Center for Independent Experts Independent Peer Review

of

Puerto Rico Port Sampling and Catch Validation Project

Jon Helge Vølstad[1]

[1] Strangehagen 22, 5011 Bergen, Norway; jon.helge.volstad@gmail.com

Executive Summary

My review focused on the evaluation of methods, statistical design, implementation, catch expansion algorithm, and recommendations, as provided in the comprehensive study report that documented the study. I am generally impressed by the statistical rigor of the survey design and estimation methods, and of the execution of this study given the many logistical challenges – particularly related to the effects of hurricanes Irma and Maria. The survey and expansion methods as implemented are viable options for a future monitoring program with minor modifications in my opinion.

I also generally agree with the recommendations for future studies. I here provide some suggestions for improvements that might be considered. The general port sampling survey for daytime sampling employed an access-point design described as two-stage sampling with days as primary sampling units and sites (landing ports) as secondary sampling units. The landings sites were stratified into high- and low-usage based on mapping of effort, and the year was stratified by weeks. Trips were subsampled within sites in some strata, thus adding a third sampling stage. Since data were collected from most trips, it is reasonable to ignore this variance contribution in the estimates of total catch. The general design involved random sampling of days each week, and random sampling of one site per day weekly, complemented with random selection of two sites for a random number of days per month. Although this is a sound survey design that supports unbiased variance estimators, I recommend that future surveys eliminate this component of two sites per day. I recommend that a simpler two-stage survey design be employed in the future, using a sampling frame defined by the matrix of days and sites, with site-days as primary sampling units (PSUs) and trips as secondary sampling units. I also recommend employing restricted random sampling of PSUs to ensure only one site per day for all weeks. This minimizes clustering of site-days and will likely increase the precision in catch estimates for fixed cost. The variance estimator for this two-stage design can be used to assess the effects of subsampling trips within PSUs. The subsampling of trips, instead of a census, can free up time for other data collections and provide more time for species identification and will likely have minimal impact on the precision of total catch estimates.

If the current highly stratified design is continued, it may be advantageous to optimize the allocation of samples (site-days) across strata, if feasible, given staffing and logistical constraints. I also agree that an option is to collapse the usage strata and select site-days proportional to expected effort. Both options can be investigated through simulation studies using empirical data. If an alternative design is viable, I propose that it be tested through a built-in experiment within a standard monitoring program. Alternative survey designs could be tested for random or systematic-random weeks over the year. I see few immediate options for using complemented methods with independent effort surveys combined with CPUE from access-point surveys. Possibly, the use of camera or video from drones, and automated image analysis through machine learning can become viable. For the bus-route design, even waiting time among sites may be beneficial, and the use of a bootstrap variance estimator is an option.

Description of the Individual Reviewer's Role in the Review Activities

This Center for Independent Experts (CIE) desk review of the proposed design for a Puerto Rico commercial fishery port sampling survey was conducted independently by Dr. Cynthia Jones and me. We have collective expertise in statistical survey sampling methods, including survey design and catch/effort expansion for fishery-dependent surveys. We have comprehensive knowledge of both theoretical and applied sampling survey methods applied to marine fisheries. Dr. Jones is a world-renowned expert on the design and analysis of marine angling surveys and quantitative fisheries biology and has published peer reviewed papers on methods related to roving creel and bus-route surveys that are particularly relevant for this review.

I bring international research and management experience in quantitative fisheries biology and ecological statistics, specializing in statistical survey sampling methods. I have broad hands-on experience in the development and optimization of fisheries-dependent and fisheries-independent surveys and monitoring programs to support stock assessments and ecosystem-based fisheries management.

My extensive experience from committee work in the International Council for the Exploration of the Sea (ICES) on the development of practical implementation of statistically sound catch sampling programs is particularly relevant for this review. My experience with the design and analysis of artisanal and recreational fisheries surveys, including access-point surveys, aerial surveys and roving creel surveys of recreational fisheries are also relevant. In this review, I have focused especially on aspects related to the survey design of the multi-stage access-point surveys and provide some thoughts on possible future improvements. I defer to Dr. Jones for more detailed evaluation of the bus-route surveys.

