UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL MARINE FISHERIES SERVICE
Silver Spring, MD 20910

DEC 20 2012

MEMORANDUM FOR: Samuel D. Rauch, III
Acting Assistant Administrator for Fisheries

FROM: Ned Cyr, Ph.D.
Director, Office of Science and Technology

SUBJECT: Approval of Implementation of a New Method for Conducting Intercept Surveys of Anglers – DECISION MEMORANDUM

I request that you approve the implementation of a new design for the Access Point Angler Intercept Survey as developed by the Marine Recreational Information Program (MRIP.)

BACKGROUND

The Marine Recreational Fisheries Statistics Survey (MRFSS), initiated in 1979 as a requirement of the Magnuson-Stevens Fishery Conservation and Management Act of 1976 (MSA), has been the primary source for national recreational fishery statistics in the United States. The MRFSS was a complementary survey design that includes the Coastal Household Telephone Survey (CHTS) and the Access-Point Angler Intercept Survey (APAIS). The CHTS collects data on angler fishing effort from a random sample of coastal residential households in each state. The APAIS is a shore-side survey that that collects data on angler catch through probabilistic sampling of fishing access points and days. Data from the two independent surveys are combined to estimate total fishing effort, participation, and catch by species. The MRFSS has been conducted in all regions except the Pacific Coast, Alaska, Texas, the Western Pacific Territories, and the U.S. Virgin Islands.

In response to the growing demand for an improved recreational fishing data collection program, NMFS commissioned the National Research Council (NRC) of the National Academies of Science to conduct a high-level, scientific review of the existing survey methods used by NMFS and its partners to monitor catch and effort in marine recreational fisheries throughout the U.S.

Specifically, the NRC was asked to:

- assess existing surveys and their suitability in monitoring effort and catch in the shore-based, private boat, and for-hire boat recreational fisheries;
- evaluate how well these methods were providing the quality of information required to support accurate stock assessments and responsible fisheries management decisions; and
- recommend improvements to ensure more accurate and precise estimates of recreational effort and catch.

The NRC's Ocean Studies Board formed a 10-member committee of experts in sampling design and statistics to conduct the requested review independent of NMFS. A final report of their findings (*Review of Recreational Fisheries Survey Methods*) was published in April 2006. The committee identified number of potential problems with the sampling and estimation designs, and questioned the adequacy of existing surveys in providing the statistics needed to support stock assessments and the kinds of fishery management decisions required by current law and practice. The report includes recommendations to redesign current surveys to improve: their effectiveness, the appropriateness of their sampling procedures, their applicability to various kinds of management decisions, and their usefulness for social and economic analyses.

Section 401(g) of the MSA, added via the 2006 reauthorization, established statutory requirements for improving recreational fisheries data collection:

- § 401(g)(3)(A): "Within 24 months after the date of enactment of the Magnuson-Stevens Fishery Conservation and Management Reauthorization Act of 2006, the Secretary, in consultation with representatives of the recreational fishing industry and experts in statistics, technology, and other appropriate fields, shall establish a program to improve the quality and accuracy of information generated by the Marine Recreational Fishery Statistics Survey, with a goal of achieving acceptable accuracy and utility for each individual fishery."
- § 401(g)(3)(B): "The program shall take into consideration and, to the extent feasible, implement the recommendations of the National Research Council in its report *Review of Recreational fishing Survey Methods* (2006), including...redesigning the survey to improve the effectiveness of sampling and estimation procedures, its applicability to various kinds of management decisions, and its usefulness for social and economic analyses..."

NMFS established the Marine Recreational Information Program (MRIP) in 2007 to address the findings and recommendations of the NRC report and to carry out the requirements of § 401(g) of the MSA. MRIP is a collaborative effort among NOAA Fisheries, regional fisheries managers and stock assessment scientists, and the recreational fishing community to develop and implement an improved recreational fisheries statistics program. The new program consists of a system of regional surveys that provides the best possible scientific information for use in the assessment and management of the Nation's marine fisheries. Decisions to implement new data collection methods are informed by a technically-sound scientific process that includes testing of new or enhanced survey methods, peer reviews of survey methods and project results, and reviews by stakeholder groups.

An Executive Steering committee (ESC) is overseeing MRIP. The ESC has established four MRIP leadership teams that are developing an improved data collection program for recreational fisheries, as well as promoting communication between and among NMFS partner organizations, and constituents. Leadership teams include representatives from a broad range of organizations, expertise, and interests, and have the flexibility to establish work groups and/or project teams to address topical, national, and regional issues, as needed. The Operations Team (OT) is charged with developing improvements to data collection programs under the auspices of MRIP. Each year, the OT members identify and prioritize research needs for the coming year and commission work groups or project teams to develop projects that address the top priorities. Work groups and project teams include experts in survey design and management, natural resource management, and stock assessments, and obtain input from knowledgeable recreational fishing stakeholders.

In its *Review of Recreational Fisheries Survey Methods* (2006), the NRC expressed several major concerns regarding the traditional design of the MRFSS' APAIS:

- "..., the estimation procedure for information gathered onsite does not use the nominal or actual selection probabilities of sample design and therefore has the potential to produce biased estimates for both the parameters of interest and their variances."
- "The statistical properties of various sampling, data-collection, and data-analysis methods should be determined. Assumptions should be examined and verified so that biases can be properly evaluated."
- "The statistical properties associated with data collected through different survey techniques differ and are often unknown. The current estimators of error associated with various survey products are likely to be biased and too low. It is necessary, at a minimum, to determine how those differences affect survey results that use differing methods."
- "The onsite sampling frame for the MRFSS should be redesigned. The estimation procedure critically depends on the assumption that catch rate does not vary according to the nature of the access point. In particular, small or private access points that most likely are missed might have different catch rates than larger access points, which would lead to bias in the resulting estimators. In addition, the sampling process requires greater quality control (less latitude on the part of the samplers) than it has at present."

In response to these concerns, the members of the OT established a working group to evaluate survey designs and test ideas for improvements. The first project of this Design and Analysis Work Group was to develop technical documentation describing sampling and estimation procedures for ongoing recreational fishing surveys. The resulting report (*Surveys and Statistical Methods for Estimation of Catch and Effort in U.S. Marine Recreational Fisheries*) serves as a technical source document for assessment of survey and estimation designs.

The Work Group's next research project was to develop an estimation approach for the APAIS that more closely matches the sampling design and subsequently accounts for selection probabilities in the estimation design. The project's final report, *A Report of the MRIP Sampling and Estimation Project: Improved Estimation Methods for the Access Point Angler Intercept survey Component of the Marine Recreational Fishery Statistics Survey (Breidt et al., 2011)* (http://www.countmyfish.noaa.gov/projects/downloads/Final%20Report%20of%20New%20Estimation_Method_for_MRFSS_Data-01242012.pdf) and its recommendations were approved for implementation by the Assistant Administrator for Fisheries on March 14, 2011. In addition to its primary recommendations for adoption of a new, design-unbiased method for estimation of catch from intercept data, the report also recommended complementary improvements to the intercept survey design.

A Project Team was established to conduct a pilot study in North Carolina in 2010 that tested the feasibility of implementing such a new sampling design and assessed its effects on various measures of survey performance through side-by-side comparisons with the ongoing MRFSS APAIS sampling. Project Team members were:

F. Jay Breidt, Colorado State University
James R. Chromy, RTI International

Kelly E. Fitzpatrick, NOAA Fisheries Southeast Fisheries Science Center
Han-Lin Lai, NOAA Fisheries Office of Science and Technology
Terri Menzel, Florida Fish and Wildlife Conservation Commission
Douglas G. Mumford, North Carolina Division of Marine Fisheries
Breda Muñoz, RTI International
Jean D. Opsomer, Colorado State University
Ronald J.Salz, NOAA Fisheries Office of Science and Technology
Kevin M. Sullivan, New Hampshire Department of Fish and Game
David A. Van Voorhees, NOAA Fisheries Office of Science and Technology
Chris Wilson, North Carolina Division of Marine Fisheries
Patricia A. Zielinski NOAA Fisheries Office of Science and Technology

Key features of the survey design in the pilot project that differ from the MRFSS survey design included:

- ***Time of Day Stratification:*** In the new design, sampling is stratified among four six-hour time intervals to ensure some coverage of fishing trips ending at all different times of day.
- ***Geographic Stratification:*** Sampling was stratified geographically in the pilot. Samplers were hired for one of three state subregions within North Carolina and only completed assignments within that particular geographic stratum.
- ***Clustering of Sites for Sampling:*** Low activity sites are clustered to form two- or three-site clusters in the new frame used for sampling. Sites expected to have a high level of activity are not clustered with other sites. The clustering of lower pressure sites into multi-site units increases their inclusion probabilities relative to the higher-pressure sites. Higher-activity sites still have higher inclusion probabilities than lower activity sites in the new sampling design, but there is generally less variability among sites in their probabilities and a greater chance that the sample is spread more evenly among sites that have similar fishing pressure. Samplers are required to visit all sites within the assigned cluster following a predetermined visitation order and times.
- ***Sampling Frame and Probability Sampling:*** The selection of all specific locations in space and time for interviewing assignments (i.e., the primary sampling units, or PSUs) is formalized based on a probability-proportional-to-size (PPS) approach. Thus, the new design uses a purely design-based approach to determining all site selection probabilities.
- ***Issuing and Completing Assignments:*** Under the new design, emphasis is placed on completing all interviewing assignments selected by probabilistic sampling. All assignments drawn have to be either completed as assigned or canceled, because rescheduling is not allowed.
- ***Interviewing limits:*** The new design removes all limits on the number of interviews obtained by samplers during an assignment. Samplers are directed to continue interviewing for the full specified duration of each site assignment.
- ***Elimination of Opportunistic Sampling:*** Sampling of fishing trips in fishing mode strata other than the one for which an assignment was selected is no longer allowed under the new design.
- ***Eligibility for Interviews:*** Under the new design, all intercepted anglers who have completed fishing for the day in the assigned fishing mode are considered eligible for an interview or "proxy" interview in the case of very young anglers.
- ***Complete vs. Incomplete Beach/Bank Interviews:*** For sampling in the beach/bank fishing mode, the new design specifies that only completed angler fishing trips are eligible for an interview.

- ***Angler Trip Counts:*** The new design strongly emphasizes the need for obtaining accurate counts of all eligible angler fishing trips ending at an assigned site during the assigned time interval.
- ***Questionnaires and Data Forms:*** With the exception of one question added to identify angler trips intercepted at tournament sites, the intercept survey questionnaire used for the new sampling design matched that used under the MRFSS design. A number of changes were made to the Assignment Summary Form (ASF) and Site Description Form (SDF) to accommodate the new design's emphasis on obtaining more accurate counts and estimates of expected fishing pressures.

A final report ("Project Report") of the results of the pilot project, *A Pilot Study of a New Sampling Design for the Access Point Angler Intercept Survey (Breidt et al., 2012),* was prepared by the Project Team. The Project Report concluded that the piloted design is feasible and desirable for implementation and recommended implementation on the Atlantic and Gulf coasts. The following were recommended for immediate action:

- **In general, the Project Team recommends use of the new access point survey sampling design tested in this pilot study for conducting future access point surveys on the Atlantic coast and in the Gulf of Mexico.** The pilot study demonstrated that the new design is feasible to implement and has many advantages over the MRFSS design as described in this report.
- **The allocation of sampling among sampling strata should be changed as needed to maximize sampling efficiency and statistical precision.** Sampling could be allocated very differently among geographic strata, fishing mode strata, and time block strata than how it was allocated in this pilot study. Without introducing any bias, other sampling allocations will likely provide higher proportions of sampling assignments that obtain at least one interview and may also provide higher average numbers of interviews per positive assignment than were observed in the pilot study. The goal should be to find the "optimal" allocation that will provide the highest level of statistical precision for the dollar spent.
- **The formal PPS sampling of sites and site clusters should be controlled to ensure all drawn assignments can be completed by existing staff.** Staffing levels for the access point surveys should always be set to match the sampling levels required to deliver desired levels of statistical precision on resulting estimates of mean catch per trip. Once those staffing levels are established, a controlled selection program that incorporates staffing constraints can be used to ensure the draw of a probability sample of assignments that can be covered by the available staff.
- **Provide clearer instructions to samplers about how to handle the catch of charter boat captains and crew.** Samplers should include any catch by the captain and crew that were mixed in with the observed catch recorded for a group of charter boat anglers, but they should not count the captain and crew as contributors to the mixed group catch.
- **Collect total catch data for any intercepted angler who just completed a multi-day fishing trip.** In addition, ask for the number of waking days that the angler fished during the trip. This will allow accurate calculation of the angler's mean catch per day for use in the mean catch estimates for the total population of angler trips.
- **To increase on-site productivity and reduce driving time, instruct samplers to stay up to 3 hours (rather than only two hours) at the first site when a two-site cluster is assigned.**

The Project Report also includes a number of recommendations for future evaluation and consideration of additional modifications to the intercept survey design and operations.

The Project Report was submitted for peer review by two independent scientists. The peer reviews are supportive of the recommendations and included suggestions for improving the Project Report. The final Project Report, Peer Reviews and the Project Team's response to the Peer Reviews were submitted to the MRIP Operations Team on November 23, 2012.

At its meeting on November 27, 2012, the OT determined that it would recommend that the Project Report be accepted and the recommended intercept survey design be certified as an accepted method for implementation. On November 30, 2012, the OT communicated its recommendation to the ESC. On December 17, 2012, the ESC determined that it will support certification of the new intercept survey design and its use to conduct the APAIS beginning in calendar year 2013.

I conclude that implementation of the revised intercept survey design recommended for implementation in the Project Report will result in substantially improving the accuracy of recreational catch estimates calculated using data collected by the APAIS, and recommend that this method be implemented as recommended in the Project Report for the Atlantic and Gulf coasts beginning in January, 2013. I also conclude that this survey design is appropriate for certification as an accepted method and design for collecting angler catch data from intercept surveys, and can be appropriately used for that purpose in other regions when selected for that purpose by appropriate regional survey managers. I endorse pursuing the Project Report's **Recommendations for Future Consideration and Evaluation**, and implementing such further improvements as they are successfully demonstrated. I further conclude that implementing this new survey design implements recommendations of the NRC in its 2006 report, *Review of Recreational Fisheries Survey Methods*, and is consistent with the provisions of §401 (g)(3)(B) of the MSA.

RECOMMENDATION

I recommend that you approve the implementation of the new APAIS design described in the attached Project Report, and certification of the design as an accepted methodology for collection of angler data via intercept surveys.

I concur ______________________________ 1/9/13
Date

I do not concur ______________________________ __________
Date

Attachments:

*A Pilot Study of a New Sampling Design for the Access Point Angler Intercept Survey*
Independent Peer Reviews (2)
Responses to Reviewers

E-mail dated December 2, 2012, from Preston Pate, Chair of the MRIP Operations Team, to John Boreman, Ph.D., Chair of the MRIP Executive Steering Committee

E-mail, dated December __, 2012, from John Boreman, Ph.D., Chair of the MRIP Executive Steering Committee ....

