

**Independent External Peer Review Report**  
**Marine Recreational Information Program (MRIP)**  
**Fishing Effort Survey (FES) Calibration**

*Calibration Model Accounting for a  
Recreational Survey Design Change*

**Sheraton Hotel, Silver Spring, MD**

**June 27-29, 2017**

**Reviewer Report to NMFS MRIP**

**Prepared By:**

**Fredric M. Serchuk**

**July 2017**

## Executive Summary

- a) This report is an independent peer review of the *Calibration Model Accounting for a Recreational Fishery Survey Design Change* presented at the MRIP Fishing Effort Survey (FES) Calibration Model Review meeting held 27-29 June 2017 at the Sheraton Hotel in Silver Spring, Maryland.
- b) About two weeks prior to the review meeting, the Peer Review Panel—comprising six independent reviewers—was provided with the Terms of Reference (ToRs) for the Peer Review, as well as with four pre-review background documents. One of these documents was a working paper entitled *A Small Area Estimation Approach for Reconciling Mode Differences in Two Surveys of Recreational Fishing Effort* (by F. Jay Breidt, Teng Liu, and Jean D. Opsomer, Colorado State University, June 10, 2017). This paper provided a description of the proposed model and statistical approach developed to calibrate the time series of recreational fishing effort estimates derived from the Coastal Telephone Survey (CHTS) during 1982-2016 with the effort estimates derived from the mail-based Fishing Effort Survey (FES) available in 2015 and 2016. A comparison of the CHTS and FES effort estimates from the contemporaneous 2015 and 2016 surveys (which will also continue in 2017) revealed large differences, with the mail survey estimates very much higher (2 to 11-fold) than the telephone estimates.
- c) Three presentations were given to the Panel on the first day of the review meeting to provide additional background information on (1) the MRIP fishing effort survey; (2) the importance of calibrated catch for stock assessments; and (3) the importance of calibrated catch for fisheries management. Two other presentations were also given: one of these focused on the general issue of calibrating survey estimates over time, while the second provided an in-depth explanation of the development, design structure, analytical methodologies, estimators, and testing/performance of the proposed fishing effort calibration model (*i.e.*, the Breidt *et al.* model).
- d) The second day of the review primarily involved follow-up discussions and dialogue with the calibration modelers to gain a fuller understanding by the Panel of the calibration model, particularly regarding variable selection and model parameterization. Several additional analyses were performed by the modelers and provided to the Panel in response to specific questions and concerns by the reviewers.
- e) The calibration model is a statistically valid approach to obtain calibrated estimates of recreational fishing effort during 1982-2016, even though the casual mechanism(s) for the differences between the CHTS and FES effort estimates remain unknown. The model uses standard and highly respected methodologies (e.g., the Fay-Harriot small area estimation procedure) and can be implemented with off-the-shelf software. Although many other modeling approaches could have considered (and indeed a few of these were evaluated by the developers), the Breidt *et al.* model is certainly an appropriate and scientifically credible statistical approach for calibrating CHTS/FES effort data.

- f) An additional year of contemporaneous data telephone and mail survey effort data will be available at the end of 2017. It is highly recommended that a series of cross-validation analyses be conducted to evaluate the calibration modeling results based on the first, second, and third years of data to ensure that the modelling framework—and the model parameter estimates and predictions errors—are stable. As but one approach, the current model (based on the 2015 and 2016 surveys), should be used to predict the 2017 FES effort given the actual 2017 CHTS effort estimate (and/or vice-versa) – and then compare this to the actual effort obtained from the FES survey. Because the calibration procedure should work equally well whether converting from CHTS to FES or FES to CHTS, this exercise should be illuminating.
- g) It is important to effectively communicate the calibrated effort results and their impacts (as well as to clearly describe the model used in the calibration) to a variety of user and stakeholder groups as the calibrated data will have significant downstream effects on future stock assessments and on various fishery management programs and activities. A variety of pro-active communication approaches should be used to dispel any misconceptions that may currently exist regarding the legitimacy of the calibration and the transition to the FES system.
- h) Finally, it is recommended that an updated report/timetable/chart be prepared illustrating current progress in meeting the tasks and timelines identified in the FES Transition Plan. This undertaking should take note of how the recommendations tendered in the current peer review, as well as those in all previous peer reviews of the MRIP Program (including the 2006 and 2016 NAS Reviews), have been addressed.

## Background

This document reports on an independent peer review of a calibration model proposed for use in revising statistics produced by surveys of marine recreational fishing effort on the Atlantic coast and in the Gulf of Mexico. This calibration model is considered by the Marine Recreational Information Program (MRIP) to be very important to adjust historical time series of recreational effort and catch estimates to account for biases in past sampling and estimation methods that have become apparent with the development of a new, more statistically sound method. The calibration model is intended to account for past biases in private boat and shore fishing effort estimates that have resulted from the continued use of a random-digit-dial telephone survey design (known as the “Coastal Household Telephone Survey” [CHTS]) that has degraded over time and will be replaced with the implementation of a new mail survey design (the “Fishing Effort Survey”, or FES) in 2018. During 2015-2017, a side-by-side benchmarking of the FES against the CHTS has been occurring to facilitate the development and application of a calibration model “to enable adjustment of past estimates that account for biases in historical effort and catch statistics after the second year.”

