Transactions of the American Fisheries Society 151:42–49, 2022

ISSN: 0002-8487 print / 1548-8659 online
DOI: 10.1002/tafs.10342

REVIEW

A Review of Nonprobability Sampling Using Mobile Apps for Fishing Effort and Catch Surveys

J. Michael Brick*
Westat, 1600 Research Boulevard, Rockville, Maryland 20850, USA

William R. Andrews and John Foster
National Oceanic and Atmospheric Administration, Fisheries, 1315 East-West Highway, Silver Spring, Maryland 20910, USA

Abstract
Technology and computational advancements have caused us to rethink data collection and sampling methods that have been the standard for gathering information to guide policy decisions. The availability of new technology such as mobile apps has made using nonprobability samples more attractive because of the speed and low expense that are associated with this approach. We review how the use of nonprobability sampling using mobile apps affects the quality of inferences in fishing effort and catch surveys. We present an approach for evaluating the potential biases that arise from both probability and nonprobability sampling methods. The approach shows that well-conducted probability samples have major advantages compared with nonprobability samples. We conclude that the application of nonprobability sampling for fishing surveys faces serious challenges and should prove that it is fit for use before being adopted more widely.

During the last few decades, a number of developments have greatly affected how we think about the ability of data that are collected by different means to inform decision making and policy development across a wide range of fields. Technological advancements have reshaped our ideas of data and data sharing. "Big data" have become commonplace, and methods for processing massive data sources have enabled users to efficiently capture and visualize the data (Rodríguez-Mazahua et al. 2016). The web, mobile device apps, and social media are common sources of these data.

Another development is a new emphasis on administrative data that are collected by government agencies and private-sector organizations as a source of data for analysts and researchers. The Foundations for Evidence-Based Policymaking Act of 2018 requires the federal government to rethink its approach to data acquisition for making policy decisions and to consider alternative data sources. For example, in its review of the National Oceanic and Atmospheric Administration's fishing effort and catch surveys, the National Research Council (2006) recommended the development of administrative records to support recreational saltwater fisheries data collection and management.

In most U.S. federal statistical agencies, the longstanding approach to policy making has been to define the information that is needed to guide a policy decision and then develop a plan to acquire the relevant supporting data and statistics. For fisheries policy, this often involves designing probability-based catch and effort surveys and making inferences from the respondents. For example, the National Oceanic and Atmospheric Administration designed and implemented the Marine Recreational Information Program (MRIP) using probability sampling so

*Corresponding author: mikebrick@westat.com
Received March 31, 2021; accepted October 16, 2021

that it could produce point estimates as well as measures of precision resulting from the sampling process. We describe the MRIP in more detail later.

The advancement of novel data sources has disrupted this paradigm. A few examples illustrate the nature of the novel data sources. One of the most discussed early applications for big data was the use of Google search data to track flu prevalence in the United States (Pervaiz et al. 2012). The data were not "designed" specifically to study flu prevalence, leading some to refer to this type of data source as "organic." Harvesting data from an existing source to answer a policy question can result in substantial savings in terms of cost and time compared with a more traditional sampling solution. After a few years, the flu prevalence estimates from the model were found to be severe overestimates. This lead to a review of the risks that are associated with continuing to rely on a model while not considering that the environment that produces the data changes over time (Lazer et al. 2014). Toole et al. (2015) give another example of using organic data. They discuss the capture of mobile data from cell towers to evaluate travel demand and infrastructure needs. Shi and Abdel-Aty (2015) use a different novel data collection technology that is not organic but is "passive" in the sense that it does not require any efforts from respondents. They discuss using data from microwave radar sensors to study traffic safety and congestion. All of these examples collect massive amounts of data by using modern data-processing methods.

When existing data are not available to answer a policy question, a hybrid approach using nonprobability sampling has been suggested to retain some of the cost and timeliness advantages that may come from organic data. For example, nonprobability samples of persons who volunteer (opt-in), or are otherwise identified from their web activity, are asked to complete a web survey. In these nonprobability samples, the data items are designed specifically to support a policy question and the data are collected through an advanced technology such as a mobile app or web survey. The main advantage of these designs is that data collection can be done quickly and more affordably than with probability samples. However, because the sampling is not controlled, inferences from the data must rely completely on mathematical models and assumptions that cannot be verified. The American Association of Public Opinion Research established three task forces to investigate nonprobability sampling, and Baker et al. (2013) describe the latest effort.

