



**Peer Reviewed**

**Title:**

Evaluation and Interpretation of Genetic Effective Population Size of Delta Smelt from 2011–2014

**Journal Issue:**

[San Francisco Estuary and Watershed Science, 15\(2\)](#)

**Author:**

[Finger, Amanda J.](#), University of California, Davis  
[Schumer, Gregg](#), Cramer Fish Sciences—Genidaqs  
[Benjamin, Alyssa](#), Dept. of Animal Science, Genomic Variation Lab, University of California, Davis  
[Blankenship, Scott](#), Cramer Fish Sciences—Genidaqs

**Publication Date:**

2017

**Permalink:**

<http://escholarship.org/uc/item/4109064c>

**Acknowledgements:**

The authors would like to thank Shawn Acuña, Stephanie Fong, Karrigan Bork, Andrea Schreier, and two anonymous reviewers for valuable comments that improved this manuscript. Funding for this work was provided by State and Federal Contractors Water Agency Agreement #201503445, and Agreement #153350 with Metropolitan Water District of Southern California.

**Keywords:**

Delta Smelt, *Hypomesus transpacificus*, effective population size, genetics, conservation, microsatellites

**Local Identifier(s):**

jmie\_sfews\_35763

**Abstract:**

<https://doi.org/10.15447/sfews.2017v15iss2art5>

Delta Smelt have collapsed demographically, but little is known about their current genetic status. We used 12 microsatellite loci to evaluate two measures of the effective population size ( $N_e$ ) of Delta Smelt.  $N_e$  is a measure that offers predictive power regarding the loss of genetic diversity in a population over time, as well as the short and long-term genetic risks for loss of fitness resulting from low diversity. We found that the  $N_e$  of Delta Smelt is too high to accurately estimate with the data (upper 95% confidence intervals were infinity), but the lower confidence intervals of  $N_{eLD}$  (linkage disequilibrium  $N_e$ ) were above 1,000, while some of the lower confidence intervals of  $N_{eV}$



(variance  $N_e$ ) were below 1,000. We interpret this to indicate that Delta Smelt are not declining because of genetic factors, and are not at immediate risk of losing genetic diversity from low  $N_e$ . We caution that these estimates are from a short-term data set estimated from a population that has already been declining for decades, and that it is likely that Delta Smelt have lost diversity. We suggest continuing efforts to maximize abundance to prevent further loss of genetic diversity.

**Supporting material:**

Appendix A: Allele Frequencies

Appendix B: Results with 10 Microsatellite Loci

**Copyright Information:**



Copyright 2017 by the article author(s). This work is made available under the terms of the Creative Commons Attribution 4.0 license, <http://creativecommons.org/licenses/by/4.0/>



**eScholarship**  
University of California

eScholarship provides open access, scholarly publishing services to the University of California and delivers a dynamic research platform to scholars worldwide.

## RESEARCH

# Evaluation and Interpretation of Genetic Effective Population Size of Delta Smelt from 2011–2014

Amanda J. Finger<sup>1\*</sup>, Gregg Schumer<sup>2</sup>, Alyssa Benjamin<sup>1</sup>, and Scott Blankenship<sup>2</sup>

Volume 15, Issue 2 | Article 5

<https://doi.org/10.15447/sfews.2017v15iss2art5>

\* Corresponding author: [ajfinger@ucdavis.edu](mailto:ajfinger@ucdavis.edu)

1 Department of Animal Science, Genomic Variation Lab  
 University of California, Davis  
 Davis, CA 95616 USA

2 Cramer Fish Sciences—Genidags  
 West Sacramento, CA 95691 USA

## ABSTRACT

Delta Smelt have collapsed demographically, but little is known about their current genetic status. We used 12 microsatellite loci to evaluate two measures of the effective population size ( $N_e$ ) of Delta Smelt.  $N_e$  is a measure that offers predictive power regarding the loss of genetic diversity in a population over time, as well as the short and long-term genetic risks for loss of fitness resulting from low diversity. We found that the  $N_e$  of Delta Smelt is too high to accurately estimate with the data (upper 95% confidence intervals were infinity), but the lower confidence intervals of  $N_{eLD}$  (linkage disequilibrium  $N_e$ ) were above 1,000, while some of the lower confidence intervals of  $N_{eV}$  (variance  $N_e$ ) were below 1,000. We interpret this to indicate that Delta Smelt are not declining because of genetic factors, and are not at immediate risk of losing genetic diversity from low  $N_e$ . We caution that these estimates are from a short-term data set estimated from a population that has already been declining for decades, and that it is

likely that Delta Smelt have lost diversity. We suggest continuing efforts to maximize abundance to prevent further loss of genetic diversity.

## KEY WORDS

Delta Smelt, *Hypomesus transpacificus*, effective population size, genetics, conservation, microsatellites

## INTRODUCTION

It is widely recognized that preventing loss of genetic diversity is critical for long-term persistence of a population (e.g., Frankham 2005). Natural selection acts on standing genetic diversity in a population, allowing adaptation to a changing environment. Small populations and populations experiencing demographic collapse or decline are subject to loss of genetic diversity through the processes of genetic drift and inbreeding. Populations that have lost genetic diversity may have also lost evolutionary potential, increasing extinction risk. Indeed, there is a link between reduced genetic diversity and extinction risk (Frankham 1995; Spielman et al. 2004). A measure frequently used to assess the level of risk a population is experiencing from genetic concerns is effective population size ( $N_e$ ), which offers predictive power regarding loss of genetic diversity in a population over time (Crow and Kimura 1970). Thus,  $N_e$  is a valuable measure for managers trying to maintain evolutionary potential by preventing loss of genetic diversity. Below, we provide a brief description of  $N_e$ , its assumptions, how it is

measured, and how estimating  $N_e$  can be applied to the case of the endangered Delta Smelt (*Hypomesus transpacificus*). Our description of  $N_e$  is not meant to be definitive or exhaustive—there is extensive literature and several comprehensive reviews of  $N_e$ . Readers may utilize those resources for a deeper understanding (e.g., Leberg 2005; Luikart et al. 2010; Palstra and Fraser 2012).

