DR. GORDON LUIKART (Orcid ID : 0000-0001-8697-0582)

DR. ROBIN S WAPLES (Orcid ID : 0000-0003-3362-7590)

Article type : Resource Article

Detecting population declines via monitoring the effective number of breeders (N_b)

Gordon Luikart[1], Tiago Antao[2]*, Brian K. Hand[1], Clint C. Muhlfeld[3], Matthew C. Boyer[4], Ted Cosart[1], Brian Trethewey[1], Robert Al-Chockhachy[5], Robin S. Waples[6]

[1]Flathead Lake Biological Station, Montana Conservation Genomics Laboratory, Division of Biological Sciences, University of Montana, Polson, MT 59860, USA

[2]The Welcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom

[3]U.S. Geological Survey, Northern Rocky Mountain Science Center, Glacier National Park, West Glacier MT 59936 USA

[4]Montana Fish, Wildlife & Parks, Kalispell, MT 59901, USA

[5]U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, USA

[6]Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, 98112 USA

Corresponding author: Gordon Luikart (Gordon.luikart@umontana.edu); (406) 982-3301

*First two authors contributed equally

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1755-0998.13251

Abstract

Estimating the effective population size and effective number of breeders per year (N_b) can facilitate early detection of population declines. We used computer simulations to quantify bias and precision of the one-sample *LDNE* estimator of N_b in age-structured populations using a range of published species life history types, sample sizes, and DNA markers. N_b estimates were biased by ~5–10% when using SNPs or microsatellites in species ranging from fishes to mosquitoes, frogs, and seaweed. The bias (high or low) was similar for different life history types within a species suggesting that life history variation in populations will not influence N_b estimation. Precision was higher for 100 SNPs (H≈0.30) than for 15 microsatellites (H≈0.70). Confidence intervals (CI's) were occasionally too narrow, and biased high when N_b was small (N_b<50); however, the magnitude of bias would unlikely influence management decisions. The CI's (from *LDNE*) were sufficiently narrow to achieve high statistical power (≥0.80) to reject the null hypothesis that N_b=50 when the true N_b=30 and when sampling 50 individuals and 200 SNPs. Similarly, CI's were sufficiently narrow to reject N_b=500 when the true N_b=400 and when sampling 200 individuals and 5,000 loci. Finally, we present a linear regression method that provides high power to detect a decline in N_b when sampling at least five consecutive cohorts. This study provides guidelines and tools to simulate and estimate N_b for age structured populations (https://github.com/popgengui/agestrucnb/), which should help biologists develop sensitive monitoring programs for early detection of changes in N_b and population declines.

Keywords: effective population size, conservation genetics, population decline, genetic monitoring, population fragmentation, connectivity, viability, computer simulations, power analysis

Introduction

The effective population size (N_e) is among the most important parameters in conservation and evolutionary biology because N_e influences the efficiency of natural selection and gene flow, as well the rate of inbreeding and loss of genetic variation (Frankham 2005; Charlesworth 2009; Jamison and Allendorf 2012). Unfortunately, N_e is notoriously difficult to estimate, especially for species with age structure. In age-structured populations, we are often interested in both the effective size per generation (N_e) and the effective number of breeders per year or reproductive cycle (N_b). N_e is a crucial metric in conservation because, for example, if N_e is less than ~50, inbreeding often leads to substantial inbreeding depression (Jamison and Allendorf 2012). While much of population genetic theory uses N_e per generation, N_b can be a more relevant parameter than N_e in age-structured species. For example, N_b is important when studying seasonal or annual processes, reproduction events, or sexual selection in age-structured species.

The effective number of breeders (N_b) per reproductive cycle or cohort is advantageous to monitor because it allows early detection of a population decline. If N_b sharply declines for multiple reproductive cycles, then N_e

and N_c (population census size) will also likely decline (but see Whiteley et al. 2015). An N_b decline might be detectable by monitoring N_b for as few as 4 or 5 consecutive reproductive cycles in a species with a long generation interval of >10-20 years (Leberg 2005; Wang 2005; Antao et al. 2010). Early detection of a decline in N_b can help prevent loss of genetic diversity, population extirpation, and subsequent loss of ecosystem services (Schwartz et al. 2007; Luck et al. 2003; Schindler et al. 2010). N_b monitoring also allows early detection of population growth or expansion following species restoration, recovery, or spread of an invasive species (Kamas et al. 2016; Tallmon et al. 2012).

N_b estimation per breeding cycle, using a single-sample estimator, provides advantages over the estimation of effective population size per generation (N_e). First, estimating N_e may require waiting several years between sampling events, for example, when using the temporal method (Waples and Yokota 2007). However, estimating N_b allows frequent (annual) monitoring of population status, which is helpful for early detection of population trends in species with long generation intervals (Waples et al. 2013). For example, samples from newborns allow estimation of N_b a few weeks or months after the birthing season, which facilitates the assessment of population threats such as reproductive failure or cryptic population bottlenecks (Luikart et al. 1998). Sampling newborns can facilitate the sampling of single cohorts because in many species only newborns (or yearlings) can be aged. In some taxa, such as fishes, plants, and amphibians, we can sample several age classes during a single collection event, which allows testing for trends in N_b (Tallmon et al. 2012).

It recently has become feasible to estimate N_b in age-structured populations, using the single sample (one time point) genetic estimator *LDNE* (Waples and Do 2010). Waples et al. (2014) quantified the bias of the *LDNE* estimator of N_b related to age structure, using 100 microsatellite loci across a range of species and relatively large N_b estimates (N_b = 200-5000)(see also Robinson and Moyer 2013). However, the precision of this N_b estimator has not been extensively quantified for age-structured populations (but see Robinson and Moyer 2013), and the bias and the precision of the estimator are poorly understood when considering populations with different or variable life histories. For example, it is not known how changes in age-specific survival, fecundity, or longevity will bias or change N_b estimates, even if the true (deterministic) N_b remains constant; this is important because these age-specific vital rates influence the N_b/N_e ratio which can influence or bias N_e estimates obtained from the *LDNE* method (Waples et al. 2014).

Finally, we know little about bias and precision at small N_b ($N_b < 200$) in age-structured populations, or when using SNP loci. Here, we focus mainly on small N_b's, because effective size estimators perform best for small population sizes and because small populations have the greatest need for monitoring to prevent extirpation (Luikart et al. 1998; Leberg 2006). We also conduct simulations with a larger N_b, ranging from 300 to 2,000 to help quantify the power to identify populations with an N_b around 500. An N_b or N_e of 500 is important in conservation because the "50/500" rule states that N_e must be larger than ~500 to maintain evolutionary potential (Jamieson and Allendorf 2012). Frankham et al. (2014) recommend changing the 50/500 rule to 100/1000 justifying use of larger N_b for some simulations here.

When using hundreds of loci, confidence intervals (CIs) can be excessively narrow because not all the pairwise comparisons between loci are independent (Waples and Do 2008). Thus, a new CI estimation method was recently produced to provide wider and more reliable CI's (Jones et al. 2016). Finally, little is known about the effects of pooling cohorts on the magnitude of bias when using SNPs and small N_b, so we quantified the effects of pooling 2 or 3 cohorts.

Many species, including threatened trout, have substantial life history variation within and among populations (Shepard et al. 1984; Fraley and Shepard 1989; Northcote 1997; Al-Chokhachy and Budy 2008). It is important to quantify the effects of life history variation on the bias of N_b estimators because if the magnitude of the bias changes, the N_b estimates could change even when true N_b remains constant. For example, trout populations can have substantially different fecundities and age-specific survival rates if mortality increases in older fish (e.g., migratory individuals), due to predation or fishing mortality. Similarly, a population's average fecundity can also change rapidly if migratory fish are constrained due to a new barrier, fragmentation, overharvest of the large migratory females, or increased predation (e.g., by introduced species) along their migration pathway to spawning or feeding areas (Al-Chokhachy and Budy 2008). Migratory fish are often far larger than non-migratory fish and thus produce far more eggs.

Our overarching goal is to improve our ability to estimate and monitor N_b in natural populations by evaluating the performance of the *LDNE* estimator using many simulated populations with known N_b and age-structured populations. This is novel and important because most population genetics theory assumes discrete generations, but the vast majority of species have overlapping generations and age structured populations. Our five objectives are to: (1) quantify effects of life history variation (vital rate differences) on the bias and precision of N_b estimates; (2) quantify bias and precision for a range of microsatellite and SNP loci (15-5,000) and heterozygosities (H = 0.25 to 0.7); (3) assess the effects of pooling cohorts on N_b estimator bias; (4) quantify our ability to compute precise and reliable confidence intervals when N_b is near 50 or 500, and (5) quantify the power of a novel linear regression approach to detect a declining N_b when sampling 5 to 10 consecutive cohorts.

We conduct most simulations using 100-400 SNPs because this number is commonly used and easily feasible in many species thanks to recent SNP chip and genotyping-by-sequencing technologies such as GTseq and Rapture (Ruegg et al. 2014; Kraus et al. 2015; Narum et al. 2015; Ali et al. 2016). For simulations with larger N_b, we used 800 - 5,000 independent SNPs to achieve reasonable precision. Finally, we provide guidelines and a computer simulation program to compute, interpret, and simulate N_b estimates and confidence intervals over a broad range of taxa (https://github.com/popgengui/agestrucnb/). This study and simulation program will help researchers and managers develop and improve genetic monitoring programs for natural and managed populations (Schwartz et al. 2007; England et al. 2010).

Methods

Simulations and life tables

We simulated age-structured populations using the forward-time, individual-based simulator simuPOP (Peng and Kimmel 2005, Peng and Amos 2008). Each simulation tracked demographic and genetic processes in an age-structured population up to 1000 reproductive cycles (years). Demographics were governed by vital rates (age-specific survival and fecundity) and longevity provided in life tables. The life table data from mosquito, wood frog, and seaweed were reported in Waples et al. (2013). We used vital rate information published for life stages of the westslope cutthroat trout (Shepard et al. 1984; Fraley and Shepard 1989) and bull trout (Al-Chokhachy and Budy 2008) and converted them into age class data to construct life tables for simulating populations (see Table S1 and Appendix 1).