The purpose of MRIP—FEC Calibration Model Review held during 27-29 June 2017 was to provide an independent peer review of a statistical model for calibrating CHTS and FES effort estimates so that a single time series of effort (from 1981 onward) could be used in the future. The statistical model developed by F. Jay Breidt, Teng Liu, and Jean D. Opsomer (all from Colorado State University) was the subject of the Peer Review. The model was described in a working paper entitled *A Small Area Estimation Approach for Reconciling Mode Differences in Two Surveys of Recreational Fishing Effort* provided to the peer reviewers about two weeks before the meeting.

The Review Panel meeting was chaired by Paul Rago (a member of the Mid-Atlantic Fishery Management Council Scientific and Statistical Committee) and the Panel included six other scientists: Robert Hicks, Cynthia Jones, and Ali Arab (all appointed by the Center for Independent Experts [CIE]), and Patrick Sullivan, Fredric Serchuk, and Jason McNamee (selected, respectively, as representatives from the New England and South Atlantic Fishery Management Council Scientific and Statistical Committees, and from the Atlantic States Marine Fisheries Commission).

Four background documents were provided to members of the Review Panel approximately two weeks prior to the meeting. These included the Breidt *et al.* working paper, the MRIP Transition Plan for the Fishing Effort Survey, a MRIP Fishing Effort Survey Transition Progress Report (dated October 28, 2016), and a report by Rob Andrews, J. Michael Brick, and Nancy A. Mathiowetz entitled *Development and Testing of Recreational Fishing Effort Surveys, Testing a Mail Survey Design* Final Report (dated July 31, 2014). Panel members were also given electronic access to a PowerPoint presentation on the Breidt *et al.* calibration model about a week prior to the review meeting.

The reviewer's Statement of Work is provided in Annex 1, the Terms of Reference (ToRs) for the Peer Review in Annex 2, a Bibliography listing Background and Working Papers for the Peer Review (as well as the Presentations and Hyperlinks provided at the Peer Review) is found at Annex 3, attendees at the Peer Review meeting are listed in Annex 4, and the draft Agenda for the Peer Review meeting is provided in Annex 5.

## **Review Activities**

This reviewer independently read all documents provided in preparation of the review, participated actively in the review meeting (and in the Panel closed sessions at the end of each day and on the last day of the meeting), identified key issues and concerns during the review, contributed to the drafting and editing of the summary report (at the closed session held on the last day of the meeting, by email correspondence several days after the meeting, and during a Panel teleconference held on Friday, 7 July), and authored this review report. As well, this reviewer interacted with the Panel Chair (in person and via email) prior to the review seeking clarification of several of the ToRs and discussing several aspects of the Breidt *et al.* working paper.

## The Peer Meeting and Peer Review Process

The Peer Review meeting encompassed 2½ days from 9 am, 27 June 2017 to 1:30 pm, 29 June 2017. The meeting opened with welcoming comments by Dave Van Vorhees (NMFS MRIP) who provided background on the Agency's planned transition from the telephone survey approach (CHTS) to obtaining estimates of marine recreational fishing effort to a mail survey (FES) for obtaining such estimates. He stated that a 3-year benchmarking process was underway (2015-2017) in which the two surveys are being conducted contemporaneously to provide the requisite data to facilitate the development and application of a calibration model to generate a single historical series of fishing effort (from 1981 onwards) that would be expressed in FES equivalents. The FES mail survey has greater coverage and higher response rates than the CHTS and is considered to represent a major improvement over the CHTS (see the 2016 review of the MRIP program conducted by the National Academy of Sciences). The FES is also much less susceptible to potential sources of bias than the CHTS. Initial examination of the data from the side-by-side 2015 CHTS and FES surveys indicate that the FES overall response rate was about 5X higher than CHTS, and that the overall FES effort estimate was 4.7X larger than the CHTS estimate. Hence, the FES is thought to be a more much efficient and inclusive survey approach than the CHTS, and is believed to produce more accurate information.

*The MRIP Transition Plan for the Fishing Effort Survey (May 2015) calls for the development and evaluation of "one or more calibration models . . . for possible use in correcting past catch statistics. Alternative models should be considered and one should be selected and defended as the most appropriate validated by external peer review."*

The Peer Review Panel was accordingly tasked (see ToR 1 for the Peer Review) *to evaluate the proposed [Breidt et al. calibration] model for converting historical estimates of private boat and shore fishing effort produced by the CHTS design to estimates that best represent what would have been produced had the new FES design been used prior to 2017.*

Following up on the introductory remarks by Van Vorhees, the Review Panel Chair, Paul Rago, also welcomed participants and meeting attendees (both those who were physically present and those who joined the meeting via a webinar) and requested that everyone introduce themselves. The draft meeting agenda was then reviewed by the Panel Chair and adopted by the Panel without change. The Chair encouraged lively and friendly debate among meeting participants and attendees, and then briefly reviewed the TORs and several administrative details relating to the responsibilities of the Panel members.

The remainder of the first day of the meeting was devoted to five PowerPoint presentations with Panel discussions following each of these. Rob Andrews (NOAA Fisheries, MRIP) provided an overview of the MRIP CHTS and FES surveys. He noted a number of significant shortcomings with the CHTS (e.g., susceptibility to non-sampling errors, including non-coverage of cell-phone only households, declining response rates, and inaccurate reporting of fishing activity) and indicated that the CHTS was inefficient for sampling recreational anglers. He briefly described the development and sampling design of the FES and highlighted that the FES had been tested in 2012 in four states before being implemented in 2015. The FES is much less susceptible to non-sampling error than the CHTS and has resulted in greater coverage, higher response rates, and given sufficient time for anglers to consider their responses before mailing back their questionnaires. The use of license lists to screen and stratify the address-based sampling has significantly increased survey efficiency and helped target the sampling to fishing households.

