Design and Estimation for Recreational Fisheries Surveys

Jean D. Opsomer[1,2], F. Jay Breidt[1]
[1] Colorado State University, Fort Collins, CO, USA
[2] Corresponding author: Jean D. Opsomer, email: jopsomer@stat.colostate.edu

Abstract

We describe new design and estimation approaches being implemented for US recreational marine fisheries surveys. The goal of the surveys is to estimate total recreational catch and its characteristics for fish species of interest. This multi–year development project is a collaborative effort with US National Marine Fisheries Service scientists and included a large–scale field pilot study in 2010. We discuss some of the challenges associated with implementing statistically valid angler intercept surveys for the full coastline of the US, and the solutions proposed to address them.
Key words: intercept survey, weighting, relative efficiency comparison.

1 Introduction

Many coastal countries implement surveys to estimate the marine catch of species of commercial, recreational or ecological importance, with the ultimate goal of monitoring the health of the stock in these species. In the US, data on both the commercial and recreational catch are collected by the Fisheries Service of the National Oceanic and Atmospheric Administration (NOAA). Here, we discuss the recreational side, which is obtained through the Marine Recreational Fisheries Statistics Survey (MRFSS) since 1979.

In 2006, an expert review conducted by the US National Research Council (Sullivan et al. 2006) identified problems in the Access Point Angler Intercept Survey (APAIS) that the NOAA Fisheries Service has conducted for many years as a component of the MRFSS. The survey estimators and measures of precision were not accounting for the complex sampling design, the data collection protocols were combining formal randomization with subjective decision-making in ways that made it difficult to develop statistically valid estimators, and the spatiotemporal sampling frame was incomplete. NOAA initiated work in 2008 to address these concerns with the help of outside consultants, including the two co-authors of this paper. In this paper, we describe some of the new design and estimation methods that are being implemented as part of the new MRFSS, and the results of large-scale field pilot studies that were recently completed.

The MRFSS is based on a complementary survey design that includes a telephone survey to estimate effort (i.e. total number of trips taken by recreational anglers) and a shoreside intercept survey to estimate catch per trip. The MRFSS is conducted in all Atlantic and

Gulf Coast US states except Texas, which conducts its own surveys. Data from the two independent surveys are combined to estimate total fishing effort, participation, and catch by species. The effort survey is currently implemented as an RDD (random-digit dialing) survey of coastal counties, although investigation of a replacement by a dual-frame survey, with a list of angling license holders as the primary frame, is on-going. In this paper, we concentrate on the intercept survey, where most of the efforts at improving the design and estimation have been focused.

The APAIS is a stratified, multistage unequal probability survey. The strata are defined as combinations of waves (two-month periods), states (or geographical areas within states), mode of fishing (shore fishing, fishing from private boat or fishing from a rented/chartered boat) and weekend/weekday. The stages of sampling are the fishing site-day combinations, and either boats or angler groups depending on the mode. The site-days are selected from a list of public access sites with unequal probability according to angling pressure (a measure of the amount of angling expected at a site), while the selection within the site-days is supposed to be equal probability among the anglers having completed their fishing activities that day.

2 Correcting past estimates

Sullivan et al. (2006) identified three key issues with the intercept surveys. First, the sampling design included a number of aspects that were not following formal sampling protocols, including allowing the interviewers discretion to switch sampling site or mode during their assignments, depending on the presence of sufficient numbers of anglers to achieve their interview number targets. Second, even though sampling of sites was done with unequal probabilities, the estimation was based almost exclusively on simple averaging of the observations, and ignored the multi-stage nature of the sampling in the computation of the standard errors. Third, even though intercepted anglers at a site were supposed to represent the fishing activity of the entire 24-hour period at the site, the interviewer assignments exclusively targeted the busiest time of the day, resulting in undercoverage of trips.

Going forward, new design and estimation procedure have been developed that avoid all three of these issues, as will be discussed in the next section. However, NOAA staff and stakeholders also required that previously reported catch estimates be corrected, so that consistent time series of catch data going back 5-10 years are available to conduct stock assessments. Therefore, we developed a "pseudo-weighting" approach that attempts to mimic the sampling process that resulted in the observed sample. It starts from the historical data on the site angling pressure, and then adjusts for the observed assignment switches. While the resulting weights are only partly design-based, they can be justified based on reasonable modeling assumptions and they are able to account for the large differences in the observed frequencies of visits to different angling sites.

