



# A replication approach to controlled selection for catch sampling intercept surveys

Katherine J. Papacostas<sup>a,b</sup>, John Foster<sup>a,\*</sup>

<sup>a</sup> National Atmospheric and Oceanic Administration Fisheries, Office of Science and Technology, 1315 East-West Highway Silver Spring, MD, 20190, USA

<sup>b</sup> ECS, 2750 Prosperity Ave, Fairfax, VA, 22031, USA

## ARTICLE INFO

Handled by: A.E. Punt

### Keywords:

Intercept survey  
Recreational fishing  
Probability sampling  
Controlled selection  
Catch estimation

## ABSTRACT

Intercept surveys following probability sampling designs are widely used for fisheries data collection, but limited resources often make implementing such designs challenging. Adjustments that field staff make after sample selection may create deviations from the initial sampling design, introducing data quality concerns. Currently available controlled selection techniques designed to handle logistical constraints can be complicated to develop and implement, and limited in versatility and practicality. We examined the utility of a novel constrained draw replication approach for selecting sampling units for intercept surveys, which first uses standard probability sampling to draw a large set of initial replicated samples. This set is then screened against customizable constraints to create a survivor subset prior to selecting a final sample for the survey. Using a simulation study, we tested the performance of this method in comparison to a standard probability sampling design for selecting site-day sampling units. Estimates obtained from the new method were as accurate and precise as the standard probability sampling design at high levels of replication. This flexible method may help fisheries scientists maintain better control of the sample distribution while allowing for calculation of appropriate design-based sample weights, and increase the design options available for fisheries intercept surveys.

## 1. Introduction

Catch sampling intercept surveys, where anglers or vessels are intercepted upon return from fishing trips, are commonly used in commercial catch sampling and recreational fisheries data collection programs worldwide (e.g., Griffiths et al., 2014; ICES, 2013; NAS, 2017; Strehlow et al., 2012), and are even being considered for small-scale fisheries monitoring in developing countries (e.g., Vølstad et al., 2014). Such surveys are typically used to produce catch and/or effort estimates, which are key sources of data used in stock assessments and fisheries management. Ensuring that these statistics are precise, unbiased and of the highest quality possible is of utmost importance in being able to make accurate stock status predictions and formulate effective strategies to ensure the sustainability of fisheries resources (NMFS, 2013).

Catch sampling intercept surveys are frequently conducted onshore at fishing access sites. These sites could include anything from docks or beaches (e.g., to intercept recreational or small-scale anglers or vessels), to ports or markets where commercial catch is landed and sold (e.g., to intercept commercial anglers or vessels). There are two standard design options for these onshore intercept surveys (ICES, 2013): 1) using site as

the primary sampling unit, or 2) using site-day as the primary sampling unit when desiring day-to-day variation to be reflected in the sample. Intercept surveys also tend to involve standard survey techniques like stratified sampling, (i.e., dividing the population into subpopulations before sampling, [Kish, 1965]), cluster sampling (i.e., sampling groups of units from a population that are linked in a logical or convenient way – household members, fishing trips returning to the same port or access site, etc. [Kalton, 1979]), and multi-stage sampling, (i.e., sampling in hierarchical levels/stages, using increasingly smaller sample units at each stage, [Kuno, 1976]). For example, when site-day is the primary sampling unit (stage 1), fishing trips may be subsampled within site-days (stage 2), and then catch may be subsampled within fishing trips (stage 3). In some cases, days may also be subsampled by splitting times into manageable shifts (e.g., Vølstad et al., 2006).

No matter what primary sampling unit is chosen, intercept surveys will often follow probability sampling designs, also referred to as random sampling designs, where inclusion probabilities, or the chances of each population element being selected as part of the sample, are known. Probability sampling designs have a number of significant benefits over censuses (full enumeration) and non-probability sampling designs (convenience/purposive/subjective/non-random sampling).

\* Corresponding author.

E-mail addresses: [katherine.papacostas@noaa.gov](mailto:katherine.papacostas@noaa.gov) (K.J. Papacostas), [john.foster@noaa.gov](mailto:john.foster@noaa.gov) (J. Foster).

First, probability sampling tends to be much more cost-effective than censuses (NRC, 2009), requiring less time and less resources to complete, with a much lower total reporting burden on respondents. Further, since the inclusion probability of each sampling unit is known in probability designs, units can be weighted appropriately (e.g. equally in a simple random sampling design, or unequally in more complicated probability sampling designs), which allows for the production of design unbiased estimates. As such, probability sampling methods are statistically superior to non-probability sampling methods, where selection bias is usually a concern. Selection bias often occurs in non-probability sampling since the selection mechanism (i.e., how sampling units are selected for the sample) cannot be accounted for in estimation, and thus the sample is less likely to be representative of the population intending to be analyzed (Baker et al., 2013). Another factor unique to probability surveys is the ability to quantify the precision of estimates (Fuller, 2009a). Such survey designs with known inclusion probabilities for sampling units are internationally considered 'best practices' for statistically sound fisheries sampling programs (ICES, 2013). Further, probability-based sampling is a key design component listed in the Quality Assessment Tool, an international standard used to evaluate recreational fisheries data quality (ICES, 2020).

While theoretically ideal, designing and adhering to probability sampling designs in practice can be a technical and logistical challenge for many fisheries sampling programs with limited resources for survey operations (staff, funds, etc.). To varying degrees of success, there are a number of methods that fisheries survey designers can use to ease implementation challenges. For instance, the standard design techniques commonly used in intercept surveys (stratification, multi-stage sampling, and cluster sampling) can incorporate some logistical constraints into the base design. However, these approaches are not necessarily flexible, and are developed on a case-by-case basis. They may therefore have limitations that make them unworkable outside of the one particular use for which the design was developed. Further, certain logistical issues might require additional controls that are not possible to incorporate by altering the base design alone (Rao and Nigam, 1990). For example, in designing an intercept survey for a large region with sites or site-days as the primary sampling unit, the survey designer may stratify the sample geographically (e.g. by state or county), and/or temporally (e.g. by month or week) to subdivide the sample into more manageable parts. However, even with stratification, trying to schedule potentially large numbers of sampling assignments for a relatively small group of samplers to complete in a specific timeframe can be very difficult (since standard designs cannot take things like the number of available samplers at a given time into account). In this instance, an angler intercept survey with a pure probability design could easily lead to the selection of either too many fishing sites, or sites that are too geographically distant from one another (e.g. on the opposite sides of a state or county) for the sampling team to realistically visit in a given day or week. Faced with an unmanageable schedule, compromises may be made to adjust or not complete assignments, but any spontaneous changes or adjustments to sampling assignments in the field will cause deviations from the inclusion probabilities that may be difficult or impossible to account for when calculating final sample weights. Inaccuracies in the sample weights create the potential for bias, bringing the statistical integrity of subsequent estimates into question (NAS, 2017; NRC, 2006).

