Using Multistage Design-Based Methods to Construct Abundance Indices and Uncertainty Measures for Delta Smelt

Article *in* Transactions of the American Fisheries Society · April 2019

DOI: 10.1002/tafs.10166

3 authors:

Leo Polansky
U.S. Fish and Wildlife Service

Lara Mitchell

Ken Newman
Biomathematics and Statistics Scotland; University of Edinburgh

Transactions of the American Fisheries Society 148:710–724, 2019

ISSN: 0002-8487 print / 1548-8659 online
DOI: 10.1002/tafs.10166

ARTICLE

Using Multistage Design-Based Methods to Construct Abundance Indices and Uncertainty Measures for Delta Smelt

Leo Polansky*
U.S. Fish and Wildlife Service, Bay–Delta Field Office, Sacramento, California 95814, USA

Lara Mitchell
U.S. Fish and Wildlife Service, Lodi Fish and Wildlife Office, Lodi, California 95240, USA

Ken B. Newman
U.S. Fish and Wildlife Service, Lodi Fish and Wildlife Office, Lodi, California 95240, USA; and Biomathematics and Statistics Scotland and School of Mathematics, University of Edinburgh, Edinburgh, UK

Abstract
Population abundance indices and estimates of uncertainty are starting points for many scientific endeavors. However, if the indices are based on data collected by different monitoring programs with possibly different sampling procedures and efficiencies, applying consistent methodology for calculating them can be complicated. Ideally, the methodology will provide indices and associated measures of uncertainty that account for the sample design, the level of sampling effort (e.g., sample size), and the capture or detection probabilities. We develop and demonstrate consistent methodology for multiple monitoring programs that sample different life stages of Delta Smelt *Hypomesus transpacificus*, a critically endangered fish species endemic to the San Francisco Estuary, whose abundance indices have been at the center of much controversy given the regulatory consequences of their listed status. Current indices use different and incomparable methods, do not account for gear selectivity, and do not provide measures of uncertainty. Using recently available information on gear-specific, length-based conditional probabilities of capture given availability, we develop new abundance indices along with measures of uncertainty by means of a single methodological approach. These new indices are highly correlated with existing ones, but the approach taken here illuminates different sources of bias and quantifies between-year variation using probabilistic statements where the previous indices cannot. Decomposition of uncertainty into its constituent sources reveals that early life stage uncertainty is dominated by gear inefficiency while later life stage uncertainty is dominated by sample size, thus providing guidance for improvements to existing surveys. An additional result of general methodological interest is a demonstration, via simulation intended to reflect realistic data properties, that a lognormal distribution is preferable to the normal distribution for making probabilistic statements about the indices. The work here facilitates the fitting of models attempting to identify factors associated with the dynamics and decline of the species.

Quantitative measures of life stage–specific fish species abundance over time are important starting points for understanding life history, assessing species' status, and doing population modeling (e.g., stock synthesis). Fish monitoring programs provide the data for constructing such measures, referred to here as abundance indices. In status assessment and population modeling, abundance indices are used to identify relative or absolute abundance trends and drivers of population dynamics.

There are many approaches to deriving abundance indices, including design-based statistical approaches (Thompson 2002), model-assisted design-based approaches

*Corresponding author: leo_polansky@fws.gov
Received December 7, 2018; accepted March 28, 2019

(Maunder and Punt 2004), and model-based approaches, e.g., geospatial models (Thorson et al. 2015). Fundamentally, these approaches differ in their assumptions about the sources of variability in the data (Gregoire 1998). The approach taken for a given species depends on the species' biology, survey methodology, the methods of analysis that can be employed given data limitations, and expectations about the applications of the resulting indices. Although model-based approaches can accommodate greater spatial variation in density between sites than design-based approaches can (Thorson et al. 2015), design-based approaches are often simpler, make fewer assumptions, can be constructed when the data cannot support the estimation of complex models, and can still be modified to account for processes such as gear selectivity (Newman 2008). These are the reasons that motivated our choice of a design-based method here.

Regardless of the method used to calculate abundance indices, the associated measurements of uncertainty are essential. First, they are necessary for determining whether apparent changes in abundance are significant according to some statistical criteria. Extending this concept to sampling, abundance indices can be described as true abundance multiplied by some bias factor plus additional sampling noise (Hilborn and Mangel 1997:60), and biologically implausible changes in abundance indices can point to changes in the bias parameter. Finally, measures of uncertainty can facilitate the fitting of population dynamics models to identify factors that impact a population's vital rates (Knape et al. 2013; Newman et al. 2014).

A species currently lacking indices with uncertainty measures is Delta Smelt *Hypomesus transpacificus*, a small-bodied (adults are 50–90 mm FL) osmerid endemic to the upper "Delta" portion of the San Francisco Estuary (Moyle and Herbold 1992). The Delta Smelt is a near-annual species in this area: spawning occurs in late winter and early spring, and individuals in the resulting cohort develop through several intermediate life stages before maturing into the spawning life stage by the subsequent winter (Bennett 2005). Delta Smelt monitoring has been ongoing since the late 1950s, although not until the mid-1990s were surveys specifically designed for Delta Smelt regularly deployed.

Abundance indices from the 1980s and early 1990s, including two California Department of Fish and Wildlife (CDFW) long-term fish monitoring programs—the Summer Tow Net (STN) and Fall Midwater Trawl (FMWT) surveys—indicated a precipitous decline during this time period (Moyle and Herbold 1992). In 1993, both the U.S. Fish and Wildlife Service (USFWS) (pursuant to the Endangered Species Act of 1973) and the state of California (under the California Endangered Species Act) listed the species as threatened (USFWS 1993; CDFW 2010). The Delta Smelt is currently one of the highest-profile endangered fishes in the United States because its habitat coincides with a water supply that supports approximately 8 percent of the country's population and a large agricultural economy, resulting in major resource conflicts between environmental and human needs (Delta Stewardship Council 2018). Despite these listings and the issuance in 2008 of a biological opinion by the USFWS to mitigate the impacts of water operations, Delta Smelt abundance indices indicate that the population has continued to decline (Moyle et al. 2016; Polansky et al. 2018). In 2010, the state of California uplisted the species' status to endangered (CDFW 2010) under the California Endangered Species Act and the USFWS warranted the uplisting; it remains critically endangered according to the International Union for Conservation of Nature (NatureServe 2014).

The Delta Smelt abundance indices most frequently used for assessing trends and conducting population modeling have been derived by CDFW and use the STN and FMWT surveys along with two other surveys, the 20-mm and the Spring Kodiak Trawl (SKT) surveys. Generally, these indices are sums of catch per unit effort (CPUE) calculated for different subregions of the Delta, with the level of spatial stratification and weighting of subregion water volumes varying between surveys. However, these indices do not have associated measures of uncertainty and implicitly assume that the probability of catching Delta Smelt is a constant throughout the survey period. For this reason, it is difficult to make direct comparisons between the different survey indices to assess where bias correction factors may be needed in population modeling and impossible to incorporate information about the uncertainties of the indices for trend analyses or modeling.

Here we develop a design-based method for calculating Delta Smelt abundance indices and the associated uncertainties that incorporates estimates of gear selectivity probabilities and assumptions about fish availability. The method is designed to be applied to data from multiple surveys, irrespective of the type of sampling gear and deployment protocols used, to produce comparable abundance indices and measures of their uncertainty. We apply the method to Delta Smelt catch data from the four previously mentioned surveys (20-mm, STN, FMWT, and SKT) as well as the Spring Midwater Trawl (SMWT) survey (which preceded the SKT), to generate abundance indices for four life stages of Delta Smelt: postlarval, juvenile, subadult, and adult. We use these results to assess recent changes in abundance and investigate potential biases in the data that might lead to unrealistic estimates of survival between life stages. Viewing the surveys as intrinsically a multistage sampling design (Hankin 1984; Newman 2008) enables us to quantify the relative contributions of different sources of variance, which provides insight into (1) features of abundance trends in recent years beyond the clear multidecadal changes and (2) strategies for improved monitoring. Finally, we use simulations to test whether describing the abundance index distribution using a lognormal distribution, which is commonly applied in state-space population models (e.g., de Valpine and Hastings 2002), is preferable to using a normal distribution, the one that arises in large-sample theory descriptors of estimate distributions (Thompson 2002).

METHODS

Survey data.— Delta Smelt abundance indices for four different life stages were derived from data collected by the five CDFW fish monitoring programs mentioned previously. These surveys differ in terms of their duration, time of year (and thus the life stage sampled), and sampling intensity (Table 1). For each survey the same sampling locations (sites) are visited each year (see Figure S1 in the Supplement to the online version of this article). These locations were not randomly chosen, however, but were purposively selected with the aim of being geographically dispersed across the Delta (Chadwick 1964). All surveys are conducted by pulling nets of varying mesh sizes through the water behind or between boats, where the net mesh size decreases from the net opening to the closed tapered end (the cod end). The ordering of cod end mesh size, from smallest to largest, for the different surveys is 20-mm, STN, SKT, FMWT, and SMWT (same as the FMWT). The 20-mm and STN surveys, which usually make three tows at each sample site, use a rigid opening net that is dropped behind the boat, allowed to sink to varying depths and then gradually pulled to the surface as the boat moves forward. The FMWT and SMWT surveys both use a midwater trawl, which has a 12 ft × 12 ft (1 ft = 30.5 cm) mouth opening held open by planing doors, that is dropped into the water, allowed to sink, and then gradually towed to the surface. The SKT uses two boats to pull a Kodiak trawl net through the water, slightly below and parallel to the surface. Further details on the surveys along with CDFW-derived indices can be found at https://www.wildlife.ca.gov/Conservation/Delta.

For each survey, the samples taken at a given site provide information on the spatial location, date, time of sampling, number and lengths of the Delta Smelt caught, and estimates of the volume of water sampled. Of relevance to the adjusted catch estimation procedure used in the index calculations (see section "Sample catch adjustments"), the STN, FMWT, and SMWT surveys did not originally take length measurements or record the volume of water sampled, but over time this became routine. Length measurements and volume calculations have always been made by the 20-mm and SKT surveys. Partly due to the lack of length and volume measurements in earlier years, the abundance indices reported here are for 1990 onward.

Several things are worth pointing out at the outset in order to contextualize the subsequent choices and assumptions of the method. At the survey data location resolution, the catch data display frequencies of zero recorded catch ranging from 74% for the SKT survey to 92% for the FMWT survey, sometimes with high spatial clustering in the regions where fish were caught. These observations motivated the use of a poststratification (described in the next section) and pure design-based approach rather than a spatial modeling approach.

Additional remarks about the 20-mm and STN surveys, which conduct repeated tows, are also necessary. To evaluate any evidence of fish depletion after the first tow, negative binomial regression models controlling for effort with and without a tow effect between the first and second tows were compared using likelihood ratio tests. No evidence was found for either survey (20-mm: $\chi^2 = 0.14$, df = 1, $P = 0.71$; STN: $\chi^2 = 0.38$, df = 1, $P = 0.54$), thus supporting the assumption of catch independence across tows and an absence of any depletion effect.

Geographic stratification and stratum volume calculations.— The design-based abundance indices calculated for the different Delta Smelt life stages are in all cases stratified random ratio estimates, where the ratios are (gear-selectivity-adjusted) catches divided by (adjusted) volume sampled that are then multiplied by estimates of stratum volumes. In this section we describe the stratification and in the next two sections discuss the sample catch and sample volume adjustments.

The Delta was partitioned into 29 subregions (Figure 1). The basis for the stratification was partly historical

TABLE 1. Summary of the CDFW fish monitoring programs that provided data for Delta Smelt abundance estimation. The number of sites sampled (n) has varied over time, and the numbers shown are approximate. The column labeled "Analyzed" shows the years used in this study.