### Effective Population Size

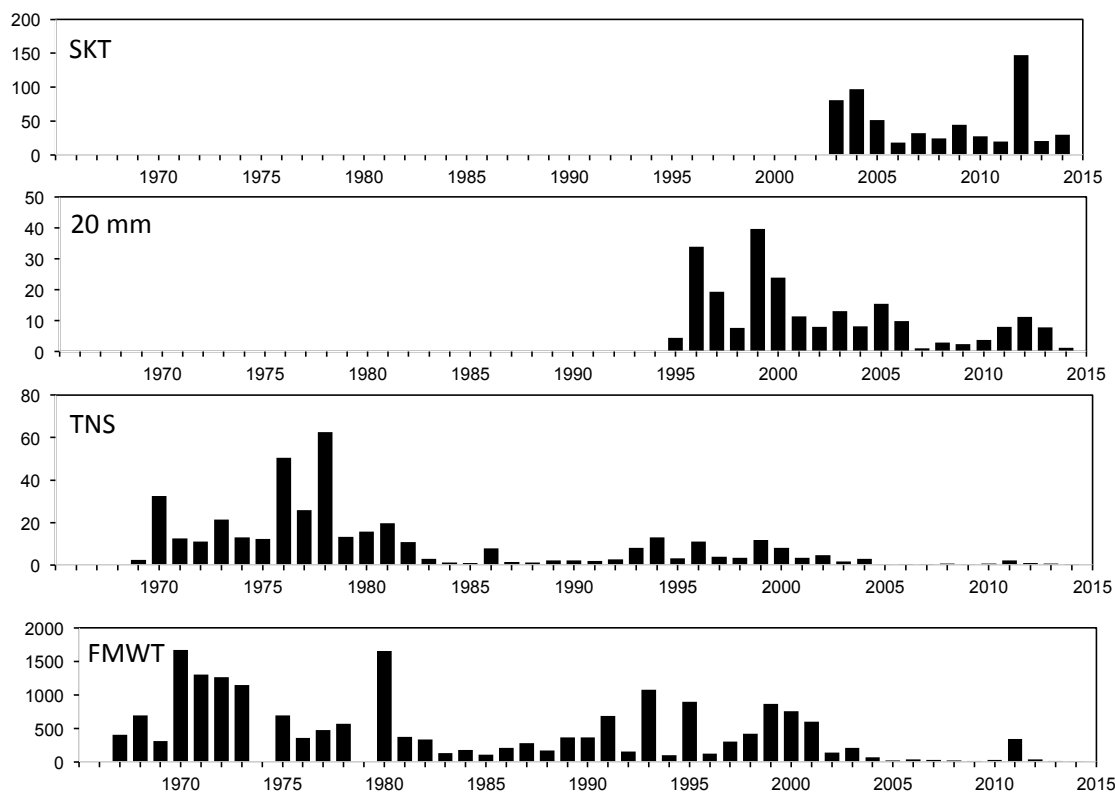
At its most basic,  $N_e$  can be defined as the number of individuals in an ideal population (non-overlapping generations, random mating, and no migration, mutation, or selection) that has the same rate of genetic drift acting on it as the population in question. There are two basic methods for using genotype data to estimate  $N_e$ : temporal methods and single-sample methods. Temporal methods use the mathematical relationship between  $N_e$  and genetic drift to estimate  $N_e$  (Kimura and Crow 1963; Nei and Tajima 1981; Lande and Barrowclough 1987; Crow and Denniston 1988; Waples 1989; Wang 2001; Anderson 2005), and require a population to be sampled twice, ideally with several intervening generations. Temporal methods estimate the variance in effective population size ( $N_{eV}$ ) of the population during the intervening generations between the two sample points (see Waples 1989 and citations therein).

Because it is often difficult to sample a population twice, researchers developed single-sample  $N_e$  estimators (Nomura 2008; Tallmon et al. 2008; Waples and Do 2008; Pudovkin et al. 2009; Wang 2009). The most common single-sample  $N_e$  estimator is the linkage disequilibrium (LD) estimator, which uses non-random correlations between allele frequencies to estimate  $N_e$  (hereafter  $N_{eLD}$ ; see Waples 2005 for a detailed explanation of the LD method). Of note is that  $N_{eLD}$  is estimating the effective number of breeders ( $N_{eb}$ ) of the parents of the individuals that were sampled, somewhat like the  $N_{eV}$  measured when the two samples are parents and offspring. In other words, if a single-sample method is used to estimate the  $N_{eLD}$  of Delta Smelt born in 2015, it is measuring the  $N_{eb}$  of the parents of the 2015 cohort, not the  $N_{eLD}$  of the 2015 cohort itself (Waples 2005). When populations are large and

stable,  $N_{eLD}$  and  $N_{eV}$  may not differ much. However, when a population rapidly declines,  $N_{eV}$  is likely to be reduced more quickly than  $N_{eLD}$  because of the influences of variance in reproductive success (Crow and Morten 1955). Conversely, after a population recovers,  $N_{eV}$  is likely to recover more quickly than  $N_{eLD}$  (Kimura and Crow 1963). When possible, it is recommended that both methods be used.