For cutthroat trout, one life table was constructed using data from Shepard et al. (1984), and a second life table (with the fecundity increasing with age) using data from Fraley and Shepard (1989) (Appendix 1). For bull trout, we constructed three different life tables that we termed "standard", "predation", and "long-lived" to span a realistic range of life histories and vital rates. The standard table was derived from migratory bull trout that exhibit an adfluvial life history (with rearing and foraging in lacustrine habitat) in the Flathead River system (Fraley and Shepard 1989) and elsewhere throughout their current range (Downs et al. 2006; Weaver 2006; Johnston and Post 2009). For the bull trout "predation" life table, we modified vital rates from the standard vital rates to simulate the effects of high predation on those age classes (ages 4 and 5) that migrate from their natal spawning streams to lakes (e.g., Flathead Lake, Montana). In many lakes, mortality caused by predation and competition is elevated by the introduced lake trout in the lake (Martinez et al. 2009; Ellis et al. 2011). For the "long-lived" life table, we used bull trout information from large lakes where individuals live longer than bull trout in the Flathead drainage (Johnston and Post. 2009).

For each life table, and for a given N_b, we computed demographic N_e by using the program *AGENE,* which is a deterministic discrete-time model (Waples et al. 2011). We used *AGENE* to determine the stable age distribution, total population size (N_T), and adult population size (N), given the life table vital rates and the number of offspring produced per year that survived to age 1 (N_1), as in Waples et al. (2014). The values of N_T, N, N_e, and N_b all scale linearly with N_1, so when a different N_1 is used, the ratios of these variables do not change. To initialize year 0 and to generate N_T individuals, the age of each individual was drawn randomly from the stable age distribution and the sex was randomly assigned (male or female) with equal probability. The total population size, and the number of individuals in each age class (by sex), varied randomly around the mean values expected in a stable population. The adult sex ratio varied randomly around 0.5 and could differ substantially from 0.5 due to sex-specific survival rates and ages at maturity.

To produce each newborn individual, one male and one female parent were drawn randomly from the pool of potential parents (those with ages for which $b_x > 0$). All potential parents of the same sex and age had an equal opportunity to be the parent of each newborn, but that was not necessarily true for individuals of different ages

or sex. That is, the probability that an individual of age x was chosen to be the parent of a newborn was proportional to b_x for that sex. We used the N_b /N_e ratio to assess the expected direction of bias and the approximate magnitude of bias in N_b estimates (as in Waples et al. 2014).

Loci

To compare microsatellites to SNPs, we simulated a set of 100 microsatellites (H ≈0.7), although most of our analyses use only 15 microsatellites as is typically used in many studies, including nearly all genetic studies of bull trout (Ardren et al. 2011; DeHaan et a. 2011). We simulate 100, 200, and 400 SNPs, which are the approximate numbers of loci often in studies using SNP chip and amplicon sequencing approaches (Hemmer-Hansen et al. 2011; Amish et al. 2012; Narum et al. 2010; Seeb et al. 2007; Seeb et al. 2012; Ali et al. 2016). To test for effects of heterozygosity on bias and precision, we simulated sets of SNPs with a range of mean heterozygosity (H ≈0.25. 0.30, 0.35, 0.40, and 0.45).

Allele frequencies for each locus in each replicate were separately initialized using a Dirichlet distribution, which is widely used in population genetics and has little influence on allele frequency distributions after a simulation burn-in of many generations (below). Multilocus genotypes in offspring were generated randomly assuming simple Mendelian inheritance from the two randomly chosen parents.

Data analysis and N_b estimation

After a simulation burn-in period of 50 years (which achieved an approximate demographic and genetic equilibrium; see Waples et al. 2014), we waited for the mean SNP heterozygosity to drop to 0.40 to achieve allele frequencies realistic for natural population and then tracked demographic and genetic parameters for another 50 years before starting a replicate. To quantify the effect of mean SNP heterozygosity as specified above, we also tracked results with other heterozygosity values (H = 0.45 - 0.25) for a subset of scenarios (e.g., N_b =50 or 100, and samples of 50 individuals and 100 SNPs). For each simulation scenario, we generated a total of 1,000 replicate samples. There was little/no difference between the low versus highest heterozygosity simulations so we present results from only one mean heterozygosity typical of many SNP studies (H = 0.30). For microsatellites, we waited until mean heterozygosity was near 0.75 (with ~8 alleles per locus).

We used four different sampling strategies useful in natural populations: (a) only newborns (that is, a single cohort), (b) two consecutive cohorts (50% newborns, 50% age 1 fish), (c) three consecutive cohorts (33% newborns, 33% age 1, and 33% age 2), and (d) all individuals in the population. In each case, individuals were sampled randomly without replacement from these targeted groups. For each strategy, we took samples of 15, 25, 50 and 100 individuals and evaluated them for 15 microsatellites and 100, 200, and 400 SNP loci. For model validation, we also conducted longer runs to track the loss of heterozygosity over time and compared (validated) the loss rate to that expected (from theoretical equations) and the rate estimated from values from *AGENE*.

In each simulation sample, we estimated effective size using the program *LDNE* (Waples and Do 2008). Because we were initially interested in assessing bias, we used $P_{crit} = 0.05$. P_{crit} is the lowest allele frequency allowed in the analysis. Waples and Do (2010) found that this P_{crit} value minimized bias with small sample sizes. Negative and infinite values of N_b estimates were converted to 10^6 as in previous related studies (Waples et al. 2014); negative N_b estimates can result when the LD signal (i.e., gametic disequilibrium signal) from sampling error noise is larger than the LD signal from the small number of parents and drift. For results reported below, unless otherwise stated (e.g., Fig. 1 left side panels), the estimates from *LDNE* were adjusted to reduce bias by applying the N_b/N_e bias adjustment from Waples et al. (2014).

The realized N_b from each replicate simulation was calculated using a standard formula for the inbreeding N_e (equation 2 in Waples et al. 2014). The realized N_b varies stochastically among simulation replicates with variance ~ N/2; therefore, the coefficient of variation in realized N_b increases as the population size decreases (Waples and Faulkner 2009). Because this simulation-induced stochasticity in N_b among simulation replicates is relatively large for small N_e and does not occur in natural populations (each of which has a single true trajectory of N_e over time) (Waples and Faulkner 2009), we constrained the realized N_b to vary only by < 1% above or below the expected (deterministic value), e.g. $N_b = 50$, or 100. Thus, all simulations of $N_b = 50$ included simulation replicates with a realized N_b of $49.5 \leq N_b \leq 50.5$.

Violin and box plots

For easy comparison among simulated scenarios (life histories, numbers of loci and individuals sampled) we produced violin plots (Fig. 1) and box plots visualizing the distribution of N_b point estimates (e.g., Figs. 2 and 3), as well as the distribution of the upper and lower confidence interval limits (Figs. 4 and 5). Each box plot shows the median, box edge percentiles (20th and 80th percentiles), and 5th and 95th percentiles of the point estimate from each of 1,000 simulation replicates for each simulation scenario.

Linear f method

To quantify our ability to detect a declining N_b by sampling multiple consecutive cohorts, we simulated 1000 independent declines of 5%, 7%, 10% and 15% per year (or cohort) and using a 0% decline as a control. This was an exponential decline because, for a 10% decline, each year N_b lost 10% of what was left: N_b =100, 90, 81, 73, 66, 59, 53, 48, 43, 39, and 35. This type of decline is close to linear for 5-10 years, linear in log space, and never reaches 0 but gets arbitrarily close. We then conducted linear regressions through 5, 7 or 10 consecutive cohort N_b point estimates (from *LDNE*) and tested whether the slope of the line was negative as expected for a declining N_b. Statistical tests for a significant negative slope (and thus a population decline) were computed using least-squares linear regression (Neter 1985). The test statistic (t*) for the slope of a linear regression can be calculated using the equation below for a normally distributed regression with a null hypothesis that b_1 is equal to zero, where b_1 is the slope of the regression and $s(b_1)$ is an estimate of the variance of the slope (Neter 1985).

227 $$t^* = {b_1}/{s(b_1)} \qquad (1)$$

228 Using t* and the degrees of freedom of the regression (n - 2 where n is the number of points used in the linear
229 regression) we calculated the p-value for that line using a Cumulative Density Function (CDF) on the T
230 distribution (Neter 1985).

231 **Results**

232 We first computed N_b and N_e for each life table using the deterministic model in the program *AGENE*, as in
233 Waples et al. (2014). For example, the N_b/N_e ratio was 0.79 for the standard bull trout life table. This ratio
234 dropped to 0.66 for the "predation" bull trout life table, which had higher mortality rates for the 4 and 5-year-
235 old age classes. The N_b/N_e ratio for bull trout with a longer life span (BT-Long) was 0.78. The N_b/N_e ratio for
236 a mosquito, wood frog, and seaweed, were 0.27, 0.60, and 1.26, respectively (Waples et al. 2013).

237 Our stochastic simulations with random demographic variability (using simuPOP) yielded populations with the
238 same N_b/N_e ratios as the deterministic model *AGENE* and agreed closely with theoretical expectations of the
239 rate of loss of heterozygosity given the N_e from *AGENE*. Thus, we next looked for potential bias in the
240 genetically based *LDNE* N_b estimates for each sample of individual genotypes simulated with simuPOP by
241 comparing these N_b estimates with the *AGENE* true N_b values.

242 *Bias, cohort pooling*

243 Our bias in *LDNE* estimates of N_b due to age structure was similar in magnitude (3% to 15%) to previous
244 evaluations that considered relatively large N_b's (N_b > 200) and microsatellite loci (Waples et al 2014). The
245 direction of the bias was generally upward for the species with $N_b < N_e$ (bull trout, cutthroat trout, mosquito,
246 and wood frog), as expected (Waples et al. 2014). The direction of bias was downward for the species with N_b
247 $> N_e$ (seaweed), also as expected (Fig. 1). The results reported below include the bias correction using the
248 N_b/N_e ratio adjustment (as in Waples et al. 2014), which generally reduced the magnitude of bias by a few
249 percent, as in previous studies (Waples et al. 2014). Many of the results below are also reported for only one
250 life history (BT-Stnd, i.e., standard bull trout), unless otherwise stated, because the magnitude of bias and the
251 precision were similar for the range of life histories considered here (Fig. S1 in supplementary materials).