Survey	Duration	Analyzed	Frequency	Months	n
20-mm	1995–present	1995-2017	Biweekly	Apr–Jul	60+
Summer Tow Net (SNT)	1959–present	1990–2017	Biweekly	Jun–Aug	30
Fall Midwater Trawl (FMWT)	1967–present	1990–2017	Monthly	Sep–Dec	100+
Spring Midwater Trawl (SMWT)	1990–2001	1991–2001	Monthly	Jan–May	100+
Spring Kodiak Trawl (SKT)	2002–present	2002–2017	Monthly	Jan–May	40+

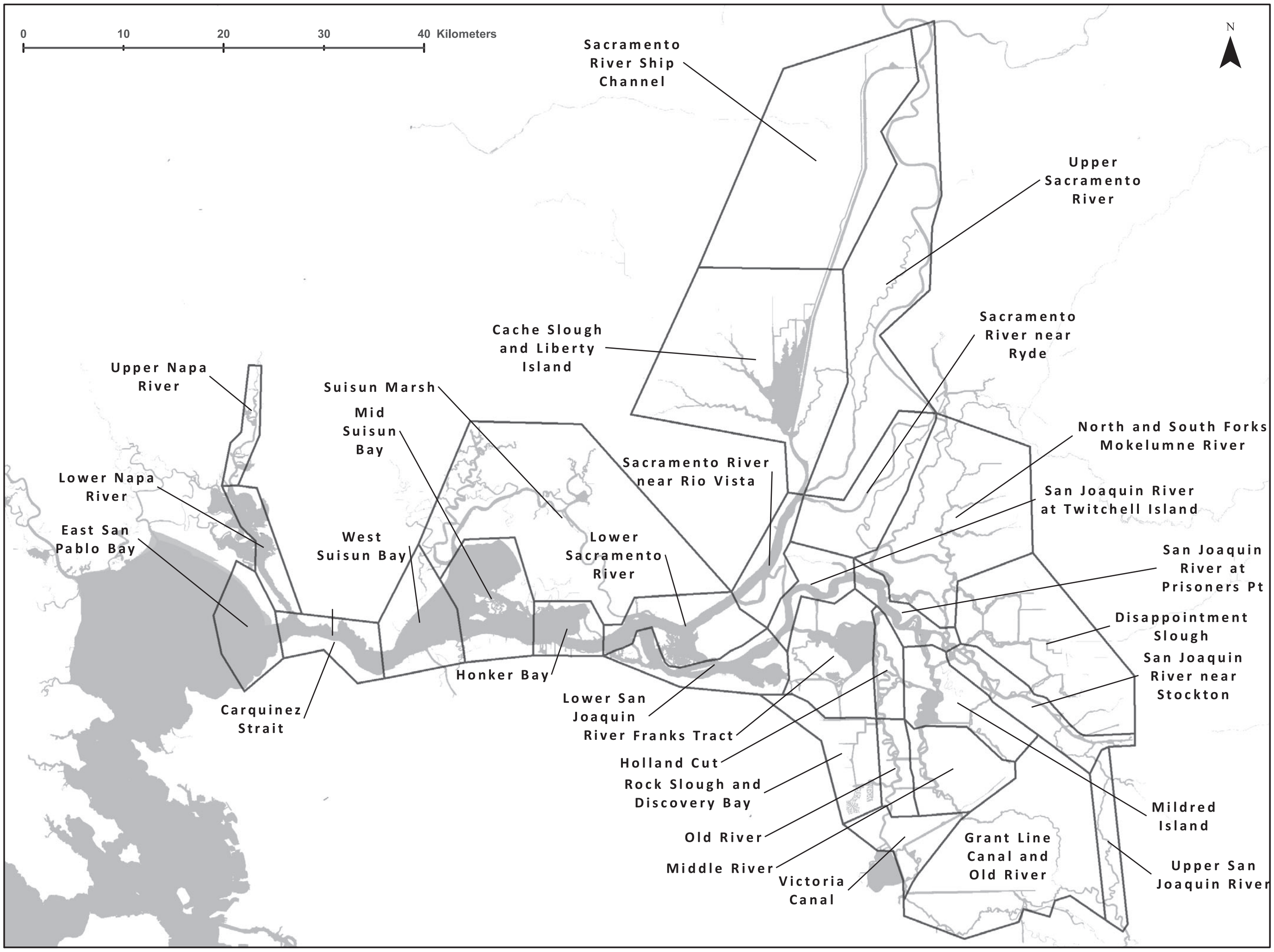

FIGURE 1. Geographic stratification of the Delta into 29 subregions (geographic strata).

(being similar to the stratification used for some of the fish indices calculated by CDFW) and partly based on similar environmental conditions within strata. Additionally, post-stratification of the sampling locations into smaller geographic regions can lessen the amount of selection bias due to nonrandom selection of sampling locations.

For each stratum, the volume of water likely to be occupied by Delta Smelt was calculated from raster files describing the bathymetry of the Delta (Fregoso et al. 2017). Two sets of volume calculations were made, one for the volume between the surface and 10 m depth (labeled the early life stage volume) and one for the volume between 0.5 and 4.5 m depths (labeled the later life stage volume; Table S1). The early life stage volume was applied to the 20-mm survey catches and the later life stage volume was applied to all other surveys. The selection of volumes is somewhat speculative, as definitive measurements of occupancy by depth are lacking. Support for the early life stage volume specification is provided by Rockriver (2004), who found that younger fish appeared to be relatively evenly and deeply distributed throughout the water column. Support for the juvenile and later life stages' being more surface oriented comes from observations that surface tows done during the summer, fall, and winter result in higher catch densities than oblique tows done during the same seasons (Souza 2002; Mitchell et al. 2017).

Sample catch adjustments.— Fish capture probabilities can be viewed as a product of two probabilities: a (marginal) probability that a fish is present and initially available for capture by the gear and a conditional probability of catching or retaining the fish given that it is available to the gear (e.g., that it is present in the volume of water passing through the net) (Crone et al. 2013). Including a length aspect to the retention probability, this probability can be expressed as $\Pr(\text{Catch fish}_L) = \Pr(\text{Fish}_L \text{ Available}) \times \Pr(\text{Catch}|\text{Fish}_L \text{ Available})$ where $\Pr(\text{Catch}|\text{Fish}_\text{L} \text{ Available})$ is contact selectivity (Crone et al. 2013) and each caught fish of length L represents $1/\Pr(\text{Catch Fish}_L)$ fish.

For the abundance index calculations made here, the catches of fish in individual tows from each survey were upwardly adjusted using only estimates of gear-specific, length-based estimates of contact selectivity. If the probability of availability was exactly 1.0 for all fish present (per gear, sampling location, and occasion), then such expansions could yield estimates of absolute abundance. However, this is almost certainly not true and is one reason the values constructed are labeled "indices" and not estimates of the true abundance.

Length-based, gear-specific contact selectivity functions were obtained from Mitchell et al. (2017, 2019). In Mitchell et al. (2017), a cover was placed over the cod end of the FMWT (and SMWT) gear and the assumption was made that all fish that slipped through the cod end mesh were retained by the cover. In Mitchell et al. (2019), different combinations of 20-mm, STN, and SKT gear were deployed more or less simultaneously in the same area. In this case, because direct information on the length distribution of the population is not available, the estimated curves are relative selectivity curves (Millar and Fryer 1999). For practical purposes, relative selectivity means that the scaling of the selectivity functions cannot be determined and thus provides another reason for the label "index."

Catch by a given gear g was adjusted as follows: Let $c_{g,o}$ be the number of Delta Smelt caught by gear g on occasion o (where o denotes an arbitrary year, month, stratum, sampling location, or [in the case of the 20-mm and STN surveys] an arbitrary tow). Let $L_{g,o,i}$ be the length of the ith fish in that catch and $\hat{p}_g(L_{g,o,i})$ be an estimate of the contact selectivity probability for that fish (where p_g is a true but unknown function). The adjusted catch, denoted $c^*_{g,o}$, is

$$c^*_{g,o} = \sum_{i=1}^{c_{g,o}} \frac{1}{\hat{p}_g(L_{g,o,i})}. \tag{1}$$

The range of fish lengths recorded in the catch data in some cases exceeded the range lengths used to estimate the selectivity curves. For fish outside the range, we assigned captured probability values from the nearest endpoint of the curve.

Sample volume adjustments.— The volume of water towed during a survey often included portions of the water column assumed to be unoccupied by Delta Smelt, namely, depths outside of the depths defined as the early life stage or later life stage volume. Effective volume v^* was defined as the portion of a tow volume that intersected the relevant life stage stratum (calculation below).

The geometry of the effective volumes can be approximated by rectangular prisms, with oblique tows (used by the 20-mm, STN, FMWT, and SMWT surveys) described by non-right prisms and surface tows (used by the SKT survey) described by right prisms. For oblique tows, the non-right prism volume is a function of tow depth and the net mouth height. Because tow depths were not routinely recorded, tow depth was estimated using the angle at which the trawl was deployed, the length of the cable released, and the block height (the height from the water surface to the block from which the cable is released). Per survey protocols, an increase of 25 ft in the length of cable released corresponds to an approximately 1.2 m increase in the depth of the trawl. We used average block heights (calculated across different boats) of 2.53 m for the 20-mm surveys (T. Morris, CDFW, personal communication), 2.48 m for the STN surveys (F. La Luz, CDFW, personal communication), and 2.03 m for the FMWT and SMWT surveys (S. Finstad, CDFW, personal communication). Given the estimated tow depth, measures of net mouth height, total sample volume, and the upper and lower bounds of the fish stratum, the effective volume was calculated as the intersection of the volume swept by the trawl and the volume occupied by the fish. For the SKT surface tows and right prism geometry, the effective volume calculation was simply the intersection of the rectangular prism parallel to the water surface (calculated from tow volume and net mouth height) and the vertical band between 0.5 and 4.5 m:

$$v^*_{\text{SKT}} = v_{\text{SKT}} \times \left(\frac{\text{Net height} - 0.5}{\text{Net height}}\right) = v_{\text{SKT}} \times \left(\frac{1.8 - 0.5}{1.8}\right) = v_{\text{SKT}} \times 0.722,$$

where 1.8 m is the height of the net mouth.

Abundance indices and variances.— The equations for abundance indices parallel the following expression for the true abundance of life stage (ls) fish during year y and month m, $N_{ls,y,m}$:

$$N_{ls,y,m} = \sum_h^H N_{ls,y,m,h} = \sum_h^H V_{ls,h}\delta_{ls,y,m,h}, \tag{2}$$

where h denotes a given geographic stratum (and H is the total number of strata), and the stratum abundances, $N_{ls,y,m,h}$, are products of (true) stratum-specific densities $\delta_{ls,y,m,h}$ and habitat water volumes $V_{ls,h}$. The general form for the abundance indices for all life stages is a stratified ratio-of-means estimator (Thompson 2002):

$$I_{ls,y,m,g} = \sum_h^H I_{ls,y,m,g,h} = \sum_{h=1}^{29} V_{ls,h}\hat{\delta}_{ls,y,m,g,h}, \tag{3}$$

with

$$\hat{\delta}_{ls,y,m,g,h} = \frac{\sum_{j=1}^{n_{y,m,g,h}} c^*_{ls,y,m,g,h,j}}{\sum_{j=1}^{n_{y,m,g,h}} v^*_{ls,y,m,g,h,j}}, \tag{4}$$

where $n_{y,m,g,h}$ is the number of tows by gear g in a year, month, and stratum, c^* is the adjusted catch (equation 1), and v^* is the adjusted volume.

For each cohort, four different life stage abundance estimates—postlarval, juvenile, subadult, and adult—were calculated based on May 20-mm, July–August STN, October–November FMWT, and February–March SMWT and SKT data, respectively (the supplemental material includes additional indices for other choices of months). When multimonth pooling was done, primarily to increase the number of sampling locations, the indices ostensibly reflect some average abundance over the sampling period that implicitly includes mortality or recruitment, though the latter is thought negligible by the month of June. In some cases sampling periods for a given survey spanned 2 months, e.g., some sampling locations in the 20-mm "June" survey were actually sampled in July. In these cases, we assigned the label m based on the month in which most samples were taken.

The variance of $I_{ls,y,m,g}$ is the sum of the variances of the stratum-specific indices, $I_{ls,y,m,g,h}$:

$$Var(I_{ls,y,m,g}) = \sum_{h=1}^{29} V_{ls,h}^2 Var(\hat{\delta}_{ls,y,m,g,h}). \tag{5}$$

If the fishing gear were 100% efficient, the variance of $\hat{\delta}$ could be estimated using standard design-based formulas for an estimated ratio (Thompson 2002) that account for between-sample variation in the ratio estimate of the number of fish within a stratum. Because the true number of fish is in fact being estimated at each location by imperfect gear, two more sources of variation need to be accounted for, which we accomplished using the notion of multistage sampling and the law of total variance (Hankin 1984; Thompson 2002; Newman 2008). For each stratum-specific estimate, there are three sources of variation: (1) between–sample location variation in fish density (the ratio of fish to volume), (2) randomness in catching fish that are available to the gear, which for a fish of length L occurs with probability $p_g(L)$ (assuming 100% availability), and (3) uncertainty in the estimated probabilities of fish capture $\hat{p}_g$. Abbreviating the estimated probability of capture of the ith fish on the jth tow in stratum h by $\hat{p}_{j,i}$ (omitting the notation identifying the gear and the length-specific dependency of this probability), the estimated variance of $I_{ls,y,m,g}$ is

$$\widehat{Var}(I_{ls,y,m,g}) = \sum_{h=1}^{29} \frac{V_{ls,h}^2}{(\bar{v}^*_{y,m,g,h})^2} \times \left(\frac{1}{n_{y,m,g,h}^2} \sum_{j=1}^{n_{y,m,g,h}} \sum_{i=1}^{c_{ls,y,m,g,h,i}} \left[\underbrace{\left(\frac{1-\hat{p}_{j,i}}{(\hat{p}_{j,i})^2} \right)}_{\text{source 2}} + \underbrace{\frac{1}{\hat{p}_{j,i}^4} \widehat{Var}(\hat{p}_{j,i})}_{\text{source 3}} \right] + \underbrace{\frac{\hat{s}^2_{ls,y,m,g,h}}{n_{y,m,g,h}}}_{\text{source 1}} \right), \tag{6}$$

where $\bar{v}^*_{y,m,g,h}$ is the mean effective tow volume within the stratum and $s^2_{ls,y,m,gh}$ is the within-stratum, between-tow variability in ratio estimates:

$$\hat{s}^2_{ls,y,m,gh} = \frac{\sum_{j=1}^{n_{y,m,g,h}} (c^*_{ls,y,m,g,h,j} - \hat{\delta}_{ls,y,m,g,h} \times v^*_{y,m,g,h,j})^2}{n_{y,m,g,h} - 1}. \tag{7}$$

Details of the derivation are given in the appendix, where the finite population correction factor is assumed to be negligible (Thompson 2002). The estimated standard error $\widehat{SE}_{ls,y,m,g}$ is the square root of equation (6), and the estimated coefficient of variation $\widehat{CV}_{ls,y,m,g}$ is the ratio of the standard error to the index.