IQA Pre-Dissemination Review form for the Project Report

Copy of the Peer Review Plan for the Project Report

**[THIS INFORMATION IS DISTRIBUTED SOLELY FOR THE PURPOSE OF PRE-DISSEMINATION REVIEW UNDER APPLICABLE INFORMATION QUALITY GUIDELINES. IT HAS NOT BEEN FORMALLY DISSEMINATED BY NOAA. IT DOES NOT REPRESENT AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.]**

# A Pilot Study of a New Sampling Design for the Access Point Angler Intercept Survey

**Submitted by the**

**MRIP Design and Analysis Workgroup:**


F. Jay Breidt, Colorado State University

James R. Chromy, RTI International

Kelly E. Fitzpatrick, NOAA Fisheries Southeast Fisheries Science Center

Han-Lin Lai, NOAA Fisheries Office of Science and Technology

Terri Menzel, Florida Fish and Wildlife Conservation Commission

Douglas G. Mumford, North Carolina Division of Marine Fisheries

Breda Muñoz, RTI International

Jean D. Opsomer, Colorado State University

Ronald J.Salz, NOAA Fisheries Office of Science and Technology

Kevin M. Sullivan, New Hampshire Department of Fish and Game

David A. Van Voorhees, NOAA Fisheries Office of Science and Technology

Chris Wilson, North Carolina Division of Marine Fisheries

Patricia A. Zielinski NOAA Fisheries Office of Science and Technology

# Contents

# 1. Executive Summary

An expert review conducted by the National Research Council (2006) identified problems in the Access Point Angler Intercept Survey (APAIS, or "intercept survey") that the NOAA Fisheries Service has conducted for many years as a component of the Marine Recreational Fisheries Statistics Survey (MRFSS). The survey estimators and measures of precision were not accounting for the complex sampling design, the data collection protocols were combining formal randomization with subjective decision-making in ways that make it difficult to develop statistically valid estimators, and the spatiotemporal sampling frame was not providing coverage of fishing trips ending on private property or at night.

The Marine Recreational Information Program's Design and Analysis Work Group (DAWG) initiated work in 2008 to address these concerns with the help of expert consultants. A first project completed in 2011 produced a new weighted estimation method that appropriately accounts for the MRFSS sampling design (Breidt et al., 2011). The NOAA Fisheries Service subsequently applied this method to produce design-unbiased annual estimates of 2004-2011 total finfish catches for the Atlantic and Gulf of Mexico. A second project initiated in 2009 focused on developing a new sampling design for the intercept survey that would address additional NRC concerns about the data collection protocols and temporal coverage of sampling, as well as specific recommendations provided by Breidt et al. (2011) to further improve its statistical validity and accuracy. This report describes the results of a 2010 pilot study conducted in North Carolina that tested the feasibility of implementing this new sampling design and assessed its effects on various measures of survey performance through side-by-side comparisons with the ongoing MRFSS APAIS sampling. This study did not aim to evaluate the relative merits of the two designs for the purpose of determining which one is better to use in future years, but rather it focused on developing a better understanding of how the changes to the new design would potentially affect sampling efficiency, statistical accuracy, and statistical precision going forward. This information is needed for assessing any possible needs for further modification that would ensure efficient and effective coastwide implementation of the new sampling design.

*SAMPLING METHOD CHANGES:*

The new sampling design tested in the pilot study incorporated a number of methodological changes needed to significantly improve the survey's statistical validity and accuracy.

***Time of Day Stratification:*** In the new design, sampling is stratified among four six-hour time intervals to ensure some coverage of fishing trips ending at all different times of day. In the original MRFSS sampling design, samplers were instructed to visit each assigned site during the "peak" hours when most fishing trips would be ending. In the new sampling design, samplers are assigned to a specified time interval, and the start and stop times for interviewing at each assigned site are fixed. Variability among samplers in the time intervals chosen for data collection is now eliminated. This change eliminates a potential bias when mean catch rates or proportions of coastal resident trips differ between peak and off-peak periods of fishing activity.

***Geographic Stratification:*** Sampling was stratified geographically in the pilot. Samplers were hired for one of three state subregions within North Carolina and only completed assignments within that particular geographic stratum. North Carolina sampling under the MRFSS design had never been stratified in this manner. This change allowed for more representative coverage of different management areas and also made it easier to manage staffing of the interviewing assignments.

***Clustering of Sites for Sampling:*** Low activity sites are clustered to form two- or three-site clusters in the new frame used for sampling. Sites expected to have a high level of activity are not clustered with other sites. The clustering of lower pressure sites into multi-site units increases their inclusion probabilities relative to the higher-pressure sites. Higher-activity sites still have higher inclusion probabilities than lower activity sites in the new sampling design, but there is generally less variability among sites in their probabilities and a greater chance that the sample is spread more evenly among sites that have similar fishing pressure. Samplers are required to visit all sites within the assigned cluster following a predetermined visitation order and times. Samplers are instructed to spend two hours at each site within the cluster before moving to the next site. By contrast, the MRFSS sampling frame consisted of individual sites only. Samplers were given discretion to visit "alternate" sites and to determine how long to spend at each site visited.

***Sampling Frame and Probability Sampling:*** The selection of all specific locations in space and time for interviewing assignments (i.e., the primary sampling units, or PSUs) is formalized based on a probability-proportional-to-size (PPS) approach. Thus, the new design uses a purely design-based approach to determining all site selection probabilities. Sampling under the MRFSS design also used a formal PPS approach to select primary sites (based on expected fishing pressure), but did not use a formal probability-based approach to select alternate sites. The formalization of a probability sampling approach for the selection of all interviewing locations allows more accurate determination of the correct sampling weights to be used in the estimation process.

***Issuing and Completing Assignments:*** Under the new design, emphasis is placed on completing all interviewing assignments selected by probabilistic sampling. All assignments drawn have to be either completed as assigned or canceled, because rescheduling is not allowed. By contrast, with the MRFSS design the emphasis was on attaining specified interview quotas rather than completing all drawn assignments. Eliminating assignment rescheduling greatly reduces the possibility of a nonresponse bias that could result from a failure to obtain observations from some of the selected assignments. It also eliminates possible temporal undercoverage biases that could result from the rescheduling of assignments.

***Interviewing limits:*** The new design removes all limits on the number of interviews obtained by samplers during an assignment. Samplers are directed to continue interviewing for the full specified duration of each site assignment. The MRFSS design instructed samplers to end an assignment when they reached an established cap on the number of interviews.

***Elimination of Opportunistic Sampling:*** Sampling of fishing trips in fishing mode strata other than the one for which an assignment was selected is no longer allowed under the new design. The MRFSS design traditionally allowed samplers to obtain interviews in "alternate" modes as a means of increasing the overall numbers of interviews, although alternate mode interviews were not allowed under the MRFSS design either in 2010 when this pilot study was conducted.

***Eligibility for Interviews:*** Under the new design, all intercepted anglers who have completed fishing for the day in the assigned fishing mode are considered eligible for an interview or "proxy" interview in the case of very young anglers. The MRFSS sampling design excluded anglers less than five years old, as well as any anglers returning to a site where a fishing tournament is in progress.

***Complete vs. Incomplete Beach/Bank Interviews:*** For sampling in the beach/bank fishing mode, the new design specifies that only completed angler fishing trips are eligible for an interview. Under the MRFSS design, samplers were allowed to obtain "incomplete trip" interviews in beach/bank mode. This change removes a potential source of bias because anglers who fish for longer durations would have a higher probability of being intercepted for an "incomplete trip" interview and would likely have higher mean numbers of fish caught per trip.

***Angler Trip Counts:*** The new design strongly emphasizes the need for obtaining accurate counts of all eligible angler fishing trips ending at an assigned site during the assigned time interval. Although the MRFSS design required counts of completed trips not intercepted for interview since 1990, these counts were not used in the estimation process to determine appropriate sample weights until the recent implementation of the new MRIP weighted estimation method. The greater emphasis in the new design to obtain accurate counts of all completed angler fishing trips while on site is very important to assure greater accuracy in the calculation of the secondary stage sampling fractions needed for proper weighting of the data.

The new sampling design effectively spreads the sampling of angler trips during any assignment to represent a larger temporal slice of fishing. Intercepted trips represent a much larger proportion of the total count of completed angler trips in the sampled time intervals. This results in smaller expansion factors for estimating total count for any sampled time period from the observed counts.

***Questionnaires and Data Forms:*** With the exception of one question added to identify angler trips intercepted at tournament sites, the intercept survey questionnaire used for the new sampling design matched that used under the MRFSS design. A number of changes were made to the Assignment Summary Form (ASF) and Site Description Form (SDF) to accommodate the new design's emphasis on obtaining more accurate counts and estimates of expected fishing pressures.

***ESTIMATION METHOD CHANGES:***

The access point intercept survey collects data needed to estimate the mean number of fish caught on marine recreational fishing trips. In addition, intercept survey data are used to estimate the proportion of fishing trips made by coastal county residents with a landline phone who could be contacted by the Coastal Household Telephone Survey of fishing effort. The inverse of this proportion comprises the "fishing effort adjustment ratio" that is used as a multiplier to account for fishing trips by non-coastal and out-of-state residents or anglers without landline phones. The total adjusted effort estimate is then used to expand mean catch estimates into total catch estimates. Therefore, total catch is estimated as (total trips by coast county residents) *(mean catch per angler fishing trip) *(1/proportion of trips by coastal county residents).

The weighted estimation method developed by Breidt et al. (2011) was used to estimate catch rate and effort adjustment ratio statistics from data collected under the MRFSS sampling design. This method utilizes a mix of design-based and model-based approaches to determine the appropriate sampling weights used in estimation. A new weighted estimation method that is strictly design-based was developed to estimate the catch rate and effort adjustment ratio statistics from data collected under the new sampling design.

***COMPARISONS BETWEEN MRFSS and PILOT DESIGNS:***

The MRFSS design was run side-by-side with the new pilot design in North Carolina for a full year to facilitate direct comparisons between the two.

***Sampling Yield Comparison:***

Several measures of sampling yield were selected to compare the relative sampling efficiency and effectiveness of the new design with that of the MRFSS design. Overall, the MRFSS sampling obtained a greater mean number of interviews per assignment (7.56) than the sampling under the new design (3.44), as well as a much higher mean number of interviews per hour (1.97 vs. 0.57). The greatest differences in the number of intercepts obtained per assignment, per site, and per hour occurred in the beach/bank and charter boat fishing modes. The MRFSS also obtained higher mean counts of completed trips per assignment (9.71) than the new design(3.45). However, the MRFSS sampling observed fewer sites per assignment (2.09) than the new sampling design (2.46).

In terms of sampling efficiency, the MRFSS design yielded a much lower percentage of assignments resulting in no interviews (32%),as more than one-half (51%) of assignments completed under the new design obtained no interviews. Comparisons of the temporal distributions of interviews predictably showed that sampling under the new design obtained proportionately more interviews in the nighttime and morning hours than the MRFSS sampling design obtained. There was no clear trend found in comparing the average numbers of reported fish per assignment between the new design and the MRFSS.

***Comparison of Estimators:***

In general, the two estimators of the proportion of fishing trips made by coastal county residents who could be contacted by the Coastal Household Telephone Survey produced very similar results. The only exception was in the beach/bank mode, where effort ratio estimators for MRFSS were higher than those for the new design. Although there is some suggestion that this difference could be attributable to the elimination of incomplete trip interviews or the inclusion of nighttime sampling under the new design, it was not possible to show a statistically significant difference in this proportion between complete and incomplete trip beach/bank interviews or between nighttime and daytime beach/bank trip interviews in this study. The possibility of a length of stay bias under the MRFSS design warrants further study.

Overall, no clear trends or systematic differences were found when comparing mean catch rate estimators. This was true for estimators of mean catch per trip for both removals (fish kept or released dead) and catch released alive. Removal estimates for seven of the 15 most commonly caught species were higher under the new design than under the MRFSS design. For the other eight species, the estimates based on the MRFSS design were higher. Confidence intervals overlapped for 13 out of the 15 landings estimates comparisons, suggesting that, for the large majority of cases, weighted annual catch estimates were not statistically different between the two sampling designs. In general, we expect that weighted catch estimates based on the new sampling design will be pretty similar to those based on the MRFSS sampling design for most species. However, there is some indication in this study that catch rate estimates for common night fishing targets will be higher under the new design due to the addition of formalized nighttime sampling assignments.

The estimates generated from the MRFSS sampling design were more precise than the estimates generated from the Pilot design largely because a greater number of sampling assignments were completed under the MRFSS design. It should be noted, however, that the potential for non-sampling errors is greater under the MRFSS sampling design than under the new design. Under the MRFSS design, there is a greater chance that errors can occur due to undercoverage(almost no coverage of nighttime and off-peak daytime fishing trips) and nonresponse (failure to complete many assignments as drawn for sampling). Although sampling under the new design in this study yielded a much smaller percentage of completed assignments with at least one angler trip interview and a much smaller mean number of interviews on such assignments, changes in the allocation of sampling across sampling strata could greatly reduce these differences.

## *RECOMMENDATIONS*

The Project Team identified specific recommendations based on results of this pilot study. In addition, we provide a number of recommendations for additional changes not implemented in this pilot study but that should be addressed prior to implementation of the new sampling design. Most of these recommendations focus on further improving the new sampling design to increase statistical precision without increasing costs. Finally, we identified several recommendations that require additional information and should be considered or evaluated in further studies.

### *Recommendations for Immediate Action:*

1. **In general, the Project Team recommends use of the new access point survey sampling design tested in this pilot study for conducting future access point surveys on the Atlantic coast and in the Gulf of Mexico.** The pilot study demonstrated that the new design is feasible to implement and has many advantages over the MRFSS design as described in this report.
2. **The allocation of sampling among sampling strata should be changed as needed to maximize sampling efficiency and statistical precision.** Sampling could be allocated very differently among geographic strata, fishing mode strata, and time block strata than how it was allocated in this pilot study. Without introducing any bias, other sampling allocations will likely provide higher proportions of sampling assignments that obtain at least one interview and may also provide higher average numbers of interviews per positive assignment than were observed in the pilot study. The goal should be to find the "optimal" allocation that will provide the highest level of statistical precision for the dollar spent.
3. **The formal PPS sampling of sites and site clusters should be controlled to ensure all drawn assignments can be completed by existing staff.** Staffing levels for the access point surveys should always be set to match the sampling levels required to deliver desired levels of statistical precision on resulting estimates of mean catch per trip. Once those staffing levels are established, a controlled selection program that

incorporates staffing constraints can be used to ensure the draw of a probability sample of assignments that can be covered by the available staff.