The next two presentations focused on the implications of calibrated catches in subsequent science and management activities. Rick Methot (NOAA Fisheries Senior Scientist for Stock Assessments) presented information on the importance of calibrated catch for fishery stock assessments noting that changes in catch streams can significantly impact stock assessment results with respect to stock abundance and exploitation rates, and also affect biological reference points. Andy Strelcheck (NOAA Fisheries, Deputy Regional Administrator, Southeast Region) then gave a presentation (Importance of Calibrated Catch for Fisheries Management) on how MRIP data are used by fishery managers (a) in setting quotas and annual catch limit, and in quota/catch monitoring; (b) in setting sector allocations; and (c) in evaluating regulatory policies. He also noted that the MRIP data are used in a variety of biological and economic models and analyses. Any changes to the baseline catches presently used in the above activities (i.e., effected through the MRIP calibrations) will affect many user and stakeholder groups (some more than others) and therefore have significant economic and social impacts. This situation will likely be exacerbated because not all stocks with recreational fisheries will be re-assessed immediately after the calibrated MRIP data become available. Hence, some stocks will be assessed, managed, and monitored using pre-calibration data, while others will use calibrated data. As well, the calibrated data may cause shifts in existing allocations among sectors and user groups. In the years ahead before fully transition to FES, successfully addressing these issues will be a major challenge for fishery managers.

The last two presentations on day 1 of the Peer Review meeting were by Jean Opsomer, Colorado State University (Calibrating Survey Estimates Over Time) and by Jay Breidt, Colorado State University (A Calibration Methodology for CHTS to FES Transition). In his presentation, Jean provided background information on the characteristics of “good” surveys (e.g., sample populations according to a prescribed statistical sampling design; have probability-weighted estimators, and allow for design-based inference; have methodologies that minimize sampling error; and are implemented following formal, documented protocols). Surveys that rely on voluntary participation and self-reported information (such as the CHTS and FES) typically result in non-response rates, and are subject to recall and reporting errors. If these attributes change over time, interpretability and estimator consistency of the survey results can become problematic. This seems to be the case for the CHTS as nonresponse rates have continued to decrease, landline-only telephone samples are no longer representative, coastal-county sampling has known coverage problems, and the CHTS does not take advantage of fishing license databases. So changing to FES makes sense but calibration presents challenges in that any calibration model will have uncheckable assumptions and unquantified uncertainty associated with the extrapolation effect. Moreover, no factor or covariate has yet been identified that can explain the large difference between the effort estimates obtained during both the 2015 and 2016 CHTS and FES surveys. Nonetheless, the proposed calibration approach developed by Breidt *et al.* “is firmly grounded in established statistical principles and methodologies [and] allows for quantification of design and model uncertainty.”

The presentation by Jay Breidt (*A Calibration Methodology for CHTS to FES Transition*) described the methodological approach used in developing and testing the proposed calibration model to allow the construction of a new, consistent time series of recreational fishing effort estimates. The calibration issue was approached statistically by identifying sources of uncertainty, applying best analytical practices, making all assumptions explicit, and evaluating the sensitivity of the model with regard to failure to meet model assumptions.

The data used for the calibration work were the side-by-side CHTS and FES effort estimates obtained during 2015 and 2016 (by state and 2-month period) and the historical times series of CHTS effort estimates of shore and private boat fishing (1982-2016) available by state and 2-month period. The calibration model assumed that both the telephone and mail estimates target a common underlying time series of true effort, but that each survey estimate is affected by both sampling and non-sampling errors. This true effort is described by a classical time series model comprising trend, seasonal and irregular components. Although the sampling error properties (and the design variances) of the CHTS and FES are well known based on the statistical designs of these surveys, the non-sampling errors (called the “Irregular Effect”) cannot be isolated from the true effort series. However, because of the side-by-side results from the two surveys, the difference in the non-sampling errors can be estimated and then modeled with covariates to allow extrapolation backward (or forward) in time. The proposed calibration approach combines the two sets of efforts estimates using a well-known mixed model called the Fay-Harriot model. The model was run accounting for temporal dynamics through regression on population size and state-by-2 month period seasonal factors, and also accounting for changing coverage properties in the CHTS due to expanded wireless telephone usage from the 1990s onward (as the CHTS only used landline telephones in sampling the recreational anglers). A desirable attribute of the model is that it can be run using readily available software.

Several novel innovations were incorporated within the model to estimate variances and to ensure analytical consistency. A large number of exploratory analyses (including simulations and sensitivity analyses) were conducted during model development to assess model structure and performance, to select appropriate covariates, and to evaluate alternative hypotheses regarding the distribution of the “Irregular Effect”.

Although the Review Panel posed many questions for the modelers about various aspects of the calibration model and its development and performance (which led to a second presentation by Jay Breidt on the second day of the meeting in which all of these issues were addressed), all Panel members were in agreement that the calibration model is a statistically valid and innovative approach to obtain calibrated estimates of recreational fishing effort during 1982-2016, although the casual mechanism(s) for the differences between the CHTS and FES effort estimates remain unknown.

