An introduction to implementing LD and temporal methods with NeEstimator

Brendan Reid
July 9th, 2025
UP Cebu, Cebu City, Philippines

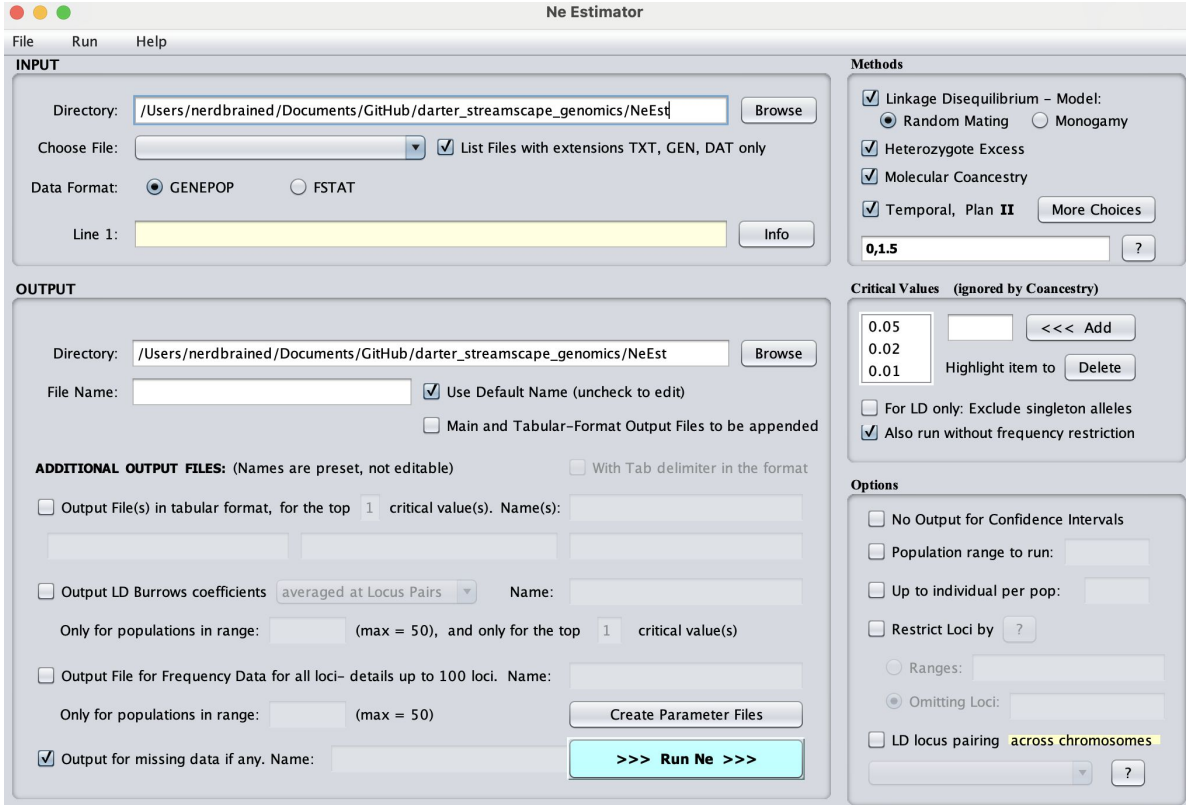
A moment of appreciation: Robin Waples

- University of Washington
- Salmonid biology
- Consultant on the PIRE project
- Master of N_e
- Co-creator of NeEstimator

NeEstimator overview

- Easy to execute wrapper for several different methods of estimating Ne from genetic data
- Several implementations
  - **Command-line version**
  - Java-based GUI
  - R (package RLDNe)
- Current version: NeEstimator v2 (Do et al., 2014)

NeEstimator GUI



Inputs: genepop format (.geno or .genepop)

```
Lower Arkansas + Rattlesnake
NC_048407_1_271437,NC_048407_1_9657082,NC_048407_1_9661459,NC_048407_1_10124451,NC_04
POP
4THST, 0102 0101 0101 0101 0101 0101 0202 0101 0101 0101 0101 0102 0101 0202 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0102 0102 0101 0101 0101 0102 0102 010
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0000 0102 0102 0101 0101 0202 0202 020
ARK10, 0000 0101 0101 0101 0101 0101 0202 0101 0000 0202 0101 0101 0101 0202 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0102 0102 0102 0101 0101 0101 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0101 0101 0101 0102 0101 0202 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0000 0101 0101 0102 0101 0202 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0101 0101 0101 0102 0101 0102 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0000 0202 0101 0101 0101 0202 0202 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0000 0102 0101 0102 0101 0102 0102 020
ARK10, 0102 0101 0101 0101 0101 0101 0202 0101 0000 0102 0101 0102 0101 0202 0202 020
```