252 The bias was generally similar for microsatellites and SNPs (Fig. 2). The bias was highest (~15%) in some
253 scenarios when using only 15 microsatellite loci (Fig. 2, N_b = 100). The magnitude of bias was similar across
254 the range of the number of loci used (up to 400) and of individuals (25-100) considered here.

255 Heterozygosity of markers had little effect on bias. For example, as the mean heterozygosity decreases from
256 0.40 to near 0.25 for 100 SNPs, the distribution of N_b point estimates (from 1000 simulations) shifted only
257 slightly (data not shown).

Pooling samples from two or three cohorts increased the magnitude of upward bias to ~30-40% higher than the deterministic (true) N_b (*AGENE*). Combining cohorts increases the upward bias when the true N_e is larger than N_b, as here (Waples et al. 2014). For example, pooling two consecutive cohorts gave a median estimate of N_b = 65 from 1000 simulations when the actual deterministic N_b was only 50. Pooling three cohorts further increased the magnitude of bias, such that the mean N_b increased to ~70 when the deterministic N_b per cohort was only 50 (Fig. 3). This bias high agrees with the upward bias reported by Waples et al. 2014 when pooling of cohorts from populations with N_b/N_e ratios less than 1.0. Pooling can be more appropriate when estimating N_e, not N_b, because the estimates obtained from *LDNE* for pooled cohorts often approach N_e (see figure 4a in Waples et al. 2014).

Precision, confidence intervals, and power

Precision was higher for 100 bi-allelic SNPs than for the 15 microsatellites having ~8 alleles per locus. For example, the range of the N_b point estimates was 90-165 for microsatellites versus 95-130 for 100 SNPs, when the deterministic (true) N_b (from *AGENE*) was 100 (Fig. 2). These N_b estimates included the N_b/N_e bias adjustment from Waples et al. (2014), which used an assumed true (deterministic) N_b/N_e computed in program *AGENE* using life history parameters. When the N_b (from *AGENE*) was only 50, the range of point estimates was ~45 to 75 for 15 microsatellites versus only ~46 to 65 for 100 SNPs, when sampling a single cohort and 50 individuals. Precision increased substantially such that the distribution of point estimates narrowed when using 200 SNPs compared to 100 SNPs, in all the species evaluated (trout, wood frog, seaweed, mosquito); however, precision only slightly improved for 400 SNPs compared to 200 SNPs (Fig. 2).

Confidence interval estimates (95% CIs, from *LDNE* jackknife method) performed well when using 100 SNP loci and 50 individuals as they contained the deterministic N_b for 94% of simulation replicates for the bull trout (BT-Stnd) (Table 1; Fig. 4). When N_b =50, only 90% of independent CI's contained the deterministic (true) N_b, when sampling 100 SNPs and 50 individuals. For example, bull trout had only 90% of independent CI's that contained the deterministic N_b, when simulating an N_b = 100 and when sampling 50 individuals and 100 SNPs (Table 1). CI's tended to be biased high, which contributed to only 90% of CI's containing the true deterministic N_b.

Confidence intervals for a true N_b = 30 were below N_b = 50 in 80% of simulations when 200 loci and 50 individuals were genotyped; thus the power was ~0.80 to detect that N_b was below 50. The CI distributions and power were similar (~0.80 to 0.90) for other species including mosquitos, westslope cutthroat trout, and wood frog (Fig. S1). Similarly, confidence intervals for a true N_b=400 were below N_b =500 in approximately 80% of simulations (power ~ 0.80) when 400 loci and 100 individuals were genotyped (Fig. 5); the power increased to >0.95 when genotyping 800 loci and 500 individuals (Fig. 5). Finally, when the true N_b=2,000, ~80% of simulated CI's allowed rejection of the null hypothesis that N_b = 2,300 when sampling at least 500

individuals and 5,000 loci (Fig. S2 in supplementary materials). Importantly, for the larger N_b values of 500 or 2,000, the size of CI's was reduced more by doubling the number of individuals than by doubling the number of loci sampled (Fig. S2).

Power to detect a declining N_b via linear regression

The linear regression method for detecting a declining N_b is visualized in Fig. 6. The benefit of doubling the number of cohorts from 5 to 10 increased power from 0.55 to 1.0 (Fig. 6, 7) when sampling 100 individuals and 100 SNP loci during a 10% annual decline in N_b. In another example, power to detect a 15% decline per year in N_b was only ~0.53 (53%) when sampling 50 individuals and 100 SNP loci from each of five consecutive cohorts and testing for a negative slope (Fig. 7). Power increased to ~0.73 and ~0.80 when doubling the number of loci and individuals, respectively. Power increased to near 100% when doubling the number of consecutive cohorts that were sampled from 5 to 10 cohorts (Fig. 7, dashed arrow).

We conducted an extensive power analysis for detecting different rates of N_b decline (5%, 7%, 10%, and 15%) when using different sample sizes of individuals, SNPs, and number of cohorts. This analysis showed that doubling the number of cohorts from 5 to 10 increased power far more than doubling the number of loci or individuals (Fig. 7). Sampling more than 5 cohorts was often required to achieve power >0.80 to detect N_b declines, given the range of N_b values and sample sizes considered here. Finally, a power analysis in wood frogs revealed very similar power for detecting N_b declines as in bull trout (see Fig. 7 versus Fig. S3 in supplementary materials).

Discussion

We evaluated the effects of life history variation and sampling strategy on estimates of N_b to help biologists plan genetic monitoring programs and obtain more reliable estimates of N_b in natural and managed populations. We found that life history variation, such as changes in survival or fecundity within a species did not cause substantial variation of N_b estimates, for the scenarios studied here. This observation is important for researchers interested in monitoring N_b in species with variable vital rates because it demonstrates that changes in N_b estimates do not likely reflect changes in vital rates. We also report that the bias in N_b estimates is generally small (<5-10%) for relatively small population sizes and SNP marker sets. These scenarios (e.g., N_b < 200; SNPs), along with sampling of pooled cohorts have not been thoroughly investigated. Our simulations and discussions below regarding the behavior of confidence interval estimates and power for detecting N_b differences and population declines will help researchers understand how to monitor N_b in age-structured populations.

Bias

The bias correction, based on a species' N_b/N_e ratio (Waples et al. 2014) reduced the magnitude of bias slightly for all five species, which had a wide range of N_b/N_e ratios. This result is similar to that reported for larger

populations and microsatellite loci (Waples et al. 2014). The results here are useful because they consider relatively small N_b (25 to 200) typical of threatened species, and they consider different marker types (15-100 microsatellites and 100-400 SNPs). The greatest proportional bias occurred at small N_b. For example, when true $N_b = 25$, point estimates after applying the N_b/N_e correction were still approximately 10-12% biased-low ($N_b = 22.2$) for mosquitos, and 10% biased-high ($N_b = 28.0$) for bull trout. This downward bias occurs for mosquitos because their N_b is greater than N_e, unlike the trout that have an N_b less than N_e (Waples et al. 2014). This magnitude of bias is only ~5% when N_b becomes large ($N_b \geq 200$), which is consistent with the findings of Waples et al. (2014). The bias is generally small and unlikely to cause biologists to make erroneous management conclusions. For example, the N_b estimate of 28 (instead of 25) for the bull trout likely would not prompt a different management decision.

The cause of bias in a single-cohort sample has been discussed by Waples et al. (2013) and Waples et al. (2014). Briefly, there are two main sources of the LD influencing the estimate of N_b: the N_b per year that reflects new LD produced by the effective number of breeders (N_b), and the N_e per generation that reflects residual LD that has not yet broken down. The sampling process also generates LD. The *LDNE* estimator, in effect, assumes N_b equals N_e. However, if N_e is larger than N_b, there is less residual LD signal from N_e than is assumed by the estimator, and the N_b estimate is biased high, as we observed in trout (e.g., Fig. 1). Conversely, if N_e is smaller than N_b, there is more LD signal from N_e than assumed by the estimator and the N_b estimate is biased low (Fig. 1; and see Fig. 2 in Waples et al. 2014). Importantly, estimating the expected bias in magnitude and direction, which is predictable from the N_b /N_e ratio, can help researchers interpret N_b estimates and avoid potentially erroneous inferences.

With microsatellite loci, the bias was occasionally slightly higher than with SNPs, likely because of the larger proportion of low-frequency alleles for microsatellites compared to SNPs (Waples and Do 2010), and perhaps because the initial bias corrections for *LDNE* were derived from simulations of two-allele loci (Waples 2006). SNPs are becoming more widely used than microsatellites for most conservation applications and taxa. A set of 15 microsatellites have been widely used to assess population genetic structure and diversity in bull trout populations (Ardren et al. 2010; DeHaan et al. 2011). However, sets of ≥100 SNPs are increasingly used because this number of SNPs can be genotyped for less cost than 10 microsatellites (Amish et al. 2010; Campbell et al. 2015; Ali et al. 2016).

An advantage of SNPs is that thousands can be screened to find hundreds with relatively high heterozygosity (e.g., $H > 0.2$) for use in SNP chip or other genotyping technologies, which improves accuracy and power for N_b estimation. Another advantage of SNP chips, GTseq, or Rapture is they can include marker loci from all chromosomes, sex identification loci, mitochondrial loci, and species-diagnostic loci for detection of hybrids (Amish et al. in press).

Importantly, changing the mean heterozygosity of SNPs ranging from 0.25 to 0.40 had little effect on bias. However, if many loci have very low heterozygosity (H < 0.1) and thus have low-frequency alleles, the N_b point estimates could become less precise and more biased (Waples and Do 2010).

Sampling multiple cohorts

Occasionally it is not feasible to sample enough individuals (n > 20-30) from a single cohort, and thus cohorts must be pooled to achieve sufficient sample sizes. Pooling samples from two or three cohorts increased the magnitude of bias to near a 20% and 30% overestimation of N_b, respectively (Fig. 3). This bias from pooling cohorts is expected only when the N_b/N_e ratio is not near 1.0 (Waples et al. 2014). This high-bias could result from less LD signal in a sample of multiple cohorts due to more individual parents contributing offspring to the sample. The bias from pooling could result from the pooling leading toward estimating the total N_e per generation, which is larger than N_b in these trout (for which $N_b/N_e \approx 0.78$); Recall that the algorithm for estimation assumes $N_b/N_e \approx 1.0$ (Do et al. 2014). The direction of the bias (high versus low) depends on whether the N_b/N_e ratio is low versus high, respectively; a bias-low occurs for an N_b/N_e ratio > 1.0 (Waples et al. 2014).