For the purpose of completing a survey request, web questionnaires and mobile apps are examples of data collection modes that are similar to traditional survey modes such as face-to-face interviews, mail questionnaires, and telephone interviews. While mobile apps have advantages over hard copy surveys, they are not the primary distinguishing feature of the new era of data collection.

Our focus is on the method of deciding who should be surveyed, not on the mode of data collection. The new development we examine is the move away from probability samples to samples of volunteers or persons who are recruited without randomizing the selection process, a transition that has been exacerbated by the proliferation of social media and smartphone apps. The concern is whether the transition from probability samples to convenience or volunteer samples allows the type of inferences that are needed to support policy decisions.

This issue is not new. Baker et al. (2013) describe data collection and analysis methods for nonprobability samples such as quota sampling, matched sampling, and alternative statistical methods. The role of sampling in making inferences has been debated in statistics (Stephan and McCarthy 1958; Royall 1970), social science (Coleman 1958; Callegaro et al. 2014), and in official statistics (U.S. Office of Management and Budget 2006). Our review concentrates on the relatively new and specific interest in using nonprobability samples to estimate recreational, saltwater fishing effort and catch. Specifically, we focus on nonprobability samples that collect data by using mobile apps to provide fishing effort and catch estimates to support fisheries policy making.

The next section describes designs for recreational fishing surveys including current probability designs, nonprobability designs that have been suggested or are being developed, and hybrid designs that combine probability and nonprobability samples. The nature of the potential biases arising from both probability and nonprobability sampling methods are examined in the following section. The fourth section examines the potential for using nonprobability samples by using mobile apps for addressing data needs that supplement those from probability samples. The final section is a discussion and summary of our view of the issues that face the application of nonprobability sampling for fishing surveys.

RECREATIONAL FISHING DATA COLLECTION DESIGNS

The approach that the MRIP has used to obtain information for managing recreational saltwater fishing has been to estimate fishing effort and catch using separate probability surveys. The MRIP and state natural resource agencies administer several different surveys to estimate fishing activity for different recreational fishing sectors. Here we focus on the generalized surveys that are used to estimate shore and private boat fishing activity on the Atlantic coast and Gulf of Mexico. The MRIP Fishing Effort Survey estimates the total number of fishing trips taken during a specific period, and the Access Point

Angler Intercept Survey (APAIS) estimates the average catch per trip, or catch rate, for each species during the same period. The product of the two estimates gives the desired estimate—total catch for each species. The Fishing Effort Survey selects a random sample of households and asks household residents to report the number of saltwater recreational trips taken during the past 2 months. The APAIS intercepts anglers at the conclusion of fishing trips and collects data on the type and number of fish caught on the trip. Both the effort and catch surveys use probability sampling and base inferences on those probabilities of selection (Papacostas and Foster 2018). The MRIP surveys have been examined by independent reviewers several times, and the latest review is by the National Academies of Sciences, Engineering, and Medicine (2021).

With the proliferation of smartphones and apps within the last decade, some researchers have been exploring volunteer samples of recreational anglers by using mobile apps as the data collection mode. Venturelli et al. (2017) discuss the utility of angler apps and the challenges associated with using them. They define angler apps as mobile apps "that allow anglers to record, share and network" their activities.

From a fisheries management perspective, it would be ideal if everyone who participates in recreational fishing used an app to record their catch and shared their data, resulting in a census of recreational fishing activity. However, Venturelli et al. (2017) identify recruitment and retention as major barriers to useful app data. For example, Pappenfus et al. (2015) reported that angler apps underestimated freshwater fishing in Canada by a factor of 254, while Lui et al. (2017) estimated trip reporting rates (percentage of total trips reported) of approximately 4% for iSnapper, an app designed to monitor Red Snapper *Lutjanus campechanus* fishing activity in the Gulf of Mexico. Similarly, Venturelli et al. (2017) note that only 5% of those who begin to use a specific app still used it after 3 months, and Ahrens (2013) found angler retention for iAngler, an app focused on saltwater fishing in south Florida, was very poor, with the majority of users only using the app once. This rate of "stickiness" or persistence with using an app is not unusual. For example, Yang et al. (2020) found that 74% of health app users stopped using an app by the tenth use and 26% used it only once. To increase the percentage of anglers who volunteer to use the apps, Venturelli et al. (2017) suggest making the apps "easy, fun, and social." Keusch and Zhang (2017) review the efforts of "gamification," a variety of techniques intended to increase respondent engagement in completing web surveys, and show that the benefits, while generally positive, are not very clear.