4. **Provide clearer instructions to samplers about how to handle the catch of charter boat captains and crew.** Samplers should include any catch by the captain and crew that were mixed in with the observed catch recorded for a group of charter boat anglers, but they should not count the captain and crew as contributors to the mixed group catch.
5. **Collect total catch data for any intercepted angler who just completed a multi-day fishing trip.** In addition, ask for the number of waking days that the angler fished during the trip. This will allow accurate calculation of the angler's mean catch per day for use in the mean catch estimates for the total population of angler trips.
6. **To increase on-site productivity and reduce driving time, instruct samplers to stay up to 3 hours (rather than only two hours) at the first site when a two-site cluster is assigned.**

***Recommendations for Future Consideration:***

1. **Consider requiring samplers to obtain counts of all boat trips on which anglers have finished fishing for the day.** The cluster of returning boat trips encountered at a site represents a secondary stage of sampling, and the cluster of anglers who fished on each intercepted boat represent a tertiary stage of sampling. This would allow determination of appropriate sampling fractions at both the secondary (boat level) and tertiary (angler level) stages of the multi-stage sampling design.
2. **Consider collecting catch data at the boat trip level rather than at the angler trip level for the boat modes of fishing.** This would eliminate a stage of sampling, thereby reducing both sampling error and the potential for sampler errors (i.e., non-sampling errors) in the selection of boat anglers for interviews.
3. **Consider including for-hire "guide boats" in the private/rental boat mode instead of the charter boat mode.** For-hire "guide boats" may have more in common with private boats than with charter boats in terms of size, access sites used, transiency, and target species. Adding guide boats to the private boat stratum may address an undercoverage issue associated with these trips and may also increase sampling efficiency.
4. **Evaluate options for combining boat mode trips (private/rental, guide boats, and charter boats) into a single stratum.** Sites with boat mode fishing activity often include a combination of private boats and for-hire boats. Combining these modes into a single stratum could result in more efficient sampling and fewer assignments

resulting in zero intercepts obtained. If needed for management purposes, separate catch estimates could still be calculated for private boat and for-hire sectors by treating these as "domains" within the boat mode stratum.

5. **Consider implementing more rigorous protocols to ensure random sampling of observed fish for weight and length measurements.** The project team discussed ways to improve the MRFSS sub-sampling fish procedures and developed a more rigorous random sampling protocol that would be feasible for field implementation. We recommend testing of this protocol.
6. **Consider basing rules for clustering sites more strictly on how geographic strata are defined.** In the Pilot Study, sites were only clustered together if they were within the same county. It would be more appropriate to allow clustering of sites across county boundaries if you are not stratifying sampling by county.
7. **Evaluate how best to use "confirmed" and "unconfirmed" counts of trips in calculating the secondary and tertiary stage sampling fractions used to weight the data.**
8. **Consider modifying the rules for clustering sites to use a total fishing pressure threshold as a basis for determining the number of sites in a multi-site cluster.** In the Pilot design, sites below a certain pressure threshold were clustered to form three-site clusters whenever possible. However, creating more two-site clusters would reduce the amount of time spent driving between sites. If a selected two-site cluster exceeds an established total pressure threshold similar to the one established for stand-alone sites, then it should not be necessary to add a third site to the cluster.
9. **Evaluate the feasibility of sampling beach/bank shore mode fishing trips in all states using a strict access point survey design as tested in the pilot.** In some states access to this type of shore fishing may be very diffuse, and well-defined access points may be hard to establish. In such cases, a "roving creel" sampling design that allows the collection of data for "incomplete trips" may be necessary.
10. **Evaluate the possible use of access point survey data to produce estimates of total fishing effort at sites included in the sampling frame.** Although such estimates would be incomplete because they would not account for fishing effort at sites with private access, they could serve as an independent means of monitoring trends relative to those observed in off-site telephone or mail surveys with more complete coverage.
11. **Consider splitting sites rated to have very high fishing pressure to create more total sites in the highest pressure category.** This could provide more high-pressure

alternatives to assign when the number of available days for sampling is limited, such as for weekend assignments.

12. **Consider conducting separate "frame maintenance assignments" that would survey sites and provide site register updates without attempting to collect any interviews.** Such assignments could be focused on improving the quality of the site register and the accuracy of site pressure ratings. The more accurate the pressure ratings, the more efficient the sampling can become.
13. **Consider alternative ways to define size measures and weights for sites and site clusters in the sampling frame.** The size measure for a site and time interval could be based on the expected number of fish landed rather than the expected number of angler fishing trips. Consideration should also be given to the categorization of sites with respect to their size measures. More categories or fewer categories may be better than the eight categories used in this study. In addition, more weight could be given to the sites and site clusters with higher pressure estimates in the PPS sampling. As long as lower pressure PSUs have some non-zero probability of being selected, an increase in the inclusion probabilities for higher pressure PSUs would not introduce any bias.
14. **Consider alternative ways to implement the desired stratification of sampling.** Consideration should be given to using some combination of "explicit" and "implicit" stratification. Explicit stratification creates disjoint subpopulations (in space and time), each of which is allocated a particular sample size and is sampled independently. This explicitly controls sample size within these spatio-temporal domains. An example of implicit stratification would be systematic sampling of sites within a spatiotemporal stratum after ordering by latitude. The sample size within a given latitude band would not be explicitly controlled, but there would be good representation of sites across latitudes. In particular, it would not be possible to have only southern sites within a latitude band, which could occur by chance without the implicit stratification.
15. **Consider defining different time intervals for the temporal stratification of sampling in other states.** Time interval sizes and boundaries should be chosen to ensure reasonable sampler productivity while maintaining representative sampling.

## 2. Introduction and Background

An expert review conducted by the National Research Council (2006) identified problems in the Access Point Angler Intercept Survey (APAIS, or intercept survey) that the NOAA Fisheries Service has conducted for many years as a component of the Marine Recreational Fisheries Statistics Survey (MRFSS). The APAIS had been using a stratified, multi-stage cluster sampling design to collect catch data from anglers at fishing access sites, but the current survey estimators and measures of precision did not account for the design complexity. For this reason, the estimators were potentially biased and the measures of precision were overly optimistic. In addition, the data collection protocols for the intercept survey had combined formal randomization with subjective decision-making in ways that further complicated the development of statistically valid, defensible estimators and corresponding measures of uncertainty. Finally, the spatiotemporal sampling frame used for the survey was incomplete and did not provide adequate coverage of angler fishing days ending either on private property or at night.

The Marine Recreational Information Program (MRIP) of the NOAA Fisheries Service initiated work in 2008 to address these concerns with the help of expert consultants. The first project initiated by the Design and Analysis Work Group (DAWG) produced a new weighted estimation method that accounts for the intercept survey sampling design (Breidt, et al., 2011). Some components of the sample weights needed for this method could be calculated directly from available data on sample selection probabilities and cluster sizes, but other components had to be approximated using modeling techniques. The resulting estimator of mean catch per angler fishing day is approximately design-unbiased, and appropriately incorporates the sampling design information as well as the sampling weights. The NOAA Fisheries Service subsequently applied this new method to produce more accurate annual estimates of 2004-2011 total finfish catches for the Atlantic and Gulf of Mexico. The new estimates confirmed that the statistical precision of the intercept survey was worse than previously thought. Although comparisons between the new and old estimates confirmed that the old MRFSS estimators of catch were biased, the magnitude and direction of the bias varied considerably among sampling strata and estimation domains. The net effects on annual estimates of total catch were relatively minor for most fish species, and the previous MRFSS estimates appeared to be consistently biased in one direction for only a small number of species.

Although the implementation of a design-unbiased estimation method was viewed as a very important improvement by the NRC (2006), both Breidt, et al (2011) and Chromy et al (2009) recommended changes to the sampling design of the intercept survey that would address additional NRC concerns about the data collection protocols and temporal coverage of sampling while further improving its statistical validity and accuracy. Breidt et al (2011) noted the new weighted estimation method will only provide correct estimates of mean catch rates "when the sampling, data collection, and data processing for the intercept survey are conducted in accordance with the documented sampling design." Bias could be introduced into the weighted estimator if the data structure is not arranged to accurately reflect the stratified, probability-proportional-to-size (PPS) multistage sampling design, or if the field samplers misinterpret the sampling and measurement protocols. More formalized sampling protocols with stricter control of sampler behavior are needed to ensure that a probability sample is consistently obtained. Chromy, et al (2009) stressed that "it is necessary to know the probability of selection of each unit (landing site, vessel trip, angler, or fish) interviewed or observed." Breidt, et al (2011) pointed out that a re-design of the intercept survey would (1) make it much less complicated to determine the true sample selection probabilities, (2) eliminate the need for model-based weighting methods, and (3) provide a means for a strictly design-based approach to unbiased estimation.

To achieve this goal, Breidt et al (2011) made the following recommendations to consider for improving the design of the intercept survey:

1. **The intercept survey should be re-designed to eliminate sampler visits to any sites that are not pre-determined in the probability sampling design.** Breidt, et al (2011) stated, "If clusters of sites were selected as primary sampling units (PSUs) and strict procedures were developed to determine the order and timing of the interviewer's visits to the assigned sites within the cluster, then the inclusion probabilities of all sites within the cluster would be dictated by the sampling design." The traditional MRFSS procedure to allow visits to "alternate" sites that were not selected by the sampling design complicates the development of appropriate sampling weights for the angler trip interviews collected at those sites.
2. **More emphasis should be placed on the need to spread out in time the interviews obtained within a selected site-day assignment.** Intercept survey samplers have been encouraged to maximize the number of interviews obtained per hour spent on site. This emphasis has often resulted in samplers making short site visits during

which they intercept a large cluster of angler fishing trips that ended near the same time. It would be more desirable to have angler trip interviews spread across a longer time period so that they could obtain data from more distinct time intervals and/or more distinct boat fishing trips.

3. **If different modes of fishing are sampled as separate strata with their own mode-specific site sampling frames, then opportunistic sampling of fishing trips in a mode other than the one assigned should not be a survey objective.** Breidt, et al (2011) stated, "Alternate mode interviews may be useful for assessing the different kinds of fishing activity that occur at individual sites, but the data collected from such interviews should not be used in the estimation of catch rates when sampling is stratified by mode. The difficulties of determining appropriate inclusion probabilities for alternate mode intercepts will probably always far outweigh any precision benefits that would be gained by trying to include them in the estimation of mode-specific mean catch rates."
4. **A re-designed intercept survey should pay more attention to getting accurate counts of the number of angler fishing trips that are completed within each site-day assignment.** The total count of angler trips, including those not intercepted by the interviewer, plays a very important role in calculating the PSU size measure which determines its selection probability. When conducting interviewing assignments for private boat and charter boat modes for example, it should also be an objective to get an accurate count of all of the completed boat trips so that secondary sampling units (SSUs) cluster sizes can be more accurately quantified. In fact, emphasis should be shifted away from maximizing the number of intercepts obtained per site-day assignment if it interferes with the ability of interviewers to obtain accurate counts of boat trips and angler trips during an assignment.
5. **Consider developing an approach that would cover completed fishing trips throughout the fishing day.** The traditional (MRFSS) sampling procedure instructs interviewers to visit an assigned site during the assigned day's peak activity period for fishing. Consequently, nighttime and off-peak daytime fishing trips are rarely sampled and are implicitly assumed to be similar to trips ending during the peak period. Future surveys could circumvent this potential source of bias by establishing different time block strata so that at least some sampling would occur during nighttime and daytime intervals when fishing occurs.
6. **Focus on maximizing the number of site-days sampled, not the number of angler interviews obtained.** The sampling procedures for the MRFSS have incorrectly focused too much attention on the need to maximize interviews. The total number of intercepts has been considered the "sample size" that needs to be maximized in

order to maximize the statistical precision of MRFSS estimates.  The focus should instead be on maximizing the number of site-days sampled, because the primary sampling unit in the multistage intercept survey sampling design is the site-day, not the angler trip and the precision of multi-stage survey estimators depends almost exclusively on the number of primary sampling units.

To respond to these recommendations in a timely manner, the MRIP Sampling and Estimation Work Group began work in 2009 to develop and test an improved sampling design for access point surveys of marine recreational fishing.  This work started well before completion of the work to develop the new weighted estimation method for use with current and past intercept survey data.  A project team consisting of expert consultants and representatives from NOAA Fisheries and three state agencies was formed to develop appropriate changes in sampling frames, sample selection methods, and on-site sampling protocols that would support a purely design-based estimation approach.  The goal was to develop a design in which the sampling protocols are more strictly formalized and subjective decision-making by survey managers and samplers is nearly eliminated. That work led to the development of a pilot study that could be used to test the feasibility of implementing the new sampling design.  This report describes the improved sampling design and summarizes the results of a 2010 pilot study conducted in North Carolina to test it and compare its performance with that of the MRFSS sampling design. The comparisons did not aim to evaluate the relative merits of the two designs, but rather to better understand how the changes in the new design would potentially affect sampling efficiency, statistical accuracy, and statistical precision going forward.  This information was considered to be useful for assessing any possible needs for further modification that would ensure effective coastwide implementation of the new design.

## 3. Methodology

### 3.1 Pilot Survey Data Collection Methods

Methodological improvements were developed for a new intercept survey design that was tested in comparison with the traditional MRFSS design in a pilot study conducted in North Carolina from January through December 2010.  The emphasis here is on describing differences between the traditional MRFSS methods and the new methods tested in the North Carolina pilot study (Pilot). Methodological changes were implemented in response to both specific NRC recommendations and to address other potential biases or inefficiencies of the old methods identified by the project team. In addition to documenting proposed changes, this section includes rationale for each change and potential issues or trade-offs associated with the new methodology. While methodological changes were extensive, some aspects of the MRFSS methodology remained essentially unchanged (e.g., survey instrument, site fishing pressure categories, angler level trip information etc.). Pilot study methods that remained the same as the MRFSS are not covered in any detail in this document but are described in other reference documents such as the North Carolina Pilot Field Procedures Manual (Appendix A) and the MRFSS 2010 Statement of Work.

Key data collection design changes (described below in more detail) that were implemented in the pilot include:

1) Sampling from four fixed 6-hour time intervals covering a full 24-hour sampling day.
2) Formalizing a probability-based approach for the selection and order of all sites to visit on a given assignment.
3) Clustering of sites for sampling.
4) Eliminating opportunistic sampling of alternate modes.
5) Attempting to complete all assignments drawn, thus reducing possible bias due to non-observation of selected elements in the sample frame.
6) Cancelling assignments that could not be completed rather than re-scheduling, which made it difficult to determine sampling probabilities.
7) Improving methods for accurately obtaining counts of eligible angler trips missed, to determine appropriate sampling weights of intercepted trips in the estimation process
8) Expanding eligible trip definition to include anglers under five years old and trips at tournament sites.
9) Disallowing "incomplete trips" in shore mode, thus eliminating potential bias associated with expanding partial trip catch to represent the entire trip.
10) Removing the interview per assignment cap which, when combined with fixed assignment time intervals, should spread the sampling to appropriately represent a larger temporal slice of fishing.