### Delta Smelt

The Delta Smelt (*Hypomesus transpacificus*) is an excellent candidate for  $N_e$  estimation for assessing long-term genetic risk, because of its annual life cycle and its weak-to-nonexistent population structure. The species is endemic to the San Francisco Estuary (estuary) and lies at the heart of the political conflict that surrounds water deliveries in California. In response to the dramatic decline of Delta Smelt in the 1980s (Sommer et al. 2007; Thomson et al. 2010), abundant research has been devoted to understanding this decline (e.g., Bennett 2005; Nobriga and Herbold 2009; Sommer and Mejia 2013; Sommer et al. 2007). Primary factors in the decline are anthropogenic: habitat alteration resulting from urbanization, water diversions, contaminants, invasive species, and engineering of the Delta for water deliveries (Nichols et al. 1986; Moyle 2008). Information regarding the status of Delta Smelt is not only important for managers trying to make appropriate decisions about how to protect Delta Smelt, but is also politically valuable because it can drive decisions about water deliveries. There are no Delta Smelt  $N_e$  estimates taken from before the decline observed in the 1980s, but Fisch et al. (2011) found that all estimates of  $N_{eLD}$  from every other year between 2003 and 2009 were at or above 969. A re-estimation of  $N_e$  and bottleneck detection in Delta Smelt is warranted based on the long-term decline and the more recent very low abundance indices of Delta Smelt (see [Figure 1](#) for abundance indices reproduced from IEP MAST 2015), as well as the historic 4-year drought in California. The objective of this study is to conduct these analyses on samples from 2011 to 2014 to determine if the low Delta Smelt abundance indices are correlated with detectable genetic effects.



**Figure 1** Abundance indices for Delta Smelt. Source: IEP MAST (2015).

## MATERIALS AND METHODS

### Sample Collection

Between 2011 and 2015, California Department of Fish and Wildlife (CDFW) personnel captured Delta Smelt from the following surveys: Interagency Ecological Program (IEP) Summer Towntnet, Fall Midwater Trawl, Spring Kodiak Trawl, and gear efficiency studies (midwater trawls, Kodiak trawls, townets). These samples were transported to UC Davis and frozen whole in liquid nitrogen. Frozen samples were thawed for an unrelated study, and fins were clipped and placed in 100% ethanol. The UC Davis Fish Conservation and Culture Laboratory (FCCL) personnel collected additional fin clips (placed directly in 100% ethanol) from Delta Smelt captured for the FCCL broodstock. Samples were divided into year classes for analysis using size of individual at sampling and date of capture. Numbers for each year class ranged from 421 to 995 individual Delta Smelt (Table 1).

### DNA Extraction and Microsatellite Genotyping

DNA was extracted and isolated from each Delta Smelt fin clip using the DNEasy extraction kit (Qiagen) following the manufacturer's protocol. We collected genotype data on each individual Delta Smelt at 12 microsatellite loci from Fisch et al. (2009; Appendix A). Polymerase Chain Reaction (PCR) protocols followed Fisch et al. (2009) Pooled PCR product (1.0  $\mu$ l) was added to 9.0  $\mu$ l of highly deionized formamide and 0.2  $\mu$ l of Liz 600. We conducted fragment analysis on an Applied Biosystems (ABI) 3730xl Genetic Analyzer. Two people independently scored individual raw microsatellite genotypes using Geneious v8 software (Kearse et al. 2012). They reconciled all observed discrepancies before the final data export. They identified and removed duplicate genotypes, but retained individuals with genotypes composed of a minimum of 10 loci (83%) for genetic analysis.

**Table 1** Sample size (N), Expected ( $H_e$ ) and observed ( $H_o$ ) heterozygosity, average number of alleles ( $N_A$ ), allelic richness with 676 genes ( $A_r$ ), number of private alleles ( $N_p$ ), and private allelic richness ( $A_p$ ) for four cohorts (years) of Delta Smelt using 12 microsatellite loci

Year (cohort)	N	$H_e$	$H_o$	$N_A$	$A_r$	$N_p$	$A_p$
2011	995	0.853	0.843	25.67	23.57	11	0.53
2012	534	0.852	0.836	24.17	23.27	1	0.38
2013	678	0.847	0.836	24.42	23.15	3	0.24
2014	421	0.853	0.846	23.42	23.09	3	0.38
Mean	—	0.851	0.840	24.42	23.27	4.5	0.38

## Genetic Diversity

MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) was used to check the genotypic data for errors in scoring and identify possible null alleles or stuttering indicated by an excess of heterozygotes. Delta Smelt from each of the four cohorts (2011 to 2014) were analyzed to assess population genetic diversity. Each year was treated as a sample, for a total of four samples. We used GENEPOP 4.2 (Raymond and Rousset 1995) to test for departures from Hardy–Weinberg equilibrium (HWE) and to identify LD between microsatellite loci. For LD, we applied a sequential Bonferroni correction ( $\alpha=0.05$ ) to correct for the increased likelihood of obtaining false positives with multiple tests. Using GenAEx 6.5 (Peakall and Smouse 2006, 2012), we calculated the expected heterozygosity ( $H_e$ ), observed heterozygosity ( $H_o$ ), average number of alleles across loci ( $N_A$ ), number of private alleles ( $N_p$ ), and allelic frequencies. We used the software HP-RARE (Kalinowski 2005) to determine allelic richness ( $A_r$ ) and private allelic richness ( $A_p$ )—genetic diversity measures which use rarefaction to correct for the increased likelihood of detecting rare alleles with larger sample sizes (Kalinowski 2004). To calculate  $A_r$  and  $A_p$ , we used  $N=676$ , the least number of genomic copies at any locus in any year class.