The review started with a webinar on May 16th. A team, consisting of Kevin McCarthy (Chief, Caribbean Fisheries Division; NOAA Southeast Fisheries Center), Shannon Calay (Sustainable Fisheries Division Director for the NOAA Southeast Fisheries Science Center), and Mark Chandler (Fishery Biologist, NOAA HQ) provided an overview of the review objectives and brief overview of the Puerto Rico port sampling project and the background material. Kevin McCarthy distributed materials to the review panel via a folder on Google Drive, where we could find the Final report on the Puerto Rico Port Sampling Project, appendices including site descriptions, and background material (Appendix 1). Manoj Shivlani (Lead coordinator for the Center of Independent Experts) informed about the review process and terms of reference (Appendix 2). During the webinar on May 16 Cynthia Jones had a question regarding the bus-route survey that was answered by NOAA the following day (Appendix 3).

Summary of Findings for each ToR in which the weaknesses and strengths are described

1. Evaluate the statistical design and implementation of the port sampling survey to estimate commercial catch by species and the methods used to develop the design.

I am very impressed by the statistical rigor of the design and the implementation of the ports sampling survey – and especially how the team has utilized the pilot study and dealt with cost- and logistical constraints. The way the team adjusted the sampling frame, survey design, and implementation after the devastating effects of hurricanes Irma and Maria on the fishing community is commendable. This study will be exceptionally useful when designing a long-term monitoring program. The survey methods employed, with small adjustments, are suitable for long-term monitoring. I only have minor comments and some suggestions that may be considered in future surveys.

Daytime Design (East, West, South)

The general port sampling survey for daytime sampling (9:00 AM to 5:00 PM) is a classic access-point survey where data on effort and landed catch are collected from completed trips. Such surveys are generally highly complex with multiple stages of sampling, with additional challenges related to survey design and implementation caused by logistical constraints (Lehtonen and Pahkinen 2004). The survey design employed is described as two-stage sampling with days as primary sampling units, and ports as secondary sampling units. Strictly, there is a third sampling stage when trips are subsampled within days and sites. The days in a year are stratified into weeks, and the ports (sites) are stratified into high-use and low-use strata. At the first stage, 5 days out of 6 (Sundays were excluded because of expected low effort) were randomly selected, and at the second stage, one port was randomly selected in each of the high-use and low-use strata. Strictly, the survey includes a third sampling stage within a site-day, since it is not always possible to collect data by census of all trips, though most trips were sampled except in the Puerto Real area (see below). There are some inconsistencies in how the survey design and sampling frame is described in the report. In some sections of the report, site-days are listed as primary sampling units, although days are sampled in the first stage, with sites as secondary sampling units.

The design (with one site per day) when followed ensures that the probability of visiting any fishing site at any day is equal within each stratum. Hence, estimates of total catch and effort will be unbiased. Strictly, this design does not ensure that any combination of two site-days have a known inclusion probability above zero, which is a requirement for unbiased variance estimation. If the addition of two-sites per day sampling ensures that any two sites within days has a known inclusion probability larger than zero in a stratum, then this design supports unbiased variance estimation using the Hansen-Hurwitz Estimator (see Cochran 1977). However, I do not think that this additional sampling is cost-effective, since such clustering of sites likely reduces the effective sample size (Lehtonen and Pahkinen 2004).

When one site is sampled per randomly selected day, this restricted random sampling minimizes clustering of sites in the sample. I argue that this is equivalent to a two-stage survey design, with primary sampling units (PSU) being site-days, and secondary sampling units (SSU) being trips. The sampling frame is then made up of all site-days in a year, which can be stratified by weeks and high- and low-usage.

If the sampling fraction of site-days is low within a stratum, say less than 10%, then it is reasonable to assume sampling with replacement and base the variance estimation on PSU (Williams 2000; Wolter 1985). To ensure good spatial and temporal coverage in each stratum, it is recommended to use restricted random sampling (Jessen, 1978), with sampling of fishing sites scheduled on random days during each calendar week from each stratum.

Puerto Real

Some departures from the general protocol for data collections were forced by logistical constraints in the Puerto Rico West high-use stratum. One site (Puerto Real Soltero, SOL) within a cluster of 5 sites in Puerto Real refused samplers permission to collect data. Effort data for this site were collected when a neighboring site was sampled. MER devised a sound approach for imputing catch estimates for SOL.