Useful for microsatellite or SNP data

Format conversion: pgdspider (java) or adegenet (R)

Methods: Linkage Disequilibrium (LD)

Received: 22 June 2023 | Revised: 26 September 2023 | Accepted: 29 September 2023

DOI: 10.1111/1755-0998.13879

INVITED TECHNICAL REVIEW

MOLECULAR ECOLOGY RESOURCES WILEY

Practical application of the linkage disequilibrium method for estimating contemporary effective population size: A review

Robin S. Waples

Measuring linkage disequilibrium

Recall Hardy-Weinberg proportions: for unlinked loci $P_{A,B} = P_A P_B$

A simple measure of **disequilibrium**: $D_{A,B} = P_{A,B} - P_A P_B$

NeEstimator uses an adjusted metric (r^2) that is independent of allele frequencies and is always positive

From linkage disequilibrium to N_e

Expected r^2 can be calculated from recombination fraction (c) and Ne…

$$E(r^2) = \frac{c^2 + (1-c)^2}{2N_e c(2-c)} = \frac{\gamma}{N_e},$$

…and N_e can be back-calculated from r^2!

$$\hat{N}_e = \frac{\gamma}{r^2 - 1/n}$$

And if we know (or assume) loci are unlinked (c=0.5), this calculation is even simpler

$$\hat{N}_e = \frac{1}{(r^2 - 1/n)} \frac{1}{3} = \frac{1}{3r^{2\prime}},$$

Waples, 2014

Confidence intervals

Parametric:

$$\mathrm{CV}\left(\hat{N}_e\right) \approx \left(1 + \frac{1}{\gamma}\frac{N_e}{n}\right)\sqrt{2/k}$$

Waples, 2014

Bootstrap: re-sample SNPs (with replacement) *x* times, calculate Ne using *x* resampled datasets, and tabulate variation among re-sampled estimates

Assumptions and considerations: sampling/loci

Loci assumed to be unlinked

Sample sizes (number of individuals) may need to be *large*, particularly to estimate reasonable confidence intervals

Increasing the number of loci independent loci can increase precision, but subject to diminishing returns

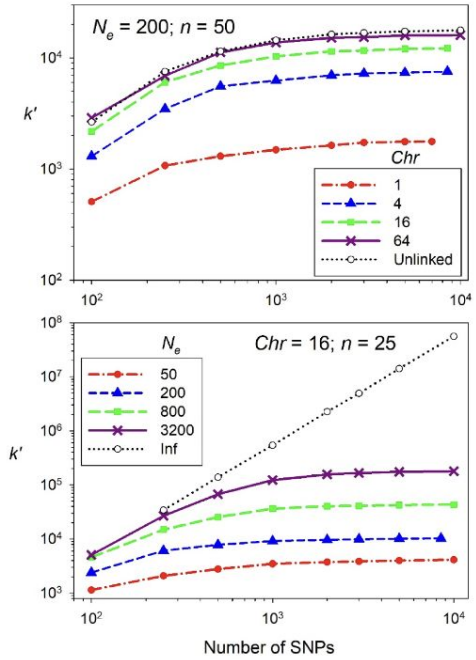


FIGURE 4 Effective degrees of freedom = effective number of locus pairs (k') for $\overline{r^2}$ as a function of the number (L) of diallelic (SNP) loci used to calculate $\overline{r^2}$. k' was calculated from simulated data based on the rate of decline in var($\overline{r^2}$) as more loci were used. Top: Effect of number of chromosomes (Chr), with N_e = 200 and n = 50. Bottom: effect of N_e, with Chr = 16 and n = 25. Modified from Waples et al. (2022).

Waples, 2014

Assumptions and considerations: biases

Mutation and selection don’t strongly bias the method

Population structure and immigration can introduce strong biases

Systems with overlapping generations, non-random mating systems, and age structure require special consideration