## Population Structure

To rule out spatial population structure for the purposes of estimating  $N_{eLD}$  and  $N_{eV}$  (spatial structure violates an assumption of  $N_e$  estimation), we checked for the presence of any population genetic structure within and between Delta Smelt cohorts using STRUCTURE 2.3.3 (Pritchard et al. 2000), software

that applies a model-based clustering method and a Bayesian approach so underlying genetic structure in a group of individuals can be inferred. We ran three iterations of  $K=1-6$  with a burn-in period of 100,000 and 1,000,000 Markov chain Monte Carlo (MCMC) replications. Using STRUCTURE HARVESTER (Earl and vonHoldt 2012), we determined the optimal  $K$  number of clusters using the highest mean log likelihood, and visualized these clusters in a bar plot produced by STRUCTURE 2.3.3. To check for temporal structure, we used FSTAT 2.9.3.2 (Goudet 1995) to calculate pairwise  $F_{ST}$  values between each year class.  $P$ -values were obtained after 200 permutations and a Bonferroni correction for multiple tests (Rice 1989).

## Population Bottlenecks and Effective Population Size

We used two methods to detect population bottlenecks in each cohort sample: (1) the Wilcoxon signed-rank test for excess heterozygosity ( $H_k$ ; Cornuet and Luikart 1996) implemented in the software BOTTLENECK 1.2.02 (Piry et al. 1999); and (2) the  $M$ -ratio test (Garza and Williamson 2001) implemented in the software  $M\_P\_Val$  (<http://swfsc.noaa.gov/textblock.aspx?Division=FED&id=3298>). First, the  $H_k$  test calculates the likelihood of more recent bottlenecks relative to the  $M$ -ratio test (between 0.8 and 4.0  $N_e$  generations ago), detecting a bottleneck when the heterozygosity expected under HWE exceeds the heterozygosity expected under mutation drift equilibrium. We performed the Wilcoxon one-tailed test for heterozygosity excess, using 5,000 replications and applying two mutational models: the stepwise mutation model (SMM) and the two-phase model (TPM). First, we used the

parameters recommended by Piry et al. (1999; 12% variance, 95% stepwise mutations and 5% non-stepwise mutations) to calculate the likelihood of bottlenecks under the TPM. Second, we applied the M-ratio test to identify the likelihood of more severe historical bottlenecks. For this test, we determined the M-ratio—the ratio of the number of microsatellite alleles (K) to the range in the allele size for each locus ( $r$ ;  $M = K/r$ )—using parameters recommended by Garza and Williamson (2001): proportion of one-step mutations ( $p_s$ )=0.9, average size of non-one-step mutations ( $\Delta_g$ )=3.5, and  $\theta = 10$ . In the event of a bottleneck, the M-ratio is expected to decrease because the number of alleles lost as a result of drift decreases more quickly than the decline in the range of alleles.

Using the software program NEESTIMATOR Version 2.01 (Do et al. 2014), we estimated  $N_e$  using three different methods. First, we first estimated  $N_{eLD}$  of each sample with the bias-corrected LD method (Waples and Do 2008).  $N_{eLD}$  was estimated with alleles below three different frequencies removed ( $P_{crit}$  values): those at 0.05, 0.02, and 0.01. Second, utilizing the temporal method in Waples (1989) with a  $P_{crit}$  value of 0.02, we used NEESTIMATOR to estimate  $N_{eV}$ . Third, we used two different methods for computing variance in allele frequency to calculate parametric confidence intervals (Pollak 1983; Nei and Tajima 1981). For this analysis, 2011 was considered generation 0, followed by 2012 as generation 1, and so forth through 2014.

## RESULTS

We detected possible null alleles using Microchecker at Htr115 and Htr126 in all Delta Smelt years; null alleles were also present for Htr116 in 2011 and Htr119 in 2012 and 2013. We detected significant deviation from HWE ( $P < 0.05$ ) in all years, yet loci out of HWE were inconsistent across years. We performed 330 tests for linkage disequilibrium, of which 24 were significant ( $P < 0.05$ ). After a sequential Bonferroni correction ( $\alpha = 0.05$ ), only seven locus pairs remained significant: Htr104/Htr109 in 2011, Htr103/Htr109 and Htr117/Htr120 in 2012, Htr103/Htr109 and Htr120/127 in 2013, and Htr109/Htr114 and Htr114/Htr120 in 2014. Given the detected null alleles, we conducted downstream analyses (1) using the full suite of 12 loci, and (2)

excluding Htr115 and Htr126. Because results with 10 loci did not differ dramatically from results with 12, analyses with 10 loci are placed in Appendix A.

## Genetic Diversity

We observed similar levels of genetic diversity among the Delta Smelt samples from different years. Expected heterozygosity ( $H_e$ ) ranged from 0.847 to 0.853, with an average  $H_e$  of 0.851; observed heterozygosity ( $H_o$ ) ranged from 0.836 to 0.846, with an average  $H_o$  of 0.840 (Table 1).  $A_r$  ranged from 23.09 (2014) to 23.57 (2011; Table 1).  $A_p$  values were highest in 2011 (0.53), and lowest in 2013 (0.24; Table 1). See Appendix A for a list of allelic frequencies,  $A_r$ , and  $A_p$  values.

## Population Structure

The STRUCTURE analysis revealed no population structure within or between years (highest mean log likelihood value was for  $K = 1$ ). Pairwise  $F_{ST}$  values were all  $< 0.001$ , and none were significant after a Bonferroni correction (adjusted  $\alpha$  after the Bonferroni correction = 0.0083) indicating no temporal or spatial structure within the data set.