Although the premise is that the apps will result in a census of fishing activity, this is extremely unlikely for recreational fishing apps. Because apps are intended to be used by anglers who volunteer to do so, instead of a census they constitute a nonprobability sample with no selection probabilities attached to those who do report. One consequence of this method of collecting data is that no techniques for controlling the respondents such as quotas or matched samples (Baker et al. 2013) are possible. Another consequence is that estimation methods that do not rely on selection probabilities must be applied for population inference. A typical scheme for household nonprobability samples is to attach a weight to each respondent so that the weighted total equals the number of persons in the population in a particular category (often demographic categories are used such as age and sex). Even this simple adjustment is not feasible for fishing surveys because the total number of anglers that could serve as the population total is not known from the U.S. Census Bureau or administrative records. For example, counts of licensed anglers would exclude exempted categories, such as kids and anglers who fish from licensed piers, and other anglers who choose to fish without a license.

Another approach that has been discussed in the literature is using some combination of probability and nonprobability samples to gain the advantages of the lower cost of nonprobability sampling but retain the rigor and quality of probability samples. For example, Chen et al. (2020) suggest using a probability sample as a reference sample to reduce the biases that are associated with using data from a nonprobability sample to produce estimates.

Liu et al. (2017) suggested a different and innovative approach to use capture–recapture sampling methods for fishing surveys. The capture–recapture scheme employs both a nonprobability sample and a probability sample, where the key to inferences is examining how many of the probability sample members were originally observed in the nonprobability sample. In Liu et al. (2017), the nonprobability sample data were obtained from apps such as iSnapper and the probability sample was an intercept survey. They found that this method could produce valid estimates and confidence intervals if the reporting rates in the nonprobability sample were sufficiently high, especially among the most avid anglers.

Stokes et al. (2021) further investigated some of the primary assumptions in the capture–recapture estimator. They identified serious departures from some of the assumptions that could result in biases in the estimates. One assumption requires the probability sample and nonprobability sample be statistically independent. They found the independence assumption was violated because some anglers only reported on the app when they were randomly sampled in the intercept survey. Another assumption, called the matching assumption, requires being able to identify whether the reports from the nonprobability sample and the probability sample are for the same trip. In practice, they found the linking of trips to be complicated and error prone, causing this assumption to fail. Stokes et al. (2021) suggested that data collection changes could help reduce matching errors, but they did not have good solutions to the independence assumption violation. As intercept survey coverage is generally limited to publicly accessible fishing access sites, noncoverage error is another potential source of bias in this hybrid design. The extent to which trips returning to private access sites differ from those returning to public sites in terms of app reporting rates and catch characteristics, as well as the proportion of total trips returning to private sites, determine the potential for bias from this noncoverage error.

It is important to understand that probability samples are also subject to violations of assumptions that can result in biases. Groves et al. (2011) summarize the sources of total survey error that include sampling error as well as other sources such as nonresponse and noncoverage. As mentioned above, the exclusion of private access fishing sites from intercept surveys is a source of noncoverage error that could affect estimates from any program that incorporates that sampling design. Stokes et al. (2021) speculated that nonresponse bias in the probability effort survey was the greatest threat to the validity of its estimates. In the next section, the potential biasing effects of nonresponse are discussed for both probability and nonprobability samples.

POTENTIAL BIASES WITH PROBABILITY AND NONPROBABILITY ESTIMATES

As described earlier, the MRIP Fishing Effort Survey estimates the number of trips that anglers took during a specific period and the APAIS estimates means of number of fish caught per angler trip. The product is an estimate of the total catch needed for managing the fisheries. While the overall estimate is an aggregate or total, an important difference between the two survey estimates is that the effort survey estimates population totals while the catch survey estimates population means.