During the morning of Day 2 of the meeting, Jay Breidt (as noted above) gave his follow-up presentation (*Followup on Comments for “A Calibration Methodology for CHTS to FES Transition”*) to the Panel that responded to the various technical concerns and questions raised by panel members the previous day. As well, analyses and figures requested by Panel members were provided and explained. A lengthy and wide-ranging discussion ensued on both the model configuration and performance, as well as on a variety of issues related to the CHTS and FES surveys themselves (particularly as related to a lack of external validation of the self-reported data obtained in both surveys and what the “wireless effect” is really aliasing). Given that the 2017 side-by-side surveys results will become available at the end of this year, the Panel recommended that a series of cross-validation exercises be conducted to ascertain whether the model and its predictive performance remain stable after the addition of the third (and final) year of contemporaneous CHTS-FES data.

The afternoon of Day 2—and all of the morning and the early part of the afternoon of Day 3, were spent by the Panel in closed session in crafting portions of the Summary Report and in exchanging views regarding individual responses to the ToRs.

## Evaluation of the Terms of Reference

1. **Evaluate the suitability of the proposed model for converting historical estimates of private boat and shore fishing effort produced by the CHTS design to estimates that best represent what would have been produced had the new FES design been used prior to 2017.**

**This TOR—and its subsections (1a, 1b, 1c, 1d, and 1e)—were satisfactorily met.** The proposed calibration model developed by Breidt *et al.* is a statistically valid approach to obtain calibrated estimates of recreational fishing effort during 1982-2016, even though the casual mechanism(s) for the differences between the CHTS and FES effort estimates remain unknown. The model uses standard and highly respected methodologies (e.g., the Fay-Harriot small area estimation procedure) and can be implemented with off-the-shelf software. Although many other modeling approaches could have considered (and indeed a few of these were evaluated by the developers), the Breidt *et al.* model is certainly an appropriate and scientifically credible statistical approach for calibrating the CHTS/FES effort data

- a) **Does the proposed model adequately account for differences observed in the estimates produced by the CHTS and FES designs when conducted side-by-side in 2015-2016?**

The proposed modeling approach uses the effort estimates obtained from the 2015-2016 concurrent CHTS and FES surveys as the foundation for developing and parameterizing the calibration model, and for estimating the difference in the non-sampling errors associated each of the two survey modes so that this difference can be modeled with covariates to allow extrapolation backward in time. The modeling approach preserves the design features of the surveys (among states, 2-month sampling periods, fishing mode [private boat fishing and shore fishing]). The proposed model is an appropriate and scientifically credible statistical approach for calibrating the CHTS/FES effort data series.

- b) **Is the proposed model robust enough to account for potential differences that would have been observed if the two designs had been conducted side-by-side in years prior to 2015 with regards to time trending biases?**

It is difficult to assess whether the proposed model is robust enough to account for potential differences in trend biases that would have been observed between the CHTS and FES had these surveys been concurrently conducted prior to 2015. There are simply no data available to evaluate this hypothesis. Some insights regarding the robustness of the calibration model may be gleaned from cross-validation exercises comparing model results based on using only the 2015-2016 side-by-side survey data vs the full three years (2015-2017) of side-by-side survey estimates. As well, estimating either one of the 2017 effort estimates based on applying the model crafted using the 2015-2016 data and the other 2017 estimate would be informative regarding model stability.

Lastly, the CHTS did not collect ancillary data on the demography (e.g., age, sex, etc.) of the survey respondents that could inform inferences concerning possible time trending biases.

**c) How does the approach used in developing the proposed FES/CHTS calibration model compare in terms of strengths or weaknesses with other potential approaches?**

The approach used in developing the proposed model was statistically well-founded and pursued in a systematic and comprehensive manner taking explicit account of the CHTS/FES methodologies, sources of variability and uncertainty, sensitivity of model assumptions, and the explanatory power of various covariates. The Fay-Harriot approach used in the model well is a highly regarded, well-established statistical methodology that easily allows for incorporation of covariates, and leads to empirical best linear unbiased predictors of either CHTS or FES effort. Performance of the model was tested through various simulations. Overall, the proposed calibration approach is an appropriate and scientifically credible statistical approach for calibrating the CHTS/FES effort data. Although no model is perfect—and while other potential modeling approaches could have been more thoroughly pursued (and a few of these approaches actually were considered during the model development phase)—the Breidt et al. calibration approach is aptly suited for modeling and for calibrating the existing time series of recreational effort estimates.

**d) Does the proposed calibration model help to explain how different factors would have contributed to changes in differences between CHTS and FES results over time?**

Although a number of factors have been identified as contributing to differences between the CHTS and FES estimates in terms of survey error (i.e., the FES survey design is less susceptible to error than the CHTS resulting from nonresponse and non-coverage issues in the CHTS; responses in the FES are likely to be more accurate than in the CHTS because the CHTS required respondents to answer on-the-spot during the phone call rather than having a sufficient time period as in the FES to more thoroughly consider their responses often using the help of memory aids such as datebooks, conversations with family members, etc.; a number of biases have been identified in the CHTS related to (a) underreporting of fishing effort due to a ‘gatekeeper effect’ (which person in the household actually answered the telephone), (b) non-coverage of wireless-only households whose members are more likely to fish than those in land-line households; and (c) insufficient sample size to detect fishing activity in some sampling strata during low-activity time waves), none of these singularly explains the temporal differences in the CHTS and FES results. The major covariate in the calibration model is population size. Although, the “wireless effect” covariate in the calibration model is statistically significant, it only accounts for a minor component of the difference between the CHTS and FES results.