## Bottlenecks and Effective Population Size

We found no significant heterozygote excess for any year using the  $H_k$  test. Similarly, there were no M-ratio values below  $P_{crit}$ ; therefore, we found no evidence of genetic bottlenecks in our data set. Ninety-five% confidence intervals (CIs) were wide for both  $N_{eLD}$  and  $N_{eV}$  (upper limit =  $\infty$ ) for all 4 years, indicating that  $N_e$  is high enough for the estimators to have low precision. The lower limits of CIs for  $N_{eLD}$  ranged from 1,893 to 23,822 across each sample year (Table 2). For  $N_{eV}$ , lower limit 95% CIs ranged from 632 to 6,619 (Table 3).

## DISCUSSION

We estimated  $N_e$  and measured the genetic diversity of four cohorts of Delta Smelt (2011, 2012, 2013, 2014) using 12 microsatellites. The bottleneck estimators did not detect bottlenecks within the data set, and we found no temporal or geographic population genetic structure. We found a small

**Table 2**  $N_{eLD}$  estimated with 12 loci. We estimated  $N_{eLD}$  with alleles at three different frequencies removed ( $P_{Crit} = 0.05, 0.02, 0.01$ ). 95% confidence intervals (CIs) using a jackknifing method reported.

Year	$N_{eLD}$ (95% CI)		
	0.05	0.02	0.01
2011	33,120 (4,854-∞)	∞ (12,957-∞)	∞ (23,822-∞)
2012	4,201 (1,893-∞)	8,955 (3,704-∞)	11,549 (4,798-∞)
2013	56,715 (3,177-∞)	14,304 (4,534-∞)	12,316 (4,030-∞)
2014	∞ (3,332-∞)	∞ (16,409-∞)	∞ (7,543-∞)

reduction in both allelic richness and private allelic richness between 2011 and 2014. Estimates and 95% CIs for both  $N_{eV}$  and  $N_{eLD}$  varied substantially; therefore, we focus on the lower CIs in this discussion, because these values are the most conservative minimum estimates of Delta Smelt  $N_e$ . Current research indicates that a minimum  $N_e$  of at least 1,000 is theoretically necessary for long-term maintenance of genetic diversity (Frankham et al. 2014). Yet, this is an area of active research and debate (Jamieson and Allendorf 2012, 2013; Frankham et al. 2013, 2014), and most research has focused on populations that are already small. Our lower 95% CIs of  $N_{eV}$  are sometimes below the recommended conservation threshold of 1,000, but the  $N_{eLD}$  lower CIs are all > 1,000. In this case, the lower 95% CIs indicate that, at present, there is little evidence to suggest that Delta Smelt have an increased short-term risk of extinction from genetic factors. Yet, at this time, we cannot rule out the possibility that the true  $N_e$  is close to the threshold where genetic risks are increased, nor can we determine if  $N_e$  has already been reduced from historical levels, though given demographic patterns, it is likely. Yet, it does not appear that their decline results from genetic factors, but rather the myriad factors already described in the literature (Moyle et al. 2016 and citations therein).

An understanding of the limitations of  $N_e$  estimation is critical for interpretation and use of our analysis in a management context. In this case, precise estimates of  $N_e$  are not possible with our data set. In part, we can interpret this as Delta Smelt having a relatively large  $N_e$  (probably > 1,000). Current  $N_e$  estimators perform better when  $N_e$  is smaller (< 200;

**Table 3**  $N_{eV}$  ( $P_{Crit} = 0.02$ ) reported for different year (cohort) combinations. Below diagonal is  $N_{eV}$  with 95% parametric CI from Pollack (1993), above diagonal is  $N_{eV}$  and 95% parametric CI from Nei and Tajima (1981).

Year (generation)	2011 (0)	2012 (1)	2013 (2)	2014 (3)
2011 (0)	—	2,363 (775-∞)	∞ (6,425-∞)	∞ (3,471-∞)
2012 (1)	2,088 (735-∞)	—	3,454 (789-∞)	6,302 (1,284-∞)
2013 (2)	∞ (6,619-∞)	3,259 (775-∞)	—	2,459 (634-∞)
2014 (3)	85,239 (3,395-∞)	4,777 (1,202-∞)	2,340 (632-∞)	—

Waples and Do 2010). Complicating this analysis, the number of markers used here is on the lower end of the recommended number of microsatellites for  $N_e$  estimation (10 to 20; Waples and Do 2010). Furthermore, our samples are from 2011 to 2014—the estimates herein are short-term estimates, using recent Delta Smelt samples, taken from a population that has been declining for decades (which may also explain why bottlenecks were not detected despite observations of population decline). We did not combine our data set with that of Fisch et al. (2011), because they used 16 rather than 12 markers, and microsatellites were not standardized between instruments. Though we cannot directly compare our results, Fisch et al. (2011) had similar findings (however they did not use the sample size correction, causing their  $N_e$  estimations to be upwardly biased). They found lower 95% CIs of  $N_{eLD}$  to be 969 or greater, and an estimate of  $N_{eV}$  to be 1,430. The difference between  $N_{eLD}$  and  $N_{eV}$  values is important: when a population rapidly decreases, genetic drift causes more random changes in allele frequencies (and sometimes loss of alleles), causing  $N_{eV}$  to decrease.  $N_{eLD}$  on the other hand, measures an increase in inbreeding, which occurs more slowly than changes in allele frequencies resulting from genetic drift. Therefore, during or after a bottleneck,  $N_{eLD}$  is likely to remain high for several generations, even when  $N_{eV}$  is decreasing. Put another way,  $N_{eLD}$  is measuring the  $N_e$  over many generations, while  $N_{eV}$  is a more contemporary estimate (Turner et al. 2002; Neigel 1996; Schwartz et al. 1999). In this case, both  $N_{eLD}$  and  $N_{eV}$  estimates and the 95% CIs vary considerably. We suggest monitoring values over time to identify longer-term trends, rather than taking an individual value in isolation. Though