Stratum-level variance estimates were undefined when only one sample was taken, and in those cases the median of the stratum-specific values of equation (6) was substituted. If the catch density was exactly the same across all sites (practically, if a stratum had zero total catch), the variance contribution for that stratum was set to 0.

Abundance indices with truncated contact selectivity functions.— A practical problem when adjusting catch using capture probabilities is that very small values of $\hat{p}_g(L)$ can lead to unrealistically large adjusted catch values. This was of particular concern for the non-monotonic 20-mm and STN selectivity curves identified by the data, which were not informed by many captures of large fish (Mitchell et al. 2019). To investigate the effects of this problem, we compared indices based on the original selectivity curves with estimates based on "truncated" curves, defined to be the same as the original curves except with the descending tail of each curve replaced by a horizontal line at 1.0 (see Figure 7 in Mitchell et al. 2019).

Measures of vital rates.— Abundance indices for successive life stages were used as measures of vital rate parameters such as recruitment (number of young produced per adult) and between–life stage survival for given cohorts. Such measures are calculated by taking the ratios of indices for successive life stages. For example, an approximate measure of the recruitment of postlarvae (pl) in cohort $t + 1$ from adults (a) in cohort t is $I_{t+1,pl}/I_{t,a}$. Similarly, a relative measure of the survival of juveniles (j) to subadults (sa) is $I_{t,sa}/I_{t,j}$. Because they are indices and not unbiased estimates of absolute abundance, such ratios are unlikely to provide estimates of actual recruitment or survival rates, but they may allow estimation of population growth rates ($I_{t+1,a}/I_{t,a}$) if all unknown scaling factors and availability probabilities are constant in time (because they will then cancel out).

Decomposition of variance components.— The three sources of variation making up the index variance estimate shown in equation (6) can be multiplied out so that the variance is the sum of terms corresponding to each source separately, i.e., $\widehat{Var}(I_{ls,y,m,g}) = s_1 + s_2 + s_3$ where s_i

is the ith source of variance (the life stage, time and gear-specific indices on the right-hand side have been suppressed for clarity). For each index, we computed the fraction of its total variance by source i, $f_i = s_i/(s_1 + s_2 + s_3)$, to describe how these changed across life stages and within life stages across years.

Lognormal distribution-based confidence intervals and a simulation study.— One approach to constructing α-level confidence intervals for the indices is to assume that the estimated indices are approximately normally distributed and to set the interval equal to

$$I \pm z_{1-\alpha/2}\sqrt{\hat{V}(I)},$$

where $z_{1\text{-}\alpha/2}$ is the $1 - \alpha/2$ quantile of a standard normal distribution. Justification for the normality assumption (the Central Limit Theorem) when sampling from a finite population without replacement is more complicated (Thompson 2002), but tows can reasonably be viewed as sampling with replacement given the extremely small sample volumes relative to the potential habitat volumes (Table S2). More critically, a practical problem with quantities like indices, which have to be nonnegative, is that such intervals can have negative lower bounds; e.g., a 95% interval will have a negative lower bound when the coefficient of variation of the estimate exceeds 0.51.

Here we used an alternative approach that ensures intervals above zero by assuming that the indices are lognormally distributed. Dropping the ls, y, m, and g subscripts, the parameters of the lognormal distribution are the log-mean $\mu = \ln\left(I/\sqrt{1 + \widehat{CV}^2}\right)$ and $\sigma^2 = \ln\left(1 + \widehat{CV}^2\right)$, which as constructed ensures that the expected value of the distribution is the index $I_{ls,y,m,g}$. Then given an α, the confidence interval is given by the $\alpha/2$ and $1 - \alpha/2$ quantiles of this lognormal distribution.

A simulation experiment (described in detail in Supplement E) was designed to gain insight into the performance of the estimation procedure and the use of the lognormal distribution as described above for constructing confidence intervals given a multistage data generation process. Nine different selectivity curves were used in combination with realistic sample sizes (i.e., very small ones). The data-generating process used a baseline abundance, N_{Tot}, of 102,000 fish, corresponding to a stratum level density of 1 fish per 10,000 m^3 of habitat, all available to be sampled. Potential catch was then simulated according to a negative binomial model, and a logistic contact selectivity curve was used to simulate a realized catch. The variation in numbers caught was purely a function of between-sample catch variation and contact selectivity, as availability was assumed to be 100%; thus, in this case, the estimated totals $\hat{N}_{Tot}$ are of the simulated baseline abundance value. A total of 1,000 simulations for each choice of gear selectivity curves were done. The bias (relative to the simulated baseline abundance) and standard errors of $\hat{N}_{Tot}$ were recorded, and the actual coverage of lognormal-based confidence intervals was compared with the nominal coverage of 95% and contrasted with normal distribution–based intervals.

RESULTS

Sample Catch Adjustments

By design, adjusted catches are always greater than or equal to the corresponding nonadjusted catches, leading to catch inflation factors (adjusted catch divided by nonadjusted catch) that are greater than or equal to 1.0. For the 20-mm catches, the mean inflation factor was 5.05 (SD, 6.12), ranging from 1.00 to 22.49; for the STN catches, the mean inflation factor was 1.70 (SD, 1.41), ranging from 1.0 to 44.01; for the FMWT catches, the mean inflation factor was 3.18 (SD, 0.75), ranging from 1.00 to 4.35; and for the SMWT catches, the mean inflation factor was 1.78 (SD, 0.59), ranging from 1.0 to 3.91. The adjusted SKT catches were identical to the nonadjusted catches because the estimated relative selectivity of the SKT gear was 1.0 over the range of lengths observed.

Sample Volume Adjustments

Effective sample volumes were always less than or equal to the corresponding raw sample volumes. For the 20-mm survey, the effective and raw volumes were identical. For the STN survey, the effective volumes were always smaller than the raw volumes, with a mean factor of 0.71 (SD, 0.17), ranging from 0.53 to 0.97. For the FMWT survey, the mean factor was 0.78 (SD, 0.10), ranging from 0.66 to 0.96, and for the SMWT survey, the mean factor was 0.78 (SD, 0.09), ranging from 0.66 to 0.96.

Abundance Indices and Variances

Declines in Delta Smelt abundance across all life stages over the past several decades are clearly evident (Table 2; Figure 2). The uncertainties in the indices, as measured by the CVs, were on average 37.04, 33.59, 45.51, 24.33, and 30.90% for the 20-mm, STN, FMWT, SMWT, and SKT-based indices, respectively. These abundance indices are highly correlated with the corresponding CDFW indices for the years when both were estimated (Figure 3). Both show similar long-term downward trends and localized periods of relatively high and low values, and with a few exceptions they track the year-over-year changes (increases or decreases). Notable differences include indices of post-larvae based on the 20-mm survey data, for which the new indices indicate higher recruitment success for 1996 and lower success for 1999 relative to the CDFW indices.

TABLE 2. Delta Smelt abundance indices (standard errors in parentheses). See Table 1 for survey abbreviations; NA denotes no available data for the given survey and year.

Year	20-mm	STN	FMWT	SMWT	SKT
1990	NA	944,890 (247,880)	485,426 (165,111)	NA	NA
1991	NA	3,947,363 (683,072)	1,178,446 (227,064)	131,260 (33,617)	NA
1992	NA	1,722,648 (287,981)	155,808 (57,644)	103,603 (26,762)	NA
1993	NA	7,957,836 (1,429,502)	1,861,967 (549,550)	55,630 (14,084)	NA
1994	NA	5,594,684 (743,458)	62,173 (24,175)	485,581 (106,027)	NA
1995	3,802,003 (1,714,167)	3,885,218 (503,118)	2,870,967 (541,800)	90,155 (24,314)	NA
1996	51,816,580 (9,680,651)	9,519,528 (3,741,334)	72,185 (24,425)	856,455 (205,410)	NA
1997	28,676,814 (5,422,401)	2,256,242 (650,399)	692,611 (204,958)	115,537 (27,425)	NA
1998	5,435,652 (2,523,076)	3,006,382 (558,410)	327,681 (70,380)	140,128 (41,206)	NA
1999	18,546,993 (4,513,613)	9,307,496 (1,464,873)	2,198,820 (484,791)	171,469 (39,449)	NA
2000	24,333,860 (5,208,331)	6,029,290 (782,124)	717,813 (166,928)	539,175 (134,012)	NA
2001	19,761,621 (4,903,592)	4,940,657 (811,880)	2,059,595 (688,896)	245,506 (41,888)	NA
2002	5,330,964 (1,608,388)	2,441,040 (368,227)	345,150 (90,302)	NA	933,982 (225,097)
2003	6,661,403 (3,668,971)	1,546,580 (238,121)	833,943 (310,214)	NA	1,167,662 (165,504)
2004	11,334,053 (3,686,194)	696,211 (165,741)	451,505 (219,759)	NA	763,619 (161,573)
2005	13,754,192 (3,625,550)	1,139,543 (263,185)	64,973 (23,449)	NA	329,722 (101,264)
2006	3,360,377 (1,586,596)	590,540 (271,746)	33,479 (16,099)	NA	301,735 (45,389)
2007	1,659,962 (1,965,671)	311,681 (133,506)	23,371 (13,005)	NA	375,070 (124,451)
2008	1,427,033 (789,623)	508,404 (160,339)	53,864 (22,792)	NA	207,930 (82,196)
2009	5,190,179 (2,021,635)	285,517 (104,853)	23,970 (13,407)	NA	217,409 (72,908)
2010	4,870,088 (1,503,243)	1,170,651 (405,384)	43,910 (29,287)	NA	278,255 (90,568)
2011	4,205,030 (1,762,812)	3,589,513 (832,610)	279,154 (95,298)	NA	232,899 (83,947)
2012	16,626,279 (4,556,209)	611,230 (139,708)	112,339 (34,401)	NA	1,105,082 (388,559)
2013	5,379,031 (1,090,197)	715,704 (183,800)	20,975 (11,389)	NA	316,806 (93,219)
2014	1,868,430 (502,275)	266,270 (92,006)	11,781 (10,316)	NA	250,095 (80,597)
2015	525,597 (177,227)	3,201 (4,517)	2,886 (3,872)	NA	162,446 (74,258)
2016	426,070 (131,597)	11,676 (15,488)	19,348 (12,632)	NA	21,730 (8,901)
2017	690,469 (250,915)	320,293 (176,681)	7,502 (8,270)	NA	30,888 (9,561)

Very recent (2013–2017) adult abundance indices also show a decline. The upper confidence limits in 2016 and 2017 are lower than the lower confidence limits for 2013–2015, suggesting a continued downward trend in recent years (Figure 4). In particular, the decline after 2015 reflects a record-low population growth rate of 0.13 for the 2015 cohort.

Abundance Indices with Truncated Contact Selectivity Functions

The indices based on truncated contact selectivity curve results can be considerably smaller (Supplement F and Figure S4). For the 20-mm survey, the nontruncated point indices ranged from about 1 to 2 (June data) or 10 (July data) times the truncated indices, while for the STN survey the nontruncated indices were between 1 and 2 times the truncated indices (Table S8). The nontruncated and truncated indices are highly correlated (Table S9; Figure S4). As expected, the proportion of the variance of an abundance index attributable to catch randomness decreased when truncated selectivity curves were used to adjust catch (Figure S5).

Measures of Vital Rates

Estimates of (relative) recruitment (postlarvae per adult) are reported separately for the 1995–2001 and later adults because of a likely change in the adult abundance index bias from 2001 to 2002, when the adult sampling gear changed from a midwater trawl (the SMWT survey) to a Kodiak trawl (the SKT survey). The mean estimated recruitment for cohorts in the earlier period is 89.07 postlarvae per adult (SD, 74.43), ranging from a minimum of 38.80 per adult in 1998 to 248.20 per adult in 1997. Mean estimated recruitment for cohorts in the later period is 14.66 postlarvae per adult (SD, 9.82), ranging from a minimum of 3.24 in 2015 to a maximum of 41.72 in 2005.