A real-world example of this exact issue was the National Marine Fisheries Service's Marine Recreational Fisheries Statistics Survey (MRFSS). The MRFSS was the federal recreational fisheries monitoring program in the United States, operating from 1981 – 2008. The MRFSS consisted of various component surveys, including a complex angler intercept survey with a probability proportional to size (PPS) sampling design; in PPS sampling, each primary sampling unit is assigned a size measure (which in the case of the MRFSS Intercept Survey, was expected levels of fishing activity at each site) and each sampling unit's inclusion probability is proportional to its size-measure (Nigam et al., 1984; Yates and Grundy, 1953). Among a variety of other issues

identified in the program (see NRC, 2006), the MRFSS Intercept Survey struggled with adhering to design-based inclusion probabilities. The MRFSS Intercept Survey allowed field samplers to make certain judgments about where and when to sample, and they often visited sites that were not originally selected to be part of the sample in an effort to reduce costs and improve productivity in terms of numbers of completed intercepts. As such, the implementation of the probability design included non-probability deviations, which had unknown impacts to estimates of catch and effort and introduced the potential for bias (NRC, 2006). Significant resources and approximately a decade of work were needed to address these issues through a redesign of the entire program (NAS, 2017).

When standard design components are not enough to ease implementation challenges, survey designers may turn to controlled selection. Controlled selection is a broad class of sample selection methods, first introduced by Goodman and Kish (1950), that increases the probability of selecting a desirable sample, while preserving the properties of a probability sample (i.e. retaining the ability to calculate appropriate sample weights to permit unbiased estimation). The idea behind the method is to minimize or even eliminate the selection of undesirable samples (i.e., samples that consist of combinations of sampling units that are costly or logistically difficult sample). Already mentioned, PPS is one standard example of controlled selection, where size measures are used to limit selection of undesirable samples (in the case of the MRFSS Intercept Survey, undesirable samples were those that frequently included sites with little to no fishing activity [NRC, 2006], since it was not a good use of funds to send samplers to low-activity sites too often). Nearest PPS is a more extreme version of standard PPS, where all undesirable combinations of sampling units have a zero probability of selection (Gabler, 1987). Other controlled selection methods include the use of dependent selection (i.e., the choice of a particular sampling unit influences the selection of another sampling unit: see Hess, 1961; Waterton, 1983). In more recent years, sophisticated linear (e.g., Rao and Nigam, 1992) and nonlinear (e.g., Tiwari et al., 2007) programming approaches have also been developed that essentially apply a suite of constraints to one or more objective functions (e.g., costs). Rejective sampling, a special kind of controlled selection, can further be used to prevent undesired combinations of sampling units (Fuller, 2009b). In rejective sampling, single sampling units are selected in a set of successive sample draws, each with its own probability, and any sampling units that do not fit a specified criteria, based on the sampling units already selected, are rejected.

Many of these established approaches, although useful, have a number of limitations and disadvantages. For example, PPS methods are usually based on one primary size measure, making it difficult to incorporate additional constraints if needed. The current linear and non-linear programming techniques can be extremely computationally intensive, and have limited applications in sampling large populations and/or if a survey has large target sample sizes (Arnab, 2013). In certain sampling schemes, there is no a priori guarantee that the set of desired controls can be achieved (e.g., Waterton, 1983) although replacement sampling, where sampling units can be drawn multiple times to be part of the sample, can be used (Hartigan, 1971). Further, more complex controlled selection designs require more complicated estimation methods (Tiwari and Chilwal, 2014). Exploration of how to improve the versatility and practicality of controlled selection methods has been limited in both the fisheries and survey statistics literature more broadly.

Here, we present a simulation study testing a novel, constrained draw replication approach to probability sampling in intercept surveys using site-day as the primary sampling unit. The approach closely approximates a standard probability sampling design, but incorporates logistical constraints directly into the sample selection process. First, standard probability sampling methods are used to draw, or generate, a large set of initial replicated samples, each consisting of different combinations of sampling units. Then, these replicated samples are

screened against a set of constraints to create a survivor subset of replicated samples. From this survivor subset of replicated samples, a final selection is made as the official sample to be used for the survey. The method works with any common probability-based sampling design; takes less computing power than other modern controlled selection techniques; is adaptable to different spatial scales; allows for additional constraints to be added to the base survey design, and; allows for straight-forward estimate production using any standard weighted estimator. Further, it can remove the logistical need for post-hoc adjustments, allowing samplers to more easily adhere to the probability sampling design in the field. Such a method may be of widespread utility to fisheries scientists seeking to improve their control of sample distributions for easier calculation of appropriate inclusion probabilities and unbiased, design-based estimates when implementing complex intercept surveys.

## 2. Methods

### 2.1. Constrained draw replication approach

The general approach we describe involves drawing a large number of independent, replicated samples ( $S_u$ ), using an unconstrained base probability sampling design, and then filtering them through desired constraint(s) to produce a survivor subset of replicated samples ( $S_c$ ). Thus,  $S_c$  consists of those samples that happen, by chance, to satisfy the constraint(s). From  $S_c$ , one replicated sample ( $a$ ) is randomly selected and becomes the official sample to be used for the intercept survey.

For replicate sample  $a$ , inclusion probabilities for each sampling unit is calculated as:

$$\pi'_i = P(i \in A | S_c) = \sum_{a \in A_i | S_c} p(a) \quad (1)$$

where the inclusion probability ( $\pi'_i$ ) of the  $i$  th sampling unit is the fraction of surviving replicated samples ( $A_i$ ) that contain sampling unit  $i$  out of the survivor subset of replicated samples ( $S_c$ ). This calculation is simply a modification of the standard definition of inclusion probability ( $\pi_i$ ), where  $\pi_i$  is the sum of the sampling probabilities  $p(a)$  for the set of samples that contain sampling unit ( $A_i$ ) (Fuller, 2009a):

$$\pi_i = P(i \in A) = \sum_{a \in A_i} p(a) \quad (2)$$

Estimates can then be produced using a standard weighted estimator such as the Horvitz-Thompson (Horvitz and Thompson, 1952) estimator for a total ( $\hat{Y}$ ):

$$\hat{Y} = \sum_{i=1}^n w_i y_i \quad (3)$$

where  $y_i$  is the value observed at an individual sampling unit (for  $i = 1 \dots n$  sampling units), and  $w_i$  is the sample weight, which is simply the inverse of the inclusion probability ( $1/\pi_i$ ).

### 2.2. Simulation study

We used angler intercept survey simulations, run in SAS 9.4 (SAS Institute Inc., 2016; see [https://github.com/NOAA-Fisheries-OST-ST1/Sample\\_Draw\\_Replication](https://github.com/NOAA-Fisheries-OST-ST1/Sample_Draw_Replication) for the full simulation code), to test the constrained draw replication approach, comparing its performance to that of a standard, unconstrained probability-based sample draw process. For both simulations (constrained vs. unconstrained), our sample frame, or the list of all sampling units in our survey population, consisted of fishing site-days (crossing 30 sites and 30 days, for a total of 900 sampling units in the survey population). The sample size for the survey was set to  $n = 30$  units, resulting in a 3 % sampling fraction. This number was selected purely for illustrative purposes, since in practice, sample sizes and sampling fractions can vary widely,

determined by factors such as funding levels, the per-unit cost of data collection, the scale of the survey population, and desired precision levels of estimates (e.g., NRC, 2012; Wright, 2014). For comparison, the National Marine Fisheries Service's Access Point Angler Intercept Survey, the new, redesigned survey that replaced the MRFSS Intercept Survey, had sampling fractions ranging from < 1 % to > 5 % in 2019 (NMFS, 2019a). The simulated surveys had a simple, single stage design (i.e., no sub-sampling of site-days), and relied on probability proportional to size without replacement (PPSWOR) sample draw processes, where fishing pressure was the size measure used to draw sampling units. PPSWOR is a type of PPS sampling where sampling units are not replaced when drawn from the sample frame, and therefore will only be selected once for a given sample. Site-days were assigned three general fishing pressure categories in order to set inclusion probabilities for the PPSWOR draw process: 5 % of the fishing sites were set to high activity, 35 % were set to medium activity, and 60 % were set to low fishing activity. These were set to be generally representative of a plausible proportion of activity levels across sites (e.g., NMFS, 2019b). The arbitrary size measures ( $m_i$ ) used for the PPSWOR inclusion probabilities were  $m_i = 5$  for low activity sites,  $m_i = 10$  for medium activity sites, and  $m_i = 20$  for high activity sites. This design essentially ensured that sites with high activity were drawn to be part of the sample more frequently, and sites with low activity were drawn to be part of the sample with minimal frequency.