we find no evidence of short-term genetic risks based on our  $N_e$  estimates, there may be increased long-term genetic risk. The fact remains that the Delta Smelt abundance indices show an alarming collapse (Figure 1). From a genetic perspective, this is of concern because the relationship between average number of alleles, census size ( $N_c$ ), and  $N_e$  is different for large and small populations at their equilibriums. For example, Ryman et al. (1995) describe a scenario where there are two equilibrium populations: A and B, with census sizes of 100,000 and 10,000,000 (respectively) that are both reduced by 99%. The difference in loss of heterozygosity may be negligible between A and B, but population A will retain 90% of its alleles, while B will only retain 8%. This is because large populations retain far more alleles than small populations at their equilibrium states, based on the relationship between mutation (generates new alleles) and drift (removes alleles; see Crow and Kimura 1970). Additionally, population genetic theory predicts that declining populations will inevitably lose genetic diversity through genetic drift and inbreeding. This scenario could lead to an “extinction vortex,” a self-reinforcing iterative process where bottlenecks and lowered  $N_e$  lead to increased genetic drift, reduced heterozygosity, and increased inbreeding, which can reduce adaptive potential (Gilpin and Soulé 1986; Caughley 1994). We can expect that if the decline of Delta Smelt continues, genetic factors will increase an already high threat of extinction (Quiñones and Moyle 2014).

## CONCLUSIONS AND RECOMMENDATIONS

Our findings are important and encouraging, and suggest that the declining abundance of Delta Smelt has not led definitively to reduced maintenance of long-term evolutionary potential – if the population is recovered. Our results do not indicate that Delta Smelt are not suffering from serious declines over decades or aren't imperiled. Maintaining an  $N_e > 1,000$  may require a census population of over 100,000 based on published  $N_e/N_c$  ratios, which typically range from 0.1 to 0.5, but can be as low as 10-5 for marine species (Palstra and Ruzzante 2008; Hare et al. 2011). Management efforts should continue to focus on improving abundance rather than targeting genetic diversity. In addition, more work is needed to estimate abundance levels, an

increasingly challenging endeavor as Delta Smelt continue to decline and are harder to find. Our analyses provide additional useful insight into the status of this population, but we recommend continued estimation of  $N_e$  to monitor short-term extinction risk with utilization of newer genomic technologies, historical samples, and integration with current survey and monitoring efforts. Genomic technologies in particular are promising, because genomic data are more powerful than microsatellites for giving precise estimates of  $N_e$ . For example, Hoban et al. (2014) found that 2,500 single-nucleotide polymorphisms (SNPs) are as powerful as 250 microsatellites, and could detect genetic erosion even while 80–90% of a population's historical diversity remained. In addition, genomic data are easier to standardize across studies (making long-term data sets more feasible), and are increasingly more efficient and affordable to collect. Monitoring genetic diversity in threatened populations is widely recognized as important for assessing extinction risk and long-term evolutionary potential, and by extension, potential for recovery (Frankham 2005). Though estimated  $N_e$  values are largely  $>1,000$ , Delta Smelt are an annual species that inhabit an extremely variable ecosystem. This variability combined with typical population boom-and-bust cycles could drastically reduce  $N_e$  very quickly. The estimated  $N_e$  values in the face of overwhelming evidence of population collapse underscore the difficulty in fully understanding basic aspects of Delta Smelt life history, their conservation status, and the complex nature of the estuary. Yet, they also provide hope for recovery and the resilience of the species.

## ACKNOWLEDGEMENTS

The authors would like to thank Shawn Acuña, Stephanie Fong, Karrigan Bork, Andrea Schreier, and two anonymous reviewers for valuable comments that improved this manuscript. Funding for this work was provided by State and Federal Contractors Water Agency Agreement #201503445, and Agreement #153350 with Metropolitan Water District of Southern California.