Postlarval survival rates ranged from a minimum of 0.01 juveniles per postlarva in 2015 to a maximum plausible value of 0.85 juveniles in 2011 (and a single larger, and implausible, value greater than 1.0). Juvenile survival

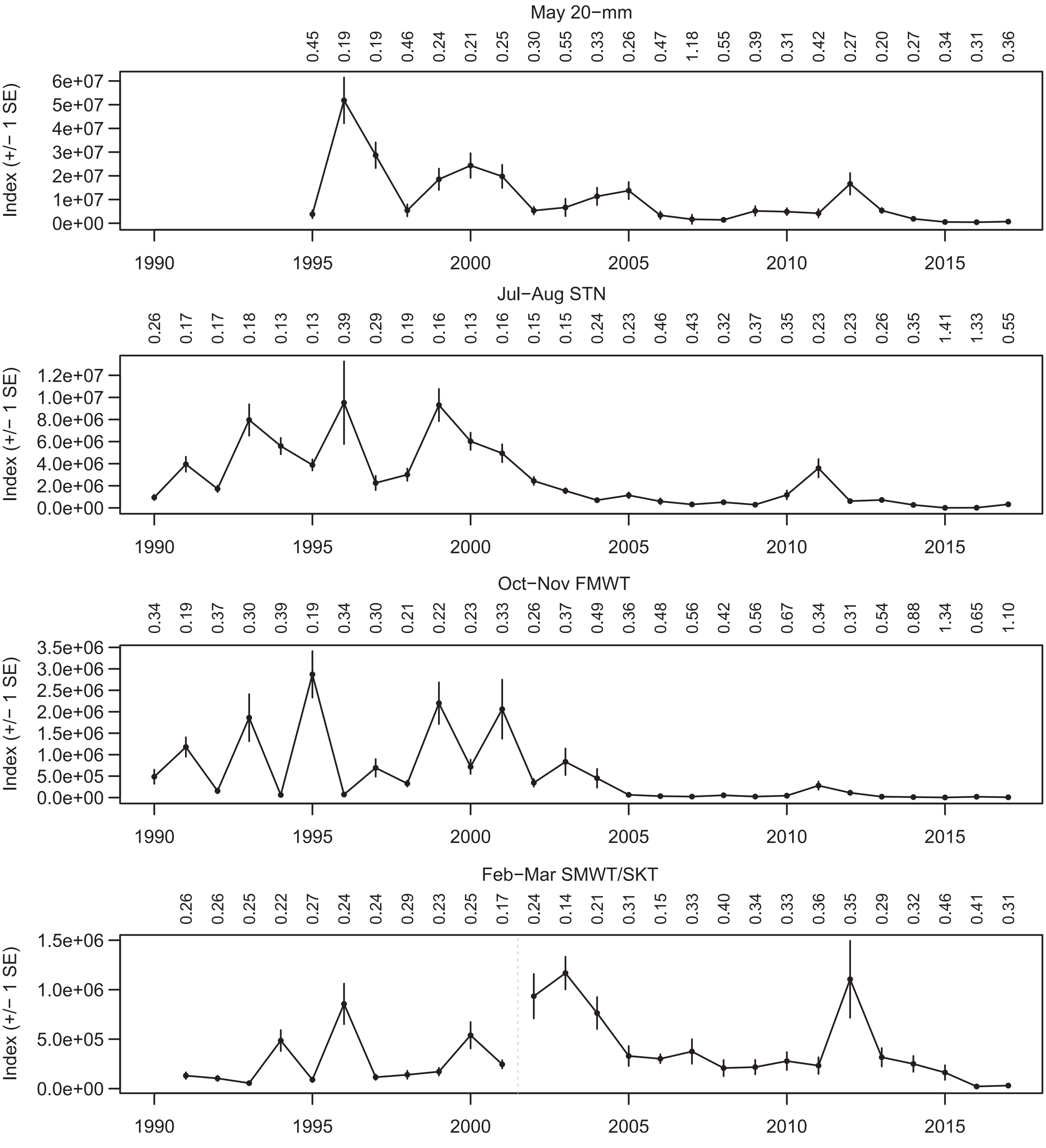

FIGURE 2. Abundance index time series for Delta Smelt, with error bars denoting ± 1 standard error. Coefficients of variation are shown along the upper axis of each panel. The gray vertical line in the bottom panel separates adult abundance indices based on the SMWT survey (earlier years) from those based on the SKT survey (later years).

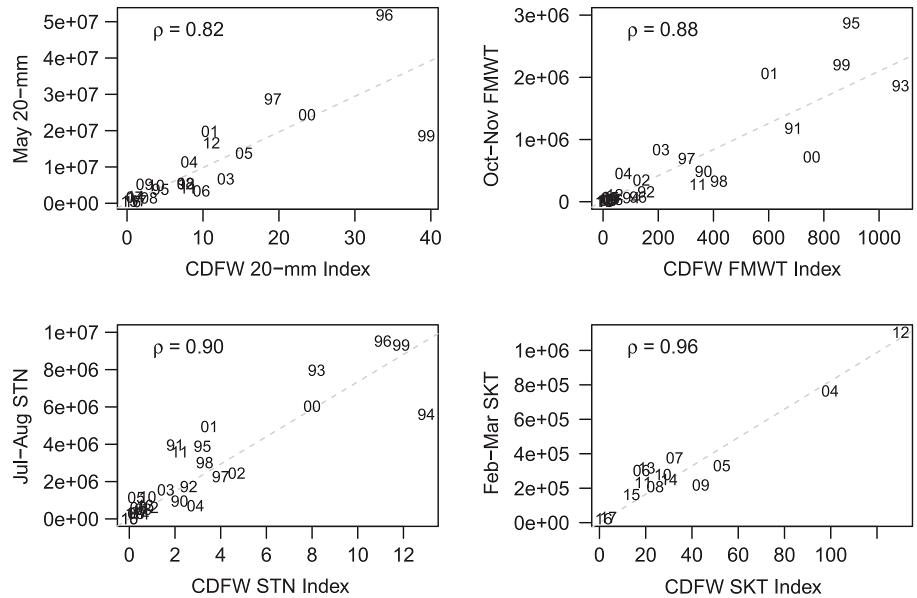

FIGURE 3. The abundance indices ($I_{ls,y,m,g}$) computed in this article versus the CDFW indices, with points indicated by the last two digits of the calendar year of the data used in index construction. The dashed gray lines are regression lines through the origin predictions. Pearson pairwise complete correlations are shown at the top left of each panel.

rates ranged from a minimum of 0.01 subadults per juvenile in 1996 to a maximum plausible value of 0.90 in 2015 (and a single value larger than 1). Subadult survival rate estimates were especially problematic for the 2002–2017 time period, with 13 of the 16 based on SKT adult abundances being larger than 1. However, for the period prior to 2002, given that the SMWT and FMWT used identical gear, the unmeasured gear efficiencies (e.g., related to availability to the gear) are presumably quite similar; thus, gear-selectivity effects should be minimal. For the subset of subadult survival rates based on SMWT adult estimates (11 total), plausible values ranged from 0.09 adults per subadult in 1991 to 0.52 in 1998 (with two being implausibly larger than 1).

Cohort population growth rates, each the product of the postlarval recruitment value and the three survival rates of the subsequent life stages, ranged from 0.13 in 1996 to 9.50 in 1995 for the cohorts with adult abundance indices derived from the SMWT survey, and from 0.13 in 2015 to 4.74 in 2011 for the cohorts with adult abundance indices derived from the SKT survey. The 2012 adult abundance index is noticeably higher than other contemporary abundance indices, likely a reflection of the relatively large population growth rate in 2011—the next largest being 1.42 in 2016 when abundances where relatively very low.

Decomposition of Variance Components

The proportion of variance contributed by each of the three separate sources of variability depended on the combination of gear and life stage (Figure 5). The variance of the index based on the 20-mm survey is slightly dominated

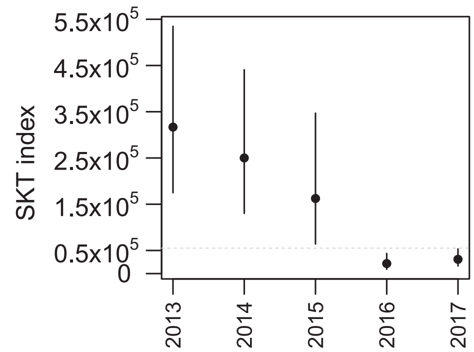

FIGURE 4. Adult abundance indices, with vertical lines showing the lower and upper confidence intervals for 2013–2017 based on the February and March SKT survey. The dotted horizontal line is drawn at 55,000, above the upper confidence limits for 2016 and 2017 and below the lower confidence limit for prior years.

by the randomness in catching fish that are available to the gear, followed by between–sample location variability in fish density, with relatively little contribution from the uncertainty in the estimated probabilities of fish capture. In contrast, the STN, FMWT, SMWT, and SKT abundance index uncertainties are all dominated by between–sample location variability.

Lognormal Distribution-Based Confidence Intervals and the Simulation Study

The simulation study showed that the distributions of the multistage estimates of abundance are right-skewed,

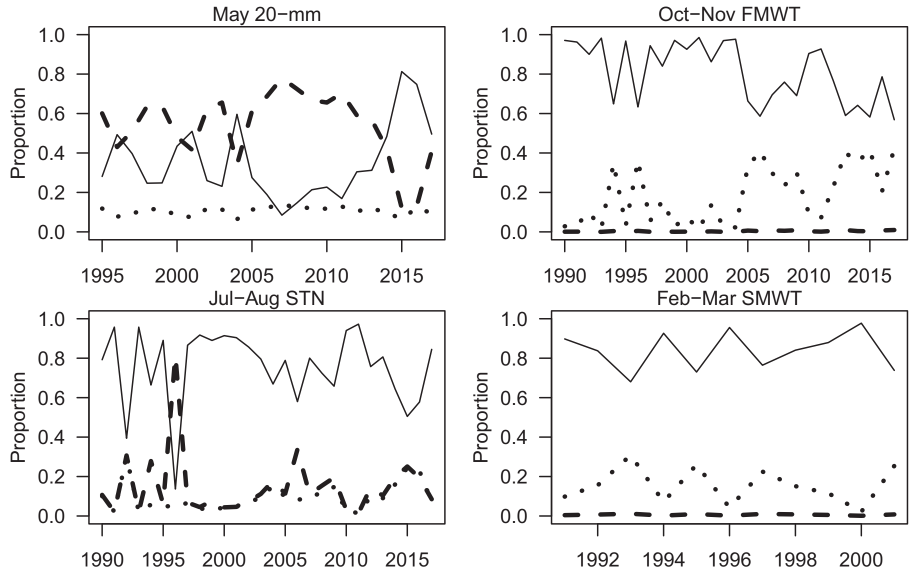

FIGURE 5. Proportions of the total variance of the abundance index accounted for by the three sources of variation: between–sample location variation (solid lines), the randomness in catching fish that are available to the gear (dashed lines), and the uncertainty in the estimated probabilities of fish capture (dotted lines). For the SKT survey (not shown), the proportion due to sampling location was always 1 because the gear selectivity was assumed to be 1 with no uncertainty.

with the degree of skewness varying as a function of the contact selectivity parameters (Figure S3). The estimates $\hat{N}_{Tot}$ have relatively small bias even for highly inefficient gear, ranging from −1% to +2% (Table S5). However, the average coefficient of variation (for indices with non-zero values) range from 37% to 91% (Table S6). Such CVs, while relatively large, are within the range of the empirical estimates from the Delta Smelt data set (Figure 2). Baseline abundance estimates ($\hat{N}_{Tot}$) equal to zero resulted only when we were using the selectivity curves with near-zero values across much of the range of fish lengths (Table S7).

Actual coverage of the 95% confidence intervals based on the lognormal distribution is affected by the contact selectivity function. For the logit models corresponding to overall intermediate selectivity ($\beta_0 = 0.5$), observed coverage equaled nominal coverage. However, with the overall high-selectivity models ($\beta_0 = -0.5$), observed coverage was slightly lower (from 90% to 94%), while for the overall low-selectivity models ($\beta_0 = 0.9$) coverage was too high (from 97% to 100%). Confidence intervals based on a normal distribution, which were also affected by the contact selectivity curves, increasingly yielded negative lower bounds as β_0 increased, from up to 4% with $\beta_0 = -0.5$, 20–26% with $\beta_0 = 0.5$, and up to 100% with $\beta_0 = 0.9$.

DISCUSSION

A single, well-established finite population sample estimation procedure, namely, stratified random sample ratio expansions (Thompson 2002), was applied to trawl catch data collected from several long-term fish monitoring programs to calculate survey-specific point estimates of relative abundance, along with variances. These abundance indices are strongly correlated with the conventional indices, with both showing substantial declines over the past several decades. Because a similar estimation procedure was applied to all the surveys, direct comparisons of the estimates were possible, revealing that despite corrections for gear selectivity the FMWT and SMWT indices continue to be biased relative to the other indices. This sort of bias identification can be useful for population modeling efforts, particularly for structuring the observation error equations.

The uncertainty measures, variances, and confidence intervals provided insights beyond those possible from point estimates alone. First, in conjunction with the lognormal assumption about the point estimate distribution, it appears that in the past few years abundances have continued to decline significantly, something that the conventional indices could not establish given the absence of estimates of uncertainty. The ability to make probabilistic statements about year-over-year changes in abundance is critical for scientific assessments of the changing status of the population.

Second, partitioning the variation into three categories helps identify how different life stages may be distributed throughout their habitat relative to the surveys. If there are many postlarval Delta Smelt for the 20-mm survey gear to encounter, then gear-related uncertainty overshadows between-sample variability. One explanation for the apparent increase in the relative importance of between-sample uncertainty from the postlarval to the adult life stage is the inherent decline in population size from one life stage to the next. As the number of Delta Smelt available to each successive survey (STN and FMWT, then SMWT or SKT) decreases, the patchiness of their distribution could increase, and between–sample location variability becomes more important. The very high frequency of zero catches combined with sometimes very high catch totals could be evidence of such patchiness.

Third, partitioning the variance also suggests both what is working and what improvements in data collection procedures can be made. For the 20-mm survey, the largest component of variance came from the randomness that a fish in the path of the gear will be caught, supporting the current sample design of conducting multiple tows at a single location, as is currently done. The relative dominance of between–sample location variability for the STN, FMWT, SMWT, and SKT abundance indices suggests expanding the spatial coverage of these surveys. Such an expansion is a feature of a new enhanced Delta Smelt monitoring program conducted by the USFWS, which samples from an increased number of spatially random sites per stratum and to date has consistently detected Delta Smelt when the FMWT has not.