While the pseudo-weights constructed as above adjust for the effects of the site selection process, a more difficult issue was to account for the undercoverage in anglers at the sites due to the fact that only part of the day was observed. We developed a model-based adjustment for the fraction of anglers that are present at the site for a given window of interviewing time. The adjustment requires estimation of the circular (time-of-day) distribution of the departures of anglers at sites, which can be done based on large dataset of individual trips

available from the telephone effort survey. The model is a so-called *projected normal* regression model (Presnell et al. 1998), for which we developed a Bayesian hierarchical specification and estimation method.

Once the estimated time-of-day departure distribution were available, within-site weights that account for the fraction of day at the site were constructed. These weights were combined with the pseudo-weights for the site selection, and new weighted estimates and corresponding standard errors were produced for the historical catch data.

3 New sampling design and estimation procedures

Going forward, new sampling design procedures were developed to put the APAIS on a firmer methodological footing. The most obvious non-random elements of the existing APAIS, in particular the interviewer discretion to switch site or mode, were removed from the procedures. The most far-reaching practical change is that the stratification now includes 4 time-of-day strata of 6 hours each instead of a single day, completely removing the undercoverage of angling trips due to sampling only in the middle of the day. However, in order to avoid sending interviewers to sites or at times with very low angling activity for 6 hours, a method to randomly cycle between several neighboring low-activity sites during an interview shift was developed.

Another major change was due to the recognition that the number of anglers (both intercepted and not intercepted) at the fishing site is a crucial component of the weighting procedure, since they are needed to obtain the within-site weights. New procedures were therefore put in place to ensure that this number is accurately recorded as part of the on-site intercept survey protocols. Finally, the selection of anglers and of the angler's catch was made more rigorous and fully randomized, to remove some of the smaller but still potentially troublesome sources of potential selection bias in the estimates.

4 Pilot study and analysis

The new procedures were implemented in 2010 in North Carolina in a large-scale field pilot study, in parallel with the on-going traditional APAIS approach. The pilot study was implemented for a whole year and involved over 1,000 assignments by 6-10 field interviewers (depending on the season). The report on the results of the pilot study have recently been released as an internal NOAA document (Breidt et al. 2013), pending external review.

At the conclusion of the data collection, estimates were computed for both the new and the traditional design, to assess the relative efficiency of both and also look retroactively for any obvious selection biases in the existing survey. The relative efficiency was of particular interest to NOAA, since there was a concern that the new procedures, while statistically correct, would result in estimates that are either too costly to obtain or too variable to use in stock assessments. However, since the pilot study was much smaller than the existing survey and no attempt had been made to stratify it efficiently, a direct comparison was not possible. A "hypothetical pilot" was therefore created to match the sample size of the APAIS and with an allocation that is closer to optimal, so that a more appropriate comparison could be

made. This comparison showed that implementing the new sampling design and estimation procedures would not result in appreciable losses in efficiency, and might actually result in an increase.

References

Breidt, F. J., J. R. Chromy, K. E. Fitzpatrick, H.-L. Lai, T. Menzel, D. G. Mumford, B. Muñoz, J. D. Opsomer, R. J. Salz, K. M. Sullivan, D. A. Van Voorhees, C. Wilson, and P. A. Zielinski (2013). A pilot study of a new sampling design for the Access Point Angler Intercept Survey. Internal report, NOAA.

Presnell, B., S. P. Morrison, and R. C. Littell (1998). Projected multivariate linear models for directional data. *Journal of the American Statistical Association 93*(443), 1068–1077.

Sullivan, P. J., F. J. Breidt, R. B. Ditton, B. A. Knuth, B. M. Leaman, V. M. O'Connell, G. R. Parsons, K. H. Pollock, S. J. Smith, and S. L. Stokes (2006). *Review of Recreational Fisheries Survey Methods*. Washington, DC: National Academies Press.