This section is divided into the following subsections: Sampling Methods, Issuing and Completing Assignments, and On-site Interviewing Procedures.

### 3.1.1 Sampling Methods

#### 3.1.1.1Expanded Coverage and Fixed Time Intervals

This sub-section addresses two important design improvements:

1. Expanded coverage of fishing trips to include trips ending at nighttime and off-peak daytime hours eliminates potential for bias when those trips differ in mean catch rates from trips ending in peak activity periods.
2. Implementation of fixed time-block strata for sampling and fixed time intervals for interviewing makes it easier to determine appropriate cluster sampling weights (at SSU level) to be used in estimation.

In the MRFSS design, samplers determined the start and stop times of each assignment. Samplers were instructed to be at the site during the "peak" hours when most fishing trips would be ending. To remove any sampler discretion regarding selection of assignment times, clearly defined assignment time intervals were used for the Pilot. Historical MRFSS North Carolina data were used to compare trip completion times between the access point intercept survey and Coastal Household Telephone Survey. A six-hour sampling interval was selected as this would allow for a standard eight-hour workday when travel time (to the first site and from the last site comprising a selected cluster)is included. For the Pilot, assignment start and stop times for four distinct 6-hour time intervals were defined as follows: Interval A: 2AM-8AM
Interval B: 8AM-2PM
Interval C: 2PM-8PM
Interval D: 8PM-2AM

Samplers were instructed to arrive at their assigned site at the start of the assigned time interval and to only conduct interviews within that interval and selected fishing mode. In the event of late arrival, the samplers were instructed to adhere to the original ending time (i.e., they were not allowed to stay late to "make up" for being late). Establishment of assignment time intervals resulted in the following design improvements:

1. Removed sampler discretion regarding sampling times that may lead to biases that are unknown and/or unaccounted for;

2. Removed sampler discretion associated with determining “peak activity” times which resulted in improved Pilot fishing pressure estimates for each particular time interval and weekday/weekend combination;
3. Allowed for a more temporally distributed sample across the day that could be properly weighted using angler counts specific to each time interval;
4. Eliminated potential under-coverage bias from missed fishing activity during “off-peak” sampling times (i.e., night and early morning).

The master site register (MSR), a database of all saltwater recreational fin-fishing locations in each state, is the basis for the sampling frame. In the MRFSS, fishing pressure was estimated for each site, mode, kind of day (weekend or weekday), and wave, and was intended to represent the expected fishing pressure during the peak activity. In the Pilot, the fishing pressures were estimated for each of the four six-hour time intervals. Samplers provided fishing pressure updates only for the specific time interval and assigned mode observed, rather than for some undefined “peak” 8-hour interval as with the MRFSS. This eliminated the guesswork associated with estimating pressures for the whole day that was often a problem under the old approach. Previously, samplers often estimated pressures beyond the amount of time actually spent at a particular site since there was no requirement that the sampler stay on site for any particular amount of time. Table 1 shows the pressure categories and values used in both the MRFSS and Pilot.

### Table 1. Pressure Categories

| Pressure Category | Expected Number of Angler-trips |
|---|---|
| 0 | 1 – 4 |
| 1 | 5 – 8 |
| 2 | 9 – 12 |
| 3 | 13 – 19 |
| 4 | 20 – 29 |
| 5 | 30 – 49 |
| 6 | 50 – 79 |
| 7 | 80+ |
| 8 | Unable to determine |
| 9 | Mode not present at site or inactive site |

### 3.1.1.2. Clustering of Sites

For the Pilot the maximum number of sites in a given cluster was three. All sites within the cluster had to be visited in the exact order specified during the assignment draw process. In addition, the sample period was set at a maximum of two hours at each site, after which time the sampler was required to move to the next site. For two-site clusters samplers were instructed to spend two hours at the first site, two hours at the second site, and as time allowed return to the first site and sample until the six-hour time interval was up. Two hours duration was maintained at two-site clusters for consistency with three-site clusters. At single site clusters, the sampler remained at one site for the entire 6-hour time interval.

The project team developed the following constraints for clustering:

- Sites with a pressure code of "4" or greater would not be clustered with other sites (i.e. single site cluster);
- Sites with a pressure code of "3" or less could be clustered with up to two additional sites;
- Driving time between any two sites within a single cluster must be less than 60 minutes;
- Total driving time for the entire cluster should be minimized;
- Clusters will contain sites only within the same county (see Regional Stratification in section 3.1.1.5.);
- Sites will be clustered by strata (state subregion/month/mode/time interval) such that all sites within the cluster are required to have positive fishing pressure in that strata. Clusters must be time-interval specific since individual site pressures will vary across intervals (e.g., a high pressure site may be a single site cluster from 2:00PM-8:00PM but clustered with other sites from 8:00PM-2:00AM due to a change in pressure rating).

### 3.1.1.3 Clustering Method

Using the clustering constraints described above, a GIS algorithm was developed based on the concept of "simulated annealing." Simulated annealing involves establishing certain criteria (desirable or not) and assigning "costs" to those (high or low) depending on their desirability. Simulated annealing attempts to maintain low cost at all times.

For the Pilot, desirable attributes included minimizing driving distance between sites within a cluster and maintaining similar size measures (total fishing pressure or effort) across clusters. For example, a desirable clustering attribute such as two sites in close proximity to one another would have a relatively low cost compared to two sites farther apart. Similarly, a non-desirable attribute such as clustering three relatively high pressures sites would have a high cost compared to clustering a relatively high pressure site with two very low pressure sites. The algorithm developed identifies many possible clustering combinations and then ranks them such that the combination with the most desirable attributes (i.e. "lowest total cost") can be identified. High activity sites (fishing pressure 4 or greater) were automatically identified as single site clusters. Since fishing pressures are not static across waves and modes, cluster combinations also changed across waves and modes. For example, two sites may be in the same cluster during Wave 3 but not Wave 4. Similarly, two sites may be clustered for Charter boat mode assignments but not for Private Boat mode assignments.

The result is a list of clusters, each containing anywhere from 1 to 3 sites, with minimized "cost" (i.e. meeting the constraints).Project team members with considerable knowledge of North Carolina's fishing sites thoroughly reviewed and evaluated all clusters before each sample draw. Site cluster maps were produced for each cluster identified for sampling (Appendix B).

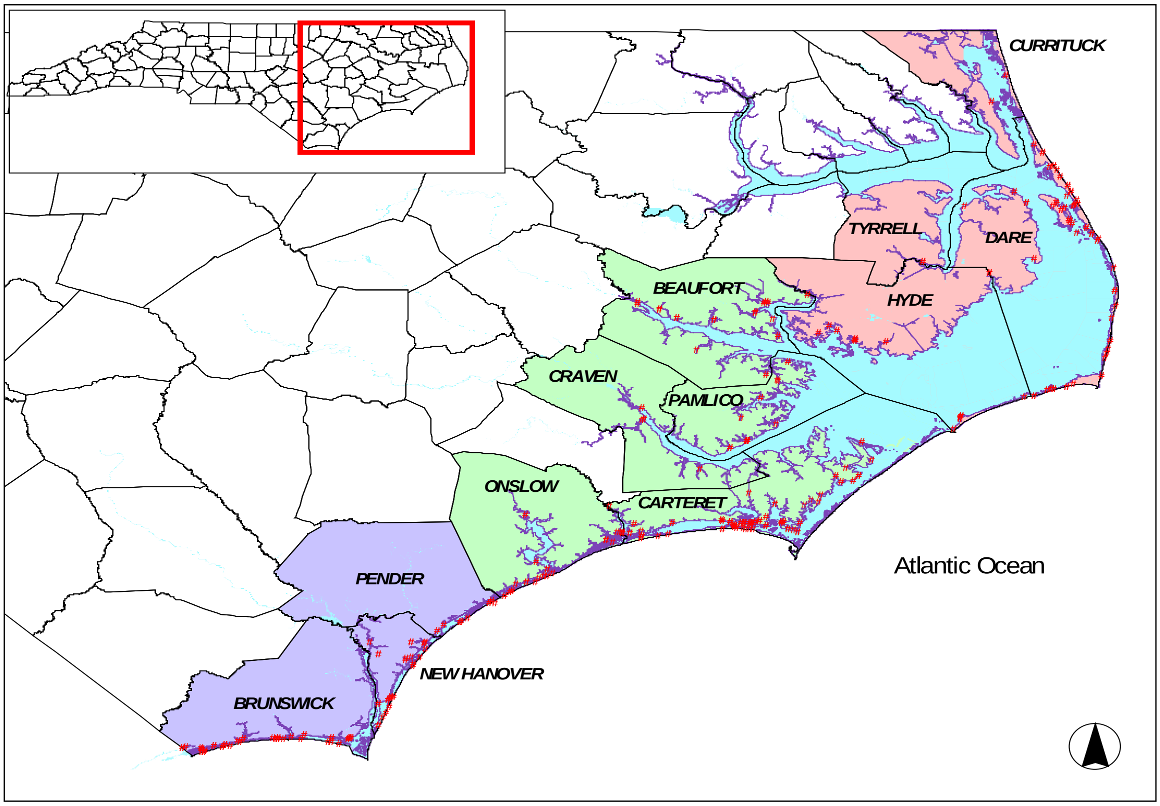
#### 3.1.1.4Formalized Probability Sampling of Sites

A new selection procedure was developed that pre-determined all site assignments through the sample draw process. Interviewers were required to collect data at a selected site for a specified time interval and were not allowed discretion regarding when to leave a site or which site to visit next.

#### 3.1.1.5 Regional Stratification

For the Pilot, the project team tested regional stratification within North Carolina. North Carolina's coastal zone was divided into three subregions (Northern, Central, and Southern) using county boundary lines based on existing state and federal fisheries management units as well as recreational fishing and geographic diversity (Figure 1).

Figure 1 Survey subregions and fishing access sites used for the NC Pilot Project




## 3.1.1.6 Sample Size and Allocation

Under the MRFSS intercept design, "sample size" referred to the total number of completed interviews obtained. Specific sampling goals or quotas were established for each strata and attainment of these goals was closely managed and monitored by contractors, state agencies and NOAA Fisheries. By contrast, for the Pilot study design, the effective sample size was defined as the total number of assignments completed or PSUs rather than the number of interviews obtained.

The total number of interviewing assignments to be selected for the Pilot was determined by the number of samplers available for the Pilot and the number of working days allowed per sampler. From January through September, 6 samplers were available for the Pilot with two samplers being assigned to each state subregion. Samplers were limited to one assignment per day for the Pilot. Since each sampler was available to work a maximum of 12 weekday days and 8 weekend days per month, the maximum number of monthly assignments per state subregion was 24 for weekdays and 16 for weekend days. Ten samplers were available for October through December, with corresponding increases in the number of maximum assignments.

For the Pilot, assignments were allocated evenly across the four modes in each state subregion: Man-made (MM), Beach Bank (BB), Private/Rental (PR), and Charter (CH). Allocation of mode-specific assignments within each state subregion and day type (i.e. kind of day)was determined monthly.

In the initial Pilot allocation a minimum of one PSU was sampled from each interval, resulting in at least two night interval assignments (A:8PM – 2 AM& D:2AM – 8 AM) selected for every month, mode, state subregion, and day type. The only exception was if there was no night fishing activity for a particular stratum. This allocation resulted in a much higher proportion of night time interval assignments selected than was warranted based on fishing pressures. With 4 modes, 3 state subregions, and 2 night time intervals the number of night time interval assignments per months can add up quickly (i.e., 4 X 3 X 2 = 24). While the actual number of night assignments selected was less than this number (i.e., not all combinations had night activity) the proportion of night assignments was still quite large in many months. For example, 34 out of a total 118 assignments (29%) drawn in May were night time interval assignments. It is anticipated that night time interval (A & D) fishing pressure estimates will improve over time once the new design is fully implemented.

To resolve the issue of night assignments being drawn too frequently, the two night intervals (A & D) were combined into one stratum for sampling purposes starting with the June sample draw. Although the two night-intervals were combined, no PSUs were removed from any of the intervals. This approach allowed for probability sampling within the combined night interval that more closely reflected the estimated pressures while still assuring that some minimal number of night assignments were drawn within each month, mode, and state subregion.

In the first five months, a minimum of one assignment was drawn and completed for each of the sampling strata under the new design, resulting in at least two night interval assignments selected for every month, mode, state subregion, and day type. The only exception was if there was no night fishing activity for a particular stratum. Starting in June, the two nighttime blocks were combined into one “nighttime” stratum requiring the minimum of one interviewing assignment.

### 3.1.1.7 Sample Frame and Assignment Draw

The North Carolina Pilot sample frame consisted of all possible combinations of clusters, calendar days, and time intervals within a given stratum, i.e. month/mode/kind-of-day/ state subregion combinations. The D:8PM-2AM time interval extends over two calendar days. For purposes of the draw, the Friday 8:00 PM to Saturday 2:00 AM time interval was considered a "weekend" assignment while the Sunday 8:00 PM to Monday 2:00 AM interval was considered a "weekday" assignment in the pilot.

The total pressure for a cluster was defined as the sum of individual site pressures calculated as the midpoint of the pressure category range. For example, if a pressure category 1 site (5-8 angler trips) is clustered with a pressure category 3 site (13-19 angler trips) the cumulative cluster pressure is 22.5 (6.5 + 16). The interval weights were calculated as the inverse of total cluster pressure for each state subregion and kind of day. Probability proportional to size (PPS) systematic sampling was used to select a random sample of assignments for each state subregion.

Several logistical constraints related to sampler availability were incorporated into the assignment draw process:

- No more than two day interval (B or C) assignments (PSUs) could be selected on the same day in a given state subregion, since only 2 samplers were available per state subregion.
- Single-site cluster assignments with pressure codes of five or higher required two samplers, one to conduct interviews and one to count angler trips.
- Eight or more hours of employee rest between assignments were required by state labor regulations. For example, if time interval A:2AM-8AM on June 4$^{th}$ is assigned to a sampler, that sampler cannot be issued the two intervals before the assignment (C:2PM-8PM or D:8PM-2AM on June 3$^{rd}$) or two intervals after the assignment (B:8AM-2PM or C:2PM-8PM on June 4$^{th}$).
- For safety reasons, an assignment in either of the night intervals (A:2AM-8AM or D:8PM-2AM) required two samplers working together in the field. Therefore, no more than one night interval assignment could be selected within a 12 hour period (i.e., two intervals) in a given state subregion since only 2 samplers were available per state subregion.
- Samplers cannot work more than 40 hours per week, including travel and editing time.

The Pilot study assignment schedule process maximized the number of assignments that could be completed by the relatively small number of samplers.