The true catch ( $Y$ ), or total catch summed across all 900 site-day sampling units ( $N$ ) in the sample frame, was set to 100,000 fish, and we simulated intercept surveys by generating three probability distribution scenarios for catch-per-unit ( $y_i$ ), meant to represent the three most probable scenarios for the distribution of catch in relation to the level of site-specific fishing activity. The first was a random Poisson distribution, representing a scenario where the number of fish caught, attributed to a site-day, is unrelated to size measure  $m_i$  (or level of fishing activity) of that site-day. In a random Poisson distribution, the probability of a particular number of fish caught per site-day ( $y_{i,poi}$ ) is (Katti and Rao, 1968)

$$P(y_{i,poi}) = \frac{e^{-\lambda} \lambda^{y_{i,poi}}}{y_{i,poi}!} \quad (4)$$

where  $\lambda = \frac{Y}{N}$ . The second distribution scenario was a Poisson distribution with a strong positive correlation of catch to a sampling unit's size measure  $m_i$ , representing a scenario where site-days with higher fishing activity account for more of the total fish caught. In this scenario, the probability of a particular number of fish caught per site-day ( $y_{i,pcor}$ ) is

$$P(y_{i,pcor}) = \lambda + \rho \sqrt{\lambda} \left( \frac{m_i - \bar{m}_i}{\sigma_{m_i}} \right) + \sqrt{\lambda - \lambda \rho^2} \left( \frac{P(y_{i,poi}) - \lambda}{\sqrt{\lambda}} \right) \quad (5)$$

where  $\sigma_{m_i}$  is the standard deviation of size measure  $m_i$ , and  $\rho = 0.95$ , which is a correlation coefficient representing the strong positive correlation with the size measure. The third and final distribution scenario was a Poisson distribution with a strong negative correlation of catch to a sampling unit's size measure  $m_i$ , representing a scenario where site-days with lower fishing activity account for more of the total fish caught. In this scenario, the probability of a particular number of fish caught per site-day ( $y_{i,ncor}$ ) is calculated using Eq. (5), but with  $\rho = -0.95$  to represent the negative correlation with size measure (see e.g., Wicklin, 2013 for more about generating correlated data).

In the constrained simulation, seven initial sets of replicated samples, ranging in size from 500 to 50,000 samples per set ( $S_u = 500, 1000, 2000, 5000, 10,000, 20,000, \text{ and } 50,000$ ) were drawn. This process was done with standard PPSWOR using the SAS SURVEYSELECT procedure (SAS Institute Inc., 2016), and then looping the procedure so that it would repeat 500–50000 times to draw the desired number of samples for each  $S_u$ . Then, each  $S_u$  was filtered through a constraint of two sampling assignments per day, so any samples with

three or more sampling units selected for the same day were excluded from the set of surviving replicated samples ( $S_c$ ). This particular constraint was chosen as an example for several key reasons. First, in order to adequately evaluate the effectiveness of the constrained draw replication approach, we needed to create highly constraining conditions that could not be addressed using standard survey design and sample selection techniques. Secondly, the number of sampling assignments that can be conducted on the same day is a common logistical constraint in catch sampling intercept surveys (that often hinges on staff availability [NRC, 2006]), and an actual two-assignments per day constraint is among the constraints currently used by the National Marine Fisheries Service's Access Point Angler Intercept Survey (see Papacostas and Foster, 2018). After filtering through the constraint, each of the seven  $S_u$  had survival rates of approximately 5%, resulting in seven  $S_c$  consisting of 24, 45, 97, 240, 483, 1000, and 2386 replicated samples respectively. From each  $S_c$ , a final sample was selected using simple random sampling. The sample weights for sampling units in the final selected sample ( $a$ ) were then calculated as  $(1/\pi_i)$ , using Eq. (1). Total catch ( $\hat{Y}_{S_c}$ ) was estimated using Eq. (3) for each of the three catch distribution scenarios ( $y_{i,poi}$ ,  $y_{i,ncor}$ , and  $y_{i,pcor}$  described by Eqs. (4) and (5)), and the process was repeated for 1000 iterations for each of the seven starting sets of replicated samples (see Fig. 1).

The unconstrained simulation served as an experimental control for the study; it had an identical design as the constrained simulation, but did not apply the constraint after using standard PPSWOR to draw replicated samples (see Fig. 1). First, seven  $S_u$  of the same sizes as the unconstrained simulation were produced. Then, sample weights for the sampling units were calculated as  $1/\pi_i$ , using Eq. (2), and total catch ( $\hat{Y}_{S_u}$ ) was estimated using Eq. (3) for each replicated sample and catch distribution scenario ( $y_{i,poi}$ ,  $y_{i,ncor}$ , and  $y_{i,pcor}$  described by Eqs. (4) and (5)). These results were expected to vary around the true catch, as standard PPSWOR is known to be design unbiased (Yates and Grundy, 1953), so they served as a useful comparison to the results of the constrained simulation.

We compared the distributions of total catch estimates, variance estimates, and sample weights from the constrained and unconstrained simulations. We also examined bias in the estimates relative to the true catch, for each catch distribution scenario, in both simulations. In the unconstrained simulation, this relative bias ( $RelBias(\hat{Y}_{S_u})$ ) was estimated as the difference of the mean of the catch totals for each set of replicated samples ( $\hat{Y}_{S_u}$ ) from the true catch ( $Y$ ), divided by the true catch:

$$RelBias(\hat{Y}_{S_u}) = \frac{\hat{Y}_{S_u} - Y}{Y} \quad (6)$$

Since the estimates from the unconstrained simulation were expected to vary around true catch, all of those relative bias estimates were expected to be approximately zero. Relative bias in the constrained simulation ( $RelBias(\hat{Y}_{S_c})$ ) was estimated in a very similar manner, as the difference of the mean of the estimated catch totals from each survivor set of replicated samples ( $\hat{Y}_{S_c}$ ), from the true catch ( $Y$ ), divided by the true catch:

$$RelBias(\hat{Y}_{S_c}) = \frac{\hat{Y}_{S_c} - Y}{Y} \quad (7)$$

The only minor difference between Eqs. (6) and Eq. (7) was in estimating the mean of the catch totals for the unconstrained ( $\hat{Y}_{S_u}$ ) and constrained ( $\hat{Y}_{S_c}$ ) simulations. In the unconstrained simulation,  $\hat{Y}_{S_u} = \frac{\hat{Y}_{S_u}}{S_u}$  (where  $S_u = 500, 1000, 2000, 5000, 10,000, 20,000$ , and  $50,000$  replicated samples). In the constrained simulation however,  $\hat{Y}_{S_c} = \frac{\sum \hat{Y}_{S_c}}{1000}$ , since the constrained draw replication approach underwent 1000 iterations for each of the seven sets of replicated samples from which catch totals were estimated (see Fig. 1).