## REFERENCES

- Anderson EC. 2005. An efficient Monte Carlo method for estimating  $N_e$  from temporally spaced samples using a coalescent based likelihood. *Genetics* 170:955–967. <https://doi.org/10.1534/genetics.104.038349>
- Bennett WA. 2005. Critical assessment of the Delta Smelt population in the San Francisco Estuary, California. *San Franc Estuary Watershed Sci* 3(2). <https://doi.org/10.15447/sfews.2005v3iss2art1>
- Caughley G. 1994. Directions in conservation biology. *J Anim Ecol* 63(2):215–244. <https://doi.org/10.2307/5542>
- Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001–2014. <http://www.genetics.org/content/144/4/2001>
- Crow JF, Denniston C. 1988. Inbreeding and variance effective population numbers. *Evolution* 42(3):482–495. <http://www.jstor.org/stable/2409033>
- Crow JF, Morton NE. 1955. Measurement of gene frequency drift in small populations. *Evolution* 9:202–214.
- Crow JF, Kimura M. 1970. An introduction to population genetics theory New York (NY): Harper and Row. 591 p.
- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovendon JR. 2014. NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective size ( $N_e$ ) from genetic data. *Mol Ecol Resour* 14:209–214. <https://doi.org/10.1111/1755-0998.12157>
- Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Fisch KM, Petersen JL, Baerwald MR, Pedroia JK, May B. 2009. Characterization of 24 microsatellite loci in Delta Smelt, *Hypomesus transpacificus*, and their cross-species amplification in two other smelt species of the Osmeridae family. *Mol Ecol Resour* 9(1):405–408. <https://doi.org/10.1111/j.1755-0998.2008.02254.x>
- Fisch KM, Henderson JM, Burton RS, May B. 2011. Population genetics and conservation implications for the endangered delta smelt in the San Francisco Bay-Delta. *Conserv Genet* 12(6): 1421–1434. <https://doi.org/10.1007/s10592-011-0240-y>
- Frankham R. 1995. Inbreeding and extinction: a threshold effect. *Conserv Biol* 9(4):792–299. <https://doi.org/10.1046/j.1523-1739.1995.09040792.x>
- Frankham R. 2005. Genetics and extinction. *Biol Conserv* 126(2005):131–140. <https://doi.org/10.1016/j.biocon.2005.05.002>
- Frankham R, Brook BW, Bradshaw CJA, Traill LW, Spielman D. 2013. 50/500 rule and minimum viable populations: response to Kalinowsk and Allendorf. *Trends Ecol Evol* 28(4):187–188. <https://doi.org/10.1016/j.tree.2013.01.002>
- 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. *Biol Conserv* 170: 56–63. <https://doi.org/10.1016/j.biocon.2013.12.036>
- Garza JC, Williamson EG. 2001. Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* 10:305–318. <https://doi.org/10.1046/j.1365-294X.2001.01190.x>
- Gilpin, ME, Soule ME. 1986. Minimum viable populations: the processes of population extinction. In: Soule ME, editor. *Conservation biology: the science of scarcity and diversity* Sunderland (MA): Sinauer Associates. p. 13–34.
- Goudet J. 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* 86(6):485–486.
- Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, Ruegg K, Palstra F. 2011. Understanding and estimating effective population size for practical application in marine species management. *Cons. Biol.* 25(3):438–449. <https://doi.org/10.1111/j.1523-1739.2010.01637.x>
- Hoban S, Arntzen JA, Bruford MW, Godoy JA, Hoelzel AR, Segelbacher G, Vila C, Bertorelle G. 2014. Comparative evaluation of potential indicators and temporal sampling protocols for monitoring genetic erosion. *Evol Appl* 7(9):984–998. <https://doi.org/10.1111/eva.12197>

- [IEP MAST] Interagency Ecological Program Management, Analysis, and Synthesis Team. 2015. An updated conceptual model for Delta Smelt: our evolving understanding of an estuarine fish [Internet]. [accessed 2016 Oct 05]. Sacramento (CA): California Department of Water Resources. Available from: [http://www.water.ca.gov/iep/docs/Delta\\_Smelt\\_MAST\\_Synthesis\\_Report\\_January%202015.pdf](http://www.water.ca.gov/iep/docs/Delta_Smelt_MAST_Synthesis_Report_January%202015.pdf)
- Jamieson IG, Allendorf FW. 2012. How does the 50/500 rules apply to MVPs? *Trends Ecol Evol* 27:578-584. <https://doi.org/10.1016/j.tree.2012.07.001>
- Jamieson IG, Allendorf FW. 2013. A school of red herring: reply to Frankham et al. *Trends Ecol Evol* 28(4):188-189. <https://doi.org/10.1016/j.tree.2013.01.012>
- Kalinowski ST. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conserv Genet* 5:539-543. <https://doi.org/10.1023/B:COGE.0000041021.91777.1a>
- Kalinowski ST. 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic diversity. *Mol Ecol Notes* 5:187-189. <https://doi.org/10.1111/j.1471-8286.2004.00845.x>
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12):1647-1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Kimura M, Crow J. 1963. The measurement of effective population number. *Evolution* 17:279-288. <https://doi.org/10.2307/2406157>
- Lande R, Barrowclough GF. 1987. Effective population size, genetic variation, and their use in population management. In: Soule M, editor. *Viable populations for conservation*. Cambridge (UK): Cambridge University Press. p. 87-124.
- Leberg P. 2005. Genetic approaches for estimating the effective size of populations. *J Wildl Manag* 69:1385-1399. [https://doi.org/10.2193/0022-541X\(2005\)69\[1385:GAFETE\]2.0.CO;2](https://doi.org/10.2193/0022-541X(2005)69[1385:GAFETE]2.0.CO;2)
- Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv Genet* 11:355-373. <https://doi.org/10.1007/s10592-010-0050-7>
- Moyle PB. 2008. The future of fish in response to large-scale change in the San Francisco Estuary, California. *Am Fish Soc Symp* 64. Available from: <https://watershed.ucdavis.edu/pdf/Moyle-AFSDelta-2008.pdf>
- Moyle PB, Brown LR, Durand JR, Hobbs JA. 2016. Delta Smelt: life history and decline of a once-abundant species in the San Francisco Estuary. *San Franc Estuary Watershed Sci* 14(2). <https://doi.org/10.15447/sfews.2016v14iss2art6>
- Nei M, Tajima F. 1981. Genetic drift and estimation of effective population size. *Genetics* 98:625-640. Available from: <http://www.genetics.org/content/98/3/625>
- Neigel JE, 1996. Estimation of effective population size and migration parameters from genetic data. In: Smith, TB, Wayne RK. *Molecular genetic approaches in conservation*. Oxford (UK): Oxford Press. p. 329-346. Available from: [https://books.google.com/books?hl=en&lr=&id=QI8sdCHszYOC&oi=fnd&tpg=PA329&tdq=Neigel+JE,+1996.+Estimation+of+effective+population+size+and+migration+parameters+from+genetic+data&ots=Dj7GT0pCqu&sig=KlpqwwXUg\\_\\_8kWrXaQe88nkmpCM#v=onepage&q&f=false](https://books.google.com/books?hl=en&lr=&id=QI8sdCHszYOC&oi=fnd&tpg=PA329&tdq=Neigel+JE,+1996.+Estimation+of+effective+population+size+and+migration+parameters+from+genetic+data&ots=Dj7GT0pCqu&sig=KlpqwwXUg__8kWrXaQe88nkmpCM#v=onepage&q&f=false)
- Nichols FH, Cloern JE, Luoma SN, Peterson DH. 1986. The modification of an estuary. *Science* 231:567-573. Available from: <http://online.sfsu.edu/models/Files/References/Nichols1986Science.pdf>
- Nobriga ML, Herbold B. 2009. The little fish in California's water supply: a literature review and life-history conceptual model for Delta Smelt (*Hypomesus transpacificus*) for the Delta Smelt Regional Ecosystems Restoration and Implementation Plan (DRERIP). [Internet]. [accessed 2016 Feb 04]. Available from: <https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=28420>
- Nomura T. 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evol Appl* 1:462-474. <https://doi.org/10.1111/j.1752-4571.2008.00015.x>