### 3.1.2 Issuing and Completing Assignments

The issuing of assignments in the Pilot differed from the MRFSS in several important ways. The MRFSS draws three different kinds of assignments in hierarchical order of importance: 1) fixed - must be issued, 2) flexible – must be issued only until the interview goal is attained for a particular stratum, and 3) reserve – only issued if anticipated that the interview goal cannot be attained with fixed and flexible assignments alone. By contrast, all drawn Pilot assignments had the same importance and were issued.

All Pilot assignments that were drawn (i.e., issued) had to either be completed or cancelled since rescheduling was not allowed. As discussed above, sampler discretion regarding sites visits (i.e., order, duration, exact time start and stop times) was removed for the Pilot. For multi-site clusters the site visitation order was circular (e.g., ABC, ABC... as time allows within the 6-hour interval) and the starting point was randomized prior to assignment.

### 3.1.3 On-Site Interviewing Procedures

Pilot survey samplers only conducted Pilot assignments to avoid confusion with MRFSS procedures. A more detailed description of the Pilot field interview procedures, including procedures that remained the same as those followed by MRFSS samplers, can be found in the NC Pilot Field Procedures Manual (Appendix A).

#### 3.1.3.1 Definition of an Eligible Angler Trip

The NRC report identified several potential under-coverage biases associated with the MRFSS intercept survey criteria for defining an eligible angler trip. The Pilot attempted to address these and other potential coverage biases through the following design changes regarding the definition of an eligible angler trip:

**1.** Anglers Under 5 Years Old

Anglers under 5 years of age are excluded from the MRFSS Intercept survey as ineligible, though they are tallied on the Assignment Summary Form. In the Pilot all anglers, regardless of age, were eligible to be interviewed either in person or through proxy interviews, as was the case with very young anglers.

**2.** For-Hire Captains and Crew

Similar to the MRFSS, Pilot survey samplers did not count the captain and crew as contributors since they were technically not fishing recreationally and their trip would not be reported as recreational trips in the For-Hire phone survey. However, unlike in the MRFSS, Pilot samplers were instructed to include any catch by the captain and crew that were mixed in with the observed catch (Type A catch) recorded for a group of charter boat anglers.

3. Tournament Trips

For the Pilot, there was no tournament restriction in place and samplers were instructed to stay and interview at tournament weigh station sites if they were part of the assigned cluster. Pilot samplers were reminded that they should not station themselves in locations that only anglers with catch would visit (e.g. the cleaning station or weigh station) as this could bias catch rates, particularly at tournament settings. A question was added to the Pilot intercept form (to be asked of every person interviewed) as to whether or not the angler fished in a tournament that day. In addition, samplers were instructed to record whether or not the site was an official tournament weigh-station for that assignment on the Assignment Summary Form (ASF).

4. Incomplete Trip Interviews

To increase intercept productivity, MRFSS procedures allow for up to half (50%) of intercepts for a beach/bank (BB) mode assignment to be conducted with anglers who are at least 1/3$^{rd}$ done with their fishing trip (i.e., “incomplete trip” interviews). The determination of whether 1/3$^{rd}$ of a trip is complete is based on asking the angler how much longer they intend to fish. Incomplete trip interviews were seen as a way to increase BB productivity because 1) BB anglers tend to fish longer periods of time than in other modes (i.e. beyond the constraints of a typical work day) and 2) at some BB sites anglers are spread out across a large distance and use multiple points of egress making it difficult for a sampler to intercept completed trips. MRFSS catch rates during the completed portion are then extrapolated to the uncompleted portion of the trip for estimation purposes.  However, this will likely biased survey estimates of the length of the fishing trip, since the assumption  catch rates for the completed portion are the same as catch rates for the uncompleted portion may be erroneous.  To eliminate this potential bias, incomplete trip interviews were not allowed in the Pilot.

### 3.1.3.2 Angler Trip Counts (SSU Cluster Sizes)

A “missed eligible” is an angler trip that was likely eligible to be interviewed, but was not due to the sampler already interviewing other anglers or some similar situation. Two main types of “missed eligible” trips were identified: 1) “Confirmed” trip - sampler was able to “screen” the angler (i.e. to speak with the angler to verify the angler fished recreationally, was targeting finfish, fished in U.S. waters, and was done fishing in that mode for that day), and 2) “Unconfirmed” trip - unable to screen the person because they left the site while the sampler was busy interviewing, screening other anglers or the sampler was otherwise unable to approach the person.

For the Pilot, samplers were instructed to attempt to screen people on all vessels, including canoes, kayaks, and even jet skis, to confirm whether or not they fished that day.  In addition, people who appeared to be shellfishing or lobstering were also screened to confirm that they did not target or incidentally catch a finfish.

The distribution of the type of “missed eligible” (confirmed versus unconfirmed) tallied was expected to be correlated with the level of fishing activity at a site on a particular day.  That is, if there is little activity at a site it should be relatively easy to either interview all eligible anglers or count the few anglers not interviewed.  By contrast, if there are many boats returning at the same time or many shore anglers leaving the site at the same time the accuracy of angler counts will likely diminish and it may not be possible to screen everyone leaving the site (i.e., the proportion of “unconfirmed” trips will tend to increase).  For the Pilot, to maintain a high level of accuracy in these situations, two samplers were assigned to sites with a pressure category of 5 (30-49 anglers) or higher.  One sampler conducted interviews while the other conducted angler counts and attempted to confirm eligible angler trips by screening anglers whenever possible.  To avoid double counting trips, the sampler doing the counts did not include interviewed anglers.  At no time did both samplers engage in the same activity at the same time.  The two samplers worked together to fill out one assignment summary form (ASF) for the assignment. Similar procedures for splitting counting and interviewing between two samplers were used for all night assignments (i.e. Intervals A and D).

Procedures were also changed in the Pilot to improve the accuracy of angler trip counts for assignments with only one sampler (i.e., pressure category 4 or less). Under normal circumstances, one sampler should be able to interview all (or virtually all) eligible anglers in the assigned mode at pressure category 4 (20-29 anglers) or smaller sites, and screen any anglers that could not be interviewed. However, on any given day fishing activity level may be higher than expected making it difficult to simultaneously conduct interviews and obtain accurate counts. The physical layout of the site (e.g., size, number of egress points) may also be a factor affecting the ability to conduct interviews and accurate counts simultaneously. If the sampler determines that fishing activity is such that they cannot effectively interview and count at the same time they should alternate between conducting interviews and conducting counts, in one hour increments for the time they are supposed to be at that site. Samplers recorded the survey method used (1=interview, 2=count, 3=both simultaneously) and the start and stop times for each method used at each site on the ASF. Since some time will be dedicated to counting and not interviewing, a reduction in the number of interviews per assignment was expected with these procedural changes.

### 3.1.3.3 Intercept Limit per Assignment

Under MRFSS intercept procedures, an upper limit was placed on the number of intercepts a sampler could obtain per assignment: 20 intercepts per assignment from Maine through Virginia; 30 intercepts per assignment from North Carolina through Louisiana. The limit served to more evenly distribute intercepts over more assignments so that a few assignments with a lot of intercepts would not fill the intercept quota for a particular wave/state/mode combination, and thus heavily influence catch rates in that stratum. These concerns were not an issue for the Pilot, since sampling goals or quotas were defined in terms of site-days rather than interviews completed and appropriate weighting of Catch Per Unit Effort (CPUE) data eliminates concerns about over-sampling a given site/day combination. Therefore, for the Pilot there was no limit on the number of intercepts that could be obtained per assignment.

### 3.1.3.4 Form Changes for Pilot

With the exception of the question added for tournament trips (3.1.3.1) the intercept survey form used for the Pilot matched that used in the MRFSS. More changes were made to the Assignment Summary Form (ASF, Appendix C) and Site Description Form (SDF, Appendix D) to accommodate new field procedures implemented in the Pilot. These changes are summarized below.

Assignment Summary Form changes:

- Added box to record second sampler code to be used for night assignments and pressure category 5 or greater assignments;
- Added boxes to record total "confirmed" and "unconfirmed" numbers of angler trips and start and stop times associated with these counts. Note: "confirmed" and "unconfirmed" boxes replaced boxes for "missed" at bottom of MRFSS ASF;
- Provided boxes to tally counts of "confirmed" and "unconfirmed" angler trips and refusals and language barriers;
- Added box to indicate the survey activity: 1 = interviewing, 2 = counting, and 3 = both simultaneously;
- Added box to indicate whether or not the site was a tournament weigh station;
- Added box to record the assignment cluster identification number;
- Reason codes for leaving a site were expanded to include: 1)two hour time interval ended, 2) six hour assignment time interval ended, 3) site closed (after hours), 4)site closed (other specify), 5) site unsafe during sampling period;
- The following reason codes for leaving site were removed as they no longer applied under the new procedures: 1) no activity in mode (weather unfavorable), 2) no activity in mode (weather favorable), 3) fewer than eight intercepts in mode, 4) got quota in mode, 5) tournament weigh station.

Site Description Form changes:

- Added box to record second sampler code to be used for night assignments and pressure category 5 or greater assignments;

- Since weather can greatly affect the fishing pressure for a given day, check boxes were added to record more detailed weather information than previously recorded. Wind speed is now recorded by category using a scale ranging in knots (e.g., breezy = 1 to 16 knots, windy = 17-33 knots etc.). This type of detailed information may be useful for adjusting for weather when setting site pressures;
- Added area to record site latitude and longitude to improve the information on the site register and make it easier for samplers to locate a site, and to verify that they are in the right location;
- Added boxes to indicate whether or not night fishing is present for all modes, not just shore (SH) and private/rental (PR) as was previously done.
- For the Pilot, samplers were asked to estimate fishing pressure only for the particular mode and six-hour time interval of the assignment for both weekend/weekday and both months of the current wave. This is different from MRFSS, where pressure was estimated for all modes and "peak productivity" (morning, mid-day, afternoon, night) was also recorded.

## 3.2 Methods used for Data Analysis and Comparisons

### 3.2.1 Sampling Yield and Effectiveness

Several measures of sampling yield and effectiveness were selected to compare the relative sampling efficiency and effectiveness between the MRFSS and Pilot sampling designs. These metrics included: 1) average number of intercepts per assignment, 2) average number of intercepts per hour, 3) average number of anglers (interviewed or missed) per assignment, 4) average number of sites visited per assignment, and 5) the ratio of actual time on site versus recorded site hours (including travel time between sites). Time of intercept was also examined to determine the number of intercepts obtained through the Pilot during times not typically surveyed in the MRFSS. Finally, the average numbers of fish reported and observed were compared between surveys for selected common fish species.

Because MRFSS sampling locations consist of both locations randomly selected using a probability sampling design(i.e. primary sites) and locations chosen by samplers (i.e. alternate sites), two sets of measurements were produced for MRFSS when possible for comparison with the Pilot. Difference between methodologies for each metric was calculated as the percent change from MRFSS to Pilot.

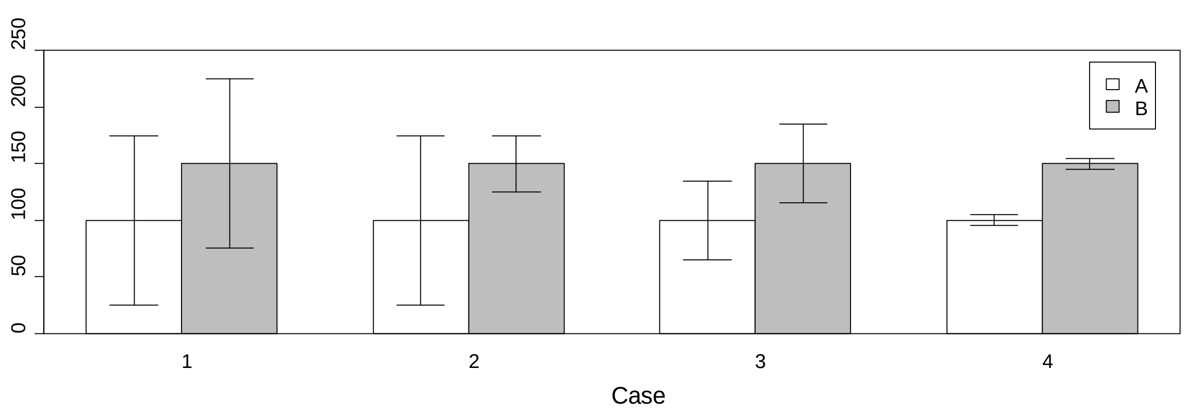
Because staffing levels and number of completed assignments differed between the MRFSS and Pilot surveys, all metrics presented use either averages (e.g. intercepts per assignment or per hour) or ratios to allow for more meaningful comparisons.

### 3.2.2 Catch Rates and Estimates Comparison

For each estimate, a 95% confidence interval (CI) was calculated as the estimate plus and minus 1.96 times the standard error. The CIs may not be valid for some estimates due to sparse or skewed distributions caused by small sample size. The degree of confidence interval overlap was used to informally assess differences between estimates. Note that statistical significance does not imply biological or management significance. Four degrees of overlap were considered:

- Case 1 - Estimate of Method B falls within Method A confidence interval and estimate of Method A falls within Method B confidence interval
- Case 2 - Estimate of Method B falls within Method A confidence interval or estimate of Method A falls within Method B confidence interval
- Case 3 - Neither estimate falls within the other confidence interval, however the confidence intervals do overlap
- Case 4 - The confidence levels do not overlap

**Table 2. Illustration of four outcomes (cases) for comparison of survey estimates.**



## 4. Results and Analyses

### 4.1 Sampling Yield and Effectiveness

Table 3 below shows a monthly comparison of the total number of assignments completed, total number of sites visited, and total number of intercepts obtained in the MRFSS and the Pilot, respectively. For comparison purposes, it is important to note that in the MRFSS there were 12 samplers in January and 15 samplers in February through December. In the Pilot study, there were 6 samplers from January through September, and 10 samplers from October through December.