One gap in our knowledge that potentially affects the quality of the abundance index estimates is precisely how Delta Smelt are distributed in the water column vertically and horizontally and how this might vary geographically. Despite extensive monitoring, the percentage of the total potential habitat sampled by a survey in a given month was typically much less than 1% (Table S2), limiting our ability to infer the distribution of density in detail. Spatial distribution affects how effective sample volumes should be calculated for estimating fish density within a stratum as well as how the stratum water volumes used for density expansions should be calculated (and ultimately affects the probability that fish are available to the gear). Evidence that fish availability to sampling gear depends on spatiotemporally dynamic habitat characteristics, particularly tide (Feyrer et al. 2013; Bennett and Burau 2015; Polansky et al. 2018) and turbidity (Feyrer et al. 2007; Nobriga et al. 2008; Polansky et al. 2018), further complicates the problem of identifying what portion of the potential habitat is actually occupied at any given moment.

How Delta Smelt are spatially distributed also has implications for whether catch densities should be further adjusted because a given survey may disproportionately sample from higher- or lower-density portions (both vertically and horizontally) of the habitat. While density estimates can be corrected to account for biased sampling, without precise knowledge of spatial distributions any such corrections are assumption laden. However, spatial poststratification of survey data can ameliorate some of the large-scale consequences of spatial density variation when expanding local catch densities.

Another issue affecting the quality of the abundance indices is that none of the sampling locations visited here were randomly selected. The sites were instead purposively selected, with the same sampling locations being visited over time, both within and between years (i.e., an always-revisit monitoring design; McDonald 2012). In fact, the surveys share many of the same sampling locations, many of which were selected when the earliest survey (the STN survey) was originally established in the late 1950s and the (fall and spring) MWT survey was established in 1967. Thus, in principle, the failure to choose sampling locations randomly could result in selection bias, e.g., if the sites were selected because of a priori knowledge that fish were more likely to be present. Further, because the chosen sites were located where the trawl gear could be safely and practically deployed, nearshore portions of the Delta volume are systematically excluded from the sample frame. This, in turn, could bias indices (up or down) if Delta Smelt densities change systematically in these areas, although the fraction of total habitat these areas represent is small.

Two factors that may partially alleviate the lack of randomness in sample site selection are the tidal dynamics of the Delta and spatial poststratification. The spatiotemporal distribution of Delta Smelt is strongly affected by the tides (Bennett and Burau 2015). The volume of water at the same fixed location is constantly changing, and pelagic fish (particularly relatively small fish like Delta Smelt) are thought to be constantly changing their position, in some cases volitionally and in other cases due to hydrodynamics. Thus, if one did continuously sample at a fixed geographic location throughout a single day, one would be sampling a body of water that covers several kilometers (Bennett and Burau 2015). Spatial poststratification can also help, in that sampling locations purposively selected because they were thought to have relatively high fish densities will have less effect on estimated totals as the densities for such locations only affect the strata in which they are located.

A somewhat more complicated situation arises if gear deployment elicits a behavioral response by the fish, causing them to either disperse or aggregate. For example, when nets are dragged behind boats and the boat displaces the fish below it, that would cause an immediate change in availability that is not easily measured with the available trawl data alone. Alternatively, the use of two boats in the deployment of the Kodiak trawl in the SKT survey could act to herd the fish toward the net. In that case, one could not say that the probability of availability is now greater than 1.0 (i.e., meaningless) but rather that the volume sampled has in fact increased. There is some evidence for such herding from the gear evaluation studies, as the two-boat surface tow method used by the SKT survey generally resulted in larger catch densities than the single-boat oblique method used by the STN and FMWT surveys (Mitchell et al. 2017, 2019). More generally, how the nets are deployed in the water, such as their position relative to the boat(s), speed, duration, and direction (relative to the direction the fish are swimming), have the potential to affect the relationship between the water volume sampled and catch, and that relationship can be affected by local habitat features such as turbidity, temperature, and flow.

Another caveat is that the estimated length-based contact selectivity functions, the $\hat{p}_g(L)$ (Mitchell et al. 2017, 2019), may be biased and inadequate. Skepticism about the ascending and descending limbs of dome-shaped selectivity curves led to sensitivity analysis using the truncated curves and the effects on the resulting abundance indices were sizable, e.g., up to a 10-fold decrease from nontruncated to truncated estimates. Equally critical is the fact that contact selectivity is undoubtedly a function of more than fish length alone. Polansky et al. (2018) showed that using a Poisson distribution for Delta Smelt catches, which implicitly assumes completely random spatial distributions, is inferior to the negative binomial distribution, which can reflect spatial aggregation ("patchiness"). If the probability of capture (for a fish that was available) is affected by the presence of other fish, the underlying independence assumption of the contact selectivity model is violated, which further complicates fitting and applying such selectivity models.

In conclusion, despite these challenges and the observation that the indices constructed reveal the same temporal trend as the CDFW derived ones, constructing indices and associated uncertainties using a uniformly applied method was useful in several ways. Estimates of uncertainty and the simulation study (designed to identify how to incorporate this uncertainty into trend analysis) allowed further progress in understanding trends and biases as well as recommendations for improved survey design. Further, the work here can be used to guide life cycle model formulation and the resulting abundance indices and standard errors can serve as input data for fitting such models, which can in turn be used to help identify factors associated with population dynamics and overall decline.

ACKNOWLEDGMENTS

We thank California Department of Fish and Wildlife staff, and Randy Baxter in particular, for discussions on the topics presented in this article. Matt Nobriga, Will Smith, Vanessa Tobias, the editor, the associate editor, and several anonymous reviewers provided helpful comments on an earlier version. The California Department of Water Resources and the Interagency Ecological Program provided funding and permits for this work. The findings and conclusions are those of the authors and do not necessarily represent the views of the member agencies of the Interagency Ecological Program for the San Francisco Estuary. All authors contributed equally to this work. There is no conflict of interest declared in this article.

REFERENCES

Bennett, W. A. 2005. Critical assessment of the Delta Smelt population in the San Francisco Estuary, California. San Francisco Estuary and Watershed Science [online serial] 3(2).

Bennett, W. A., and J. R. Burau. 2015. Riders on the storm: selective tidal movements facilitate the spawning migration of threatened Delta Smelt in the San Francisco Estuary. Estuaries and Coasts 38:826–835.

CDFW (California Department of Fish and Wildlife). 2010. Threatened and endangered fish. Available: http://www.dfg.ca.gov/wildlife/nongame/t_e_spp/fish.html. (May 2019).

Chadwick, H. K. 1964. Annual abundance of young Striped Bass, *Roccus saxatilis*, in the Sacramento–San Joaquin Delta, California. California Fish and Game Bulletin 50:69–99.

Crone, P., M. Maunder, J. Valero, J. McDaniel, and B. Semmens. 2013. Selectivity: theory, estimation, and application in fishery stock–recruitment models. Center for the Advancement of Population Assessment Methodology, Workshop Series Report 1. Available: https://swfsc.noaa.gov/publications/CR/2013/2013Crone.pdf. (May 2019).

de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs 72:57–76.

Delta Stewardship Council. 2018. The Delta Plan: ensuring a reliable water supply for California, a healthy Delta ecosystem, and a place of enduring value. Available: http://deltacouncil.ca.gov/delta-plan-0. (May 2019).

Feyrer, F., M. L. Nobriga, and T. R. Sommer. 2007. Multidecadal trends for three declining fish species: habitat patterns and mechanisms in the San Francisco Estuary, California, USA. Canadian Journal of Fisheries and Aquatic Sciences 64:723–734.

Feyrer, F., D. Portz, D. Odum, K. B. Newman, T. Sommer, D. Contreras, R. Baxter, S. B. Slater, D. Sereno, and E. Van Nieuwenhuyse. 2013. SmeltCam: underwater video cod end for trawled nets with an application to the distribution of the imperiled Delta Smelt. PLoS (Public Library of Science) ONE [online serial] 8:e67829.

Fregoso, T. A., R.-F. Wang, E. Alteljevich, and B. E. Jaffee. 2017. San Francisco Bay–Delta bathymetric/topographic digital elevation model (DEM). U.S. Geological Survey, Coastal and Marine Geology Program. Available: https://doi.org/10.5066/f7gh9g27. (May 2019).

Gregoire, T. G. 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Fisheries and Aquatic Sciences 28:1429–1447.

Hankin, D. G. 1984. Multistage sampling designs in fisheries research: applications in small streams. Canadian Journal of Fisheries and Aquatic Sciences 41:1575–1591.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey.

Knape, J., P. Besbeas, and P. de Valpine. 2013. Using uncertainty estimates in analyses of population time series. Ecology 94:2097–2107.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70:141–159.

McDonald, T. 2012. Spatial sampling designs for long-term ecological monitoring. Pages 101–125 *in* R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies. Cambridge University Press, New York.

Millar, R. B., and R. J. Fryer. 1999. Estimating the size-selection curves of towed gears, traps, nets, and hooks. Reviews in Fish Biology and Fisheries 9:89–116.

Mitchell, L., K. Newman, and R. Baxter. 2017. A covered cod end and towpath evaluation of midwater trawl gear efficiency for catching Delta Smelt (*Hypomesus transpacificus*). San Francisco Estuary and Watershed Science [online serial] 15(4).
Mitchell, L., K. Newman, and R. Baxter. 2019. Estimating the size selectivity of fishing trawls for short-lived fish species. San Francisco Estuary and Watershed Science [online serial] 17(1).
Moyle, P. B., L. R. Brown, J. R. Durand, and J. A. Hobbs. 2016. Delta Smelt: life history and decline of a once-abundant species in the San Francisco Estuary. San Francisco Estuary and Watershed Science [online serial] 4(2).
Moyle, P. B., and B. Herbold. 1992. Life history and status of Delta Smelt in the Sacramento–San Joaquin Estuary, California. Transactions of the American Fisheries Society 121:67–77.
NatureServe. 2014. *Hypomesus transpacificus*. The IUCN red list of threatened species 2014:e.T10722A18229095. Available: https://doi.org/10.2305/iucn.uk.2014-3.rlts.t10722a18229095.en. (May 2019).
Newman, K. B. 2008. Sample design-based methodology for estimating Delta Smelt abundance. San Francisco Estuary and Watershed Science [online serial] 6(3).
Newman, K. B., S. T. Buckland, B. J. T. Morgan, R. King, D. L. Borchers, D. J. Cole, P. Besbeas, O. Gimenez, and L. Thomas. 2014. Modelling population dynamics. Springer, Methods in Statistical Ecology, New York.
Nobriga, M. L., T. R. Sommer, F. Feyrer, and K. Fleming. 2008. Long-term trends in summertime habitat suitability for Delta Smelt (*Hypomesus transpacificus*). San Francisco Estuary and Watershed Science [online serial] 6(1).
Polansky, L., K. B. Newman, M. L. Nobriga, and L. Mitchell. 2018. Spatiotemporal models of an estuarine fish species to identify patterns and factors impacting their distribution and abundance. Estuaries and Coasts 41:572–581.
Rockriver, A. 2004. Vertical distribution of larval Delta Smelt and Striped Bass near the confluence of the Sacramento and San Joaquin rivers. Pages 97–109 *in* F. Feyrer, L. R. Brown, R. L. Brown, and J. J. Orsi, editors. Early life history of fishes in the San Francisco Estuary and watershed. American Fisheries Society, Symposium 39, Bethesda, Maryland.
Souza, K. 2002. Revision of California Department of Fish and Game's spring midwater trawl and results of the 2002 spring Kodiak trawl. Interagency Ecological Program for the Sacramento–San Joaquin Estuary Newsletter 15(3):44–47.
Thompson, S. K. 2002. Sampling, 2nd edition. Wiley, New York.
Thorson, J. T., A. O. Shelton, E. J. Ward, and H. J. Skaug. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science 72:1297–1310.
U.S. Fish and Wildlife Service. 1993. Endangered and threatened wildlife and plants: determination of threatened status for the Delta Smelt. Federal Register 58(March 5, 1993):12854–12864.

Appendix: Variance of $\hat{\delta}$

Calculation of the variance that accounts for the three sources of uncertainty is similar to the formula used for multistage sample designs (Hankin 1984; Thompson 2002; Newman 2008), which is based on the law of total variance with three levels of variation. To reduce notation, V and E correspond to variance and expected value, respectively.

$$V(\hat{\delta}) = E_1\left[E_2\left(V_3(\hat{\delta}|1,2)\right)\right] + E_1\left[V_2\left(E_3(\hat{\delta}|1,2)\right)\right] + V_1\left[E_2\left(E_3(\hat{\delta}|1,2)\right)\right]. \tag{A.1}$$

The sources of variation, labeled numerically, are (1) between–sample location variation in the ratio estimate of number of fish within a stratum, (2) the randomness in catching fish that are available to the gear, and (3) uncertainty in the estimated probabilities of capture $\hat{p}(L)$.

The equation for $\hat{\delta}$ is as follows (without subscripting for year, month, life stage, gear, and stratum):

$$\hat{\delta} = \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j}\frac{1}{\hat{p}(L_i)}\right)}{\sum_{j=1}^{n} v_j^*}.$$

In equation (A.1), the innermost expectation and variance (at level 3, the variance in the $\hat{p}(L_i)$) refer to the estimated number of fish represented by the ith fish conditional on a known gear selectivity function. The expectation and variance can be approximated as follows:

$$E_3(\hat{\delta}|1,2) \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j}\frac{1}{\hat{p}(L_i)}\right)}{\sum_{j=1}^{n} v_j^*} \tag{A.2}$$

$$\begin{aligned} V_3(\hat{\delta}|1,2) &= \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} V\left(\frac{1}{p(L_i)}\right)\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \\ &\approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j}\frac{1}{\hat{p}(L_i)^4} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2}, \end{aligned} \tag{A.3}$$

where the delta method is used to approximate the quantity $V\left(\frac{1}{p(L_i)}\right)$.

