generations. *Genetics* 92, 317–322.
- Hill, W., 1981. Estimation of effective population size from data on linkage disequilibrium. *Genet. Res.* 38, 209–216.
- Hössjer, O., 2011. Coalescence theory for a general class of structured populations with fast migration. *Adv. Appl. Probab.* 43, 1027–1047.
- Hössjer, O., 2014. On the eigenvalue effective size of structured populations. *J. Math. Biol.* <http://dx.doi.org/10.1007/s00285-014-0832-5>.
- Hössjer, O., Olsson, F., Laikre, L., Ryman, N., 2014. A new general analytical approach for modeling patterns of genetic differentiation and effective size of subdivided populations over time. *Math. Biosci.* 258, 113–133.
- Hössjer, O., Olsson, F., Laikre, L., Ryman, N., 2015. Metapopulation inbreeding dynamics, effective size and subpopulation differentiation—a general analytic approach for diploid organisms. Under revision for *Theoretical Population Biology*.
- Hössjer, O., Ryman, N., 2014. Quasi equilibrium, variance effective population size and fixation index for models with spatial structure. *J. Math. Biol.* 69(5), 1057–1128.
- Johnson, D., 1977. Inbreeding in populations with overlapping generations. *Genetics* 87, 581–591.
- Jorde, P., 2012. Allele frequency covariance among cohorts and its use in estimating effective size of age-structured populations. *Mol. Ecol. Resour.* 12, 476–480.
- Jorde, P., Ryman, N., 1995. Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics* 139, 1077–1090.
- Jorde, P., Ryman, N., 1996. Demographic genetics of brown trout (*salmo trutta*) and estimation of effective population size from temporal change of allele frequencies. *Genetics* 143, 1369–1381.
- Jorde, P., Ryman, N., 2007. Unbiased estimator for genetic drift and effective population size. *Genetics* 177, 927–935.
- Kaj, I., Krone, S.M., Lascoux, M., et al., 2001. Coalescent theory for seed bank models. *J. Appl. Probab.* 38, 285–300.
- Krimbas, C.B., Tsakas, S., 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.
- Leslie, P., 1945. On the use of matrices in certain population mathematics. *Biometrika* 33, 183–212.
- Lewontin, R., Krakauer, J., 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74, 175–195.
- Luikart, G., Ryman, N., Tallmon, D., Schwartz, M., Allendorf, F., 2010. Estimation of census and effective population sizes: the increasing usefulness of dna-based approaches. *Conserv. Genet.* 11, 355–373.
- Nagylaki, T., 1980. The strong-migration limit in geographically structured populations. *J. Math. Biol.* 9, 101–114.
- Nei, M., Tajima, F., 1981. Genetic drift and estimation of effective population size. *Genetics* 98, 625–640.
- Nunney, L., 1991. The influence of age structure and fecundity on effective population size. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 246, 71–76.
- Nunney, L., 1993. The influence of mating system and overlapping generations on effective population size. *Evolution* 47, 1329–1341.
- Nunney, L., 1996. The influence of variation in female fecundity on effective population size. *Biol. J. Linn. Soc.* 59, 411–425.
- Olsson, F., Hössjer, O., 2015. Simulation methods for age structured populations. (submitted for publication).
- Olsson, F., Hössjer, O., Laikre, L., Ryman, N., 2013. Characteristics of the variance effective population size over time using an age structured model with variable size. *Theoret. Popul. Biol.* 90, 91–103.
- Pollak, E., 1983. A new method for estimating the effective population size from allele frequency changes. *Genetics* 104, 531–548.
- Pollak, E., 2011. Coalescent theory for age-structured random mating populations with two sexes. *Math. Biosci.* 233, 126–134.
- Sagitov, S., Jagers, P., 2005. The coalescent effective size of age-structured populations. *Ann. Appl. Probab.* 15, 1778–1797.
- Serbezov, D., Jorde, P.E., Bernatchez, L., Olsen, E.M., Vøllestad, L.A., 2012. Short-term genetic changes: evaluating effective population size estimates in a comprehensively described brown trout (*salmo trutta*) population. *Genetics* 191, 579–592.
- Sjödin, P., Kaj, I., Krone, S., Lascoux, M., Nordborg, M., 2005. On the meaning and existence of an effective population size. *Genetics* 169, 1061–1070.
- Turner, T.F., Richardson, L.R., Gold, J.R., 1999. Temporal genetic variation of mitochondrial dna and the female effective population size of red drum (*sciaenops ocellatus*) in the northern gulf of mexico. *Mol. Ecol.* 8, 1223–1229.
- Wang, J., 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Mol. Ecol.* 18, 2148–2164.
- Wang, J., Brekke, P., Huchard, E., Knapp, L.A., Cowlshaw, G., 2010. Estimation of parameters of inbreeding and genetic drift in populations with overlapping generations. *Evolution* 64, 1704–1718.
- Wang, J., Caballero, A., 1999. Developments in predicting the effective size of subdivided populations. *Heredity* 82, 212–226.
- Wang, J., Whitlock, M., 2003. Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* 163, 429–446.
- Waples, R.S., 1989a. A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* 121, 379–391.
- Waples, R.S., 1989b. Temporal variation in allele frequencies: testing the right hypothesis. *Evolution* 43, 1236–1251.
- Waples, R.S., 2002. *Definition and Estimation of Effective Population Size in the Conservation of Endangered Species*. University of Chicago Press.
- Waples, R.S., Antao, T., Luikart, G., 2014. Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics* 197, 769–780.
- Waples, R.S., Do, C., Chopelet, J., 2011. Calculating n_e and n_e/n in age-structured populations: a hybrid felsenstein-hill approach. *Ecology* 92, 1513–1522.
- Waples, R.S., Yokota, M., 2007. Temporal estimates of effective population size in species with overlapping generations. *Genetics* 175, 219–233.
- Wright, S., 1931. Evolution in mendelian populations. *Genetics* 16, 98–160.