Daytime Design (North)

In the Puerto Rico North coast, the high-use stratum contained only one port (Jarealito). One port-sampler was assigned to cover both the high-use and low-use strata, and one port-sampler was assigned to the high-use stratum only. Hence, the low- and high-use strata were not sampled independently. Initial oversampling of the high-use stratum was reduced to two days per week from July 2018, the other days being allocated to the low-use stratum. By collapsing the two strata, and use of unequal probability sampling of days and sites would solve the issue of independence. This would also allow one to set an inclusion probability for Jarealito that would result in an acceptable frequency of visits.

Auxiliary Sampling

Off-hour Sampling (AM, PM, Sundays)

Bus-route surveys

The bus-route surveys employed seems reasonable for obtaining data on the coverage of the daytime surveys and are useful for the planning of future surveys. The surveys were conducted to access catch and effort in the early morning (AM; 4 hours before 9:00 AM), late evening (PM; 4 hours after 5:00 PM), and on Sundays (Daytime, 9:00 AM to 5:00 PM) and were described as a Roving design, collecting data on randomly selected days four times per month from multiple ports along a predetermined route according to protocols resembling a bus-route design. The starting point of the route was a randomly picked port. A roving design (Hoenig et al. 1997) would entail that data are collected from incomplete trips by intercepting fishers when they are actively fishing. In the surveys described here, data are collected from complete trips, so the term Roving design seems inappropriate.

Chen and Woolcock (1999) provides a theoretical framework for using bus-route surveys to estimate recreational fishing effort that is applicable for this survey. It could be useful to collect information about departure times from the trips sampled. These data could be used for the design of future bus-route surveys considering the arrival and departure distributions of the fishing trips.

An evaluation by Kinloch et al. (1997) of a large recreational bus route survey in Australia suggest that accuracy of effort estimates was not significantly affected by the size of primary or secondary sampling units or by the allocation of waiting times to ramps either equally or proportional to effort. The optimal design proved to be a combination of a smaller PSU with a larger SSU and waiting times allocated equally at all ramps. The variance was driven by the day-to-day variability in fishing effort which is influenced by factors such as weather and season. Perhaps this design is applicable for this study.

Alternative Designs

The alternative survey approaches for sites in the sampling frame where the above roving design was not practical for collecting data outside the daytime hours (due to travel costs or other constraints) focused on sites with expected high effort to provide upper bounds for catch estimates. This was reasonable, as these studies were conducted within cost-constraints. The cost-effective approach employed was to extend daytime sampling with AM or PM sampling within the standard access-point surveys for the east and north coasts. For surveys with one site per randomly selected day, four random sampling days per month were extended with AM or PM sampling. Although not specified, I assume that they extended the sampling evenly between AM and PM in the random picks (i.e., two random AM samples, and two random PM samples). For AM or PM surveys with two sites for some days, the AM and PM sampling were conducted for two random days per month for each replicate site. This approach eliminates extra travel time but may have some costs related to over-time pay for the port sampler. In some periods over the year this auxiliary sampling was more restricted area-wise due to logistical or other constraints.

2. Evaluate survey expansion algorithm for total catch by species.

<u>Estimator of catch</u>

<u>Daytime Design (East, West, South)</u>

I find the expansion algorithm for total catch by species to be sound, although the variance estimator does not consider that subsampling of trips within a day and site occurs at several ports. However, since the fraction of trips sampled generally is large, the effect of ignoring this third stage variance is likely to be minimal.

For completeness, however, it is recommended to use a variance estimator that also accounts for the subsampling of trips. In future surveys it may be an alternative to reduce the number of trips sampled since the overall variance in catch estimates mostly is driven by the first-stage sampling variance component.

Alternative estimators of catch and its variance can be provided by defining the survey as two-stage sampling with site-days as PSUs and trips within site-days as secondary sampling units (SSUs). Let c_{hij} c_{hij} represent the total catch of a species for sampled trip j j in the PSU ($j = 1,...,m_{hi}$ $j = 1,...,m_{hi}$), either observed directly or estimated. Let $\hat{C}_{hi}$ $\hat{C}_{hi}$be an estimate of the total catch for all trips for PSU i, obtained by extrapolating the mean catch per unit for the m_{hi} units in the sample to all trips M_{hi}. An estimator for total catch in stratum *h* is then obtained by extrapolating the mean catch per PSU to all units in the sampling frame,