The effort survey estimate is the number of angler trips taken during the period. It can be written as follows:

$$\widehat{y}_{eff} = \frac{\widehat{y}_{pt}}{\widehat{y}_{pt} + \widehat{y}_{np}} \overline{y}_{trip} T, \quad (1)$$

where $\widehat{y}_{pt}$ is the estimated total number of households with residents who fished during the period, $\widehat{y}_{np}$ is the estimated number with residents who did not fish in the period, $\overline{y}_{trip}$ is the estimated mean number of trips for fishing households, and T is the total number of households in the population (known from other sources). We concentrate on the first term, which is the fishing prevalence or estimated proportion of the total population that takes a fishing trip during the period. This quantity is likely to be the primary source of nonresponse bias in fishing effort surveys.

Brick et al. (2016) point out that for estimating fishing effort, probability surveys would be susceptible to nonresponse bias if fishing participants were more likely to respond than nonparticipants were. This would result in an overestimate of fishing prevalence. They describe the design of data collection instruments that appeal to a general audience, increasing the likelihood that nonangling households will respond. The key is that all households are sampled and the survey should not appear to be only for anglers. Furthermore, they suggest weighting adjustment methods to reduce the effects of differential nonresponse between anglers and nonanglers. To examine the potential for nonresponse bias due to differential response between anglers and nonanglers, nonresponse bias studies were undertaken in 2012 and 2020. Andrews et al. (2014) and Andrews (2021) report that in both nonresponse follow-up surveys the respondents to the follow up did not differ from the initial respondents with respect to effort, suggesting that the potential nonresponse bias was not large.

The mobile apps, on the other hand, have no capacity to estimate the number of nonparticipants in the ratio in equation (1), and consequently no capacity to estimate fishing prevalence. The apps are designed for anglers to record their activities while fishing; none of the suggestions of Venturelli et al. (2017) for increasing the use of the apps are intended for those who do not fish.

With this type of nonprobability sample, the estimator becomes

$$\hat{y}'_{eff} = \hat{y}_{pt} \overline{y}_{trip}. \quad (2)$$

The potential for bias becomes much greater with this estimator compared with equation (1) for two reasons. First, the estimator can no longer take advantage of the auxiliary data, T, to reduce bias. Second, unless all people who fished in the period use the app for every trip (i.e., a census) the estimate will be biased downward. Given the experiences of the use of apps that is described in the previous section, the bias will be large and consistent in direction. It is worth noting that none of the articles on using mobile apps suggest using an app to estimate fishing effort or fishing prevalence, but instead they suggest that it could be used for other purposes. As noted earlier, the lack of a reliable number or estimate of the number of people who fished during the period makes it difficult to improve on the accuracy of equation (2). As mentioned earlier, administrative data on registered anglers are incomplete due to unlicensed fishing activity, and even if it were complete the data would not provide counts of the number of anglers who took trips during the specified period.

For catch surveys, the estimates of catch rates or the number of fish caught per trip are means, where the means are like $\hat{y}_{pt}$ for characteristics such as average number of fish caught per trip or average weight of the catch. Valliant (2013) points out that estimating means generally has an advantage over estimating totals because means are ratios (number of trips taken divided by the number of persons taking at least one trip) and the biases in the numerator and denominator of the ratios may be partially offsetting. When this holds, the bias for the estimated mean is reduced compared with separate estimates of the numerator and denominator. Some researchers have considered using nonprobability samples for estimating means, proportions, and other relationships because of this observation. Notice that unlike effort estimates, the catch rate estimates that are derived from mobile apps do not require any data from those who do not fish during the period.

While the bias for estimating means is generally lower than for totals, it does not mean that the estimates are unbiased regardless of the source. In probability samples, the bias for a mean is

$$bias\left(\bar{y}_{pr}\right) \cong (1 - R)\left(\overline{Y}_r - \overline{Y}_m\right), \quad (3)$$

where R is the percentage of the sample responding, $\overline{Y}_r$ is the mean of the respondents, and $\overline{Y}_m$ is the mean of those not responding. Bias is greatest when the response rate is low and the difference between the respondents and nonrespondents is high. The APAIS has low nonresponse rates, and nonresponse bias is not likely to be a source of substantial bias in that survey, although noncoverage bias due to excluding private fishing access sites could be more problematic.