Because we now know the magnitude of bias from pooling, our results and those from Waples et al. (2014) suggest we could correct for the bias when interpreting or estimating N_b (and N_e) from pooled cohorts. For example, for bull trout, the bias for two pooled cohorts is approximately 20% high, and thus we can subtract approximately 20% from any N_b point estimate calculated from samples of pooled cohorts for bull trout. Bias correction for any given species will depend on the N_b/N_e ratio and the effects of cohort pooling, which can be quantified as we did here using simulations in the program AgeStrucNe.

Previous work showed that pooled cohort samples can yield *LDNE* estimates that reflect the generational N_e more accurately than the cohort N_b (Robinson and Moyer 2013, Waples et al. 2014). Our results for bull trout simulations, which yielded cohort-pooled estimates of ~65 and 70 from two and three cohorts, respectively, indicate a relatively accurate estimation of the generational N_e with pooled cohorts (true generational $N_e \sim 64$ based on $N_b = 50$ and $N_b/N_e = 0.78$). Thus, these results suggest that biologists can use the *LDNE* output to obtain approximate estimates of N_e from pooled cohorts. N_e can also be inferred from an N_b estimate of a single cohort if you know the N_b/N_e ratio (e.g. $N_b/N_e = 0.78$); For example, an N_b/N_e ratio of 0.78 would correspond to an N_e that was 28% higher than an estimated N_b value (1/0.78 = 1.282); Thus, if $N_b = 100$, $N_e = 128$.

Precision and Confidence Intervals

Poor precision is usually the main limitation for the application of N_e estimators to natural populations (Leberg 2005; Wang 2006; Luikart et al. 2010). The precision and the width of confidence intervals for the LDNE method improves rapidly (geometrically) with the number of loci (L) because the degrees of freedom is based

on a multiple of L as follows: n = [(K-1)^2]*L(L-1)/2, where K is the number of alleles per locus and K-1 is the number of independent alleles. There are L(L-1)/2 pairs of loci. For each pair of loci, there is the equivalent of (K-1)^2 independent comparisons of alleles (Waples and Do 2010). Ironically, in this genomics age, high precision (narrow confidence limits) can be problematic because when thousands of loci are used, CI's can become excessively narrow.

Two major factors determine the performance of confidence intervals: a) whether the point estimate is unbiased, and b) whether the correct degrees of freedom are used to generate the width of the CI's. If the point estimate is strongly biased, even CI's with the proper width will perform poorly, and if the degrees of freedom are too large the CI's will be too narrow and will include the true value less than the expected fraction of the time, even if the point estimate is unbiased. For the LD method, the number of pairwise comparisons increases with the square of the number of loci. If all of these pairwise comparisons provided independent information, precision would be very high with 1000s of SNP loci, and resulting CI's would be very tight. In reality, however, physical linkage and overlapping pairs of loci in the comparisons mean that the effective (true) degrees of freedom is considerably less than the number of pairwise comparisons (e.g., see Figure 7 in Waples et al. 2016).

This reduction in effective degrees of freedom is less of an issue in most of our evaluations, which use no more than 100-800 loci. Furthermore, we used the Jones et al. (2016) improved jackknife method (which is implemented in NeEstimator V2.1) to generate realistic confidence intervals that reflect the true effective degrees of freedom for each dataset (Do et al. 2014). Therefore, any deviations in the performance of the CI's can be attributed to bias in the estimates of N_b. Ironically, if enough data are used, even a small bias can translate into poor CI performance in terms of covering the true N_b, because off-centered CI's will become narrower as precision increases and less likely to contain the true parameter value.

Our observation of lower precision for 15 microsatellites compared to 100 SNPs was expected from the lower degrees of freedom (fewer pairwise locus comparisons) for the microsatellites. For 100 SNPs, approximately 96% of CI's contained the deterministic N_b (known from *AGENE*) when the N_b was 200 and when sampling 50 individuals. This 96% containment is close to the 95% coverage expected when computing 95% CI's for standard statistical tests. CI's were extremely wide when N_b =200 with samples of only 25 individuals (and 100 SNPs), likely because of the low signal to sampling-noise ratio (Waples et al. 2010). When N_b =200 and sampling 25 individuals, the upper CI limit was usually greater than 600 and often was infinity when genotyping 100 loci. Thus, a larger sample size (>75-100) or more loci (> 200-400) will be needed to achieve reasonable precision when N_b is 200 or larger.

Confidence intervals contained the true N_b less often than expected when N_b was relatively small. For example, when N_b = 50 and when using 100 SNPs, approximately only 90% of CI's contained the deterministic N_b (known from *AGENE*), when sampling 50 individuals (Table 1). Thus, in this scenario, CI's

contained the deterministic N_b 5% less often than the expected 95% of CI's. There are two main causes for this. First, the N_b estimator is slightly biased high (even after using the N_b/N_e bias correction from Waples et al. 2014), thus CI's tend to be shifted high and therefore contain the deterministic N_b less often than expected. Second, CI's become relatively narrow as the number of loci increases to >100-200, even after using the recent correction to widen CI's (Jones et al. 2016).

Power to Determine When N_b is Small

It is crucial for population assessment and monitoring programs to have high power to identify populations with a small N_b. For example, according to the "50/500 rule" if N_e is smaller than 50, inbreeding can occur at a high rate and cause reduced fitness, i.e., inbreeding depression (Jamieson and Allendorf 2012). Similarly, an excessive loss of evolutionary potential can occur if $N_e < 500$ (Jamieson and Allendorf 2012). Therefore, we simulated N_b values that were 20%-30% below 50 (and also below 500) to determine how many loci and individuals are required to identify populations with N_b less than 50, or 500. Recall that N_b might nearly equal N_e for some taxa like bison, red deer, mole crabs, fruit flies, sagebrush lizards, dolphins, Atlantic cod, razorback suckers (see supplementary materials in Waples et al. 2013). However, the N_b/N_e results for these analyses assumed that males had random reproductive success within each age class. If this is not true, N_b will be affected more than N_e, and thus N_b might not equal N_e.

Confidence intervals were narrow enough to identify a population with N_b approximately 20-30% less than 50. For example, when the true N_b equaled 30, we could reject the hypothesis that $N_b = 50$ in approximately 80% of simulations when sampling 50 individuals and 200 loci (Fig. 4, middle panel); thus the power was ~0.80. The power was higher (~0.95) to reject $N_b = 50$, when the true $N_b = 30$ and when sampling 100 individuals and 200 loci (Fig. 4, middle-right panel).

Similarly, confidence intervals were narrow enough to identify a population with N_b ~20% lower than 500. That is, when the true $N_b = 400$, ~80% of CI's were less than $N_b = 500$ if 5,000 loci and 200 individuals are sampled (Fig. 5). These examples and results in figures 4 and 5 will help biologists develop genetic monitoring programs to precisely estimate N_b and detect when $N_b < 50$ or $N_b < 500$ (see also Fig. S1).

Power to detect a declining N_b using linear regression

Genetic monitoring programs need high power to detect a population decline. Many biologists would like to detect a decline in N_b of 10% per year (or reproductive cycle). Power was too low to detect a 10% decline when using the linear regression test for a negative slope when regressing a line through estimates of N_b for each of five consecutive cohorts and using 50-100 individuals with 100-200 SNPs (Fig. 6, 7). Power to detect a 10% decline increased to >0.80 when sampling 400 SNPs, 100 individuals, and only 5 consecutive cohorts (Fig. 7, red diamond). Statisticians generally recommend a power of > 0.80 to make a study worth conducting or a monitoring program worth implementing. In another example, the power to detect a 15% decline (starting

at N_b = 50) was 0.80 when sampling 100 SNPs and 100 individuals from each of 5 consecutive cohorts (Fig. 7). Power to detect a 15% decline was >0.80 when 7 cohorts, 25 individuals and 200 SNPs were sampled to test for a linear decline in N_b across cohorts. Power to detect a 15% decline was similar for bull trout and wood frog as is shown by comparing Fig. 7 versus Fig. S3 (see dash line ovals).

For a comparison with microsatellite loci, we quantified the power to detect a 15% decline in N_b using 30 microsatellites. We discovered that 50 individuals from each of 5 consecutive cohorts provide power of 0.59 when using 30 loci (Figure S3). Power increased to 1.00 when 10 cohorts were sampled. These microsatellite power results are similar to power from 100 SNPs for the same 15% decline when also sampling 100 individuals for 5 and then 10 consecutive cohorts (Fig. 7). Researchers can quantify power to develop sensitive monitoring programs using the simulation program AgeStrucNe that is freely available at https://github.com/popgengui/agestrucnb/

Our power analysis provides guidelines for the number of cohorts, loci, and individuals needed to achieve high power to detect a linear or exponential N_b decline of 5% to 15% per reproductive cycle (Fig. 7). The results suggest that we must generally sample >5 consecutive cohorts to achieve power > 0.80 unless >100 individuals are sampled per cohort. It can be difficult in threatened species or small populations to sample >25 individuals, which will make it difficult to achieve power > 0.80, even with 400 SNPs and samples from 10 consecutive cohorts. Future research is needed to test if a thousand SNPs might increase power above 0.80 to detect a 5% decline when sampling small numbers of individuals. Biologists can address this and other questions using the AgeStrucNe simulation package.

Limitations and future research

Future research is needed using simulations and empirical datasets with larger N_b and thousands of loci to understand the limitations of N_b estimation using genomic approaches. Marandel et al. (2018) simulated populations with N_e of 1,000 to 1,000,000 and ~200 loci and concluded that large samples of individuals (thousands to millions) must be sampled to obtain useful *LDNE* estimates of N_e. Using many thousands of loci can improve precision. However, loci often are not independent when many pairs of loci are from the same chromosomes (Larson et al. 2014). Use of loci from different chromosomes is facilitated by using program NeEstimator and inputting the chromosomal map position of loci. Restricting comparisons to loci residing on different chromosomes will eliminate linkage bias but does not make all the pairwise comparisons of loci independent (Waples et al. 2016).