$$\hat{C}_h = \frac{N_h}{n_h} \sum_{i=1}^{n_h} \hat{C}_{hi} \qquad (1).$$

The restricted random sampling (one site per day) within strata does not ensure that any combinations of two site-days have a joint inclusion probability above zero, which is a requirement for unbiased variance estimation. A conservative estimator for the variance of (1) under the assumption of simple random sampling of PSUs within strata is:

$$v\left(\hat{C}_h\right) = \frac{1}{n_h - 1}\left\{\frac{N_h^2}{n_h}(1-f_h)\sum_{i=1}^{n_h}\left(M_{hi}\bar{c}_{hi} - \frac{\hat{C}_h}{N_h}\right)^2\right\} + \frac{N_h}{n_h}\sum_{i=1}^{n_h} M_{hi}^2(1-f_{hi})\frac{s_{hi}^2}{m_{hi}} \qquad (2)$$

where $\bar{c}_{hi}$ $\bar{c}_{hi}$ is the mean catch of the m_{hi} trips sampled in PSU i in stratum h, $f_h = n_h/N_h$ is the sampling fraction of PSUs, $f_{hi} = m_{hi}/M_{hi}$ $f_{hi} = m_{hi}/M_{hi}$ is the sampling fraction of trips within the i^{th} PSU, and

$$s_{hi}^2 = \frac{1}{m_{hi} - 1}\sum_{j=1}^{m_{hi}}\left(c_{hij} - \bar{c}_{hi}\right)^2$$

is the estimated (population) variance in catches between fishing units within PSU i (Cochran 1977, pp. 300-303; Wolter 1985). Wang (2017) provides methods for estimating variance for this type of complex design based on bootstrapping.

If the sampling fraction of site-days (PSUs) is low in some strata, say less than 10%, then it is reasonable to assume sampling with replacement (Berger 2004; Williams 2000; Wolter 1985), and a much simpler variance estimator can be computed based on the variation in estimated total catch for the PSUs in the sample (see Cochran, 1977, p 279; Pollock et al., 1994, p. 42; Vølstad et al., 2006, 2014; Williams, 2000):

$$V(\hat{C}_h) = \frac{N_h^2}{n_h}\sum_{i=1}^{n_h}\frac{(\hat{C}_{hi} - \bar{C}_h)^2}{n_h - 1} \qquad (2)$$

If the sampling fraction at the primary level is small ($\frac{n_h}{N_h} < 10\%$), the bias in the variance estimates is likely to be insignificant and slightly positive (Wolter, 1985, p. 34), and thus are conservative.

The accuracy of the total catch estimates clearly depends on the coverage of the landings sites for the sampling frame employed. The MER team have done a thorough mapping and evaluation of the sites before the study was done, so I expect the bias due to pruning of sites to be small.

Daytime Design (North-coast)

The survey design slightly departs from the standard design with one site sampled each random sampling day. For a small number of days, two sites were sampled. I agree with the simple averaging of catches for days where more than one site is sampled, and then using the same estimators of catch as for the standard design with one site per day.

The proposed variance estimator for two-stage sampling, where within-day variance component is based on replicate observations for two days per month, is reasonable.

Bus-route designs

The estimator for total catch for off-hours (AM, PM, and Sundays) seems reasonable. Since only a small fraction of days (< 10%) per month are sampled, it is reasonable to assume sampling of days with replacement and employ a simple variance estimator of total catch on the first sampling stage (day-to-day) variance in catches (Cochran 1977).

3.3. Evaluate the approach for estimating total catch considering other approaches that utilize CPUE and estimates of total effort.

The current approach and sampling effort yields quite accurate results (high precision and low bias). It is difficult to envisage cost-effective and practical methods based on CPUE from access-point surveys and independent estimates of total effort.

Camera monitoring could in theory be used to count the total number of trips by week in selected ports. This would facilitate the estimation of total landed catch by combining estimates of CPUE from the access-point or bus-route surveys with total effort estimates from the camera surveys. The use of camera monitoring to collect independent effort data (trip counts) or auxiliary effort data (traffic data) has proved to increase precision and reduce bias in catch estimates substantially in several surveys of recreational fisheries in Australia (Hartill 2015, 2020; Steffe 2017). These methods still have high data collection and entry costs due to the time necessary to manually count anglers. Automatic image analysis using Artificial Intelligence (AI) will likely facilitate counts at a lower cost soon. However, camera monitoring was tested at selected sites in this study and was abandoned due to vandalism and maintenance issues.