A similar expression holds for nonprobability estimates of a mean

$$bias\left(\bar{y}_{np}\right) \cong (1 - P)\left(\overline{Y}_R - \overline{Y}_M\right), \quad (4)$$

where P is the percentage of the population responding, $\overline{Y}_R$ is the mean of the respondents, and $\overline{Y}_M$ is the mean of those not responding.

We begin by examining the components of equation (3) and equation (4). The quantity P refers to the entire population, whereas R refers to the sample. In a probability sample where equation (3) applies, data collection efforts are concentrated on getting the sampled units to respond by making multiple contact attempts and using other tools like monetary incentives. In a nonprobability sample where equation (4) applies, any efforts to increase the first term, making it closer to equation (1), requires doing the same type of work for the whole population of anglers—encouraging more anglers to use a fishing app, for example.

The second terms in equation (3) and equation (4), respectively, are more related to Valliant's (2013) comment. These terms become relatively small when a conditional independence assumption holds. For example, the assumption is that catch rates are equivalent for app users and nonusers. The requirement is that catch rates are independent of using the app.

The relative sizes of $\left(\overline{Y}_r - \overline{Y}_m\right)$ and $\left(\overline{Y}_R - \overline{Y}_M\right)$ depend on the specific application. For a catch survey, we suspect that typically $\left(\overline{Y}_r - \overline{Y}_m\right)$ is less than $\left(\overline{Y}_R - \overline{Y}_M\right)$. This would hold if the likelihood of responding and catch rate were not as highly correlated as the likelihood of using the app and the catch rate. While we suspect that this is the case, the main contributor to the bias mostly likely arises due to the differences between R and P.

Throughout this discussion, we have described the bias due to differences between the respondents and the full population in terms of nonresponse bias. It is more typical to classify this bias in nonprobability samples as selection bias. Bethlehem (2010) refers to selection bias as the difference between the responding sample and the population being the result of self-selection rather than using probability sampling. An example of self-selection bias is if anglers only reported successful fishing trips. Bethlehem indicates that selection bias is the greatest problem in nonprobability samples. We used the nonresponse formulation to be able to use the same structure for both types of samples. Meng (2018) and Kalton (2019) also discuss some of the issues with nonprobability sampling, but they primarily show that the effective sample size in nonprobability samples can be dramatically lower than the nominal sample size.

Even without resolving the magnitudes of the second terms in the biases, the differences between R and P would almost certainly result in considerably smaller biases in catch estimates for probability samples. One situation where anglers might use an app on a regular basis and make P closer to R is when reporting is mandatory. A mandatory reporting of Red Snapper in Alabama using Snapper Check for private boat captains did have higher reporting rates, but they were still only approximately 50% (Alabama 2019). On the other hand, if a probability sample is conducted with little emphasis on response rates, the value of R could also be low. The difference depends on the specific circumstances.

SUPPLEMENTARY USES OF NONPROBABILITY APP DATA

A common theme in the early research into big data and nonprobability sampling methods is that probability samples are too costly and slow and collecting inexpensive data from nonprobability samples fills a void. Several researchers have suggested that apps might provide data to supplement existing programs that do not collect information on a topic of interest.

Papenfuss et al. (2015) proposed that apps could be used to provide fine-scale movement data that are not available from traditional surveys as well as a platform for on-demand angler surveys and for quickly launching projects to collect data passively. Venturelli et al. (2017) discussed apps as a source of information on topics such as bait and tackle, depth, fish kills, invasive species, and water conditions. Jiorle et al. (2016) considered apps for collecting data on discarded fish characteristics, greater spatial resolution, and for low-effort or rarely encountered fisheries where small sample sizes lead to unstable estimates in probability samples.