We also need future research to advance the use of linked sets of loci with known recombination rates because the use of recombination information can increase power to detect and date historical bottlenecks (e.g., Hill 2001; Tenesa et al. 2007; Lehnert et al. 2019). The use of runs of homozygosity (RoH) to estimate N_e is becoming feasible for non-model species (Browning and Browning 2015; Grossen et al. 2018). However, this

will remain difficult for many species because it requires the mapping of tens of thousands of loci and genotyping the loci in many individuals.

Future research should go beyond simply detecting an N_b decline to also determine the cause of a decline. For example, if the slope of an N_b decline can be inferred from the linear regression method, this slope could be tested for correlations with environmental variables that might be driving declines (or increases). Environmental variables such as temperature, habitat availability, invasive species, diseases or predators, are increasingly available from public databases from NASA and other sources (e.g., Table II in Grummer et al. 2019). Interestingly, Whiteley et al. (2015) suggested that inter-annual variation in streamflow could be driving inter-annual variation in N_b in native trout populations.

Importantly, biologists must know the N_b/N_e ratio before estimating N_b (or N_e) in species with age structured populations because the interpretation of N_b estimates (from *LDNE*) requires knowledge of this ratio. This is because N_b estimates are biased if $N_b \neq N_e$. Fortunately, estimation of the N_b/N_e ratio is easily feasible using the program AgeNe or AgeStrucNe (Waples 2011; and see https://github.com/popgengui/agestrucnb/).

Finally, we need future research to understand the temporal stability of the N_b/N_c ratio in natural populations. If the ratio remains stable over many generations, the population census size (N_c) could be inferred from N_b which would facilitate monitoring of population abundance from N_b (Pierson et al. 2018).

Conclusions

We show how N_b point estimates and confidence intervals from the one-sample *LDNE* method can be reliably computed and used to estimate and monitor N_b in age-structured populations. The bias adjustment method, based on the N_b/N_e ratio (Waples et al. 2014), produced N_b estimates biased by 5-10%. This magnitude of bias is relatively small and unlikely to influence conservation or management decisions. Life history and vital rate variation within species had little effect on the magnitude of bias of N_b estimates, suggesting managers can monitor N_b with little concern that a change in vital rates (survival or fecundity) would strongly shift the estimates of N_b when the true N_b has not changed. Our results showed that confidence intervals (CIs) for N_b estimates are generally reliable. However, the CI's were occasionally narrow and biased-high when N_b was small (<30) and hundreds of loci were used. *LDNE* CI's were sufficiently narrow to reject the hypothesis that $N_b = 50$ when the true N_b was only 40 and when sampling >100 individuals and 400 SNPs. Similarly, CI's were sufficiently low to reject the hypothesis that $N_b = 500$ when the true N_b was only 400 and when sampling >300 individuals and ~5,000 independent SNPs. Power to detect a declining N_b was high (>0.80) when using the linear regression test across ≥ 7 consecutive cohorts (breeding cycles) and when sampling at least 50 individuals and 100 loci. The guidelines and simulation approach presented here, along with the software AgeStrucNb, will help biologists develop sensitive genetic monitoring programs to detect changes in N_b and thereby help to conserve populations and prevent extinctions.

Acknowledgments: GL and BKH were supported in part by funding from NASA grant number NNX14AB84G. GL, BKH, and CCM were supported in part by funding from U.S. National Science Foundation grants DEB-1258203 and DoB-1639014. Montana Fish, Wildlife & Parks provided support through Bonneville Power Administration project #199101903. CCM was also supported by the Great Northern LCC of the US Fish and Wildlife Service.

Literature cited

Al-Chokhachy R, Budy P (2008) Demographic characteristics, population structure, and vital rates of a fluvial population of bull trout in Oregon. *Trans Am Fish Soc* **137**: 1709–1722.

Ali, OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jefferes C, Miller MR (2016) RAD Capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics* **202**:389–400

Amish SJ, Hohenlohe PA, Leary RF, Muhlfeld CC, Allendorf FW, Luikart G (2012) Next-generation RAD sequencing to develop species-diagnostic SNPs chips: An example from westslope cutthroat and rainbow trout. *Mol Ecol Resour* **12**: 653–660.

Amish SJ, Miller M, O'Rourke S, Boyer M, DeHaan P, Bernall S, Muhlfeld CC, Luikart G (2019) Improved relatedness estimation, hybrid detection, and sex identification using a SNP-chip developed from next generation RAD sequencing in threatened bull trout. In revision.

Antao T, Perez-Figueroa A, Luikart G (2011) Early detection of population declines: high power of genetic monitoring using effective population size estimators. *Evol Appl* **4**: 144–154.

Ardren WR, DeHaan PW, Smith CT, Taylor EB, Leary R, Kozfkay CC, Godfrey L, Diggs M, Fredenberg W, Chan J, Kilpatrick CW, Small MP, Hawkins DK (2011) Genetic structure, evolutionary history, and conservation units of bull trout in the continuous United States. *Trans Am Fish Soc* **140**: 506–525.

Browning SR, Browning BL (2015) Accurate non-parametric estimation of recent effective population size from segments of identity by descent. *Am J Hum Genet* **97**:404-418.

Campbell NR, Harmon SH, Narum SR (2015) Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Mol Ecol Resour* **15**: 855-67.

Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* **10**: 195-205.

DeHaan PW, Bernall SR, DosSantos JM, Lockard LL, Ardren WR (2011) Use of genetic markers to aid in re-establishing migratory connectivity in a fragmented metapopulation of bull trout (*Salvelinus confluentus*). *Can J Fish Aquat Sci* **68:** 1952–1969.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2.0: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol Ecol Resour* **1**: 209–214.

England PR, Luikart G, Waples RS (2010) Early detection of population fragmentation using linkage disequilibrium estimation of effective population size. *Conserv Genet* **11**: 2425–2430.

Fraley JJ, Shepard BB (1989) Life history, ecology and population status of migratory bull trout (*Salvelinus confluentus*) in the Flathead Lake and River System, Montana. *Northwest Sci* **63**: 133–143.

Frankham R, Bradshaw CJA, Brook BW (2014) Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation* **170:** 56–63.

Grummer, J. et al. 2019. Aquatic landscape genomics and environmental effects on genetic variation. *Trends in Ecology and Evolution*, In press. DOI: 10.1016/j.tree.2019.02.013

Hemmer-Hansen J, Nielsen E, Meldrup D, Mittelholzer C (2011) Identification of single nucleotide polymporhisms in candidate genes for growth and reproduction in a non-model organism, the Atlantic cod, *Gadus morhua*. *Mol Ecol Resour* **11** (Suppl. 1): 71–80.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* **38**: 209–216.

Johnston FD, Post JR (2009) Density-dependent life-history compensation of an iteroparous salmonid. *Ecological Applications* **19**: 449–467.

Kanda N, Allendorf FW (2001) Genetic population structure of bull trout from the Flathead River basin as shown by microsatellites and mitochondrial DNA markers. *Trans Am Fish Soc* **130**: 92–106.

Kovach RP, Muhlfeld CC, Boyer M, Lowe W, Allendorf FW, Luikart G (2014) Dispersal and selection mediate hybridization between a native and invasive species. *P Roy Soc B* **282**: 20142454.

Kraus RHS, vonHoldt B, Cocchiararo B, Harms V, Bayerl H, Kühn R, Förster DW, Fickel J, Roos C, Nowak C (2015) A single-nucleotide polymorphism-based approach for rapid and cost-effective genetic wolf monitoring in Europe based on noninvasively collected samples. *Mol Ecol Resour* **15**: 295–305.

Larson WA, Seeb LW, Everett MV, Waples RK, Templin WD, Seeb JE (2014) Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, **7**: 355-69.

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

Lehnert SJ, Kess T, Bentzen P, Kent MP, Lien S, Gilbey J, Clément M, Jeffrey NW, Waples RS, and IR Bradbury (2019) Genomic signatures and correlates of widespread population declines in salmon. *Nature Communications* 10:2996 (doi: 10.1038/s41467-019-10972).

Luck GW, GC Daily, PR Erlich (2003) Population diversity and ecosystem services. *Trends Ecol Evol* **18**: 331–336.

Luikart G, Sherwin W, Steele B, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* **7:** 963-974.

Martinez PJ, Bigelow P, Deleray MA, Fredenberg WA, Hansen BS, Horner NJ, Lehr SK, Schneidervin RW, Tolentino SA, Viola AE (2009) Western lake trout woes. *Fisheries* **34**: 424–442.

Marandel F, Lorance P, Berthelé O, Trenkel VM, Waples RS, Lamy JB (2018) Estimating effective population size of large marine populations, is it feasible? *Fish Fish* **20**:189–198.

Muhlfeld CC, Kovach RP, Al-Chokhachy R, Amish SJ, Kershner JL, Leary RF, Lowe WH, Luikart G, Matson P, Schmetterling DA, Shepard BB, Westley PAH, Whited D, Whiteley A, Allendorf FW (2017) Legacy introductions and climatic variation explain spatiotemporal patterns of invasive hybridization in a native trout. *Global Change Biology*, **23**: 4663-4674.

Narum SR, Campbell NR, Kozfkay CC, Meyer KA (2010) Adaptation of redband trout in desert and montane environments. *Mol Ecol* **19**: 4622–4637.

Neter J, Kutner, MH, Wasserman W (1985) Applied linear statistical models: regression, analysis of variance, and experimental designs (2nd ed.). Homewood, IL: Irwin.

Palstra FP, Fraser DJ (2012) Effective/census population size ratio estimation: a compendium and appraisal. *Ecol Evol* **2**: 2357–2365.

Peng B, Kimmel M (2005) simuPOP: a forward-time population genetics simulation environment. *Bioinformatics* **21**: 3686–3687.

Peng B, Amos CI (2008) Forward-time simulations of non-random mating populations using simuPOP. *Bioinformatics* **24**: 1408–1409.

Pierson JC, Graves TA, Banks SC, Kendall KC, Lindenmayer D B (2018) Relationship between effective and demographic population size in continuously distributed populations. Evolutionary Applications, 1**1:** 1162–1175.