The use of airplanes and high-resolution video or digital photography, and development of automatic image analysis through machine learning is an alternative that could provide snapshot counts of effort at landings sites, or at sea, but this may be too costly. Because each flight is expensive, the number of flights per year is limited and precision of effort estimates may be poor. The use of drones has more promise and may become feasible and cost effective in the future if the boat counts can be automated using AI. For example, US Geological Surveys, Western Ecological Research Center (WERC), are now using machine learning techniques to automate the detection and counting of seabirds and marine mammals from imagery collected in photographic aerial surveys. AI could likely also eliminate privacy concerns by removing info that can identify people or boats. Institute of Marine Research (IMR), Norway is using drones to conduct photographic surveys for the abundance estimation of ice breeding seals (harp and hooded) and coastal seals (grey and harbor seals). The images have been analyzed manually by trained experts. This is time consuming and costly and involves subjective human interpretation. IMR in collaboration with the Norwegian Computing Center is now developing methodology for automatic processing of aerial images (https://www.nr.no/en/projects/uavseal).

Guiardo et al. (2019) propose a large scale generalizable deep learning system for automatically counting whales from satellite and aerial images. They demonstrate proof of concept by applying the method to free Google Earth coastal imagery in 10 whale-watching hotspots. Possibly, similar methods could be used to count trips at landing sites.

Independent effort estimates may also be based on mandatory trip reporting, if deemed reliable.

4. Evaluate recommendations for future work as provided in the survey report. Provide any additional recommendations for future research/improvements for the survey design and expansion methods.

I generally agree with the recommendations for future work as provided in the survey report. I have a few recommendations related to the survey design, extending the impact of the survey program, and quantification of effort.

Survey design

I agree that the survey design employed for daytime sampling appears to be an attractive choice for future daytime sampling work. However, I recommend that the design be simplified to only sample one site per day in a two-stage design with site-days as PSUs, and trips as SSUs. I do not think it is beneficial (or practical) to select two sites on some random days. The restricted random sampling, with one site per randomly selected day, minimizes clustering and spreads out the sampling efficiently in time and space. This will likely result in more precise estimates than for more clustered sampling, where more than one site is selected some days.

One major benefit of a two-stage design with site-days as PSUs, and trips as SSUs, and of using the associated variance estimator for total catch for this design, is that the data from this study can be used to assess the effects on the variance of reducing the number of trips sampled per PSU. I expect minimal loss in precision by sub-sampling trips within PSUs. A reduction in sampling effort within PSUs will free up time for other data collections (e.g., length-age, maturity, genetics, etc.) and provide more time for accurate species identification.

The stratification scheme, and simple random sampling of days and sites within strata proved effective. I agree that the elimination of low- and high-usage strata, and instead sample site-days with probability proportional to expected total catch or effort (proportional to size, *pps*), could be an alternative to the highly stratified scheme that was tested, and that this be evaluated based on simulations.

Another alternative to *pps* sampling is to use the current stratification and optimize the allocation of sample sizes (PSUs) across strata, if feasible within logistical and staffing constraints. Estimates of catches and associated variances by stratum from this study could be used in a Neyman type of allocation (Cochran 1977) where the number of PSUs to be sampled in each stratum is directly proportional to the product of the catch estimate in that stratum and its standard deviation using data from the MER study.

I recommend evaluating such alternative sampling schemes using simulations, based on data from the current study and the pilot study. Such alternative sampling schemes, if deemed practical, could also be tested in an experiment built into a standard monitoring survey, where sampling designs are alternated over the weeks randomly or systematically.

Extending impact of survey program

I agree with the consideration to enhance future port sampling work with the collection of biological samples for age, growth, maturity, and stock identification work. This could easily be achieved if the sampling fraction of trips (SSUs) were reduced.

Quantification of effort

See section 3. I believe the use of drones, and possibly satellite data, may become viable for obtaining independent effort estimates. Possibly, drone flights could cost-effectively be scheduled to provide snapshot counts during the time of expected maximum effort, using effort distributions from the access-point surveys to schedule flight times and adjust for bias (see Vølstad et al. 2006).