As noted by Jay Breidt, there is no estimated regression coefficient in the model that is the “smoking gun” accounting for the differences the two survey estimates over time, and hence the causal mechanism(s) resulting in the large disparities in the survey estimates remain elusive.

- e) Is it reasonable to conclude that revised 1981-2016 private boat and shore fishing effort estimates based on the application of the proposed FES/CHTS calibration model would be more accurate than the estimates that are currently available? Does evidence provided for this determination include an assessment of model uncertainty?**

Because both the CHTS and FES effort estimates are based on self-reported information that has never been externally validated, the accuracy of any of the estimates cannot be ascertained. There are known shortcomings and biases in the CHTS estimates (see comments in subcomponent [d] above) because of design and coverage issues that are not present with the FES estimates. The FES is clearly the superior approach for obtaining estimates of private boat and shore fishing, and calibrating the 1981-2016 effort estimates to FES equivalents is sensible if only the FES approach will be used in the future.

- 2. Briefly describe the panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.**

The review panel proceedings went smoothly. Operationally, the meeting room had sufficient space for the Panel, presenters, and meeting attendees. The sound and projection systems worked well, as did the webinar link. Four background documents were provided to Panel members two weeks prior to the meeting, and all additional documents and presentation were made available to the Panel during the meeting via a web-site (i.e., Confluence).

The Panel Chair provided each of the reviewers with a proposed meeting Agenda a day prior to the start of the meeting, requesting that any comments and possible changes be provided back to him before the meeting opened. As the proposed Agenda was satisfactory to all the Panel members, no changes to the Agenda were needed.

Discussions during the 2½ day MRIP Calibration Review illuminated various issues related to the results provided in the background documents and the PowerPoint presentations. Many of the concerns involved clarification of the information provided and/or requests for additional data and analyses. In all cases, these requests were satisfactorily fulfilled allowing the Panel to gain fuller insight on:

1. The sampling designs, strengths, and shortcomings of the telephone (CHTS) and mail (FES) survey methods, including their relative performance and sources of error.

2. The development, design, statistical properties, testing, and application of the proposed MRIP FES calibration model. This included consideration of alternative modeling approaches, cross-validation of the modeling framework to examine the stability of model parameter estimate (as well as prediction errors), the sufficiency and explanatory power of the model's covariates, and the possible underlying mechanism(s) affecting the distribution of the "Irregular" random effect, which is not explicitly accounted for within the proposed small-area estimation approach.
3. The potential impacts of the calibrated recreational fishing effort estimates during 1981-2016 on future stock assessments, and on subsequent fishery management policies and practices.
4. The need to effectively communicate the results of the calibration work (as well as the basis and need for continuing only the mail-based survey method in the future) to various constituency groups (i.e., the recreational and commercial fishing communities; scientists; fishery managers; the lay public) so that these groups fully understand and accept the calibration results and their subsequent use in deriving recreational catch estimates for application in stock assessments and in the fishery management process.

The Review Panel acknowledged that proposed MRIP FES calibration model developed by Breidt *et al.* was a well-suited and statistically-appropriate approach to obtain calibrated estimates of recreational fishing effort (by state and 2-month calendar quarter for shore-based and private boat anglers) during 1982-2016.

Although the Peer Review process worked very well and the Panel concluded that all of the TORs for the Review were met, I believe that there are few areas in which the process could have worked even better. These include:

1. It would have been helpful for the Panel to have been provided (several weeks before the review) additional background documents (available from the MRIP Team and/or the MRIP Website) to enhance a collaborative understanding by Panel members of (a) various aspects of the MRIP program and (2) of recent analyses using MRIP data. For example, the *MRIP Data User Handbook*, and the October 2016 report, '*Possible Effects of Calibration Scenarios on Stock Assessments Planned for the MRIP Fishing Effort Survey Transition*' would have especially useful for Panel members to have had and read before the actual peer review occurred.

2. Prior to the presentation and discussion of the Breidt *et al.* report at the Peer Review, this report was difficult to understand for anyone other than a highly trained statistician. Although a more complete understanding of this report was fostered by distribution of a PowerPoint presentation a week or so before the Review Meeting (and subsequently enhanced at the meeting by direct dialogue and interaction with the authors of the paper who clarified and responded to many issues raised by the Panel), it is recommended that in any future reviews in which a highly technical paper is seminal to the crux of such reviews that efforts be made by the paper authors to present the essence of their work in a manner that facilitates full appreciation and understanding of the import of such work by educated non-specialists. This becomes especially critical when the methods/approach provided in a paper will have significant downstream effects. This matter should be recognized in the future APAIS peer review.
3. In its comments on the various subcomponents of TOR 1 (1a, 1b, 1c, 1d, 1e), the Review Panel highlighted a number of issues related to additional work and analyses that might be undertaken to provide additional insight into the performance and robustness of the proposed CHTS/FES calibration model and the efficacy of the effort collection survey methodologies. It is recommended that the MRIP Team (in collaboration where necessary with Breidt *et al.*) develop a protocol to facilitate the timely accomplishment of the highlighted additional work.
4. Finally, it is recommended that an updated report/timetable/chart be prepared illustrating current progress in meeting the tasks and timelines identified in the FES Transition Plan. This undertaking should also take note of how the recommendations tendered in all previous peer reviews of the MRIP Program (including the 2006 and 2016 NAS Reviews) have been addressed.