Finally, we examined how the relative bias of the estimates in the constrained simulation may have been affected by sample frame units missing from the survivor subsets of replicated samples ( $S_c$ ). We decided to look at the effects of these missing sample frame units since, as the number of surviving replicated samples in  $S_c$  decreased, the number of sampling units in the sample frame that were not included in  $S_c$  increased. To explore these effects, we measured the proportion of sample frame units in each  $S_c$  against relative bias.

### 3. Results

Catch estimate distributions in the constrained simulation were most similar to those produced from the unconstrained simulation at high replication levels (i.e., the sizes of  $S_u$  [the initial number of replicated samples] and  $S_c$  [the survivor subset of replicated samples after applying the two sites-per-day constraint]). As expected, estimates from the unconstrained simulation were unbiased, varying around the true catch (Fig. 2). The constrained simulation underestimated the true catch when  $S_u \leq 2000$  and  $S_c \leq 97$ , but at  $S_u = 10,000$  and  $S_c = 483$  the constrained results began to match those from the unconstrained simulation. Thus the results from the constrained simulation were unbiased at  $S_u \geq 10,000$  and  $S_c \geq 483$  (Fig. 2). Results were very similar across all three catch-per-unit distribution scenarios (Fig. 2).

Changes in variance distributions corresponded to replication level (i.e. the sizes of  $S_u$  and  $S_c$ ) in the constrained simulation. At  $S_u = 2000, 5000$  and  $10,000$  ( $S_c = 97, 240$  and  $483$  respectively), the constrained simulation had the highest variances with the broadest distributions, including a large number of outliers particularly when  $S_u = 5000$  and  $10,000$  ( $S_c = 240$  and  $483$  respectively; Fig. 3). Variances in the constrained simulation at the lowest replication level ( $S_u = 500$ ) were actually lower than in the unconstrained simulation, an artifact of the similarity of catch-per-unit values and sample weights, given the small number of surviving replicated samples ( $S_c = 21$ ). At  $S_u = 1000, 20,000$  and  $50,000$  ( $S_c = 45, 1000$  and  $2386$  respectively), variances and their distributions were very comparable to the unconstrained simulation results (Fig. 3).

The sample weights followed a similar pattern to the catch estimates in that the sum of the sample weights in the constrained simulation were most similar to those of the unconstrained simulation at high levels of replication ( $S_u = 20,000$  and  $50,000$ ;  $S_c = 1000$  and  $2386$ ; Fig. 4). The sample weights were lowest at the lowest levels of replication ( $S_u = 500$  and  $1000$ ;  $S_c = 21$  and  $45$ ). Similar to the variance distributions, sample weights had the most variability at the intermediate levels of replication ( $S_u = 2000, 5000$  and  $10,000$ ;  $S_c = 97, 240$ , and  $483$ ; Fig. 4).

Results indicated that estimated bias relative to the true catch was directly related to the proportion of sample frame units missing from  $S_c$  in the constrained simulation (Fig. 5). At low levels of replication ( $S_u < 5000$ ;  $S_c < 240$ ), sample frame units were missing from  $S_c$  (Fig. 6), correlating with a negative relative bias of the estimates (Figs. 5 and 7). With increased replication when most sample frame units ( $> 99\%$ ) were represented in the surviving set of replicate draws (at  $S_c = 240$ , Fig. 6), relative bias was still slightly negative, but very close to zero (Fig. 7). Finally, when all sample frame units were represented (at  $S_c \geq 483$ ), constrained results matched those of the unconstrained simulation, and estimates were unbiased (Fig. 7).

### 4. Discussion

Our results indicate that the constrained draw replication approach can be used to produce unbiased estimates at sufficiently high levels of replication. Low numbers of starting, and more importantly, surviving replicated samples, however, may lead to significant underestimation of total catch. In our simulations, this bias was driven by the missing sample frame units in the smaller number of surviving replicated samples, since estimates were unbiased when all sample frame units