- Palstra FP, Ruzzante DE. 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol Ecol* 17:3428-3447. <https://doi.org/10.1111/j.1365-294X.2008.03842.x>
- Palstra FP, Fraser DJ. 2012. Effective/census population size ratio estimation: a compendium and appraisal. *Ecol Evol* 2(9):2357-2365. <https://doi.org/10.1002/ece3.329>
- Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288-295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537-2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Piry S, Luikart G, Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Heredity* 90:502-503. <https://doi.org/10.1093/jhered/90.4.502>
- Pollak E. 1983. A new method for estimating the effective population size from allele frequency changes. *Genetics* 104(3):531-548.
- Pritchard JK, Stephens M, Donnelly PJ. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959. Available from: <http://www.genetics.org/content/155/2/945.short>
- Pudovkin AI, Zhdanova OL, Hedgecock D. 2009. Sampling properties of the heterozygote-excess estimator of the effective number of breeders. *Conserv Gen* 11: 759-771. <https://doi.org/10.1007/s10592-009-9865-5>
- Quinones RM, Moyle PB. 2014. Climate change vulnerability of freshwater fishes of the San Francisco Bay Estuary. *San Franc Estuary Watershed Sci* 12(3). <http://doi.org/10.15447/sfew.s.2014v12iss3art3>
- Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 86:248-249.
- Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* 43(1):223-225.
- Ryman N, Utter F, Laikre L. 1995. Protection of intraspecific biodiversity of exploited fishes. *Rev Fish Biol Fisher* 5:417-446. <https://doi.org/10.1007/BF01103814>
- Schwartz MK, Tallmon DA, Luikart GH. 1999. DNA-based  $N_e$  estimation: many markers, much potential, uncertain utility. *Anim Conserv* 2:320-322.
- Sommer T, Mejia F. 2013. A place to call home: a synthesis of Delta Smelt habitat in the upper San Francisco Estuary. *San Franc Estuary Watershed Sci* 11(2). <https://doi.org/10.15447/sfew.s.2013v11iss2art4>
- Sommer T, Armore C, Baxter R, Breur R, Brown L, Chotkowski M, Culberson S, Feyrer F, Gingras M, Herbold B, Kimmerer W, Mueller-Solger A, Nobriga M, Souza K. 2007. The collapse of pelagic fishes in the upper San Francisco Estuary. *Fisheries* 32:270-277. [https://doi.org/10.1577/1548-8446\(2007\)32\[270:TCOPFI\]2.0.CO;2](https://doi.org/10.1577/1548-8446(2007)32[270:TCOPFI]2.0.CO;2)
- Spielman D, Brook K, Frankham R. 2004. Most species are not driving to extinction before genetic factors impact them. *PNAS* 101(42):15261-15264. Available from: <http://www.pnas.org/content/101/42/15261.full.pdf>
- Tallmon DA, Koyuk A, Luikart G, Beaumont MA. 2008. ONESAMP: a program to estimate effective population size using approximate Bayesian computation. *Mol Ecol Resour* 8:299-301. <https://doi.org/10.1111/j.1471-8286.2007.01997.x>
- Thomson JR, Kimmerer WJ, Brown LR, Newman KB, Nally RM, Bennett WA, Feyrer F, Fleishman E. 2010. Bayesian change point analysis of abundance trends for pelagic fishes in the upper San Francisco Estuary. *Ecol Appl* 20:1431-1448. <https://doi.org/10.1890/09-0998.1>
- Turner TF, Wares JP, Gold JR. 2002. Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* 162:1329-1339. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1462333/>
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotype errors in microsatellite data. *Mol Ecol Notes* 4:535-538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>

- Waples RS. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* 121:379–391. Available from: <http://www.genetics.org/content/121/2/379.short>
- Waples RS, 2005. Genetic estimates of contemporary effective population size: To what time periods do the estimates apply? *Mol Ecol* 14:3335–3352. <https://doi.org/10.1111/j.1365-294X.2005.02673.x>
- Waples RS, Do C. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* 8:753–756. <https://doi.org/10.1111/j.1755-0998.2007.02061.x>
- 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(3):244–262. <https://doi.org/10.1111/j.1752-4571.2009.00104.x>
- Wang J. 2001. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genet Res* 78:243–257. <https://doi.org/10.1017/S0016672301005286>
- Wang J. 2009. A new method for estimating effective population size from a single sample of multilocus genotypes. *Mol Ecol* 18:2148–2164. <https://doi.org/10.1111/j.1365-294X.2009.04175.x>