The expectations and variances at the second level (variability in the number of fish caught) are

$$E_3\left(E_3(\hat{\delta}|1,2)\right) = \frac{\sum_{j=1}^{n} E\left(\sum_{i=1}^{c_j}\frac{1}{p(L_i)}\right)}{\sum_{j=1}^{n} v_j^*} \approx \frac{\sum_{j=1}^{n} f_j}{\sum_{j=1}^{n} v_j^*} \tag{A.4}$$

$$V_2\left(E_3(\hat{\delta}|1,2)\right) \approx \frac{\sum_{j=1}^{n} V\left(\sum_{i=1}^{c_j}\left(\frac{1}{p(L_i)}\right)\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j}\frac{1-\hat{p}(L_i)}{\hat{p}(L_i)}\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \tag{A.5}$$

$$E_2\left(V_3(\hat{\delta}|1,2)\right) \approx \frac{\sum_{j=1}^{n} E\left(\sum_{i=1}^{c_j} \frac{1}{\hat{p}(L_i)^4} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{*_j} E[I_i] \frac{1}{\hat{p}(L_i)^4} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \quad \text{(A.6)}$$

$$E_2\left(V_3(\hat{\delta}|1,2)\right) = \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} \frac{1}{\hat{p}(L_i)^3} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} \frac{1}{\hat{p}(L_i)^4} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2}. \quad \text{(A.7)}$$

The term I_i on the right-hand side of equation (A.6) is an indicator variable for whether the ith fish out of *all* c_j^* fish at site i is caught. It has an expected value of $p(L_i)$, which cancels with one of the $p(L_i)$ terms in the denominator yielding the first expression on the right-hand side of equation (A.7). The total number of fish, c_j^*, and their respective lengths are unknown and that expression cannot be calculated. However, the total number of fish of a given length L' can be estimated by $c_{L'}/p(L')$ where $c_{L'}$ is the observed number of length L' fish. As this is the same as summing the $1/p(L_i)$ over the observed catch $1/p(L_i)$ is multiplied against $\frac{1}{p(L_i)^3} V(\hat{p}(L_i))$ yielding the final expression in equation (A.7).

Lastly, expectations and variances are calculated at the first level:

$$E_1\left[E_2\left(V_3(\hat{\delta}|1,2)\right)\right] \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} \frac{1}{p(L_i)^4} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} \frac{1}{\hat{p}(L_i)^4} V(\hat{p}(L_i))\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \quad \text{(A.8)}$$

$$E_1\left[V_2\left(E_3(\hat{\delta}|1,2)\right)\right] \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} \frac{1-p(L_i)}{p(L_i)}\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \approx \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{c_j} \frac{1-\hat{p}(L_i)}{\hat{p}(L_i)}\right)}{\left(\sum_{j=1}^{n} v_j^*\right)^2} \quad \text{(A.9)}$$

$$V_1\left[E_2\left(E_3(\hat{\delta}|1,2)\right)\right] \approx V_1\left[\frac{\sum_{j=1}^{n} f_j}{\sum_{j=1}^{n} v_j^*}\right] \approx \frac{\sum_{j=1}^{n}\left(c_j^* - \hat{\delta} v_j^*\right)^2}{\bar{v}^{*2} n(n-1)}. \quad \text{(A.10)}$$

SUPPORTING INFORMATION

Additional supplemental material may be found online in the Supporting Information section at the end of the article.

Supplemental Materials: Using multistage design-based methods to construct abundance indices and uncertainty measures for Delta Smelt

Leo Polansky, Lara Mitchell, Ken B. Newman

A Subregion Water Volumes and Substitution Orders

Table S1: Estimates of habitat volume, or volume of water occupied by Delta Smelt, by geographic stratum and fish stratum. Volumes are in cubic meters (m^3).

	Fish stratum	
Geographic stratum	Later life stage (0.5 to 4.5 m)	Earlier life stage (0 to 10 m)
Cache Slough and Liberty Island	51,786,023	90,039,906
Carquinez Strait	60,455,559	135,019,878
Disappointment Slough	14,107,778	18,995,896
East San Pablo Bay	104,537,750	175,671,563
Franks Tract	52,701,925	71,232,869
Grant Line Canal and Old River	7,313,463	9,635,826
Holland Cut	17,642,507	27,809,216
Honker Bay	55,100,817	101,141,758
Lower Napa River	24,372,905	40,588,808
Lower Sacramento River	71,561,907	147,188,708
Lower San Joaquin River	76,919,425	141,250,258
Mid Suisun Bay	134,714,482	214,551,584
Middle River	9,000,880	13,707,843
Mildred Island	35,712,993	52,829,804
North and South Forks Mokelumne River	34,680,881	52,688,223
Old River	9,991,399	14,659,405
Rock Slough and Discovery Bay	3,718,423	4,718,521
Sacramento River near Rio Vista	45,878,622	83,461,347
Sacramento River near Ryde	12,833,585	18,948,518
Sacramento River Ship Channel	14,472,933	29,744,374
San Joaquin River at Prisoners Pt	36,436,501	67,727,034
San Joaquin River at Twitchell Island	32,369,636	66,601,478
San Joaquin River near Stockton	21,986,848	39,996,300
Suisun Marsh	30,289,939	47,763,576
Upper Napa River	800,061	1,733,454
Upper Sacramento River	37,840,015	57,161,007
Upper San Joaquin River	3,537,223	4,463,237
Victoria Canal	8,238,349	11,384,303
West Suisun Bay	89,106,803	172,557,863

Table S2: Percentage of habitat volume sampled by survey and month based on effective sample volumes.

Survey	Month	Mean	Min	Max
20-mm	May	0.012	0.008	0.017
20-mm	Jun	0.013	0.010	0.021
STN	Jun	0.008	0.004	0.018
STN	Jul	0.010	0.005	0.016
STN	Aug	0.008	0.002	0.015
STN	JulAug	0.016	0.005	0.026
FMWT	Sep	0.038	0.029	0.051
FMWT	Oct	0.040	0.034	0.051
FMWT	Nov	0.038	0.029	0.049
FMWT	Dec	0.037	0.030	0.046
FMWT	OctNov	0.077	0.063	0.100
SMWT	Jan	0.038	0.032	0.047
SMWT	Feb	0.039	0.031	0.053
SMWT	JanFeb	0.065	0.035	0.100
SMWT	JanFebMar	0.104	0.073	0.146
SMWT	FebMar	0.074	0.040	0.099
SKT	Jan	0.016	0.013	0.021
SKT	Feb	0.017	0.013	0.020
SKT	Mar	0.018	0.012	0.023
SKT	Apr	0.018	0.013	0.021
SKT	May	0.018	0.012	0.022
SKT	JanFeb	0.032	0.018	0.040
SKT	JanFebMar	0.049	0.040	0.061
SKT	FebMar	0.034	0.025	0.041

Table S3: List of subregion substitutions used in constructing abundance indices. The "Missing Subregion" is the subregion without data. Density and estimates from the first available "Substitute Subregion" were used as substitutes for the missing density. Abundance indices used the volume data of the missing subregion times density from the substitute region. Similarly, the variance of a missing subregion used the volume of that missing subregion.

Missing subregion	Substitute subregion
East San Pablo Bay	Carquinez Strait
East San Pablo Bay	Mid Suisun Bay
Upper Napa River	Lower Napa River
Upper Napa River	Carquinez Strait
Upper Napa River	West Suisun Bay
Upper Napa River	Mid Suisun Bay
Lower Napa River	Upper Napa River
Lower Napa River	Carquinez Strait
Lower Napa River	West Suisun Bay
Lower Napa River	Mid Suisun Bay
Carquinez Strait	West Suisun Bay
Carquinez Strait	East San Pablo Bay
Carquinez Strait	Lower Napa River
Carquinez Strait	Mid Suisun Bay
West Suisun Bay	Mid Suisun Bay
Mid Suisun Bay	West Suisun Bay
Suisun Marsh	Mid Suisun Bay
Suisun Marsh	Honker Bay
Suisun Marsh	Lower Sacramento River
Honker Bay	Mid Suisun Bay
Honker Bay	Suisun Marsh
Honker Bay	Lower Sacramento River
Lower Sacramento River	Lower San Joaquin River
Lower Sacramento River	Honker Bay
Lower Sacramento River	Suisun Marsh
Lower San Joaquin River	Suisun Marsh
Lower San Joaquin River	Lower Sacramento River
Lower San Joaquin River	Honker Bay
Sacramento River Ship Channel	Cache Slough and Liberty Island
Sacramento River Ship Channel	Upper Sacramento River
Sacramento River Ship Channel	Sacramento River near Ryde
Sacramento River Ship Channel	Sacramento River near Rio Vista
Sacramento River Ship Channel	San Joaquin River at Twitchell Island
Sacramento River near Rio Vista	Cache Slough and Liberty Island
Sacramento River near Rio Vista	San Joaquin River at Twitchell Island
Sacramento River near Ryde	Upper Sacramento River

Table S3 (continued)

Sacramento River near Ryde	Sacramento River near Rio Vista
Sacramento River near Ryde	Cache Slough and Liberty Island
Sacramento River near Ryde	San Joaquin River at Twitchell Island
Upper Sacramento River	Cache Slough and Liberty Island
Upper Sacramento River	Sacramento River near Ryde
Upper Sacramento River	Sacramento River near Rio Vista
Upper Sacramento River	San Joaquin River at Twitchell Island
Cache Slough and Liberty Island	Sacramento River Ship Channel
Cache Slough and Liberty Island	Lower San Joaquin River
Cache Slough and Liberty Island	San Joaquin River at Twitchell Island
Cache Slough and Liberty Island	Sacramento River near Rio Vista
San Joaquin River at Twitchell Island	Lower San Joaquin River
San Joaquin River at Twitchell Island	Cache Slough and Liberty Island
Franks Tract	Holland Cut
Franks Tract	San Joaquin River at Prisoners Pt
Franks Tract	San Joaquin River at Twitchell Island
North and South Forks Mokelumne River	San Joaquin River at Prisoners Pt
North and South Forks Mokelumne River	Sacramento River near Ryde
North and South Forks Mokelumne River	Upper Sacramento River
North and South Forks Mokelumne River	Disappointment Slough
San Joaquin River at Prisoners Pt	Holland Cut
San Joaquin River at Prisoners Pt	Middle River
San Joaquin River at Prisoners Pt	Old River
San Joaquin River at Prisoners Pt	Mildred Island
Holland Cut	San Joaquin River at Prisoners Pt
Holland Cut	Middle River
Holland Cut	Old River
Holland Cut	Mildred Island
Middle River	Mildred Island
Middle River	Old River
Middle River	Holland Cut
Middle River	San Joaquin River at Prisoners Pt
Upper San Joaquin River	San Joaquin River near Stockton
Upper San Joaquin River	Disappointment Slough
Upper San Joaquin River	Middle River
Upper San Joaquin River	Mildred Island
Upper San Joaquin River	North and South Forks Mokelumne River
Upper San Joaquin River	Sacramento River near Ryde
Victoria Canal	Old River
Victoria Canal	Middle River
Victoria Canal	Grant Line Canal and Old River

Table S3 (continued)

Victoria Canal	San Joaquin River near Stockton
Victoria Canal	Rock Slough and Discovery Bay
Grant Line Canal and Old River	Victoria Canal
Grant Line Canal and Old River	Middle River
Grant Line Canal and Old River	Old River
Grant Line Canal and Old River	San Joaquin River near Stockton
Grant Line Canal and Old River	Rock Slough and Discovery Bay
San Joaquin River near Stockton	Victoria Canal
San Joaquin River near Stockton	Grant Line Canal and Old River
San Joaquin River near Stockton	Rock Slough and Discovery Bay
Disappointment Slough	North and South Forks Mokelumne River
Disappointment Slough	Upper San Joaquin River
Disappointment Slough	San Joaquin River at Prisoners Pt
Disappointment Slough	Sacramento River near Ryde
Rock Slough and Discovery Bay	Old River
Rock Slough and Discovery Bay	Victoria Canal
Rock Slough and Discovery Bay	Holland Cut
Rock Slough and Discovery Bay	Grant Line Canal and Old River
Rock Slough and Discovery Bay	San Joaquin River near Stockton
Old River	Holland Cut
Old River	Franks Tract
Old River	Mildred Island
Old River	San Joaquin River at Prisoners Pt
Old River	Middle River
Mildred Island	Old River
Mildred Island	Middle River
Mildred Island	Holland Cut
Mildred Island	San Joaquin River at Prisoners Pt
Mildred Island	Franks Tract

B Survey Station Locations

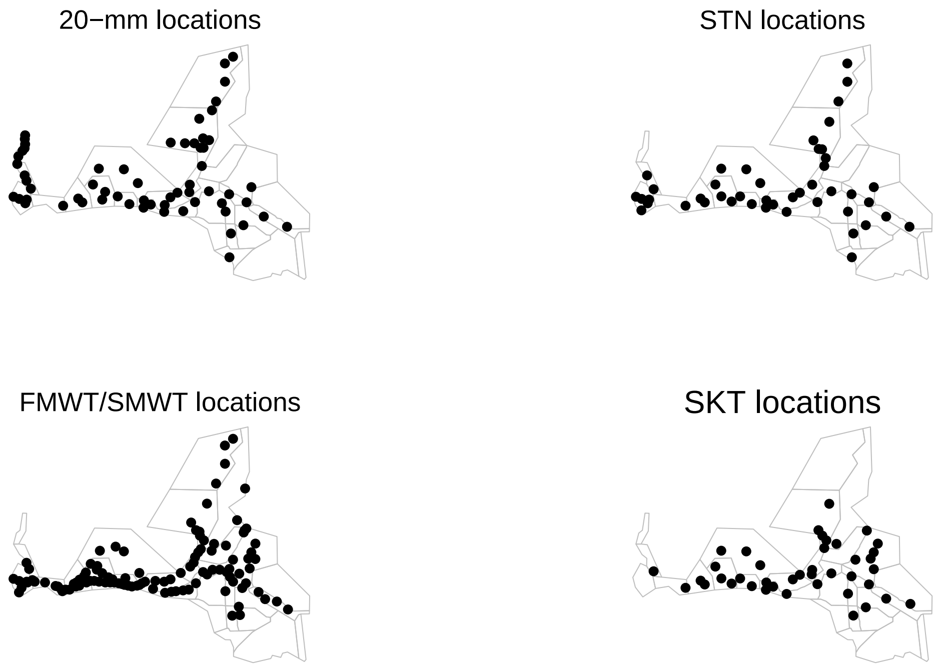

Figure S1: Station locations for each survey.