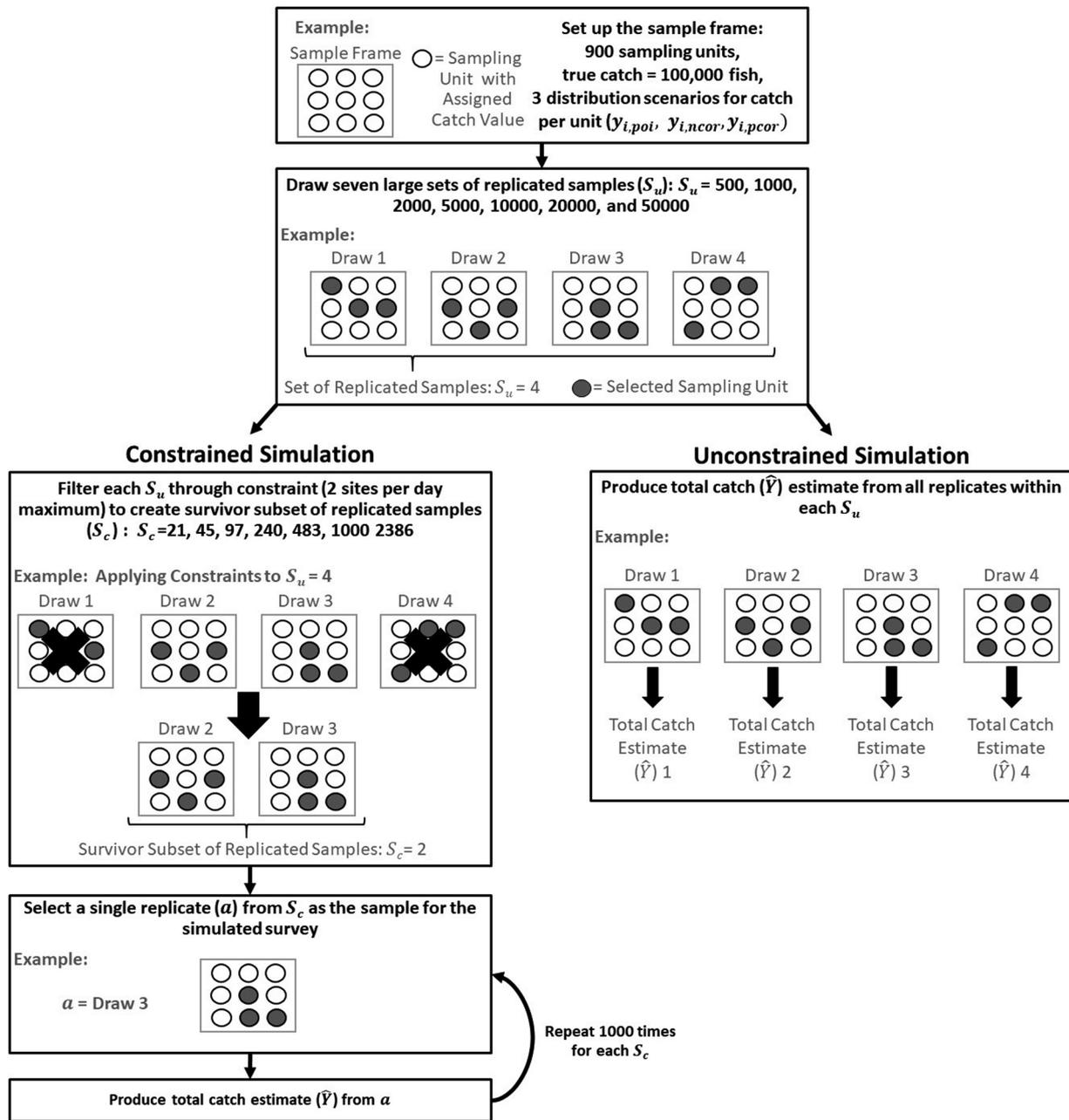


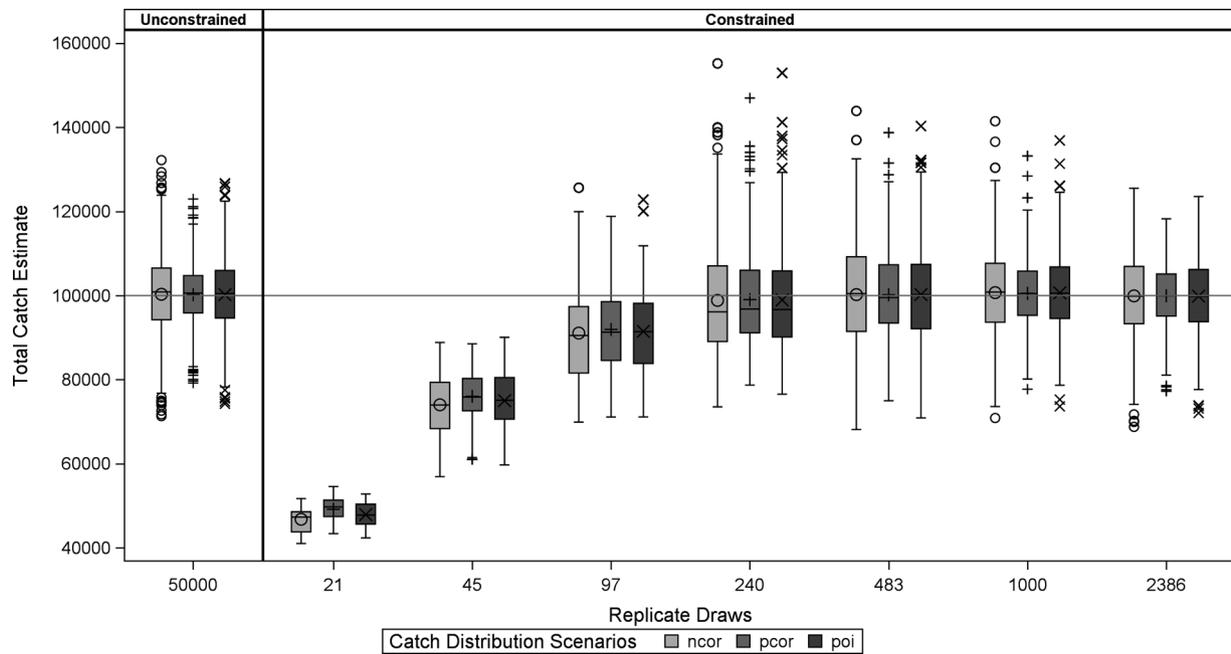
Fig. 1. A schematic of the constrained and unconstrained simulation designs, with accompanying visuals illustrating the procedures with an example sample frame consisting of 9 sampling units. (Note that the visualization is meant to serve as a basic example to aid in understanding of the general methodology, particularly the process of drawing replicated samples in both simulations; the exact study sample frame, probability proportional to size sampling scheme, and the two sites-per-day constraint are not specifically visualized).

were represented at least once in the surviving subset of replicated samples. The sum of the sample weights were lower at low levels of replication, explaining why bias always skewed in the negative direction.

It is a well-established result in sampling statistics that sample estimates are design unbiased when all units in the frame have a known inclusion probability which is greater than zero, and that those inclusion probabilities (and corresponding sample weights) are used in calculating the weighted estimator; this result holds true at least for linear estimators (Fuller, 2009a; Kish, 1965), like the standard weighted total used here in our study. From this result, we should expect the constrained draw replication approach to return unbiased estimates whenever all sample frame units appear in at least one of the surviving replicated samples. When this condition was met, we then had known,

positive inclusion probabilities for all sampling units in the sample frame, which we used in calculating the weighted total.

Even when all sample frame units were represented in the surviving replicated samples, however, the intermediate levels of replication led to reduced precision. This result was likely because the constrained draw replication process impacted the sample weights, which have a known relationship with variances; increases in sample weight variability directly contribute to increases in variance estimates (Kish, 1965). At intermediate levels of replication, the median of the total catch estimates were closer to the true catch, but variances were high and widely distributed, correlating with the high variability of the sample weights at those replication levels. At the highest level of replication, however, the sample weights became less erratic and estimates were unbiased with very comparable precision to the unconstrained