For example, length of discarded catch is not collected in the APAIS but could be of value in managing fisheries. As suggested by Jiorle et al. (2016), an app could be designed to ask questions about discarded catch such as the number, species, and length of the discarded fish. The bias from the data that are collected in this manner could be computed using equation (4) and speculating about different values for the quantities in that equation. Suppose that 20% of anglers used the app ($P = 0.2$) and the mean number of discards for those using the app is 1.5 and for those not using the app is 1.2. The true mean number of discards is 1.275, but the estimate based on just app users is 1.5 (bias = 0.225, or 18% of the mean). The question is whether an estimate with this magnitude of bias would be useful or could be misleading. A sensitivity analysis could be conducted by substituting other values for the parameters. The bias goes to 0 as P goes to 1, or the difference between app users and nonusers goes to zero.

Other supplementary uses that have been suggested might be more like observational studies such as changes in the spatial distributions (ranges) of species and the potential link to climate change. Similarly, small-scale localized studies might be mounted. These examples of supplementary uses differ from the discard example given above because population-level inferences are not the goal. Rather, these types of supplementary uses might generate hypotheses about processes that are worthy of further investigation. Apps could also provide longitudinal data for anglers over time. The main concern with this use is the high attrition rate for those who do start reporting on apps. Wenz et al. (In press) attempted to increase both the participation and retention rate of app use by providing feedback to app users in a longitudinal survey but had no success.

DISCUSSION

Fishing effort and catch surveys pose many challenges irrespective of the sampling. However, the sampling method has major implications. With probability samples, the costs of data collection are generally considerably higher than when nonprobability samples are used, and the probability samples must use methods to ensure relatively high-quality data if the advantages of probability sampling are to be realized. By high quality, we mean keeping both sampling errors and the nonsampling errors relatively small. The sampling errors can be made as small as desired by increasing the sample size, although there are cost implications. Nonsampling errors are more difficult to address, but there are known techniques to reduce these biases. For example, nonresponse bias for a fishing effort survey can be reduced by increasing the response rate so that the rate for those who do not fish is similar to the rate for those who do fish. In catch surveys, techniques can be instituted to sample and interview a sample that is balanced with respect to catch and other important trip characteristics to reduce bias.

Nonprobability fishing surveys face substantial challenges if they are going to produce high-quality estimates for statistical inferences. The challenges exist for both effort and catch surveys, but finding ways to encourage those who do not fish on a regular basis to use a fishing app to report effort seems to be an unreachable goal. It is difficult to contemplate how this would be possible. The National Academies of Sciences, Engineering, and Medicine (2021) recently stated

> The potential for voluntary reporting to enhance fishery data collection has generated much excitement, but in practice, participation in such programs has invariably been extremely low. Unless these patterns are reversed, reliance on such voluntary data collection systems is unlikely to advance MRIP over the coming years.

If probability surveys for fishing effort are to be replaced, methods other than using apps are needed.

While an app may be infeasible for estimating fishing effort, an app could be considered for estimating catch rates because catch rate estimates only require input from anglers and the quantities being estimated are means rather than totals. However, there are still considerable challenges that would have to be overcome to rely on apps for estimating catch characteristics. First, the apps would have to be used by a much larger proportion of anglers than has been observed in any of the early evaluations. The anglers would have to install and use the app and continue to use the app over time. The current evidence regarding keeping users after a single use is not very encouraging. Mandatory reporting could improve the percentage of anglers who use an app, but it is unclear whether the greatly increased response burden on anglers would be sustainable to implement mandatory reporting for all or even a subset of managed fisheries.

Greater adoption of mobile app reporting is also one of the conditions for making the capture–recapture design more effective. With this design, the problems with matching errors and the conditional independence assumption still need solutions. The role of noncoverage bias due to excluding private fishing access sites in the probability samples also requires some investigation for this application.

Some data, especially for those who are without the resources that are needed to collect high-quality data from probability samples, seems better than no data at all. Perhaps there are situations in which this holds, such as some of the supplementary uses where population inferences are not necessary. Nevertheless, the absence of data may be better than poor-quality data in many cases. For example, a decision on whether to restrict fishing for a particular species might be the outcome of the analysis of fishing survey data. If the estimates from a survey are seriously biased, it might lead to restricting fishing when it should be open. The opposite result of allowing fishing when it should be restricted could have even more grim consequences. When the policy has such consequences, it is important to have estimates and confidence intervals for those estimates that can be trusted. Currently, the best general approach to providing such estimates is by using probability samples.