## **Annex 1. Statement of Work**

### **Statement of Work**

#### **National Oceanic and Atmospheric Administration (NOAA)**

#### **National Marine Fisheries Service (NMFS)**

#### **External Independent Peer Review**

#### *Calibration Model Accounting for a Recreational Fishery Survey Design Change*

#### **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.

[http://www.cio.noaa.gov/services\\_programs/pdfs/OMB\\_Peer\\_Review\\_Bulletin\\_m05-03.pdf](http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).

#### **Scope**

The Office of Science and Technology requests an independent peer review of a calibration model proposed for use in revising statistics produced by surveys of marine recreational fishing effort on the Atlantic coast and in the Gulf of Mexico. This calibration model is considered by the Marine Recreational Information Program (MRIP) to be very important to adjust historical time series of recreational effort and catch estimates in order to account for biases in past sampling and estimation methods that have become apparent with the development of a new, more statistically sound method. The calibration model is intended to account for past biases in private boat and shore fishing effort estimates that have resulted from the continued use of

a legacy random-digit-dial telephone survey design that has degraded over time and will be replaced with the implementation of a new mail survey design (the “Fishing Effort Survey”, or FES) in 2018.

### **Calibration Model for the Fishing Effort Survey**

In 2015, MRIP formed a Transition Team to collaboratively plan a transition from a legacy telephone survey design to a new mail survey design for estimating private boat and shore fishing effort by marine recreational anglers. Since 2008, MRIP had conducted six pilot studies to determine the most accurate and efficient survey method for this purpose on the Atlantic and Gulf coasts. The most recent study, conducted in four states in 2012-2013, compared a new mail survey design with the Coastal Household Telephone Survey (CHTS) design that has been used since 1979. MRIP subjected the final report from the pilot project to external peer review in 2014 and certified the new survey design, called the Fishing Effort Survey (FES), in February 2015 as a suitable replacement for the CHTS. The FES is much less susceptible to potential sources of bias than the CHTS because it can reach more anglers, achieve higher response rates, and is less prone to possible recall errors. The pilot project results indicated that FES estimates were substantially higher than CHTS estimates for both private boat fishing and shore fishing.

MRIP recognized the FES should not be implemented immediately as a replacement for the CHTS, and a well thought out transition plan was needed to ensure that the phase-in of the FES is appropriately integrated into ongoing stock assessments and fisheries management actions in a way that minimizes disruptions to these processes, which are based on input from multiple data sources over lengthy time series. The Transition Plan developed by the Transition Team called for side-by-side benchmarking of the FES against the CHTS for three years (2015-2017) with the development and application of a calibration model to enable adjustment of past estimates that account for biases in historical effort and catch statistics after the second year. With this timeline, revised estimates can be incorporated into stock assessments during 2018 using a peer reviewed calibration model, and new Annual Catch Limits (ACLs) can then be set in 2019 for at least some stocks.

### **Requirements**

NMFS requires five reviewers to conduct an impartial and independent peer review in accordance with the SoW, OMB Guidelines, and the Terms of Reference (ToRs) below. The reviewers shall have working knowledge and recent experience in the design of sampling surveys, the evaluation of non-sampling errors (i.e., undercoverage, nonresponse, and response errors) associated with changes to survey designs over time, and the evaluation of differences between surveys using different modes of contact (e.g., mail *versus* telephone). In addition, they should have experience with complex, multi-stage sampling designs, time series analyses, regression estimators, and small domain estimation methods. Some recent knowledge and experience in current surveys of marine recreational fishing is desirable but not required.

NMFS will designate a Chair who has experience with U.S. fisheries stock assessments and their application to fisheries management. The Chair would ensure that reviewers understand the importance of maintaining a comparable time series of marine recreational fisheries catch statistics for use in stock assessments and their application to fisheries management. The Chair will not be selected by the contractor and will be responsible for facilitating the meeting, developing and finalizing a summary report and working with the reviewers to make sure that the ToRs are addressed in their independent reviews.

## **Tasks for Reviewers**

### **Pre-review Background Documents**

The following background materials and reports prior to the review meeting include:

Transition Plan for the FES:

<https://www.st.nmfs.noaa.gov/Assets/recreational/pdf/MRIP%20FES%20Transition%20Plan%20FINAL.pdf>

Report recommending the FES to replace the CHTS: *Finalize Design of Fishing Effort Surveys* ([https://www.st.nmfs.noaa.gov/pims/main/public?method=DOWNLOAD\\_FR\\_PDF&record\\_id=1179](https://www.st.nmfs.noaa.gov/pims/main/public?method=DOWNLOAD_FR_PDF&record_id=1179))

2015 Benchmarking Progress Report:

[https://www.st-test.nmfs.noaa.gov/Assets/recreational/pdf/2015\\_FES\\_Progress\\_Report-20161115.pdf](https://www.st-test.nmfs.noaa.gov/Assets/recreational/pdf/2015_FES_Progress_Report-20161115.pdf)

Report on FES/CHTS Calibration Model:

This report will be provided by ECS (via electronic mail or make available at an FTP site) to the reviewers.