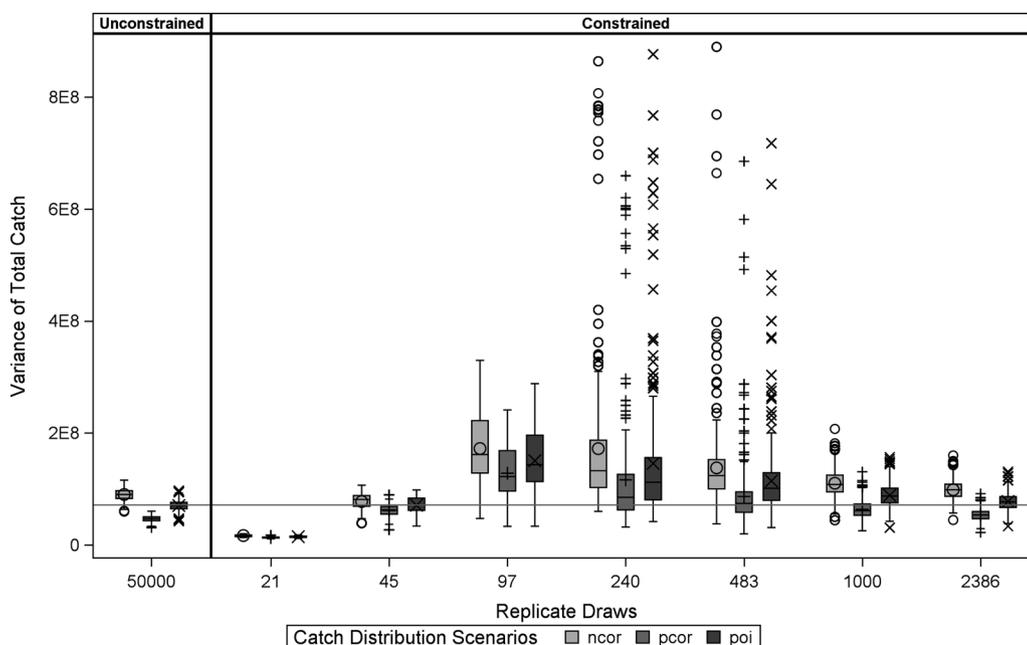


**Fig. 2.** Total catch estimates from the unconstrained (left panel) and constrained (right panel) simulations: all results from the unconstrained simulation were near identical, so only one result (the set of 50,000 replicated samples) is visualized here as an example, while all seven results from the constrained simulation are shown (displaying the number of surviving replicated samples on the x-axis). Estimates are displayed for the three catch-per unit distribution scenarios: poi, pcor, and ncor. Estimates are presented in box plots where the lower boundary of the box indicates the 25th percentile, the line near the middle of the box represents the median, the point near the middle of the box represents the mean, and the upper boundary of the box indicates the 75th percentile. Whiskers below and above the box indicate the 10th and 90th percentiles respectively. Points lying outside the whiskers indicate outliers outside of the 10th and 90th percentiles.

simulation.

The example we present here is relatively simple in applying the constrained draw replication approach using PPSWOR and just one added constraint (a two sites-per-day maximum), but since this approach is based on draw replication, it does not require one specific underlying design and can theoretically be implemented with any standard sample selection process. The survey design in the simulation was also single stage (only sampled site-days, with no subsampling of angler or vessel trips) whereas in reality most intercept surveys involve

multiple stages of sampling (ICES, 2013; NRC, 2006; NAS, 2017). However, this approach can still be applied to more complex survey designs to select primary sampling units, and the same estimation methods can be used with additional sample weighting to account for the additional stages of sampling. As referenced earlier, the National Marine Fisheries Service applies this constrained draw replication approach to their new intercept survey, the Access Point Angler Intercept Survey, which has a stratified, multi-stage cluster design, uses PPSWOR with nine different size measures for fishing pressure, and is conducted



**Fig. 3.** Total catch variance estimates from the unconstrained (left panel) and constrained (right panel) simulations: all results from the unconstrained simulation were near identical, so only one result (the set of 50,000 replicated samples) is visualized here as an example, while all seven results from the constrained simulation are shown (displaying the number of surviving replicated samples on the x-axis). Estimates are presented in box plots where the lower boundary of the box indicates the 25th percentile, the line near the middle of the box represents the median, the point near the middle of the box represents the mean, and the upper boundary of the box indicates the 75th percentile. Whiskers below and above the box indicate the 10th and 90th percentiles respectively. Points lying outside the whiskers indicate outliers outside of the 10th and 90th percentiles.

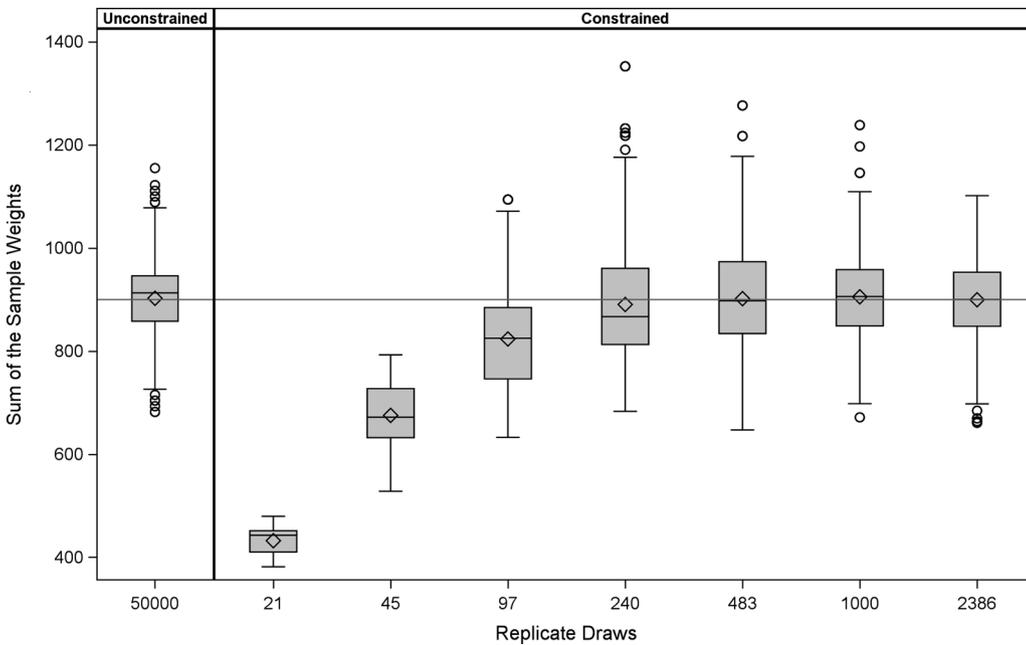


Fig. 4. Sum of the sample weights from the unconstrained (left panel) and constrained (right panel) simulations and each of the three distribution scenarios: all results from the unconstrained simulation were near identical, so only one result (the set of 50,000 replicated samples) is visualized here as an example, while all seven results from the constrained simulation are shown (displaying the number of surviving replicate draws on the x-axis). Sample weight sums are presented in box plots where the lower boundary of the box indicates the 25th percentile, the line near the middle of the box represents the median, the point near the middle of the box represents the mean, and the upper boundary of the box indicates the 75th percentile. Whiskers below and above the box indicate the 10th and 90th percentiles respectively. Points lying outside the whiskers indicate outliers outside of the 10th and 90th percentiles.

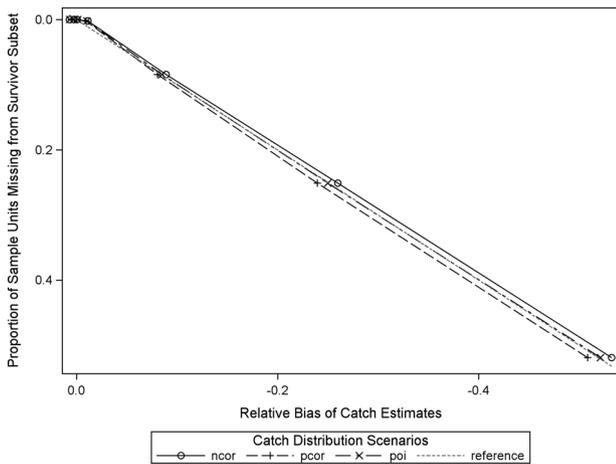


Fig. 5. Relative bias of estimates in the constrained simulation based upon the proportion of frame units missing from the surviving set of replicated samples for the three catch distribution scenarios: poi, pcor, and ncor. The gray diagonal dotted line down the center of the figure represents a reference line.