Conclusions and Recommendations in accordance with the ToRs

This study is conducted with statistical rigor. I am generally impressed by the design and implementation of the ports sampling survey – and especially how the team has utilized the pilot study and dealt with cost- and logistical constraints. This study will be exceptionally useful when designing a long-term monitoring program. I believe that the survey methods employed, with small adjustments, are suitable for long-term monitoring. I only have minor comments and some suggestions that may be considered in future surveys.

Recommendation 1

For the access-point surveys I recommend that a slightly modified survey design be considered in the future, using a two-stage survey design with the primary sampling units (PSU) being site-days, and secondary sampling units (SSU) being trips. The sampling frame is then a matrix of days crossed with sites, made up of all site days in a year, which can be stratified by weeks and high- and low-usage. For this design, restricted random sampling of site days, ensuring only one site per day, is recommended. This approach minimizes clustering of site-days and is likely to increase the effective sample size and hence precision.

Recommendation 2

Consider subsampling of trips within site-days to free up time for additional data collections and improved species identifications. By fitting the employed design into the proposed two-stage design (averaging data for days with two sites) a two-stage variance estimator can be used to assess the effects of subsampling trips, using empirical data.

Recommendation 3

In addition to assessing an alternative design with sampling of site-days (PSUs) proportional to expected catch or effort (elimination of usage strata for sites), the feasibility and benefit of the optimization of sample allocation to current strata should be assessed. Either of these options could be tested through built-in experiments in future monitoring programs, for example by testing alternative designs in random weeks.

Recommendation 4

For bus-route design, an alternative to be considered is to employ equal waiting times for each site, and to use bootstrap variance estimators with days as primary sampling units. Since few days per month are sampled, it is reasonable to assume sampling with replacement.

Recommendation 5

Consider using data drone surveys and satellite data with automated image analysis in independent effort surveys.

References

Berger, Y.G. 2004. A Simple Variance Estimator for Unequal Probability Sampling without Replacement, Journal of Applied Statistics, 31:3, 305-315, DOI: 10.1080/0266476042000184046

Chen, S. X., & Woolcock, J. L. 1999. A Condition for Designing Bus-Route Type Access Site Surveys to Estimate Recreational Fishing Effort. Biometrics, 55(3), 799-804.

Dainys, J., Gorfine, H., Mateos-González, F., Skov, C., Urbanavičius, R., & Audzijonyte, A. 2022. Angling counts: Harnessing the power of technological advances for recreational fishing surveys. Fisheries Research, 254, 106410.

Dutterer, A. C., Dotson, J. R., Thompson, B. C., Paxton, C. J., & Pouder, W. F. (2020). Estimating recreational fishing effort using autonomous cameras at boat ramps versus creel surveys. North American Journal of Fisheries Management, 40(6), 1367-1378.

Guirado, E., Tabik, S., Rivas, M.L. et al. Whale counting in satellite and aerial images with deep learning. Sci Rep 9, 14259 (2019). https://doi.org/10.1038/s41598-019-50795-9

Hartill, B. W., Taylor, S. M., Keller, K., & Weltersbach, M. S. 2020. Digital camera monitoring of recreational fishing effort: applications and challenges. Fish and Fisheries, 21(1), 204-215.

Hartill, B. W., & Edwards, C. T. T. 2015. Comparison of recreational harvest estimates provided by onsite and offsite surveys: detecting bias and corroborating estimates. Canadian Journal of Fisheries and Aquatic Sciences, 72(9), 1379-1389.

Hoenig, John M., Cynthia M. Jones, Kenneth H. Pollock, Douglas S. Robson, and David L. Wade. 1997. "Calculation of Catch Rate and Total Catch in Roving Surveys of Anglers." Biometrics 53, no. 1: 306–17. https://doi.org/10.2307/2533116.

ICES. 2014. Report of the third Workshop on Practical Implementation of Statistical Sound Catch Sampling Programmes, 19–22 November 2013, ICES HQ, Copenhagen, Denmark. ICES CM2013/ACOM:54. 109pp

R.J. Jessen. 1978. Statistical Survey Techniques. John Wiley & Sons, New York, p. 520

Kinloch, M. A., McGlennon, D., Nicoll, G., & Pike, P. G. 1997. Evaluation of the bus-route creel survey method in a large Australian marine recreational fishery: I. Survey design. Fisheries Research, 33(1-3), 101-121.