Robinson JD, Moyer GR (2013) Linkage disequilibrium and effective population size when generations overlap. *Evol Appl* **6**: 290–302.

Ruegg KC, Anderson EC, Paxton KL, Apkenas V, Lao S, Siegel RB, DeSante DF, Moore F, Smith TB (2014) Mapping migration in a songbird using high-resolution genetic markers. *Mol Ecol* **23**: 5726–5739.

Schindler DE, Hilborn R, Chasco B, Boatright CP, Quinn TP, Rogers LA, Webster MS (2010) Population diversity and the portfolio effect in an exploited species. *Nature* **465**: 609–612.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends Ecol Evol* **22**: 25–33.

Seeb LW, Antonovich A, Banks MA, Beacham TB, Bellinger MR, Blankenship SM, Campbell MR, Decovich NA, Garza JC, Guthrie III CM, Lundrigan TA, Moran P, Narum SR, Stephenson JJ, Supernault KJ, Teel DJ, Templin WD, Wenburg JK, Young SF, Smith CT (2007) Development of a Standardized DNA Database for Chinook Salmon. *Fisheries* **32**: 540–552.

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Mol Ecol Resour* **11** Suppl. S1: 1–8.

Shepard B, Pratt K, Graham P (1984) Life Histories of Westslope Cutthroat and Bull Trout in the Upper Flathead River Basin, Montana. Montana Department of Fish, Wildlife and Parks, Kalispell, Montana. 85 p.

Tallmon DA, Waples RS, Gregovich D, Schwartz MK (2012) Detecting population recovery using gametic disequilibrium-based effective population size estimates. *Conserv Genet Resour* **4**: 987–989.

Tenesa A, Navarro P, Hayes BJ, Duffy DL, Clarke GM, Goddard ME, Visscher PM (2007) Recent human effective population size estimated from linkage disequilibrium. *Genome Research* **17**: 520–526.

Waples RS, Do C (2008) *LDNE*: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resourc* **8**: 753–756.

Waples RS, Faulkner JR (2009) Modelling evolutionary processes in small populations: not as not as ideal as you think. *Mol Ecol* **18**: 1834–47.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol Appl* **3**: 244–262.

Waples, RS, 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.

641 Waples RS, Luikart G, Faulkner JR, Tallmon DA (2013) Simple life history traits explain key effective
642 population size ratios across diverse taxa. *P Roy Soc London B* **280**: 20131339.

643 Waples RK, Larson W, Waples RS. 2016. Estimating contemporary effective population size in non-model
644 species using linkage disequilibrium across thousands of loci. Heredity 117(4) . 233-40.

645 Waples RA, Antao T, Luikart G (2014) Effects of overlapping generations on linkage disequilibrium estimates
646 of effective population size. *Genetics* **197**: 769–780.

647 Waples RS, Yokota M (2007) Temporal estimates of effective population size in species with overlapping
648 generations. *Genetics* **175**: 219–233.

649 Whiteley AR, Coombs JA, Hudy M, Robinson Z, Nislow KH, Letcher BH (2012) Sampling strategies for
650 estimating brook trout effective population size. *Conserv Genet* **13**: 6625–637.

651 Whiteley AR, Coombs JA, Cembrola M, O' Donnell MJ, Hudy M, Robinson Z, Nislow KH, Letcher BH
652 (2015) Effective number of breeders provides a link between interannual variation in stream flow and
653 individual reproductive contribution in a stream salmonid. *Molecular Ecology* **24:** 3585–3602

654

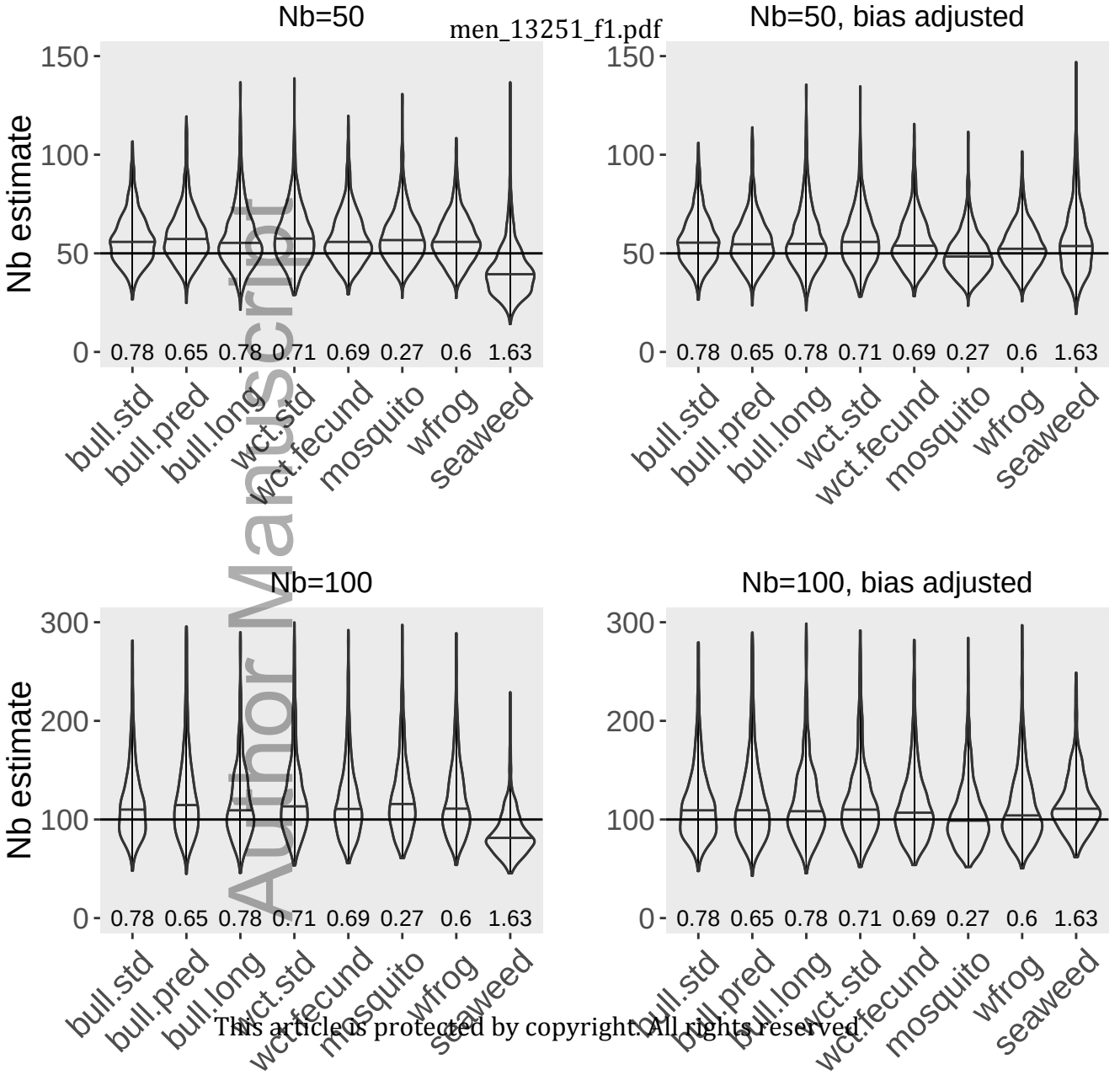
655 **Data Accessibility Statement:** Life tables and a web link to the computer program (AgeStrucNe) with user's
656 manual are available in supplementary material. The web link to the program is
657 https://github.com/popgengui/agestrucnb/ . This link is also the Abstract. Lifetables for several species
658 (seaweed, wood frog, and mosquito) were reported in Waples et al. 2013.

Table 1. Median of N_b point estimates from 1000 simulated populations with true N_b values of 25, 50, 100, and 200. The "% of CI's low" is the percentage of populations with the upper CI below the true (simulated) N_b. The "% of CI's high" is the percentage of populations with the loci CI above the true N_b. Simulations were conducted using 100 SNPs and samples of 25, 50 and 100 individuals for BT-Stnd life history. Note that CI's are often biased high, especially when the true N_b is small with larger samples of individuals (see bold numbers).

True N_b	**Median of 1000 N_b point estimates**	**Number of individuals sampled**	**% of CI's low***	**% of CI's high**
25	26.6	25	1%	6%
25	27.8	50	1%	**15%**
25	NA	**100	NA	NA
50	52.0	25	1%	4%
50	54.8	50	1%	9%
50	53.6	100	1%	**13%**
100	101.0	25	1%	2%
100	109.9	50	1%	3%
100	105.0	100	1%	5%
200	166.6	25	1%	2%
200	226.0	50	1%	3%
200	211.8	100	1%	3%

* most percentages in this column were between 0.5 to 1.04 and were rounded to 1%