Our view is that nonprobability samples should first be studied and evaluated in situations where the effects of wrong decisions are not serious. Baker et al. (2013) refer to this idea as "fit for use." Currently, the application of nonprobability sampling for fishing effort and catch surveys have not proven that they are fit for use.

One method that has been used in social science research to examine the quality of nonprobability sampling is to produce estimates from both probability and nonprobability samples and compare them with each other or with benchmark estimates from some gold standard. Callegaro et al. (2014) do a comprehensive review of such studies. Yeager et al. (2011) experimented with using the same instrument in multiple nonprobability samples and in a probability sample. By highlighting issues that would otherwise go undetected, such comparisons have value even if they do not generalize directly to other applications. The existing studies have shown that the probability samples typically produce estimates with smaller biases and that the estimates from nonprobability samples vary significantly from each other in terms of average absolute bias.

Jiorle et al. (2016) did a small comparison of estimates from a mobile app to estimates from a probability sample in a fishing survey setting. However, these comparisons were not the equivalent of those done in the social science research noted above. They made only a few comparisons and considered only a few species that have very limited variation in catch rates and in very limited geographic areas. Furthermore, the comparisons were to observations from a probability sample not estimates from the sample. Thus, this method of evaluation has not been explored much for fishing surveys. Without robust evaluation of nonprobability samples for fishing surveys, there is no evidence that they are fit for use in providing population-level inferences for recreational fishing effort and catch.

ORCID

J. Michael Brick https://orcid.org/0000-0003-3490-8925

REFERENCES

Ahrens, R. 2013. Assessing and refining the collection of app-based angler information in relation to stock assessment. National Oceanic and Atmospheric Association, Marine Recreational Information Program FY-2013, Washington, D.C.

Alabama. 2019. 2019 Alabama exempted fishing permit status report. Available: https://www.outdooralabama.com/sites/default/files/2019%20AL%20DCNR%20EFP%20Update%20No.%204.pdf. (November 2021).

Andrews, R. 2021. Evaluating nonresponse bias in the MRIP Fishing Effort Survey. Available: https://apps-st.fisheries.noaa.gov/pims/main/public?method=DOWNLOAD_FR_DATA&record_id=2018. (August 2021).

Andrews, R., J. M. Brick, and N. Mathiowetz. 2014. Development and testing of recreational fishing effort surveys: testing a mail survey design. Available: https://www.st.nmfs.noaa.gov/pims/main/public?method=DOWNLOAD_FR_PDF&record_id=1179. (February 2021).

Baker, R., J. M. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. J. Gile, and R. Tourangeau. 2013. Summary report of the AAPOR task force on non-probability sampling. Journal of Survey Statistics and Methodology 1:90–143.

Bethlehem, J. 2010. Selection bias in web surveys. International Statistical Review 7:161–188.

Brick, J. M., W. R. Andrews, and N. A. Mathiowetz. 2016. Single-phase mail survey design for rare population subgroups. Field Methods 28:381–395.

Callegaro, M., A. Villar, D. S. Yeager, and J. A. Krosnick. 2014. A critical review of studies investigating the quality of data obtained with online panels based on probability and nonprobability samples. Pages 23–50 *in* M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, and P. J. Lavrakas, editors. Online panel research: a data quality perspective. Wiley, Chichester, UK.

Chen, Y., P. Li, and C. Wu. 2020. Doubly robust inference with nonprobability survey samples. Journal of the American Statistical Association 115:2011–2021.

Coleman, J. S. 1958. Relational analysis: the study of social organizations with survey methods. Human Organization 17:28–36.

Groves, R., F. Fowler, M. Couper, J. Lepkowski, E. Singer, and R. Tourangeau. 2011. Survey Methodology. Wiley, Hoboken, New Jersey.

Jiorle, R. P., R. N. Ahrens, and M. Allen. 2016. Assessing the utility of a smartphone app for recreational fishery catch data. Fisheries 41:758–766.

Kalton, G. 2019. Developments in survey research over the past 60 years: a personal perspective. International Statistical Review 87:S10–S30.