C Data Processing

This section provides a brief overview of the data sets used to calculate design-based estimates. We started with survey-specific files containing catch and length data provided by CDFW. The 20-mm and SKT surveys periodically conduct investigative or experimental surveys; we removed data from these supplemental surveys and retained data from routine surveys, which correspond to annual survey numbers 1 through 5 for SKT and 1 through 9 for 20-mm. Each survey program (20-mm, STN, FMWT, SKT) has core stations that have been sampled since the beginning of the survey as well as non-core stations that have been consistently sampled starting in more recent years. We retained data from both core and non-core stations. We also retained stations that were sampled sporadically but were not part of a complete supplemental survey.

We imputed missing or physically unrealistic values of tow volume (i.e., volume of water sampled in a tow), station depth (i.e., depth to the bottom of the sampling location), and "cable out", which is the amount of cable let out when conducting an oblique tow (see section "Sample volume adjustments"). Mean values, calculated at the finest spatiotemporal resolution possible, were used as substitute values. The finest resolution we considered for these variables was date-station. We also imputed fork lengths for Delta Smelt that were not measured for length. If other Delta Smelt were caught and measured in the same tow, we used the mean fork length from that tow, otherwise we used the mean fork length calculated for a given year-survey number combination or for a given month (calculated across years), if necessary.

Some of the tow depth values that were calculated as described in the section "Sample volume adjustments" were physically unrealistic and in these cases we replaced the unrealistic values as follows. If a calculated tow depth was greater than station depth, we replaced the calculated tow depth with station depth. When cable out values are at the low end of the range (e.g., 75 feet), the corresponding tow depth can be less than the mouth height of the net. If the net does break the surface of the water during sampling, the crew will slow the boat and increase the cable angle to keep the net fully submerged (T. Morris, personal communication, February 23, 2016). As a result, if a calculated tow depth was less than the mouth height, we replaced the calculated tow depth with the mouth height.

D Organization of R Code and Output

Accompanying this document are input data and R (R Foundation for Statistical Computing, Vienna, Austria) code needed to run this analysis. They are contained in the directory `code_and_data`. Everything can be run from the file `run_vX.r`, where `vX` denotes a version number. The file `DataCleaner_FishSurveys_vX.r` does the initial data processing and the file

`Design_based_abund_calc_vX.r`, which depends on the file `Abund_util_vX.r`, calculates the design-based abundance indices. This analysis produces three csv files and one RData file:

`DB_abundance_long_vX_DATE.csv`

`DB_abundance_wide_vX_DATE.csv`

`DB_abundance_wide_cohort_vX_DATE.csv`

`Design_based_abund_calc_vX_DATE_Everything.RData`

where `vX` represents the version number from the `Design_based_abund_calc_vX.r` script and `DATE` represents the date on which the file was generated. The csv files contain the same data but are organized differently. The first has a separate record for each combination of calendar year, month, gear type, and Delta Smelt age class. The second has a separate row for each calendar year and different columns for different combinations of survey type, month, and age class. The third file is similar to the second file except that each row corresponds to a different cohort year, where a cohort year is defined roughly from March of the year the cohort was born to June of the following year. The .RData is a copy of the `R` workspace after all objects area loaded in and the calculations are executed.

E Simulation Study to Evaluate the Use of a Lognormal Distribution in Abundance Approximation

Catch data and indice estimates were simulated under different scenarios of gear selectivity curves. An adjustment related to the use of effective volume was not included because this is not treated as a source of variability in the estimation process. Further, availability was assumed to be 100% thus total abundances, not just indices, were estimated.

Fish lengths were scaled to lie between 0 and 1. Nine different selectivity curves, intended to cover a wide range of possible gear efficiencies and dependencies (or the near lack of) on fish length, were used to simulate catch (Figure S2). The pseudo-code in Box E1 describes how catch abundances were simulated, and parameter values shown in Table S4. These values were selected to approximate the Delta Smelt survey efforts and data. For the choice of H, V_h, and δ_h, the simulated baseline abundance was N = 102,000 (5,100 per each of the 20 strata).

A total of 1,000 simulations for each gear selectivity choice were made. The distributions were right-skewed with the degree of skewness varying as a function of the contact selectivity parameters (Figure S3). Estimates of N_{Tot} equal to zero resulted only for the selectivity curves closest to zero across much of the range of fish lengths (e.g., β_0=0.9 and β_1=10, Table S7). The bias was relatively low, ranging from -1.7% to 2.2% (Table S5). The coefficient of variation could be relatively large, ranging from 37% to 91% (Table S6). Actual coverage of the 95% confidence intervals based on the lognormal distribution was affected by the contact selectivity function with exact coverage for the mid-range intercept (β_0=0.5), slightly low for the negative intercept (β_0 = -0.5), and too high (97 to 100%) for the largest intercept (β_0=0.9, Table S4). On the other hand, confidence intervals based on a normal distribution yielded negative lower bounds with increasing probability as β_0 increased (Table S4).

Box E1: Pseudo-code to simulate indice estimates

(1) Choose a total number of strata H and stratum specific densities δ_h with which to set the simulated baseline abundances N_h and total $N_{Tot} = \sum_h N_h$.

(2) For h in $1, ..., H$

 (i) For j in $1, ..., n_{h,j}$

  (a) Simulate the baseline abundance in the sampled volume of water v_s according to a negative binomial distribution, $y_{h,j} \sim \text{NegBin}(\mu = \delta_h * v_s, \theta)$. This simulates random *potential* catch level variation.

  (b) If $y_{h,j} > 0$

   A. Assign lengths to each of the $y_{h,j}$ fish in the patch of water sampled according to a length distribution, $L_{h,j,i} \sim \text{Beta}(\alpha_1, \alpha_2)$.

   B. For i in $1, ..., y_{h,j}$ simulate a Bernoulli random variable $I_{h,j,i} \sim \text{Bern}(p = p_g(L_{h,j,i}))$ and assign these fish to the total catch $c_{h,j} = \sum_i I_{h,j,i}$. This step simulates a random total catch according to the gear selectivity function, $p_g(L)$, which was modeled with a logit transform: $\text{logit}(p_g(L)) = \beta_1(L - \beta_0)$. The lengths of the specific fish assigned to the total catch are recorded.

   C. Compute the adjusted catch $c^*_{h,j} = \sum_i^{c_{h,j}} 1/p_g(L_{h,j,i})$.

 (ii) Compute the estimated stratum density as a ratio of means, $\hat{\delta}_h = \frac{\sum_j^{n_{h,j}} c^*_{h,j}}{\sum_j^{n_{h,j}} v_{s,j}}$.

 (iii) Compute the estimated stratum total $\hat{N}_h = \hat{\delta}_h V_h$ and the estimated variance $\widehat{\text{Var}}(\hat{N}_h)$ according to the Appendix in the main text.

(3) Estimate the total abundance $\hat{N}_{Tot} = \sum_h \hat{N}_h$ and total variance $\widehat{\text{Var}}(\hat{N}_{Tot}) = \sum_h \widehat{\text{Var}}\hat{N}_h$.

(4) Use the total abundance and total variance estimates to parameterize normal and lognormal distributions for confidence interval construction, check to see if N_{Tot} falls within the confidence intervals, check if the lower conidence intervals based on a normal distribution are negative, and compute other summary statistics.

Table S4: Parameter values used for the simulation study.

Parameter	Value	Description
V_h	$5.1 \times 10^7 \mathrm{m}^3$	Stratum volume, a constant.
δ	0.005 fish/m^3	True density in each stratum.
θ	0.2	Dispersion parameter for the negative binomial distribution used to simulate stratum values. Parameterized so that $\mathrm{Var}(y_{h,j}) = \delta * v_s + (\delta * v_s)^2/\theta$.
v_s	10,000m^3	Sample volume.
H	20	Total number of strata.
$n_{h,j}$	3	Number of replicate samples per stratum.
$\alpha_1 (= \alpha_2)$	60	Shape parameters for the beta distribution assigning lengths to the fish in each stratum. The expected length of the fish in each stratum is 0.5 and the variance is $\alpha_1\alpha_2/((\alpha_1 + \alpha_2)^2(\alpha_1 + \alpha_2 + 1))$.
β_0	-0.5, 0.5, 0.9	Mid-point parameter for the selectivity function $p_g(L)$, the length at which an individual has a 0.5 probability of being captured. Negative values have the effect of making fish (with lengths between 0 and 1) have a high and nearly constant value of being captured. See Figure S2.
β_1	1, 5, 10	Slope parameter of the selectivity function $p_g(L)$. See Figure S2.
$\mathrm{Var}(\hat{p}_g(L))$	0.07	Variance of the selectivity curve estimate. This was made constant across fish lengths and chosen from the larger values of the empirically estimated ones.

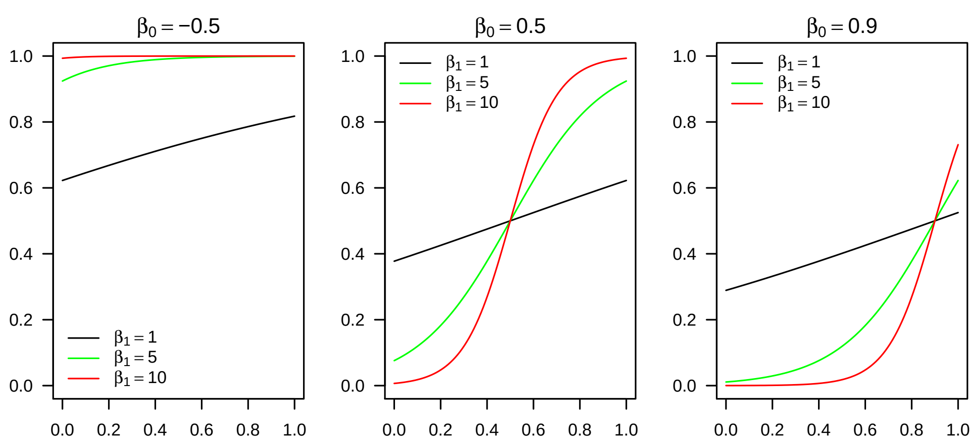

Figure S2: Selectivity curves, $p_g(L)$, used in the simulation study. Lengths (between 0 and 1) are on the x-axis and the probability of capture is on the y-axis. The mid-point parameter value is printed above each panel, with three different slope parameter values used per mid-point parameter value.