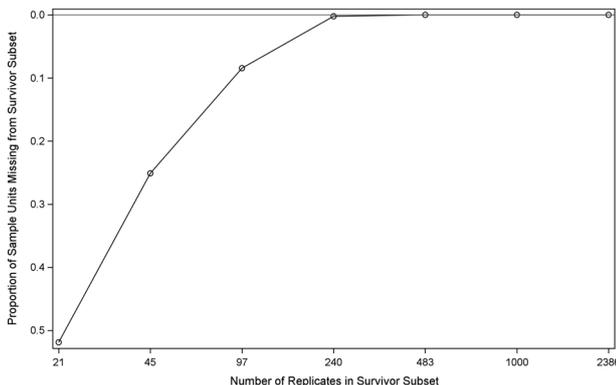


Fig. 6. The proportion of sample frame units missing from the survivor subset of replicated samples at all seven levels of replication in the constrained simulation.

across 16 states (NAS, 2017). In this survey, sampling assignments include visiting site clusters (which contain up to two fishing access sites) on specified days during specified time intervals out of the 24-h day. Multiple constraints are imposed on the Access Point Angler Intercept Survey’s sample selection, including an assignments-per-day constraint like in our example, an additional constraint that ensures that no time intervals overlap at one particular site, and a few others that are tailored at different spatial scales (state or county) depending on that area’s unique field logistical issues. The Access Point Angler Intercept Survey uses Eq. (1) to estimate the primary stage weight of samples, and then applies relatively simple additional weights to account for each stage of subsampling (see Papacostas and Foster, 2018 for these methods). The Access Point Angler Intercept Survey is therefore a good example of the potential flexibility of the constrained draw replication approach, in handling multiple constraints, being highly adaptable across spatial scales, and being applicable to a multi-stage sampling scenario.

Our methods presented here may serve as a useful, simple, and flexible approach to probability sampling in intercept surveys. Based upon the results of this study, this constrained draw replication approach can be used to produce precise, design unbiased estimates, comparable to those from a standard PPSWOR sampling design. Incorporating logistical constraints directly into the sample selection process, rather than afterwards in the field, can alleviate data quality concerns, and reduce deviations from probability-based designs in practice. This method should reliably produce unbiased estimates as long as all sample frame units are represented in the set of surviving replicated samples such that positive inclusion probabilities for all units in the frame are calculable and used appropriately in estimation. To be conservative, we recommend using this approach with the highest number of replicated samples that is computationally or logistically feasible, given available resources. For example, the default numbers for the National Marine Fisheries Service’s Access Point Angler Intercept Survey draw process is  $S_n=200,000$  with a target of  $S_c \geq 1000$ , but the exact numbers of initial replicates and any targets for numbers of surviving replicated samples will be entirely case-specific; the number of units in the sample frame, sample size, specifics of the constraints, and other design aspects will all factor into the initial number of replicated samples needed to ensure a viable survivor subset, and design unbiased estimates. Overall, our work attempts to increase design options available to fisheries scientists, and provide a possible

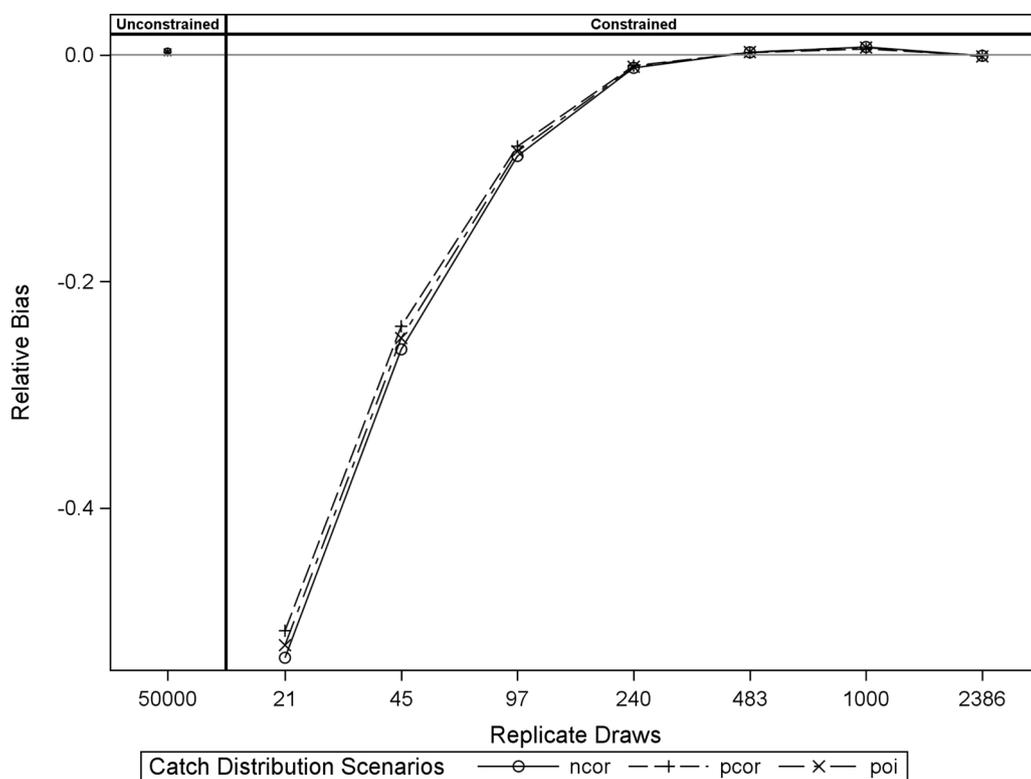


Fig. 7. The bias of catch estimates relative to the true catch in the unconstrained (left panel) and constrained (right panel) simulations: the estimates from the unconstrained simulation all had zero relative bias from the true total catch, so only one result (the set of 50,000 replicated samples) is visualized here as an example, while all seven results from the constrained simulation are shown (displaying the number of surviving replicated samples on the x-axis).

solution to logistical challenges that can affect the statistical integrity of estimates in catch sampling intercept surveys.

#### CRediT authorship contribution statement

**Katherine J. Papacostas:** Writing - original draft, Writing - review & editing, Visualization, Data curation. **John Foster:** Conceptualization, Methodology, Software, Formal analysis, Writing - review & editing.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

We would like to thank F. Jay Breidt and Jean Opsomer for helpful initial discussions on concept development. We also thank Ryan Kitts-Jensen, Alan Lowther, Tom Sminkey, André Punt and two anonymous reviewers for useful comments on the manuscript.

#### References

- Arnab, R., 2013. Controlled sampling: a review. *Stat. Appl.* 11, 127–146.
- Baker, R., Brick, J.M., Bates, N.A., Battaglia, M., Couper, M.P., Dever, J.A., Gile, K.J., Tourangeau, R., 2013. Summary report of the AAPOR task force on non-probability sampling. *J. Surv. Stat. Methodol.* 1 (2), 90–143. <https://doi.org/10.1093/jssam/smt008>.
- Fuller, W.A., 2009a. *Sampling Statistics*. John Wiley & Sons, Inc, Hoboken, New Jersey, USA.
- Fuller, W.A., 2009b. Some design properties of a rejective sampling procedure. *Biometrika* 96, 933–944. <https://www.jstor.org/stable/27798877>.
- Gabler, S., 1987. The nearest proportional to size sampling design. *Commun. Stat. Theory Methods* 16, 1117–1131. <https://doi.org/10.1080/03610928708829426>.
- Goodman, R., Kish, L., 1950. Controlled selection - a technique in probability sampling. *J. Comput. Graph. Stat.* 45, 350–372. <https://doi.org/10.1007/BF02868149>.
- Griffiths, S., Sahlqvist, P., Lyle, J., Venables, W., Pollock, K., Sawynok, W., Moeseneder,