Lehtonen, R., and Pahkinen, E. 2004. Practical methods for design and analysis of complex surveys. John Wiley & Sons.

Steffe, A. S., Taylor, S. M., Blight, S. J., Ryan, K. L., Desfosses, C. J., Tate, A. C., Smallwood, C. B., Lai, E. K., Trinnie, F. I. and Wise, B. S. 2017. Framework for Integration of Data from Remotely Operated Cameras into Recreational Fishery Assessments in Western Australia. Fisheries Research Report No. 286, Department of Primary Industries and Regional Development, Western Australia. 36pp.

Vølstad, J. H., Pollock, K. H., & Richkus, W. A. (2006). Comparing and combining effort and catch estimates from aerial–access designs as applied to a large-scale angler survey in the Delaware River. North American Journal of Fisheries Management, 26(3), 727-741.

Vølstad, J. H., Afonso, P. S., Baloi, A. P., de Premegi, N., Meisfjord, J., & Cardinale, M. 2014. Probability-based survey to monitor catch and effort in coastal small-scale fisheries. Fisheries Research, 151, 39-46.

Wang, Z., Peng, L., & Kim, J. K. 2022. Bootstrap inference for the finite population mean under complex sampling designs. Journal of the Royal Statistical Society Series B: Statistical Methodology, 84(4), 1150-1174.

Williams R.L. 2000. A note on robust variance estimation for cluster-correlated data. Biometrics, 56: 645–646.

Wolter, K. M. 1985. Introduction to Variance Estimation. Springer-Verlag. New York. 421pp.

Appendices

Appendix 1: Bibliography of materials provided for review

For Desk Review:
PR Port Sampling – MER – Part 1 Executive Summary
PR Port Sampling – MER – Part 2 Main Body
PR Port Sampling – MER – Part 3 Appendices
PR Port Sampling – MER – Part 4 Site Descriptions

Pre-review Background Documents:
SOW MER Consultants 2017
Design of a Port Sampling Program for the U.S. Caribbean Appendix 1
MER Final Port Sampling Report Appendix 2
MER Port Sampling Site Descriptions Appendix 3

All these documents were provided on Google Drive.

Appendix 2: A copy of the CIE Performance Work Statement

Performance Work Statement

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

Puerto Rico Port Sampling and Catch Validation Project

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards[2].

Scope

The National Marine Fisheries Service is seeking a desk review to evaluate the methods, results, statistical design, implementation, catch expansion algorithm, and recommendations of a commercial port sampling fishing survey in Puerto Rico. The study was carried out by an independent consulting firm. NMFS is requesting a review of the consultant's final report.

The port sampling fishing survey utilized a stratified random design. Stratification was based upon a combination of time, space, and fishing effort (high and low use sites). An expansion algorithm was used to calculate total catch by species.

The goals and objectives specific to the review are to:

1) Evaluate the survey design for the port sampling survey to estimate commercial catch by species.
2) Evaluate survey expansion algorithm for total catch by species.

[2] https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

3) Suggest future research priorities to improve the existing survey design and expansion algorithm.

The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (ToRs) for the review of the surveys are listed in **Annex 2**.

Requirements

NMFS requires two reviewers to conduct an impartial and independent peer review in accordance with this Performance Work Statement (PWS), OMB Guidelines, and the ToRs below. The reviewers shall have expertise and experience in probability survey design and the implementation of fishery independent and/or fishery dependent probability surveys. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Tasks for Reviewers

Each CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

1. Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send by electronic mail or make available at an FTP site to the CIE reviewers all necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE on where to send documents. The CIE reviewer shall read all documents in preparation for the peer review, for example:
   a. SOW MER Consultants 2017
   b. Design of a Port Sampling Program for the U.S. Caribbean Appendix 1
   c. MER Final Port Sampling Report Appendix 2
   d. MER Port Sampling Site Descriptions Appendix 3

2. Webinar: Additionally, approximately two weeks prior to the peer review, the CIE reviewers will participate in a webinar with the NMFS Project Contact and Population Evaluation Tool Subgroup members to address any clarifications that the reviewers may have regarding the ToRs or the review process. The NMFS Project Contact will provide the information for the arrangements for this webinar.

3. Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the PWS and ToRs, and shall not serve in any other role unless specified herein. Modifications to the PWS and ToRs cannot be made during the peer review, and any PWS or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor. The consultant's report includes the following documents, all are to be considered in the review.
   a. PR Port Sampling – MER – Part 1 Executive Summary
   b. PR Port Sampling – MER – Part 2 Main Body
   c. PR Port Sampling – MER – Part 3 Appendices
   d. PR Port Sampling – MER – Part 4 Site Descriptions

4. Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the PWS. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Place of Performance
Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Period of Performance
The period of performance shall be from the time of award through June 2023. Each reviewer's duties shall not exceed 10 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Schedule	Milestones and Deliverables
Within two weeks of award	Contractor selects and confirms the two reviewers.
No later than two weeks prior to the review	Contractor provides the pre-review documents to the reviewers.
April 2023	Each reviewer conducts an independent peer review as a desk review.
Within two weeks after review	Contractor receives draft reports.
Within three weeks of receiving draft reports	Contractor submits final reports to the Government.

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each ToR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
Since this is a desk review, travel is neither required nor authorized for this contract.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact

Kevin McCarthy
Supervisory Research Fish Biologist
NMFS Southeast Fisheries Science Center
75 Virginia Beach Drive, Miami FL 33149
Kevin.J.McCarthy@noaa.gov
305-361-4492

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

3. The reviewer report shall include the following appendices:
    a. Appendix 1: Bibliography of materials provided for review
    b. Appendix 2: A copy of the CIE Performance Work Statement

Annex 2: Terms of Reference for the Peer Review

Puerto Rico Port Sampling and Catch Validation Project

1. Evaluate the statistical design and implementation of the port sampling survey to estimate commercial catch by species and the methods used to develop the design.
2. Evaluate survey expansion algorithm for total catch by species.
3. Evaluate the approach for estimating total catch in light of other approaches that utilize CPUE and estimates of total effort.
4. Evaluate recommendations for future work as provided in the survey report. Provide any additional recommendations for future research/improvements for the survey design and expansion methods.

Appendix 3: Bus route question from Cynthia Jones on May 16, and answer from NOAA

Question:
There is a very proscribed way to do a bus route and I didn't see those kinds of schedules laid out with randomized-time start locations. I need clarification.

Response from MER:
We traded statistical simplicity for deliberately intercepting the most fish for interpretation ease.
Pages 50 – 51 in the report have our original write up.
Page 52 - 53 has diagrams of each bus-route with times listed and the statistical basis for our approach

<u>Excerpt from page 50 – 51 of Main Body:</u>
Roving design. Resembles a bus-route design (Robson and Jones 1989), and involves having the port sampling agent follow a precise route to visit each sampling site during a sampling day. For the AM surveys, the sampling day was four hours long; for the PM surveys it was also four hours and for the Sunday surveys it was eight hours. The procedure for designing the roving survey is as follows:
1) Lay out a route that starts at one site, connects all of the sites, and ends back at the first site.
2) Determine the total travel time required to follow the entire route and subtract from the length of the day to give the amount of time available for sampling (Table 9; Figures 25, 26 and 27).
3) Divide the amount of time available for sampling by the number of sites to determine the time to spend at each site.
4) Pick the days that are to be sampled during the month.
5) For each day to be sampled, randomly pick the starting place along the route; start at that point and end up at that point at the end of the day.

Table 9. Bus-Routes total sampling and driving times on each region by type of sampling
Table 9. Bus-Routes total sampling and driving times on each region by type of sampling

Type	Region	Number of Sites	Time per site (min)	Total sampling time (min)	Total drive-time (min)
AM	North	3	66	198	42
AM	South	4	30	120	120
AM	West	3	50	150	90
PM	West	3	46	138	102
Sunday	North	3	142	426	54
Sunday	East	3	134	402	78
Sunday	South	3	134	402	78
Sunday	West- North	6	59	354	126
Sunday	West- South	3	124	372	108

For example, suppose we want to set up a schedule for AM sampling in a region with three sites to sample and total travel time of 60 minutes. Then there are 360 minutes in four hours, minus 60 minutes needed for travel, equals 300 minutes available for sampling. And 300 minutes divided by three sites gives 100 minutes per site.