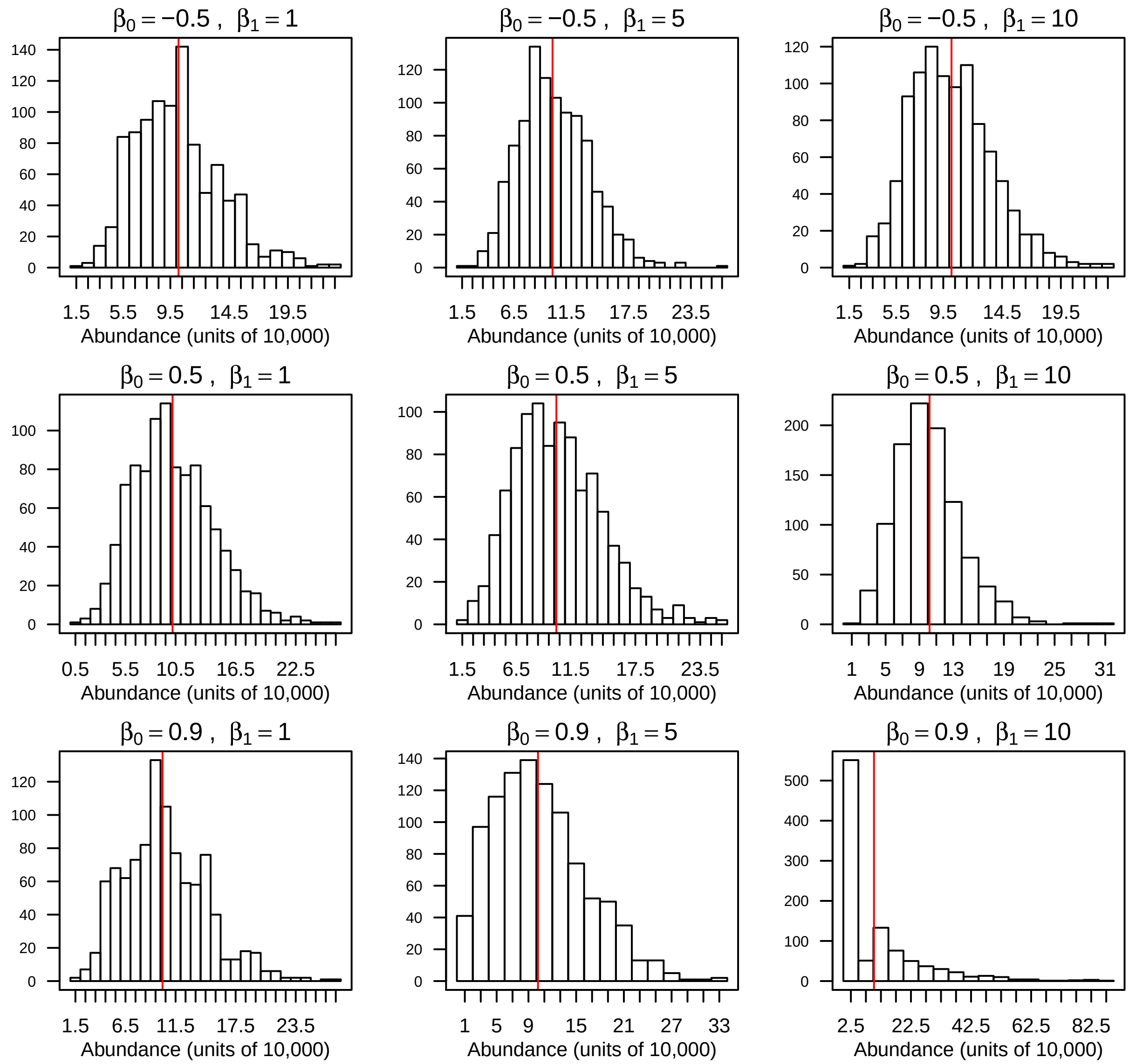

Figure S3: Histograms of abundance point estimates $\hat{N}_{Tot}$ from 1,000 simulations based on the nine different selectivity curves. The red lines are drawn at the value of the baseline total N_{Tot}.

Table S5: Relative percent bias of $\hat{N}_{Tot}$ ($[\hat{N}_{Tot} - N_{Tot}]/N_{Tot}$ * 100) by gear selectivity curve.

	$\beta_1 = 1$	$\beta_1 = 5$	$\beta_1 = 10$
$\beta_0 = -0.5$	-1.39	0.67	-1.67
$\beta_0 = 0.5$	0.25	1.08	-1.42
$\beta_0 = 0.9$	2.21	1.29	-1.44

Table S6: Mean coefficient of variation of those estimates with nonzero point estimates, i.e., those which had at least one nonzero adjusted catch value; see Table S7 for the proportions of simulations with abundance indices of zero.

	$\beta_1 = 1$	$\beta_1 = 5$	$\beta_1 = 10$
$\beta_0 = -0.5$	0.41	0.37	0.40
$\beta_0 = 0.5$	0.48	0.47	0.46
$\beta_0 = 0.9$	0.47	0.79	0.91

Table S7: Proportion of the simulations with zero abundance index (i.e., the proportion of times that no fish were caught in 60 tows.

	$\beta_1 = 1$	$\beta_1 = 5$	$\beta_1 = 10$
$\beta_0 = -0.5$	0	0	0
$\beta_0 = 0.5$	0	0	0
$\beta_0 = 0.9$	0	0.03	0.55

F The Effect of Truncation

Table S8: Ratios of non-truncated abundance to truncated abundance by survey and time period of data collection. Missing entries correspond to time periods during which no data were collected.

	20-mm		STN			
Year	May	June	June	July	August	July-August
1995	1.20	2.00		1.25	1.25	1.50
1996	1.07	1.96		1.34	1.34	1.34
1997	1.18	3.83	1.21	1.53	1.53	1.53
1998	1.04	3.44		1.58	1.58	1.66
1999	1.13	1.57		1.35	1.35	1.38
2000	1.06	2.67	1.08	1.38	1.38	1.42
2001	1.11	3.68	1.10	1.39	1.39	1.39
2002	1.27	6.94	1.26	1.37	1.37	1.49
2003	1.07	1.84	1.06	1.49	1.49	1.64
2004	1.26	4.97	1.23	1.52	1.52	1.66
2005	1.15	5.55	1.22	1.58	1.58	1.71
2006	1.10	2.65	1.14	1.21	1.21	1.23
2007	1.05	5.58	1.21	1.62	1.62	1.70
2008	1.63	10.27	1.18	1.67	1.67	1.81
2009	1.20	3.87	1.34	1.60	1.60	1.92
2010	1.30	3.08	1.45	1.70	1.70	1.76
2011	1.15	1.53	1.04	1.26	1.26	1.31
2012	1.06	1.54	1.07	1.26	1.26	1.33
2013	1.50	4.08	1.18	1.79	1.79	1.81
2014	1.81	6.05	1.18	1.70	1.70	1.79
2015	2.43	5.34	1.60			1.95
2016	2.10	10.47	1.40			1.67
2017	1.60	7.76	1.82	1.56	1.56	1.81

Table S9: Pearson correlations between non-truncated and truncated abundance indices using pair-wise complete observations.

20-mm		STN			
May	June	June	July	August	July-August
1.00	0.90	1.00	1.00	1.00	1.00

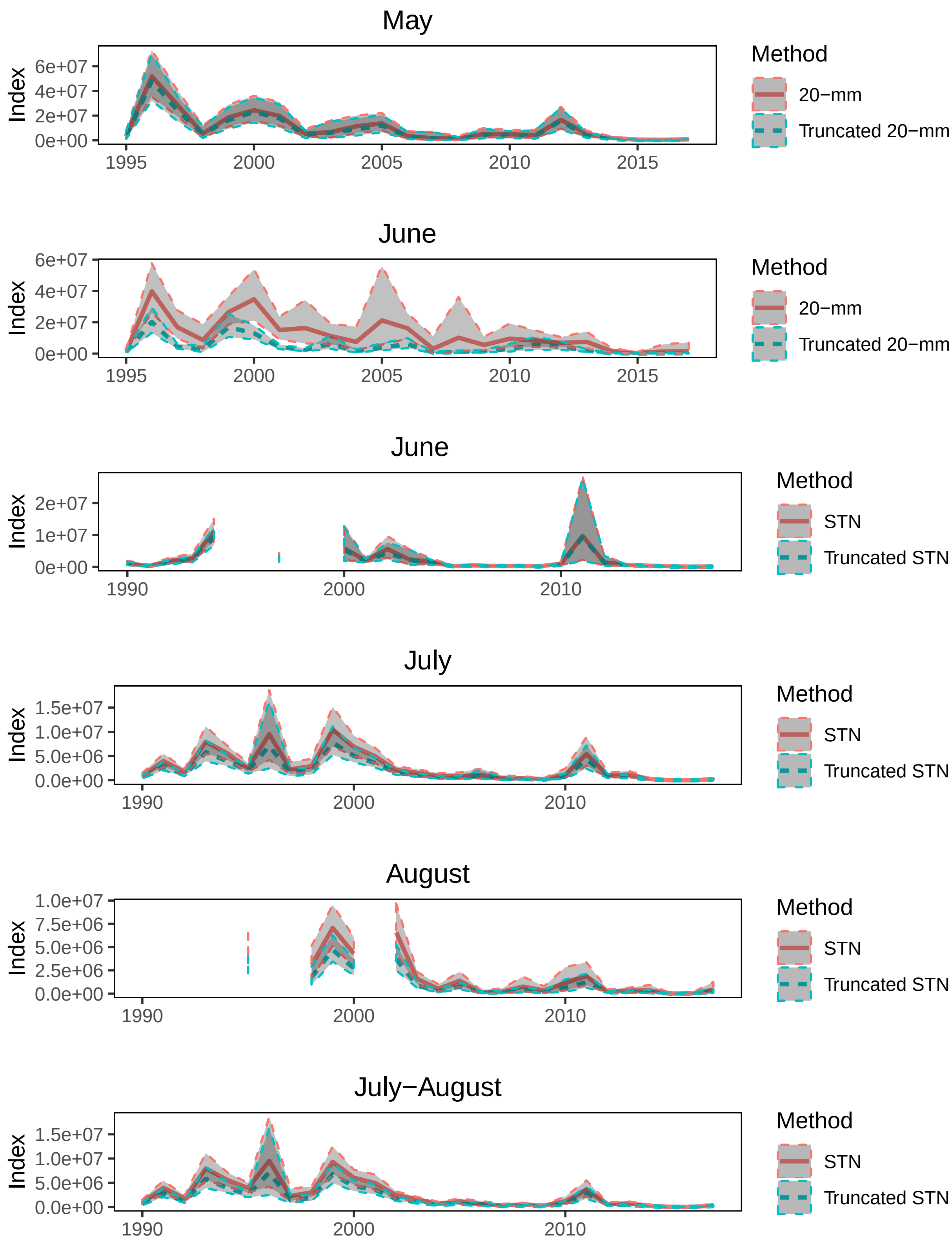

Figure S4: Abundance time series plots and 95 confidence envelopes based on non-truncated and truncated selectivity curves.

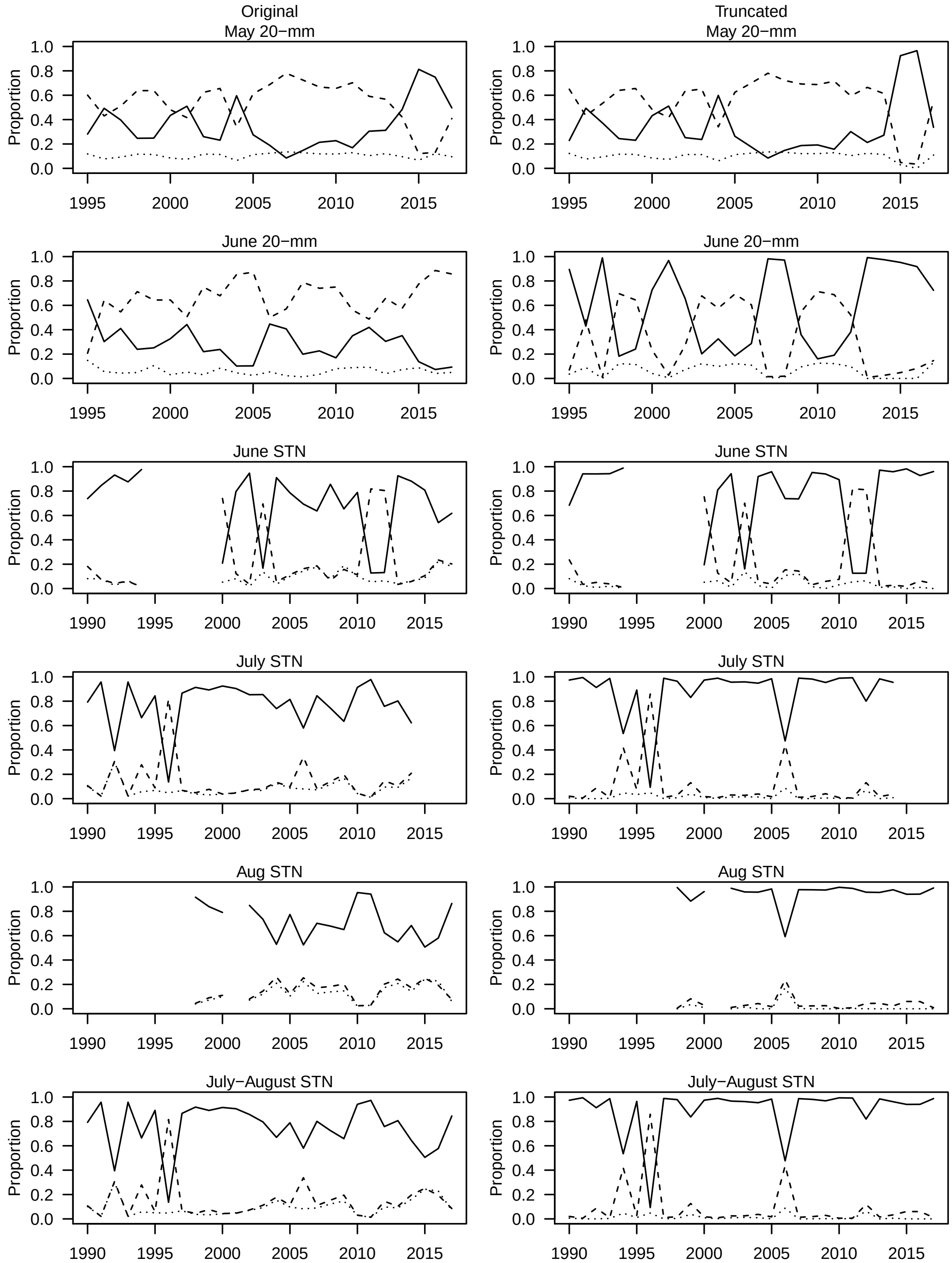

Figure S5: Proportion of the total variance of the estimated population abundance by gear type and non-truncated and truncated based catch adjustments for each of the three sources of variation: between sample variation (source 1, solid lines), the randomness that a fish present in the tow volume will be caught (source 2, dashed lines), and the variability in the estimate of selectivity curve (source 3, dotted lines). For ease of comparison the non-truncated figures are repeated here as well as in the main text (compare with Figure 5 of the main text.