- C., 2014. A Coordinated National Data Collection for Recreational Fishing in Australia. CSIRO, Brisbane, Queensland, Australia. <https://doi.org/10.4225/08/584d96f2b4b11>.
- Hartigan, J.A., 1971. Error analysis by replaced samples. *J. R. Stat. Soc. Ser. B (Methodol.)* 33, 98–110. <https://doi.org/10.1111/j.2517-6161.1971.tb00861.x>.
- Hess, I., 1961. *Probability Sampling of Hospitals and Patients*. University of Michigan, Ann Arbor, Michigan, USA.
- Horvitz, D.G., Thompson, D.J., 1952. A generalization of sampling without replacement from a finite universe. *J. Comput. Graph. Stat.* 47, 663–685. <https://doi.org/10.1080/01621459.1952.10483446>.
- ICES, 2013. Report of the Third Workshop on Practical Implementation of Statistically Sound Catch Sampling Programmes. 19–22, November, 2013. ICES CM/2013/ACOM:54, 1–109. ICES HQ, Copenhagen, Denmark. <http://ices.dk/community/groups/SiteAssets/WGCATCH-publications.aspx>.
- ICES, 2020. Working Group on Recreational Fisheries Surveys (WGRS; outputs from 2019 meeting). *ICES Sci. Rep.* 2, 1–78. <https://doi.org/10.17895/ices.pub.5744>.
- Kalton, G., 1979. Ultimate cluster sampling. *J. R. Stat. Soc. Ser. A (General)* 142, 210–222. <https://www.jstor.org/stable/2345081>.
- Katti, S.K., Rao, A.V., 1968. Handbook of the poisson distribution. *Technometrics* 10, 412. <https://doi.org/10.1080/00401706.1968.10490580>.
- Kish, L., 1965. *Sampling Statistics*. John Wiley & Sons, Inc., New York, New York, USA.
- Kuno, E., 1976. Multi-stage sampling for population estimation. *Res. Popul. Ecol.* 18, 39–56. <https://link.springer.com/article/10.1007/BF02754081>.
- NAS, 2017. Review of the Marine Recreational Information Program. National Academies Press, Washington, D.C., USA. <https://www.nap.edu/catalog/24640/review-of-the-marine-recreational-information-program>.
- Nigam, A.K., Kumar, P., Gupta, V.K., 1984. Some methods of inclusion probability proportional to size sampling. *J. R. Stat. Soc. Ser. B (Methodol.)* 46, 564–571. <https://www.jstor.org/stable/2345694>.
- NMFS, 2013. Magnuson-stevens fishery conservation and management reauthorization act; national standard 2 - scientific information. *Fed. Regist.* 78 (19 July), 43066–43090. 2013. <https://www.federalregister.gov/documents/2013/07/19/2013-17422/magnuson-stevens-act-provisions-national-standard-2-scientific-information>.
- NMFS, 2019a. Monthly Sample Frames and Draws by State for the Access Point Angler Intercept Survey. Unpublished Raw Data.
- NMFS, 2019b. Marine Recreational Information Program Public Access Fishing Site Register. (Accessed 10 April, 2020). <https://www.fisheries.noaa.gov/recreational-fishing-data/public-fishing-access-site-register>.
- NRC, 2006. Review of Recreational Fisheries Survey Methods. National Academies Press, Washington, D.C., USA. <https://www.nap.edu/catalog/11616/review-of-recreational-fisheries-survey-methods>.
- NRC, 2009. Sample Surveys: Overview, Examples, and Usefulness in Studying Aviation Safety, in: an Assessment of NASA's National Aviation Operations Monitoring Service. National Academies Press, Washington, DC., USA. <https://www.nap.edu/read/12795/chapter/5>.
- NRC, 2012. Small Populations, Large Effects: Improving the Measurement of the Group

- Quarters Population in the American Community Survey. National Academies Press, Washington, DC., USA. <https://www.nap.edu/catalog/13387/small-populations-large-effects-improving-the-measurement-of-the-group>.
- Papacostas, K.J., Foster, J., 2018. National Marine Fisheries Service's Marine Recreational Information Program: Survey Design and Statistical Methods for Estimation of Recreational Fisheries Catch and Effort. NOAA-NMFS. Silver Spring, Maryland, USA. <https://www.fisheries.noaa.gov/resource/document/survey-design-and-statistical-methods-estimation-recreational-fisheries-catch-and-effort>.
- Rao, J.N.K., Nigam, A.K., 1990. Optimal controlled sampling designs. *Biometrika* 77, 807–814. <https://doi.org/10.1093/biomet/77.4.807>.
- Rao, J.N.K., Nigam, A.K., 1992. 'Optimal' controlled sampling: a unified approach. *Int. Stat. Rev.* 60, 89–98. <https://www.jstor.org/stable/1403503>.
- SAS Institute, Inc, 2016. SAS/STAT 14.3 User's Guide. Cary, North Carolina, USA.
- Strehlow, H.V., Schultz, N., Zimmermann, C., Hammer, C., 2012. Cod catches taken by the German recreational fishery in the western Baltic Sea, 2005–2010: implications for stock assessment and management. *ICES J. Mar. Sci.* 69, 1769–1780. <https://doi.org/10.1093/icesjms/fss152>.
- Tiwari, N., Chilwal, A., 2014. A simplified approach for two-dimensional optimal controlled sampling designs. *Adv. Stat.* 2014, e875352. <https://doi.org/10.1155/2014/875352>.
- Tiwari, N., Nigam, A.K., Pant, I., 2007. On an optimal controlled nearest proportional to size sampling scheme. *Surv. Methodol.* 33, 87–94. <https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X20070019856>.
- Vølstad, J., Pollock, K., Richkus, W., 2006. Comparing and combining effort and catch estimates from aerial-access designs as applied to a large-scale angler survey in the Delaware River. *N. Am. J. Fish. Manage.* 26, 727–741. <https://doi.org/10.1577/M04-146.1>.
- Vølstad, J.H., Afonso, P.S., Baloi, A.P., de Premegi, N., Meisjord, J., Cardinale, M., 2014. Probability-based survey to monitor catch and effort in coastal small-scale fisheries. *Fish. Res.* 151, 39–46. <https://doi.org/10.1016/j.fishres.2013.11.016>.
- Waterton, J.J., 1983. An exercise in controlled selection. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* 32, 150–164. <https://www.jstor.org/stable/2347294>.
- Wicklin, R., 2013. *Simulating Data With SAS*. SAS Institute Inc., Cary, North Carolina, USA.
- Wright, T., 2014. A Simple Method of Exact Optimal Sample Allocation Under Stratification With Any Mixed Constraint Patterns. U.S. Bureau of the Census, Washington, D.C., USA. <https://www.census.gov/srd/papers/pdf/rrs2014-07.pdf>.
- Yates, F., Grundy, P.M., 1953. Selection without replacement from within strata with probability proportional to size. *J. R. Stat. Soc. Ser. B (Methodol.)* 15, 253–261. <https://doi.org/10.1111/j.2517-6161.1953.tb00140.x>.