**Table 3. Total number of assignments completed, number of sites visited, and number of intercepts obtained by survey (MRFSS and Pilot)**

| MRFSS | # of assignments completed | # of sites visited | # of intercepts |
|---|---|---|---|
| **January** | 154 | 409 | 244 |
| **February** | 139 | 352 | 235 |
| **March** | 205 | 516 | 685 |
| **April** | 159 | 362 | 1307 |
| **May** | 218 | 423 | 2384 |
| **June** | 223 | 405 | 2777 |
| **July** | 216 | 407 | 2887 |
| **August** | 237 | 429 | 2957 |
| **September** | 220 | 475 | 2677 |
| **October** | 246 | 459 | 2892 |
| **November** | 179 | 319 | 965 |
| **December** | 170 | 400 | 290 |
| | | | |
| **TOTALS** | 2366 | 4956 | 20300 |
| | | | |

| Pilot | # of assignments completed | # of sites visited | # of intercepts |
|---|---|---|---|
| **January** | 64 | 161 | 70 |
| **February** | 61 | 149 | 89 |
| **March** | 61 | 144 | 116 |
| **April** | 69 | 172 | 260 |
| **May** | 64 | 162 | 379 |
| **June** | 62 | 149 | 511 |
| **July** | 59 | 144 | 516 |
| **August** | 61 | 139 | 472 |
| **September** | 62 | 154 | 339 |
| **October** | 70 | 172 | 450 |
| **November** | 91 | 230 | 356 |
| **December** | 98 | 248 | 58 |
| | | | |
| **TOTALS** | 822 | 2024 | 3616 |
| | | | |

MRFSS samplers visited fewer sites per assignment (2.09) than Pilot samplers (2.46). Under the MRFSS sampling design, 36.7% of the interviewing assignments visited only one site, 19.5% visited two sites, and 43.8% visited three sites. Under the Pilot sampling design, 12.2% of the assignments visited only one site, 32.4% visited two sites, and 55.4% visited three sites.

The total number of completed assignments or Primary Sampling Units (PSUs) obtained for the MRFSS was larger than for the Pilot (Table 4). By contrast, the Pilot had a much larger percent of assignments that resulted in no intercepts ("empty PSUs") compared to the MRFSS. More than one-half of all Pilot PSUs were "empty."

### Table 4. Total number of Primary Sampling Units (PSUs) visited by mode and survey (MFRSS and Pilot)

| | Beach Bank | | | | Man-Made | | | |
|---|---|---|---|---|---|---|---|---|
| WAVE | Pilot PSUs | Pilot % Empty | MRFSS PSUs | MRFSS % Empty | Pilot PSUs | Pilot % Empty | MRFSS PSUs | MRFSS % Empty |
| | | | | | | | | |
| 1 | 30 | 73.3 | 59 | 67.8 | 45 | 88.9 | 0 | 0 |
| 2 | 40 | 50.0 | 87 | 43.7 | 41 | 48.8 | 56 | 17.9 |
| 3 | 43 | 25.6 | 97 | 20.6 | 41 | 4.9 | 77 | 6.5 |
| 4 | 33 | 39.4 | 103 | 13.6 | 41 | 4.9 | 86 | 7.0 |
| 5 | 44 | 40.9 | 117 | 11.1 | 38 | 10.5 | 104 | 8.7 |
| 6 | 61 | 60.7 | 118 | 38.1 | 50 | 48.0 | 91 | 36.3 |
| All Waves Combined | 251 | 48.2 | 581 | 29.3 | 256 | 35.9 | 414 | 15.2 |

| | Private/Rental | | | | Charter | | | |
|---|---|---|---|---|---|---|---|---|
| WAVE | Pilot PSUs | Pilot % Empty | MRFSS PSUs | MRFSS % Empty | Pilot PSUs | Pilot % Empty | MRFSS PSUs | MRFSS % Empty |
| | | | | | | | | |
| 1 | 62 | 62.9 | 137 | 67.2 | 29 | 86.2 | 97 | 84.5 |
| 2 | 47 | 51.1 | 159 | 45.9 | 43 | 76.7 | 106 | 67.0 |
| 3 | 48 | 33.3 | 231 | 16.0 | 35 | 48.6 | 90 | 26.7 |
| 4 | 44 | 22.7 | 255 | 11.0 | 43 | 48.8 | 72 | 19.4 |
| 5 | 46 | 45.7 | 253 | 22.5 | 42 | 78.6 | 81 | 46.9 |
| 6 | 69 | 58.0 | 126 | 36.5 | 55 | 89.1 | 95 | 71.6 |
| All Waves Combined | 316 | 47.5 | 1161 | 28.7 | 247 | 72.1 | 541 | 54.9 |

Table 5 displays average values and percent change calculated for several measures, by survey and fish mode. Percent change was calculated as the Pilot measure minus the MRFSS measure divided by the MRFSS measure (i.e., a negative percent change means that the MRFSS measure exceeded that of the Pilot).

The greatest differences in the number of intercepts obtained per assignment occurred in beach/bank (-67%) and charter boat (-65%) fishing modes (Table 5). Although differences were not as pronounced, similar results were found when comparing the number of intercepts from MRFSS primary sites with the Pilot survey (not shown in table). Geographically, the Southern region of North Carolina exhibited the smallest difference in the number of intercepts per assignment between MRFSS and Pilot for all modes except charterboat (not shown in table). Overall, across modes, the largest difference in the number of intercepts per assignment was observed in the Northern region.

Similarly, the greatest differences in the number of intercepts obtained per hour were observed for the beach/bank (-80%) and charter boat (-81%) fishing modes. Comparisons of the number of intercepts per hour at MRFSS primary sites with the Pilot survey resulted in similar differences across all modes. Overall, across modes the Northern region revealed the largest difference in the number of intercepts obtained per hour.

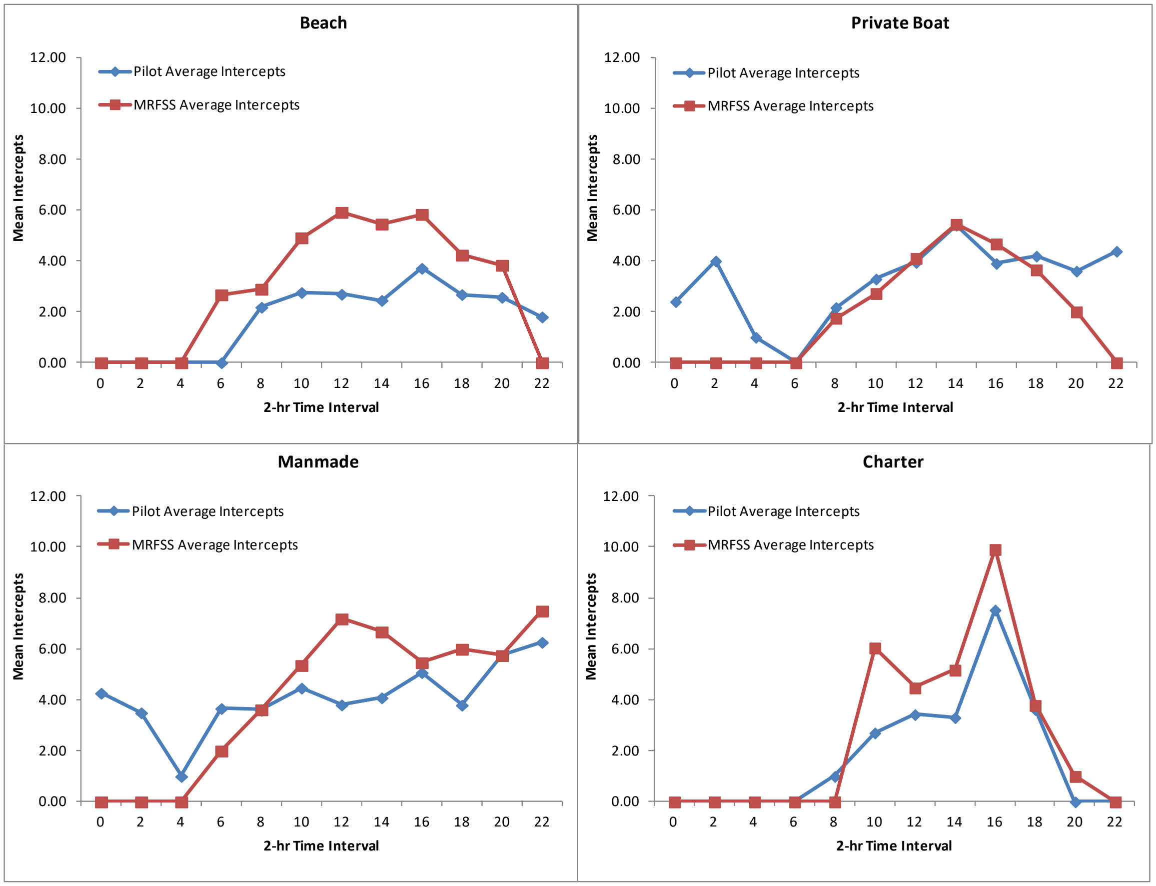
The greatest differences in the number of angler trips counted(interviewed plus missed) per assignment occurred in beach/bank and charter boat fishing modes (Table 5). Geographically, the Southern subregion of North Carolina exhibited the smallest difference between MRFSS and Pilot methodologies for all modes except charterboat. Overall, across modes, the Northern subregion generally revealed the largest difference in the number of angler trips counted (interviewed plus missed) per assignment.

Figure 2 displays the average number of intercepts per two-hour time period for both surveys methodologies. Higher numbers of intercepts were observed for pre-dawn hours for private boat and man-made fishing modes for the Pilot compared to MRFSS. The Pilot survey also had higher average intercepts from 6:00 pm through 12:00 am for the private boat mode and 11:00 pm – 12:00 am for the beach/bank mode (Figure 2).

**Table 5. Percent change of average values by measure, study and fishing mode.**

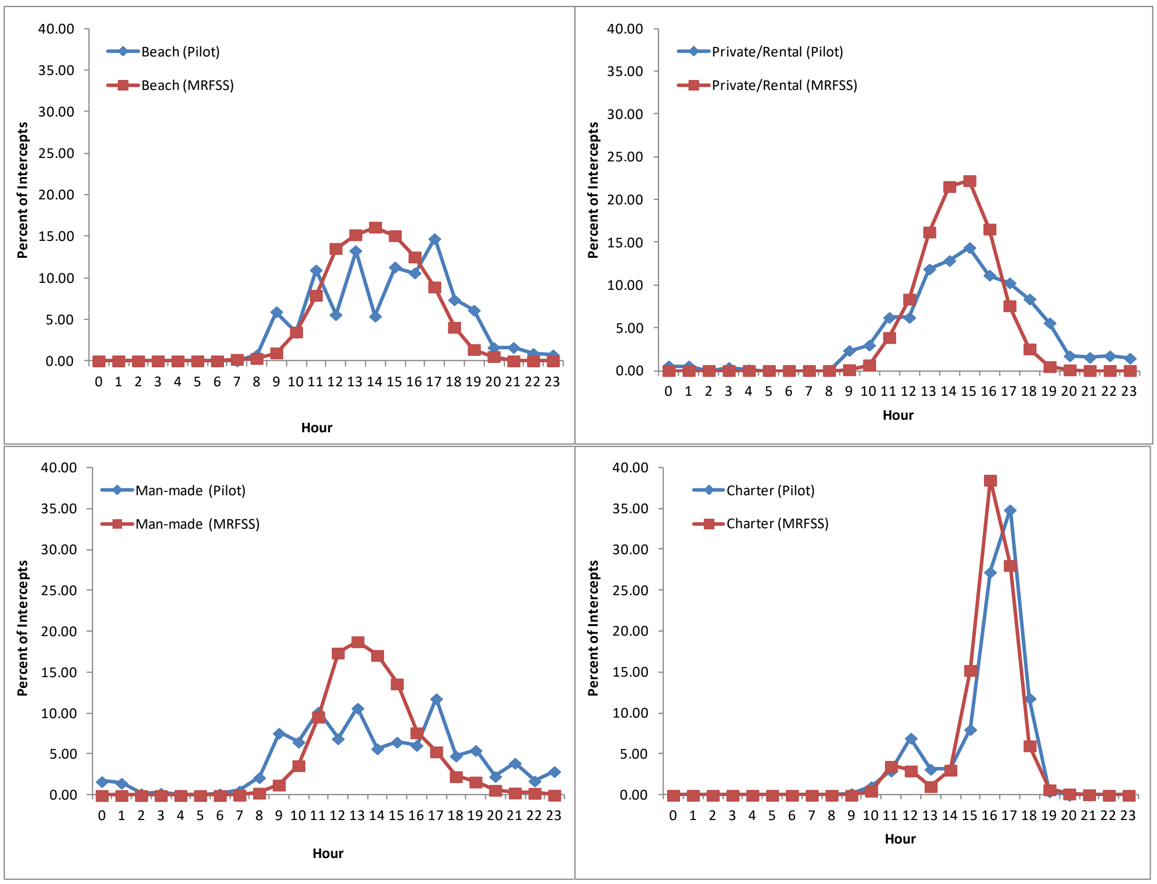
| Measure | Mode of Fishing | MRFSS | Pilot | % Difference Pilot versus MRFSS |
|---|---|---|---|---|
| Average intercepts per assignment | Beach/Bank | 7.58 | 2.48 | -67.28% |
| | Private boat | 6.98 | 3.61 | -48.28% |
| | Manmade | 11.71 | 5.97 | -49.02% |
| | Charter boat | 5.59 | 1.95 | -65.12% |
| | All Modes | 7.56 | 3.44 | -54.50% |
| Average intercepts per hour | Beach/Bank | 2.12 | 0.42 | -80.19% |
| | Private boat | 1.54 | 0.6 | -61.04% |
| | Manmade | 3.35 | 0.99 | -70.45% |
| | Charter boat | 1.69 | 0.32 | -81.07% |
| | All Modes | 1.97 | 0.57 | -71.07% |
| Average angler trip count per assignment (intercepted + missed) | Beach/Bank | 8.68 | 2.53 | -70.85% |
| | Private boat | 9.35 | 3.61 | -61.39% |
| | Manmade | 13.97 | 5.97 | -57.27% |
| | Charter boat | 8.35 | 1.95 | -76.65% |
| | All Modes | 9.71 | 3.45 | -64.47% |

**Figure 2. Average number of intercepts obtained per two-hour intervals for each mode and survey methodology.**




Within MRFSS, man-made intercepts were collected over a 17 hours time frame(7:00 am through 11:59 pm),beach/bank intercepts over 14 hours(7:00 am through 8:59 pm), and charterboat and private boat intercepts were collected over a 12-hour time frame(10:00 am through 9:59 pm and 9:00 am through 8:59 pm, respectively). The Pilot expanded intercept collection times to 24 hour coverage for man-made, beach/bank, and private boat modes. Charterboat was sampled over a 12-hour duration (8:00 am through 8:00 pm). Expansion of coverage resulted in 3.94% of man-made intercepts and 3.23% of beach/bank intercepts to be obtained outside of the time periods sampled by MRFSS. The private boat mode exhibited the greatest percentage (6.2%) of intercepts collected outside of times sampled through MRFSS. The graphs of intercepts obtained per hour through MRFSS tended to exhibit taller peaks restricted to daylight hours compared to the Pilot graphs which exhibited compressed or “shorter and wider” curves with intermittent fluctuations (Figure 3).The jagged curve for the Pilot in the shore modes (Figure 3) likely reflects times of day spent traveling from one site to another within a multi-site clusters. For example, for an 8:00 AM – 2:00PM assignment time-interval samplers would always be traveling from the first site to the second site at 10 AM and from the second site to the third site (or back to the second site) at 12 PM. Therefore, as reflected by the dips in the graphs, fewer intercepts were obtained in these hourly intervals since more time was spent traveling to the next site.