** NA = not enough individuals (100) existed to be sampled for simulations at small population size (N_b)



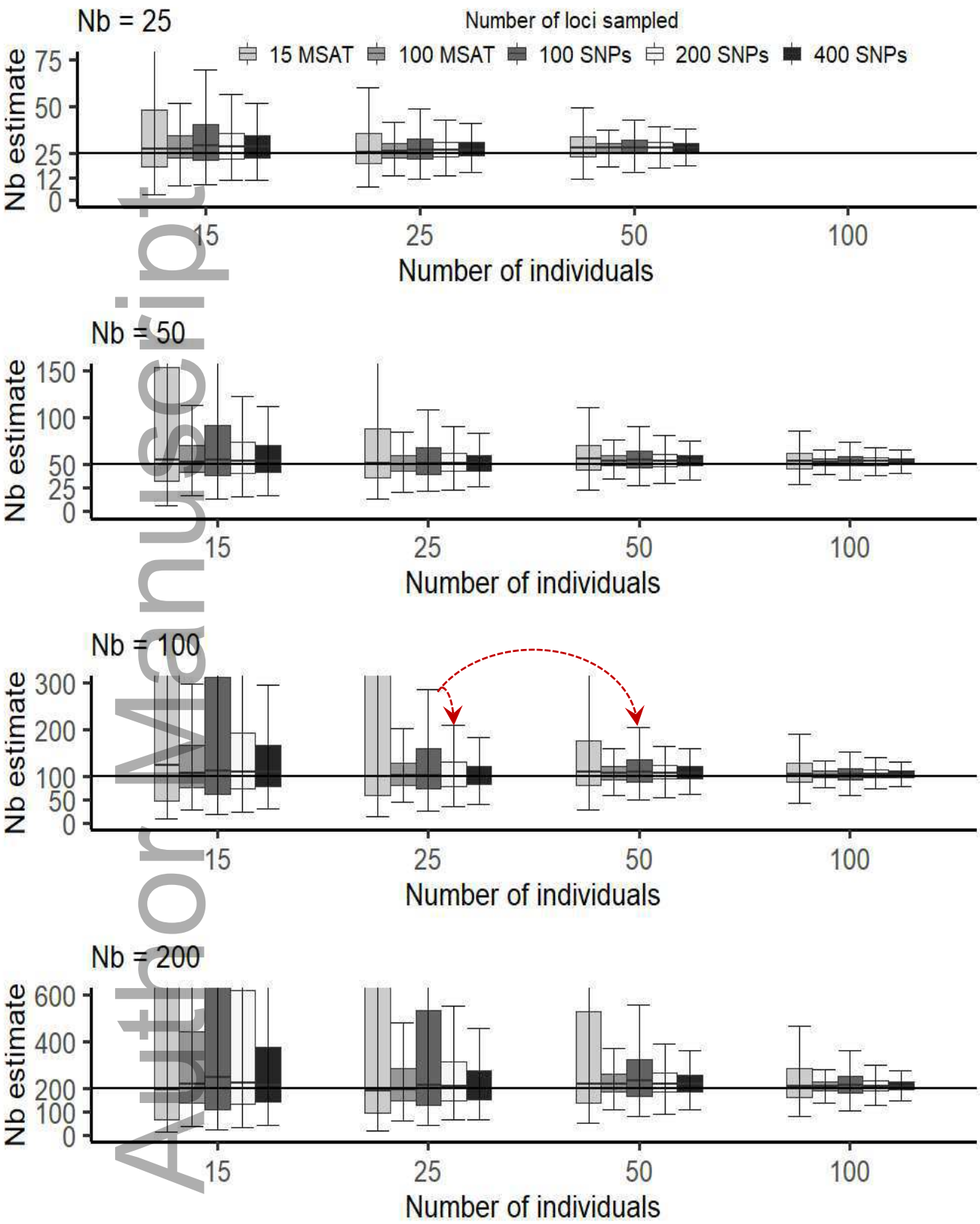
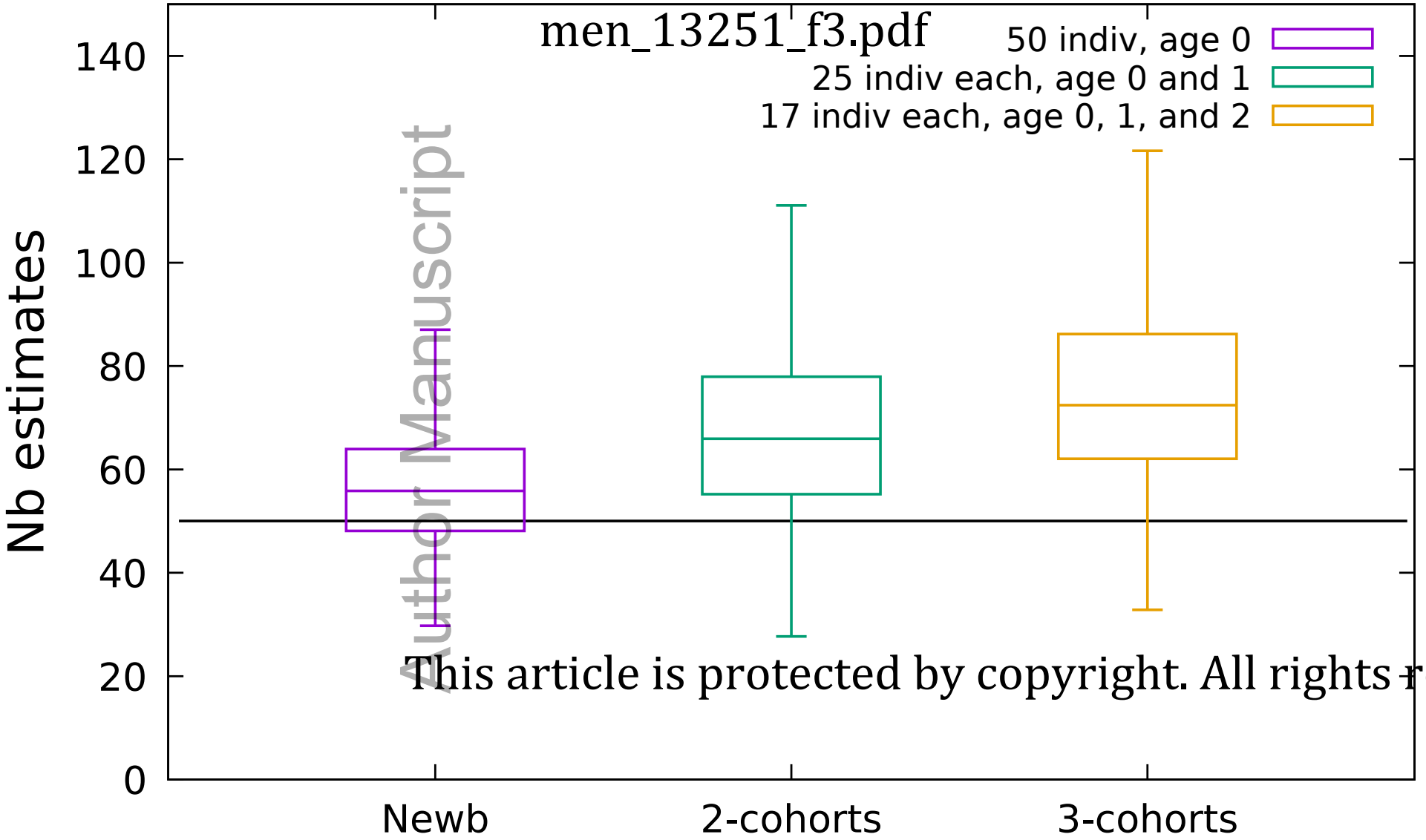
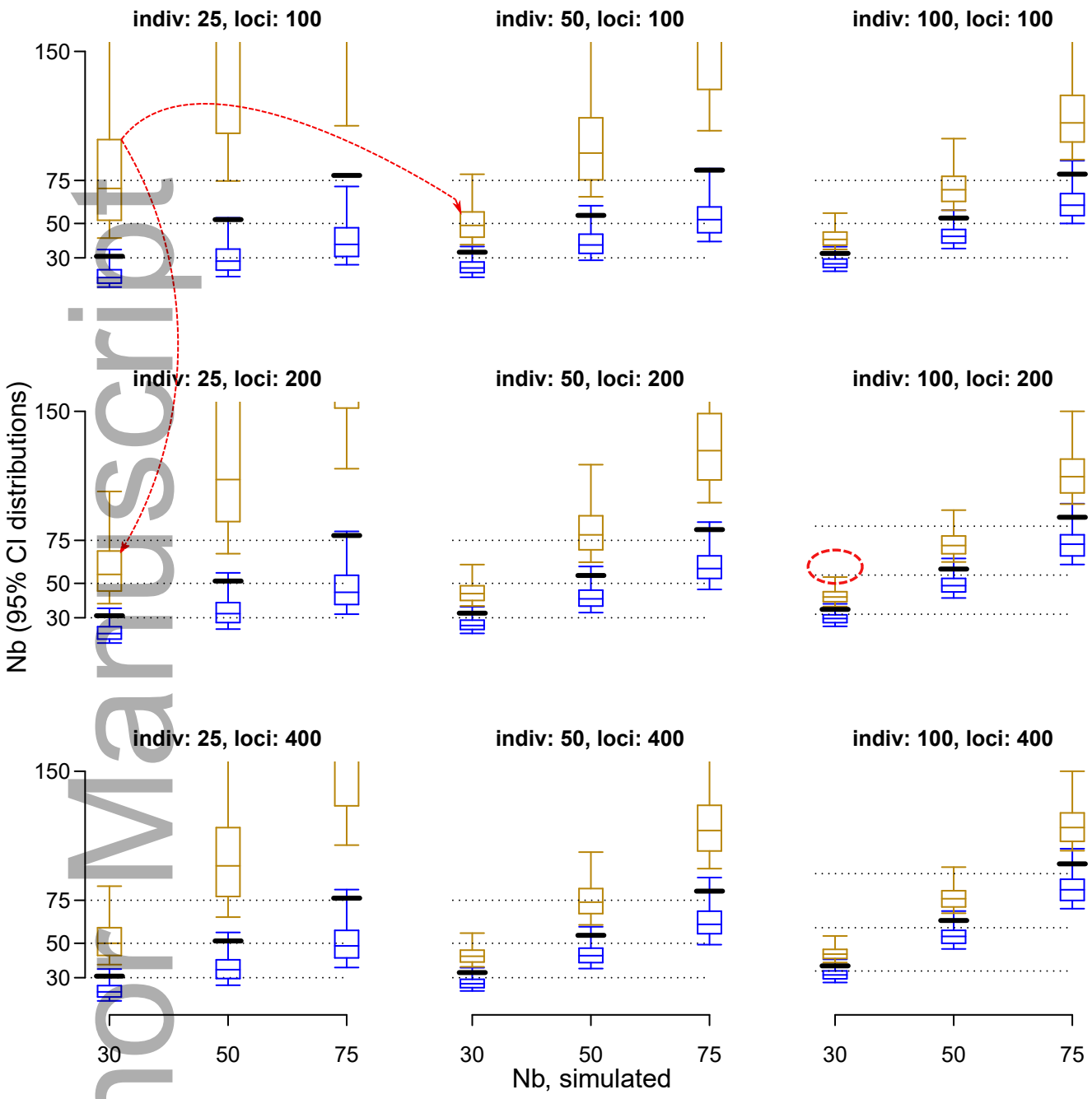
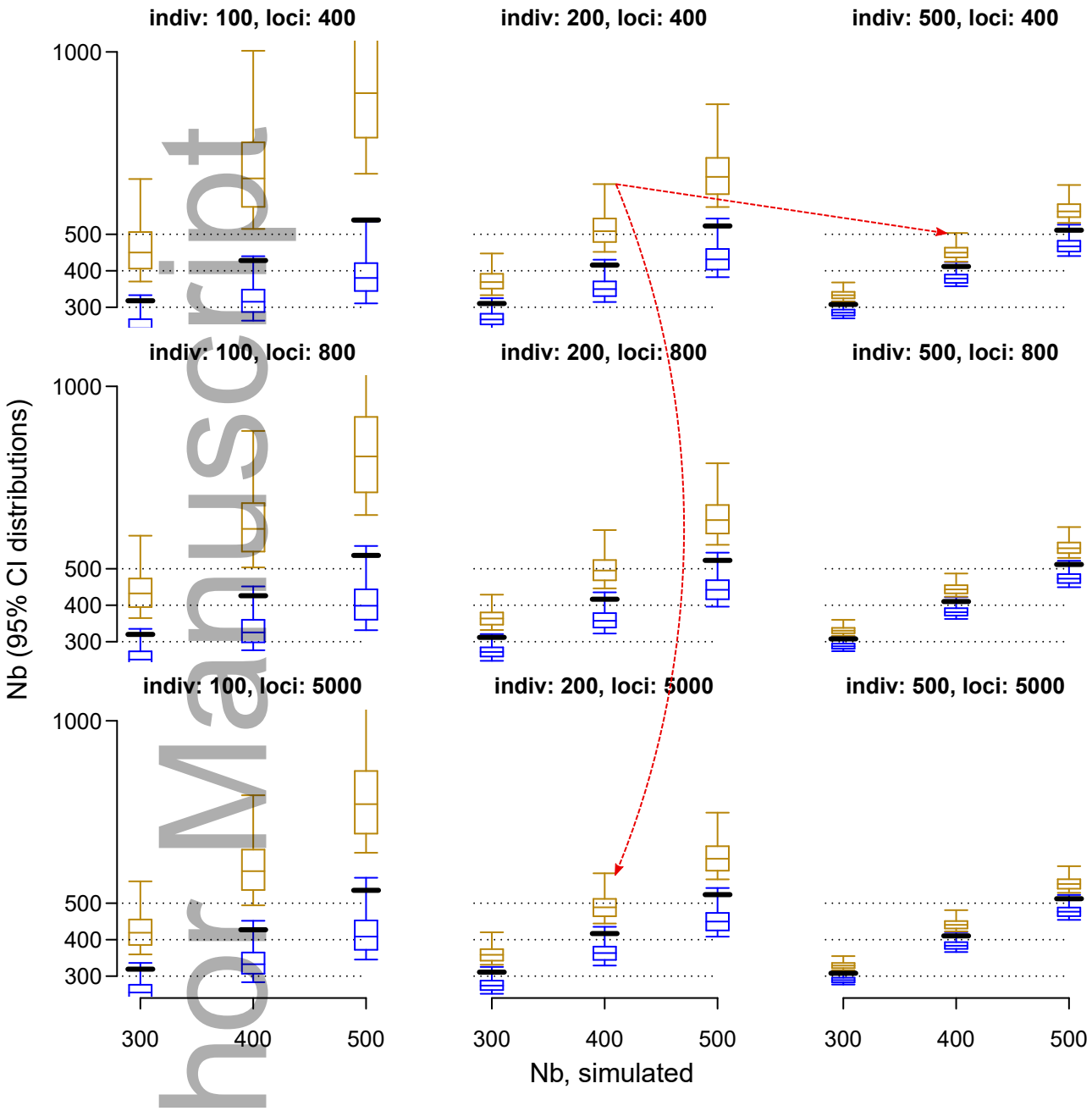