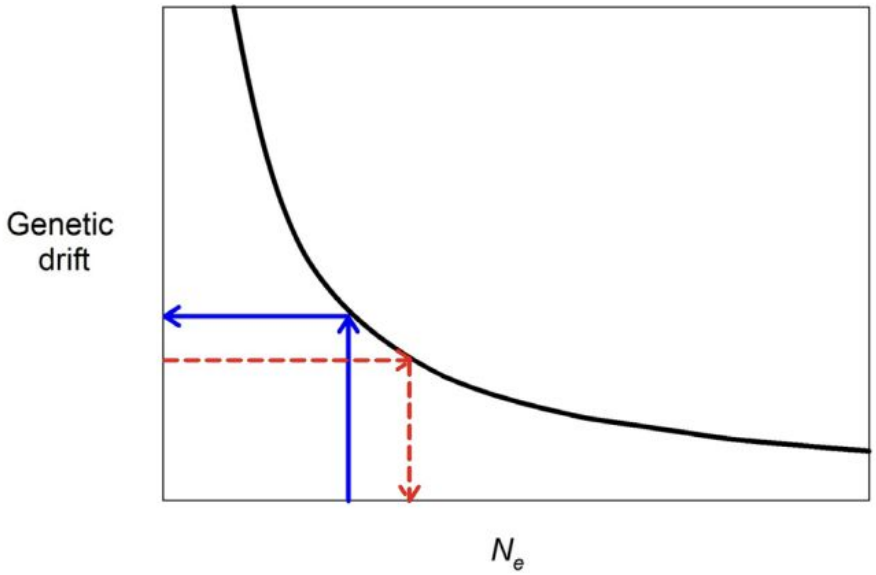
Temporal method: conceptual framework

Variance in allele frequencies (*F*) over time is related to N_e

$N_e \sim t/2F$

3 different ways of calculating *F*

- F_e (Nei and Tajima, 1981)
- F_k (Pollak, 1983)
- F_s (Jorde and Ryman, 2007)



Temporal method: sampling

Plan 1: sampling after reproduction, or sampling with replacement

Plan 2: sampling before reproduction (without replacement)

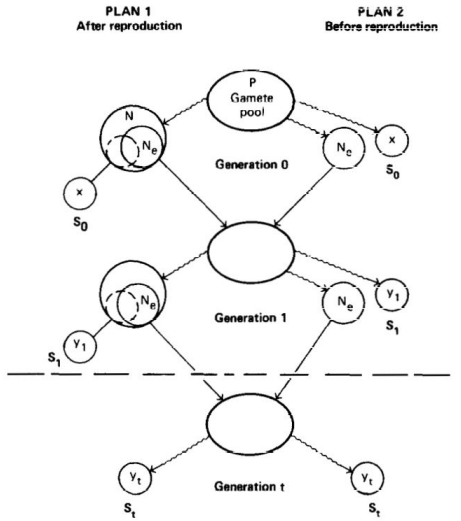


FIGURE 1.—Two sampling plans considered in the analysis. In both plans, P is frequency of an allele in gamete pool preceding generation 0, x and y_t are allele frequencies in samples (of S_0 and S_t individuals) for genetic analysis taken at generations 0 and t, respectively, N is total population size 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.

Waples, 2014

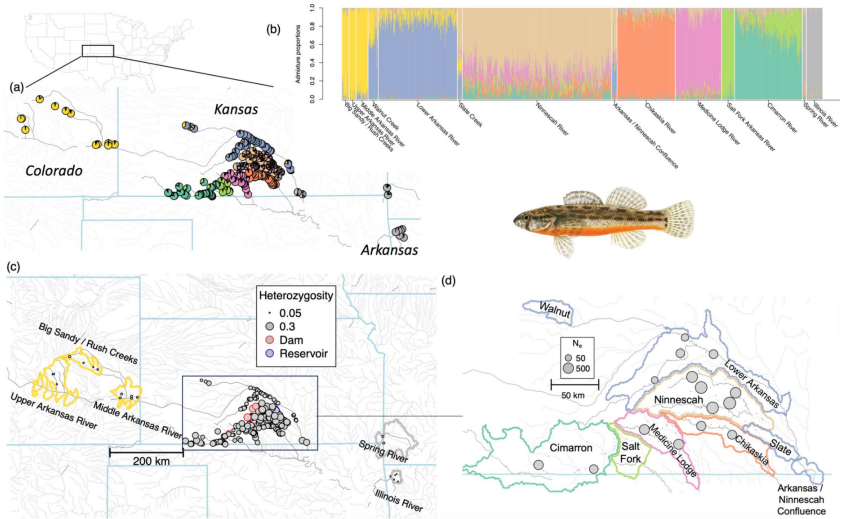
Time considerations

Temporal estimated of N_e don't represent contemporary N_e - instead, they represent a geometric average of N_e per generation over the time period sampled

Since drift accumulates over time, the number of generations between the sampling intervals can affect accuracy. More generations = greater signal of drift (to a point)...

Exercise: datasets

LD method: Arkansas darters



Temporal method: PIRE fish

S. delicatulus

H. miniatus

T. zosterophora

The Philippines PIRE Project

RESEARCH EDUCATION CONSERVATION