### Figure 3. Frequency of intercepts per hour obtained from MRFSS and Pilot



Eight species (or species groupings) were selected for comparing the average number of fish caught per assignment between the MRFSS and Pilot surveys (Table 6). These species (or groups) were selected because they are highly targeted by North Carolina anglers, or they are caught in large numbers, or both. Comparisons were made for both "reported" fish that were unavailable for inspection by the sampler, and for "observed" fish that were seen by the sampler. "Reported" includes a combination of released fish and landings. Comparisons were made only between positive assignments where at least 1 fish of that species was caught (i.e., zero catch assignments were not included in the analysis). The average numbers of reported Atlantic croaker, kingfishes, red drum, and spotted seatrout were greater in the Pilot compared to those reported in the MFRSS and slightly less for bluefish, dolphin, and flounder. The average numbers of fish observed were higher for bluefish, dolphin, flounder, and spotted seatrout under the MRFSS sampling design but the average numbers observed were higher for croaker, kingfish, and red drum under the new sampling design.

Table 6. Average numbers of fish reported and observed, and percent change by species and survey.

| | Average Number Reported | | | Average Number Observed | | |
|---|---|---|---|---|---|---|
| Species | MRFSS | PILOT | % Change | MRFSS | Pilot | % Change |
| Croaker | 4.67 | 5.66 | 21.20 | 4.94 | 6.63 | 34.21 |
| Bluefish | 3.71 | 3.60 | -2.96 | 5.78 | 4.19 | -27.51 |
| Dolphin | 5.09 | 4.92 | -3.34 | 18.99 | 13.46 | -29.12 |
| Kingfish Genus | 3.68 | 5.49 | 49.18 | 4.28 | 7.60 | 77.57 |
| Lefteye Flounder Genus | 2.96 | 2.82 | -4.73 | 2.16 | 2.04 | -5.56 |
| Red Drum | 2.47 | 3.40 | 37.65 | 1.33 | 1.38 | 3.76 |
| Spotted Seatrout | 6.40 | 10.45 | 63.28 | 2.67 | 2.55 | -4.49 |

## 4.2 Comparison of Pilot and (weighted) MRFSS Effort and Catch Estimates

The MRFSS access point survey data is used to estimate two important estimation parameters – the mean catch per angler trip and the proportion of angler trips made by coastal county residents with landline phones. The inverse of the latter estimated proportion is used to expand the Coastal Household Telephone Survey (CHTS) estimate of fishing effort to account for anglers that cannot be reached by the CHTS (i.e., non-coastal or no landline phone). The mean catch per angler trip for each finfish species is multiplied by the estimated total number of angler trips to get an estimate of the total catch of that species. Catch and effort estimates were compared between the Pilot and MRFSS. Appropriate weighting techniques were used to calculate both the Pilot and MRFSS estimates used for comparisons. North Carolina Pilot and MRFSS effort estimates were based on the same primary data sources: the Coastal Household Telephone Survey for private boat and shore modes, and the For-Hire Telephone Survey for charter boat mode. As a result, overall effort estimates were expected to be reasonably close to one another with differences being attributed to intercept survey coverage correction factors: i.e., out-of-state and non-coastal component adjustments and charter boat off frame adjustments. Differences in estimates of the proportion of trips by fishing area (ocean within 3 miles, ocean outside of 3 miles, and inland) would also be attributed to intercept survey data.

The 2010 total effort (angler trips)estimate was 4,852,349 for the Pilot and 5,677,574 for (weighted) MRFSS, with overlapping 95% confidence intervals. Nearly two-thirds of this difference was due to the beach/bank mode where effort estimates were 1,370,981 trips in the Pilot and 1,930,919 trips in the MRFSS. This difference was due to differences between the MRFSS and the Pilot in the percent of beach/bank mode intercepts conducted with coastal county residents (Table 7). However, the estimated proportion of beach/bank mode trips by fishing area did not differ between the Pilot and MRFSS.

**Table 7. MRFSS and Pilot percent of beach/bank mode intercepts with coastal residents by wave.**

| Mode | wave | Pilot % coastal | MRFSS % coastal |
|---|---|---|---|
| BB | 1 | 0.8455 | 0.6575 |
| BB | 2 | 0.3502 | 0.3339 |
| BB | 3 | 0.5252 | 0.3715 |
| BB | 4 | 0.5611 | 0.3614 |
| BB | 5 | 0.5317 | 0.3501 |
| BB | 6 | 0.4152 | 0.3997 |

There is some suggestion that the coastal resident proportion difference in the beach/bank mode could be linked to the elimination of incomplete trip interviews or the inclusion of nighttime sampling under the new design, but it was not possible to show a statistically significant differences in this proportion between complete and incomplete trip interviews or between nighttime and daytime trip interviews under the MRFSS design in this study.  The possibility of a length of stay bias under the MRFSS design warrants further study.

Pilot catch estimates were compared to revised (weighted) MRFSS catch estimates for 15 important management species.  Overall, no clear trends or systematic differences were found when comparing either landings estimates or released alive estimates for all modes combined; i.e. in some cases Pilot estimates were higher, in others, MRFSS estimates were higher.   With all waves and modes combined, Pilot landings estimates were higher than MRFSS for 7 out of 15 species, while Pilot released estimates were higher than MRFSS for 8 out of 15 species (Figures4&5).

Ninety-five percent confidence intervals were calculated for Pilot and MRFSS estimates to compare overlap and detect statistical significance.  Confidence intervals overlapped for 13 out of 15 landings estimates comparisons (Figures 4a, 4b, and 4c) and also for 13 out of 15 released estimates comparisons (Figures 5a, 5b, and 5c).  This suggests that, for the large majority of management species, Pilot and MRFSS annual catch estimates (with all modes and waves combined) were not statistically different from one another. For 21 out of the 30 comparisons (i.e. estimates for 15 species each compared for landings and for releases) at least one survey estimate fell within the confidence interval of the other survey's estimate.

**Figure 4a. 2010 weighted estimates of landings by survey and 95% confidence intervals.**

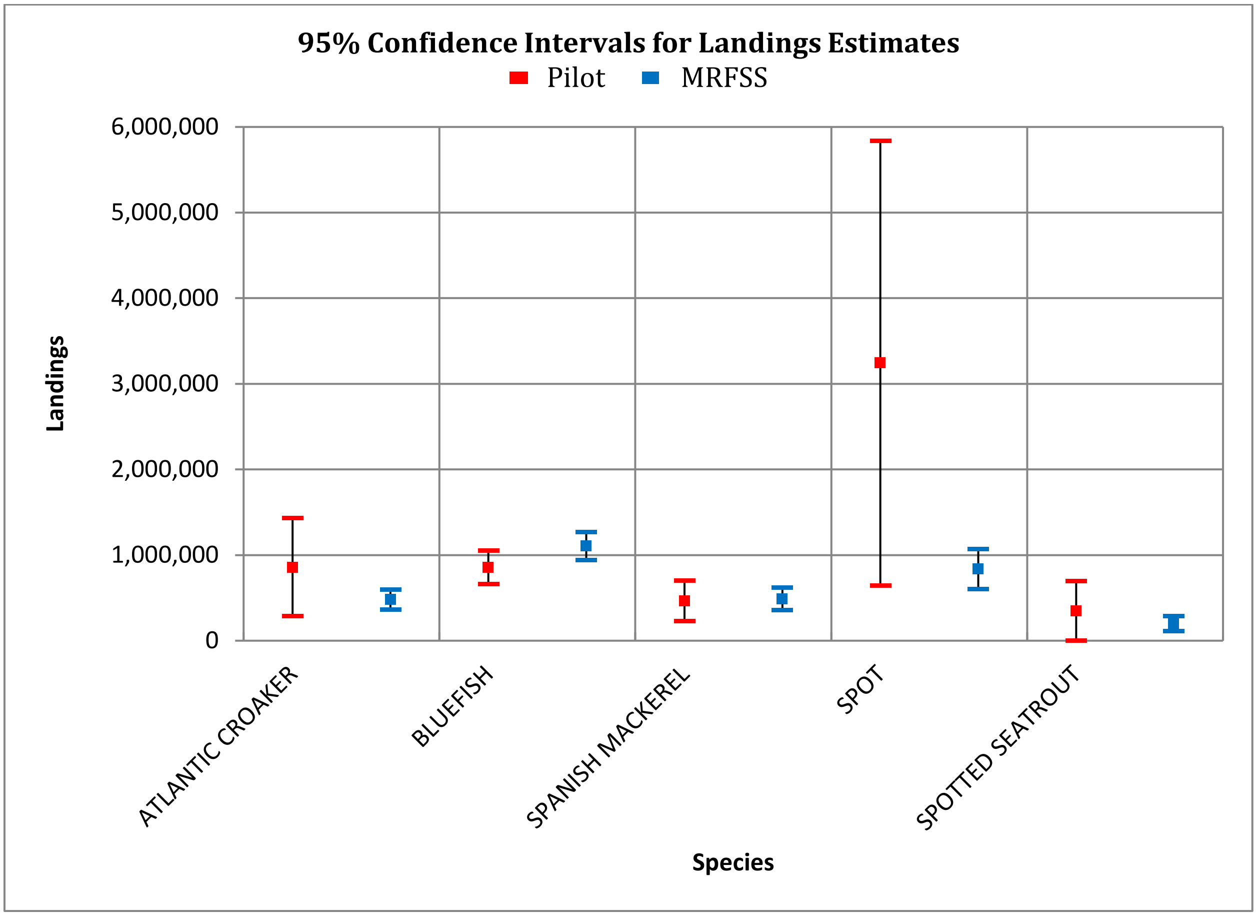


Figure 4b. 2010 weighted estimates of landings by survey and 95% confidence intervals.