Figure 2.







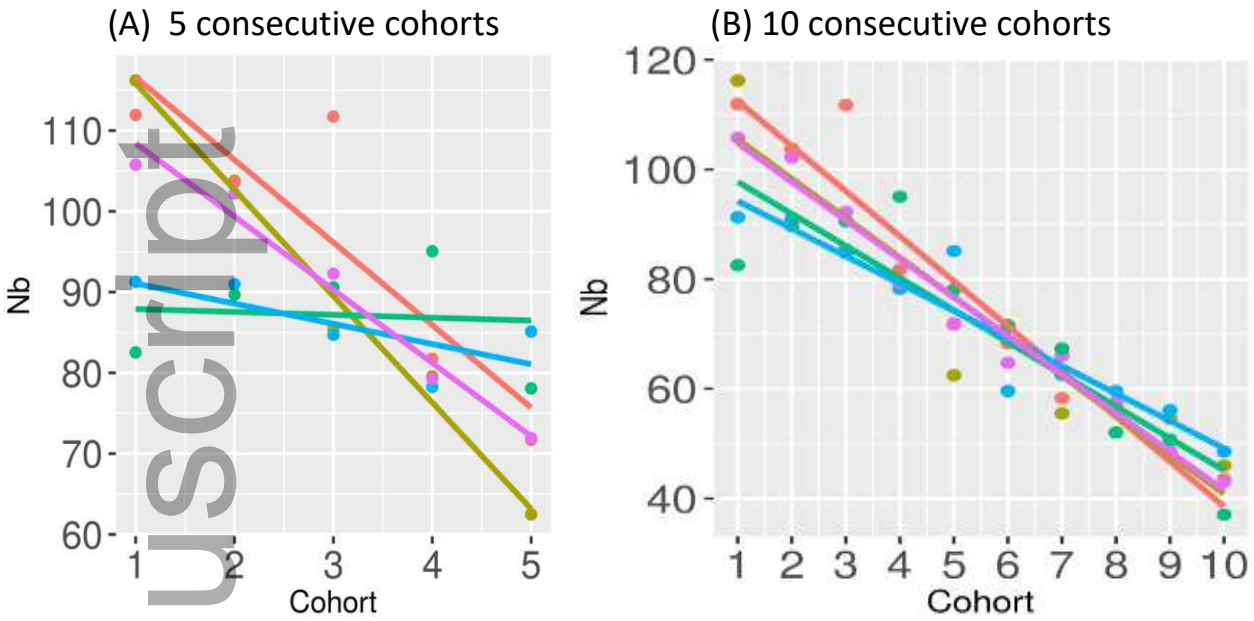


Figure 6

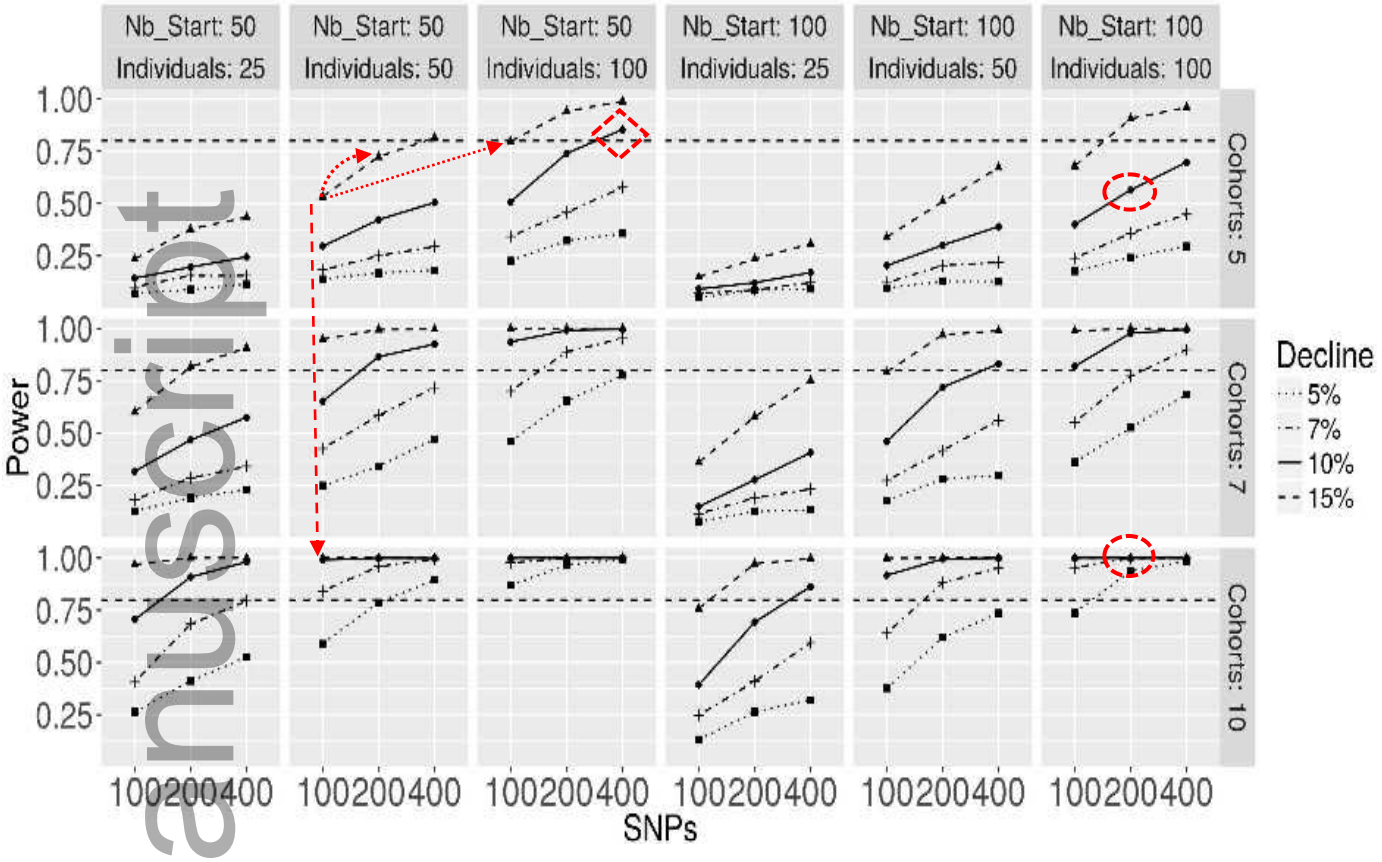


Figure 7