### **Panel Review Meeting**

Each reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. Each reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The meeting will consist of presentations by NOAA and other scientists to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers.

### **Contract Deliverables - Independent CIE Peer Review Reports**

The reviewers shall complete an independent peer review report in accordance with the requirements specified in this SoW and OMB guidelines. Each reviewer shall complete the independent peer review according to the required format and content as described in **Annex**

1. Each reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

### **Other Tasks – Contribution to Summary Report**

The reviewers may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. The reviewers are not required to reach a consensus, and should provide a brief summary of each reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

### **Place of Performance**

The place of performance shall be at the reviewers' facilities, and at the NMFS Headquarters in Silver Spring, Maryland.

### **Period of Performance**

The period of performance shall be from the time of award through July 31, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

### **Travel**

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (<http://www.gsa.gov/portal/content/104790>).

### **Restricted or Limited Use of Data**

The contractors may be required to sign and adhere to a non-disclosure agreement.

### **NMFS Project Contact:**

Dave Van Voorhees  
National Marine Fisheries Service  
1315 East West Highway  
Silver Spring, MD 20910  
[dave.van.voorhees@noaa.gov](mailto:dave.van.voorhees@noaa.gov)

## Annex 2. Terms of Reference

### Terms of Reference for the Peer Review

#### *Calibration Model Accounting for a Recreational Fishery Survey Design Change*

*The Review Panel shall assess whether or not the MRIP Working Group has reasonably and satisfactorily completed the following actions.*

1. Evaluate the suitability of the proposed model for converting historical estimates of private boat and shore fishing effort produced by the CHTS design to estimates that best represent what would have been produced had the new FES design been used prior to 2017.
  - a) Does the proposed model adequately account for differences observed in the estimates produced by the CHTS and FES designs when conducted side-by-side in 2015-2016?
  - b) Is the proposed model robust enough to account for potential differences that would have been observed if the two designs had been conducted side-by-side in years prior to 2015 with regards to time trending biases?
  - c) How does the approach used in developing the proposed FES/CHTS calibration model compare in terms of strengths or weaknesses with other potential approaches?
  - d) Does the proposed calibration model help to explain how different factors would have contributed to changes in differences between CHTS and FES results over time?
  - e) Is it reasonable to conclude that revised 1981-2016 private boat and shore fishing effort estimates based on the application of the proposed FES/CHTS calibration model would be more accurate than the estimates that are currently available? Does evidence provided for this determination include an assessment of model uncertainty?
2. Briefly describe the panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

## Annex 3. Bibliography of Documents and Presentations

### *Background Papers*

Many papers and documents on the existing and proposed survey methodology may be found at the following website:

<http://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/effort-survey-improvements>

Background on the MRIP Calibration Model Peer Review may be found at:

<https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/index.html>

The National Academies of Sciences, Engineering, and Medicine. 2016. Review of the Marine Recreational Information Program (MRIP) Washington, DC: The National Academies Press. doi: 10.17226/24640

[https://www.st.nmfs.noaa.gov/confluence/display/FESCALIB?preview=/73074985/73728799/NAS\\_MRIP\\_review.pdf](https://www.st.nmfs.noaa.gov/confluence/display/FESCALIB?preview=/73074985/73728799/NAS_MRIP_review.pdf)

National Research Council. 2006. Review of Recreational Fisheries Survey Methods. Committee on the Review of Recreational Fisheries Survey Methods, ISBN: 0-309-66075-0, 202 pages. <http://www.nap.edu/catalog/11616.html>

### *Working Papers*

Development and Testing of Recreational Fishing Effort Surveys Testing a Mail Survey Design: Final Report. Project Team Members: Rob Andrews, NOAA Fisheries, J. Michael Brick, Westat, Nancy A. Mathiowetz, University of Wisconsin-Milwaukee. July 31, 2014.

[https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/documents/Report\\_recommending\\_FES\\_to\\_replace\\_CHTS--Finalize\\_Design\\_of\\_Fishing\\_Effort\\_Surveys.pdf](https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/documents/Report_recommending_FES_to_replace_CHTS--Finalize_Design_of_Fishing_Effort_Surveys.pdf)

Marine Recreational Information Program Fishing Effort Survey Transition Progress Report. October 28, 2016.

[https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FESWorkshop/documents/2015\\_benchmarking\\_progress\\_report.pdf](https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FESWorkshop/documents/2015_benchmarking_progress_report.pdf)

Marine Recreational Information Program. Transition Plan for the Fishing Effort Survey. Prepared by the Atlantic and Gulf Subgroup of the Marine Recreational Information Program Transition Team, May 5, 2015

[https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/documents/MRIP\\_FES\\_Transition-Plan\\_FINAL.pdf](https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/documents/MRIP_FES_Transition-Plan_FINAL.pdf)

A Small Area Estimation Approach for Reconciling Mode Differences in Two Surveys of Recreational Fishing Effort draft: F. Jay Breidt, Teng Liu, Jean D. Opsomer. Colorado State University June 10, 2017

[https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/documents/DRAFT-Report\\_of\\_Calibration\\_Model.pdf](https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/documents/DRAFT-Report_of_Calibration_Model.pdf)

### ***Presentations***

[Calibration\\_Scenarios-20161115.pdf](#)

[Eblup comparisons.docx](#)

[MRIP FES website link](#)

[MRFSS Fish Hunt Comps.xlsx](#)

[FESCALIBRATIONNOTESDay2.docx](#)

[FES Errors.pptx](#)

[EBLUPS.csv](#)

[Model\\_Fits.txt](#)

[EBLUPS\\_Variable\\_Names.csv](#)

[Mode\\_3\\_logeffort\\_poly\\_fixed.pdf](#)

[FESCALIBRATIONNOTESDay1.docx](#)

[Mode\\_7\\_logeffort\\_poly\\_fixed.pdf](#)

### ***Webinar Links***

All open sections of the meeting were recorded and available for viewing at the following links.