Keusch, F., and C. Zhang. 2017. A review of issues in gamified surveys. Social Science Computer Review 35:147–166.

Lazer, D., R. Kennedy, G. King, and A. Vespignani. 2014. The parable of Google flu: traps in big data analysis. Science 343(6176):1203–1205.

Liu, B., L. Stokes, T. Topping, and G. Stunz. 2017. Estimation of a total from a population of unknown size and application to estimating recreational Red Snapper catch in Texas. Journal of Survey Statistics and Methodology 5:350–371.

Meng, X. L. 2018. Statistical paradises and paradoxes in big data (I): law of large populations, big data paradox, and the 2016 US presidential election. Annals of Applied Statistics 12:685–726.

National Academies of Sciences, Engineering, and Medicine. 2021. Data and management strategies for recreational fisheries with annual catch limits. Available: https://www.nationalacademies.org/our-work/data-and-management-strategies-for-recreational-fisheries-with-annual-catch-limits. (August 2021).

National Research Council. 2006. Review of recreational fisheries survey methods. Available: https://www.nap.edu/catalog/11616/review-of-recreational-fisheries-survey-methods. (September 2006).

Papacostas, K., and J. Foster. 2018. National Marine Fisheries Service's Marine Recreational Information Program survey design and statistical methods for estimation of recreational fisheries catch and effort. Available: https://www.fisheries.noaa.gov/resource/document/survey-design-and-statistical-methods-estimation-recreational-fisheries-catch-and. (March 2021).

Papenfuss, J. T., N. Phelps, D. Fulton, and P. A. Venturelli. 2015. Smartphones reveal angler behavior: a case study of a popular mobile fishing application in Alberta, Canada. Fisheries 40:318–327.

Pervaiz, F., M. Pervaiz, N. A. Rehman, and U. Saif. 2012. FluBreaks: early epidemic detection from Google flu trends. Journal of Medical Internet Research 145:125.

Rodríguez-Mazahua, L., C. A. Rodríguez-Enríquez, J. L. Sánchez-Cervantes, J. Cervantes, J. L. García-Alcaraz, and G. Alor-Hernández. 2016. A general perspective of big data: applications, tools, challenges and trends. Journal of Supercomputing 72:3073–3113.

Royall, R. 1970. On finite population sampling theory under certain linear regression models. Biometrika 57:377–387.

Shi, Q., and M. Abdel-Aty. 2015. Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. Transportation Research Part C: Emerging Technologies 58:380–394.

Stephan, F. F., and P. J. McCarthy. 1958. Sampling opinions: an analysis of survey procedure. Wiley, Hoboken, New Jersey.

Stokes, S. L., B. Williams, R. McShane, and S. Zalsha. 2021. The impact of nonsampling errors on estimators of catch from electronic reporting systems. Journal of Survey Statistics and Methodology 9:159–184.

Toole, J., S. Colak, B. Sturt, L. Alexander, A. Evsukoff, and M. González. 2015. The path most traveled: travel demand estimation using big data resources. Transportation Research Part C: Emerging Technologies 58(Part B):162–177.

U.S. Office of Management and Budget. 2006. Standards and guidelines for statistical surveys. Federal Register 71(184):55522–55523. Available: https://www.whitehouse.gov/omb/information-regulatory-affairs/statistical-programs-standards/. (November 2021).

Valliant, R. 2013. Comment. Journal of Survey Statistics and Methodology 1:105–111.

Venturelli, P. A., K. Hyder, and C. Skov. 2017. Angler apps as a source of recreational fisheries data: opportunities, challenges and proposed standards. Fish and Fisheries 18:578–595.

Wenz, A., A. Jäckle, J. Burton, and M. Couper. In press. The effects of personalized feedback on participation and reporting in mobile app data collection. Social Science Computer Review. DOI: 10.1177/0894439320914261.

Yang, X., L. Ma, X. Zhao, and A. Kankanhalli. 2020. Factors influencing user's adherence to physical activity applications: a scoping literature review and future directions. International Journal of Medical Informatics [online serial] 134:article 104039.

Yeager, D. S., J. Krosnick, C. Chang, H. Javitz, M. Levendusky, A. Simpser, and R. Wang. 2011. Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. Public Opinion Quarterly 75:709–747.