[0 - Intro - Paul Rago](#)

[1 - MRIP Fishing Effort Survey - Rob Andrews](#)

[2- Catch and Assessments - Rick Methot](#)

[3 - Management Implications - Andy Strelcheck](#)

[4 - Calibrating Survey Estimates over Time - Jean Opsomer](#)

[5 - Calibration from CHTS to FES - Jay Breidt](#)

[6 - Initial Calibration Review Discussion - Tuesday Afternoon](#)

[7 - Day Two, AM Discussion](#)

[8 - Day Two, PM Discussion](#)

[9 - Day Two, Initial Findings Summary](#)

## Annex 4. Attendees at the Peer Review Meeting

MRIP Calibration Model Peer Review Workshop  
 Sheraton Silver Spring Hotel  
 Silver Spring, MD  
 June 27-29, 2017  
 ATTENDANCE LIST

#	NAME	AFFILIATION
1	Paul Rago	MAFMC SSC
2	Dave Van Voorhees	NOAA Fisheries
3	John Foster	NOAA Fisheries
4	Ali Arab	Georgetown University
5	Rob Hicks	College of William and Mary
6	Cynthia M. Jones	Old Dominion University
7	Richard Cody	NOAA support ECS
8	Teng Liu	Colorado State University
9	Thomas Sminkey	NOAA Fisheries/ST1
10	Steve Turner	NOAA Fisheries SEFSC
11	Andy Strelcheck	NOAA Fisheries - SERO
12	Richard Methot	NOAA Fisheries - HQ
13	Karen Pianka	NOAA Fisheries – ST1
14	Lauren Dolinger Few	NMFS ST1
15	Chris Wright	NMFS - SF
16	Sabrina Lovell	NMFS ST
17	Patrick Lynch	NMFS ST
18	Melissa Karp	NMFS ST
19	Toni Kerns	ASMFC
20	Steve Ander	Gallup
21	Tommy Tran	Gallup
22	Melissa Niles	Fifth Estate/MRIP CET
23	Yong-Woo Lee	NOAA - Fisheries
24	Jay Breidt	Colorado State University
25	Jean Opsomer	Colorado State University
26	Rob Andrews	NOAA Fisheries
27	Ryan Kitts-Jensen	NOAA Fisheries
28	Fred Serchuk	SAFMC SSC
29	Jason McNamee	ASMFC
30	Patrick Sullivan	Cornell/NEFMC
31	Jason Didden	MAFMC
32	Daemian Schreiber	NMFS HQ
33	Laura Diederick	NOAA Fisheries

## Annex 5. MRIP Calibration Model Peer Review Draft Meeting Agenda

<i>Day</i>	<i>Date</i>	<i>Time</i>	<i>Topic</i>	<i>Rapporteur</i>	<i>Presenter</i>
Tuesday	27-Jun	9:00 AM	Welcome and Opening Remarks	TBD	Van Voorhees
		9:20 AM	Introductions		
		9:30 AM	Overview of Meeting	TBD	Rago
		9:45 AM	MRIP Fishing Effort Survey	TBD	Andrews
		10:15 AM	Importance of Calibrated Catch for Stock Assessments	TBD	Methot
		10:45 AM	Break		
		11:00 AM	Importance of Calibrated Catch for Fisheries Management	TBD	Strelcheck
		11:30 AM	Calibrating Survey Estimates over Time	TBD	Opsomer
		12:00 PM	Lunch		
		1:30 PM	A Calibration Methodology for CHTS to FES Transition	TBD	Breidt
		3:30 PM	Break		
		3:45 PM	Public Comment	TBD	
		4:15 PM	Summary of Day 1	TBD	Rago
		4:45 PM	Review Panel Coordination and Writing (closed)		
6:00 PM	Adjourn				
Wednesday	28-Jun	9:00 AM	Overview of Day 1 and Preview of Day 2	TBD	Rago
		9:10 AM	Follow-up Questions for Presenters	TBD	Various
		10:30 AM	Break		
		10:45 AM	Follow-up Questions for Presenters (cont.)	TBD	Various
		12:00 PM	Lunch		
		1:00 PM	Review Panel Coordination and Writing (closed)		
		2:30 PM	Initial Summary Findings of Review Panel (open)	TBD	Panel
		3:30 PM	Public Comment	TBD	
		4:00 PM	Review Panel Coordination and Writing (closed)		
6:00 PM	Adjourn				
Thursday	29-Jun	9:00 AM	Review Panel Coordination and Writing (closed)		
		12:30 PM	Adjourn		
Closed sessions allow the panel to discuss and clarify technical issues, and begin initial writing of reports.					
Attendance of public, staff and presenters, if at all, is by invitation only and for purposes of clarification.					