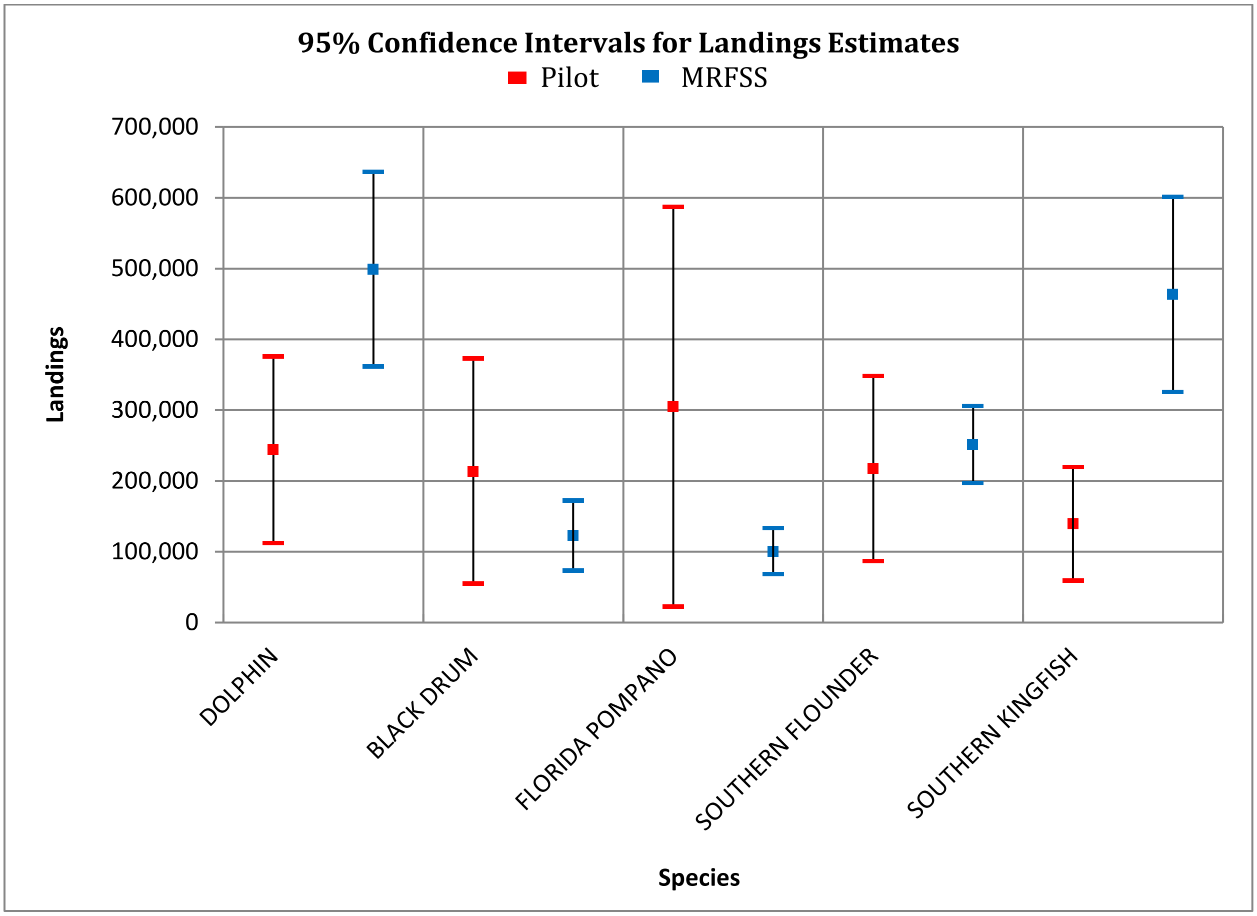


**Figure 4c. 2010 weighted estimates of landings by survey and 95% confidence intervals.**

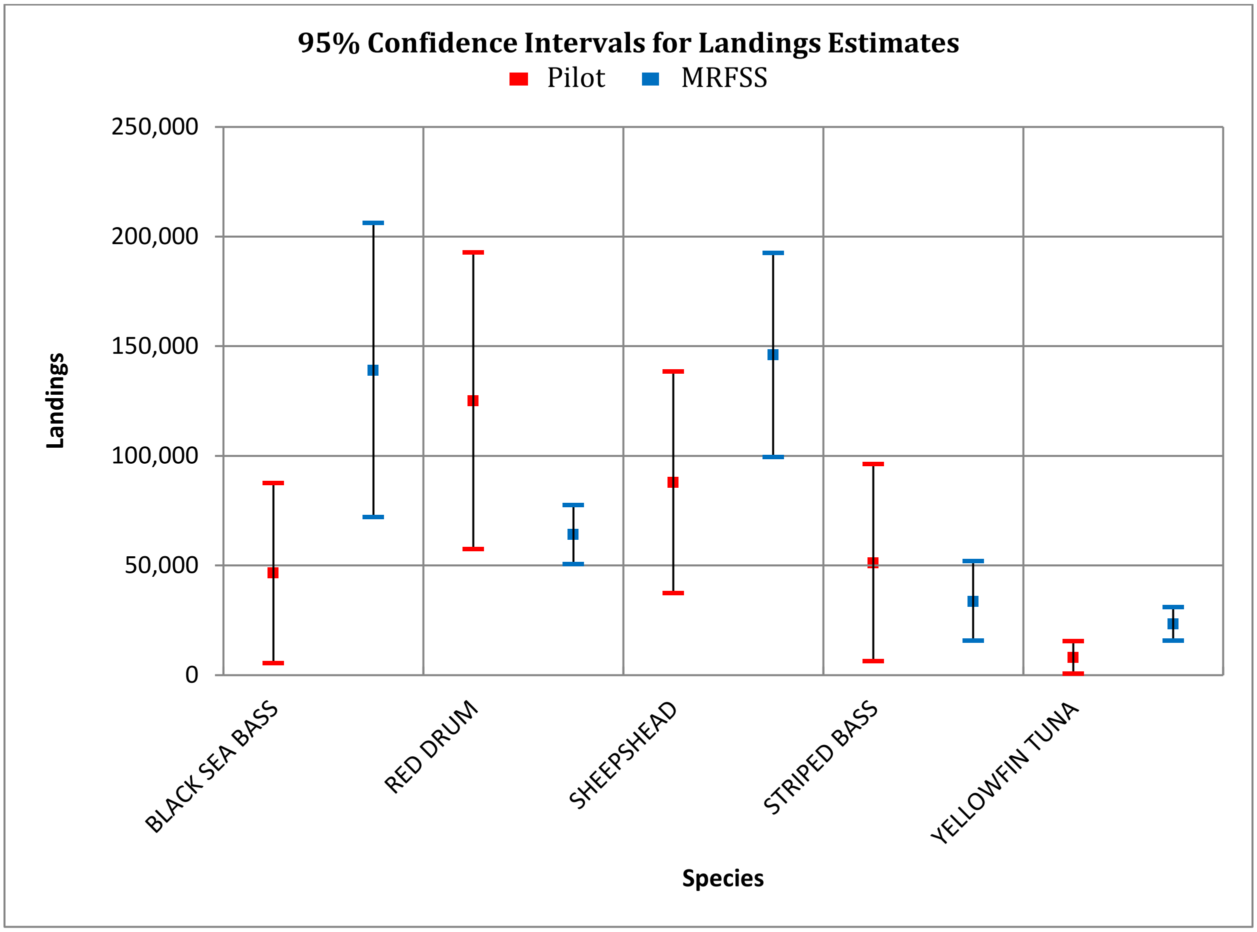


**Figure 5a. 2010 weighted estimates of fish released alive by survey and 95% confidence intervals.**

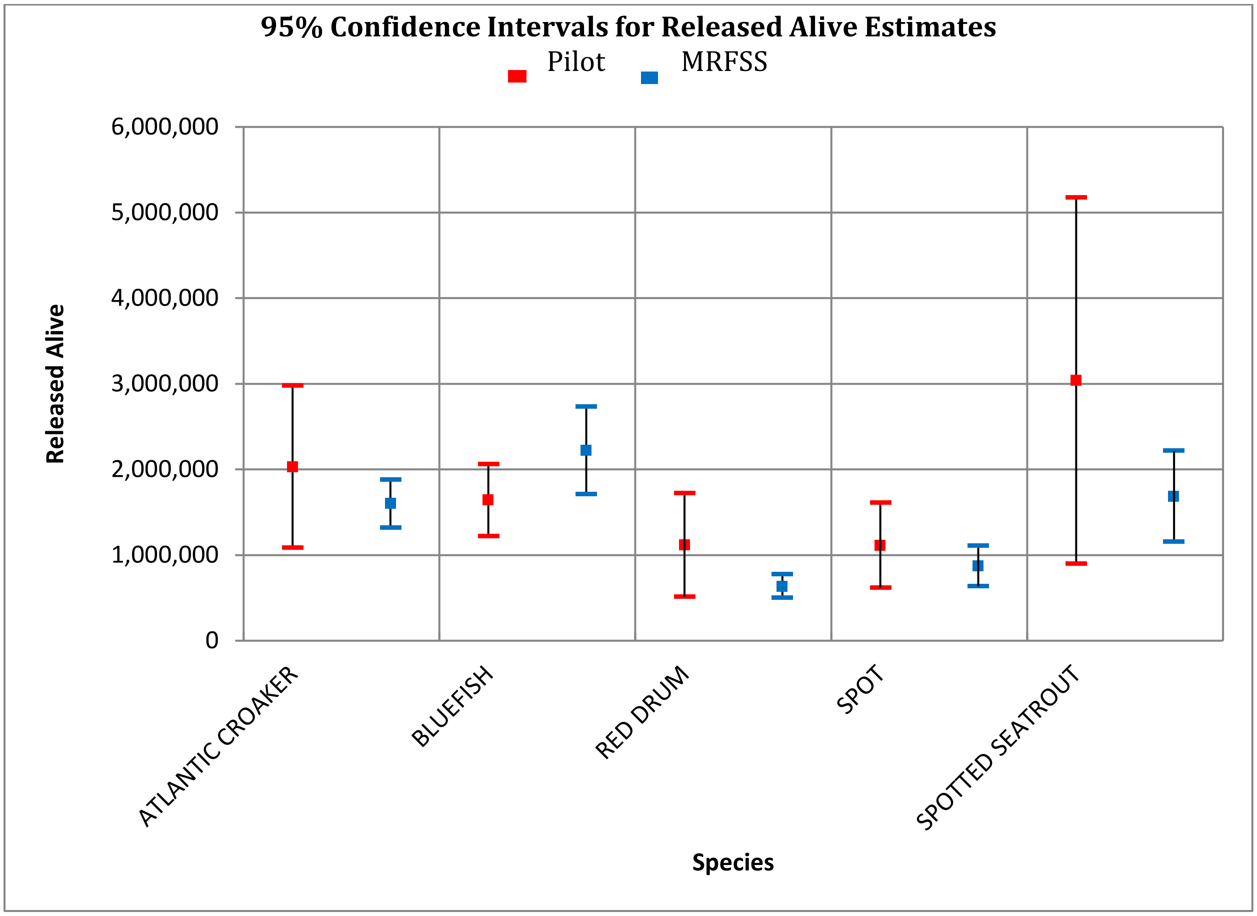


**Figure 5b. 2010 weighted estimates of fish released alive by survey and 95% confidence intervals.**

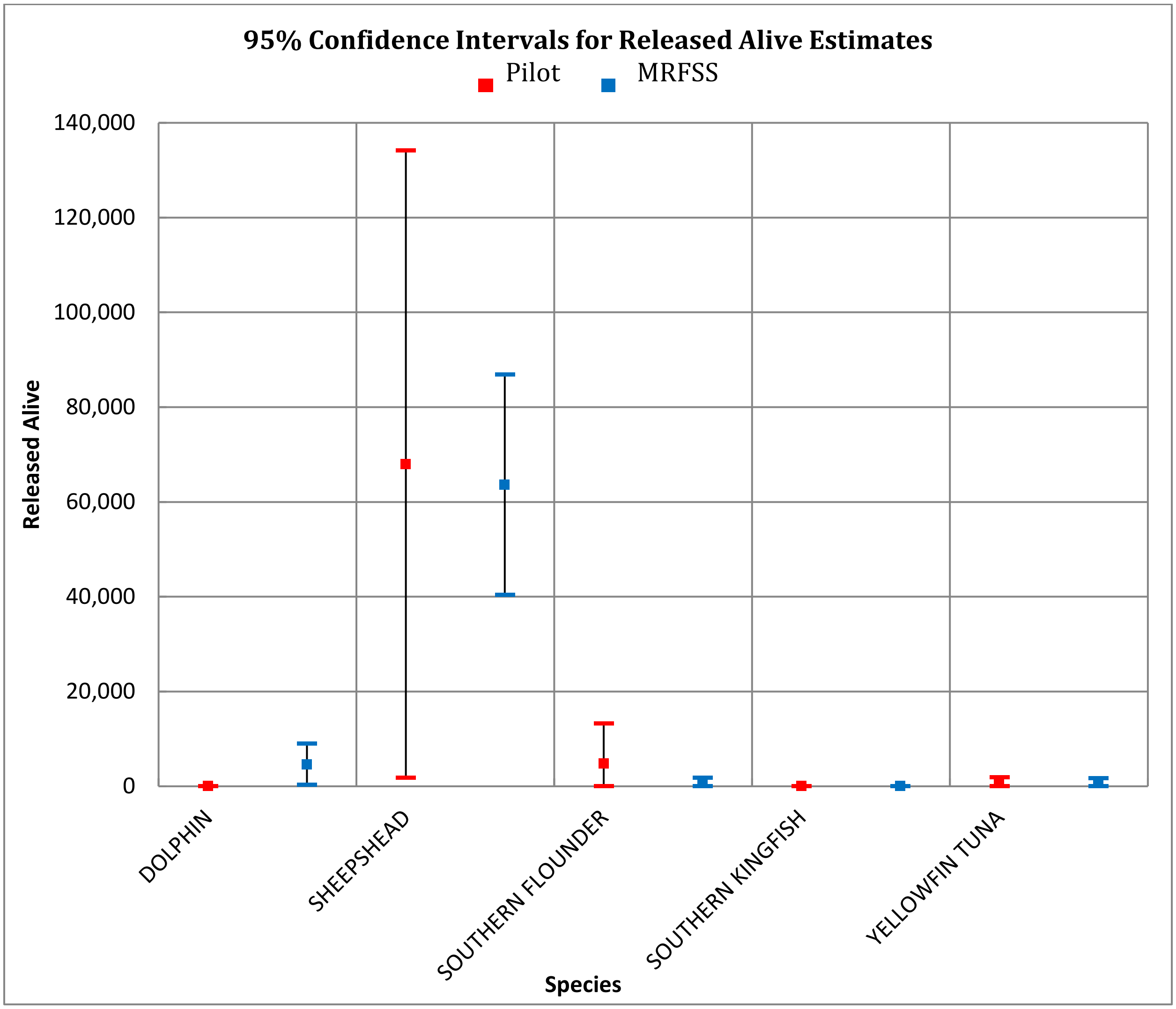
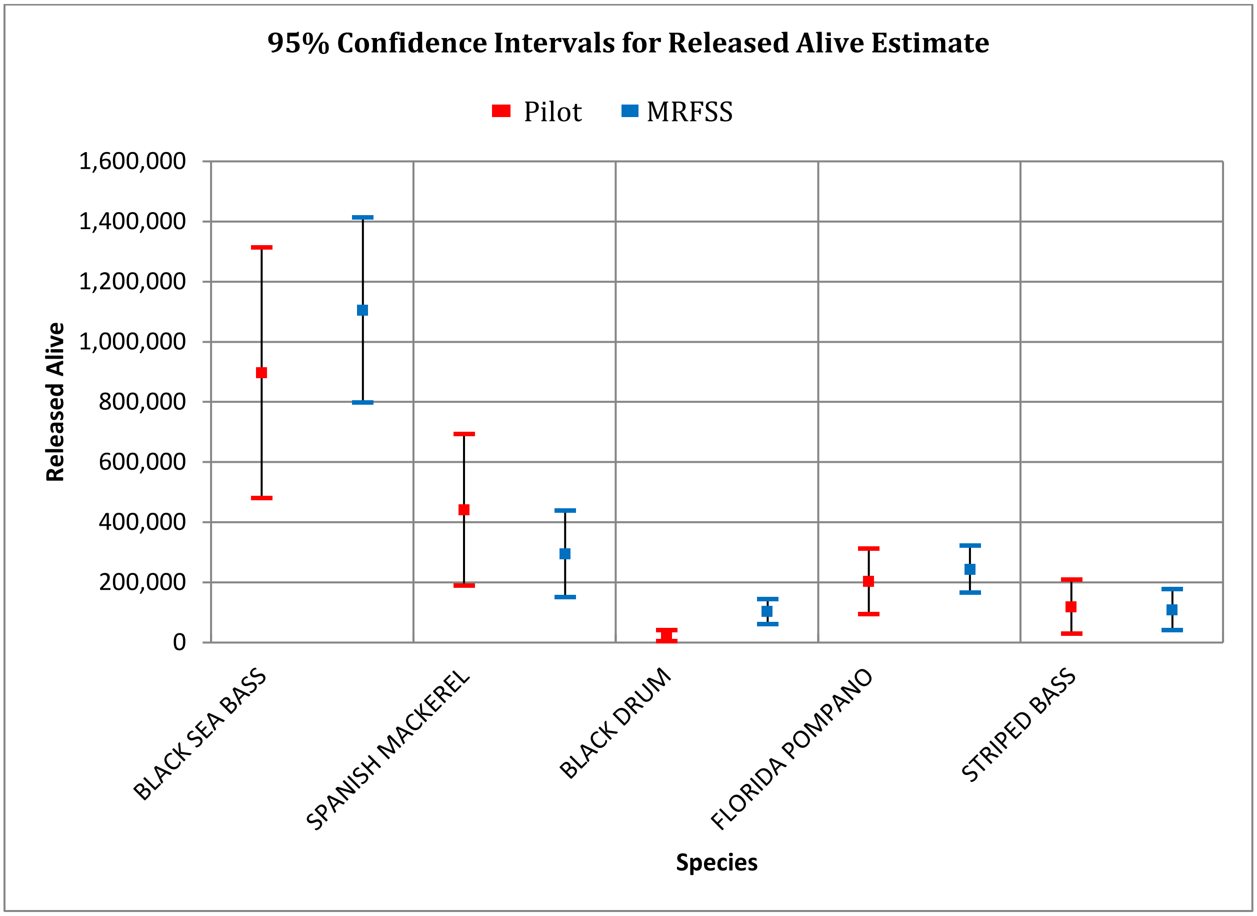


**Figure 5c. 2010 weighted estimates of fish released alive by survey and 95% confidence intervals.**



Comparisons of Pilot and MRFSS catch estimates at the mode/wave stratum level yielded similar results with $95^{th}$ percentile confidence intervals overlapping in nearly 90% of all cases for both landings and released estimates (Figure 6). The boat modes (private and charter) more frequently had non-overlapping confidence intervals compared to the shore modes. Figures 7 and 8 show the difference in landings and released estimates, expressed as pilot minus MRFSS, for wave level comparisons (with all modes combined) with non-overlapping confidence intervals. The MRFSS estimate exceeded the Pilot estimate in about 95% of all cases with non-overlapping confidence intervals. In stratum level comparisons with overlapping confidence intervals the Pilot estimate often exceeded the MRFSS estimate. Stratum level differences in catch estimates are likely due to sample size effects (i.e., small sample sizes in many Pilot stratum) rather than an identified design bias.

**Figure 6. Frequency distribution summarizing degree of overlap between NC pilot and weighted MRFSS catch estimates (landing and released) and 95% confidence intervals across all mode/wave strata for 15 important management species (see Figures 4a, 4b, and 4c for species included).**

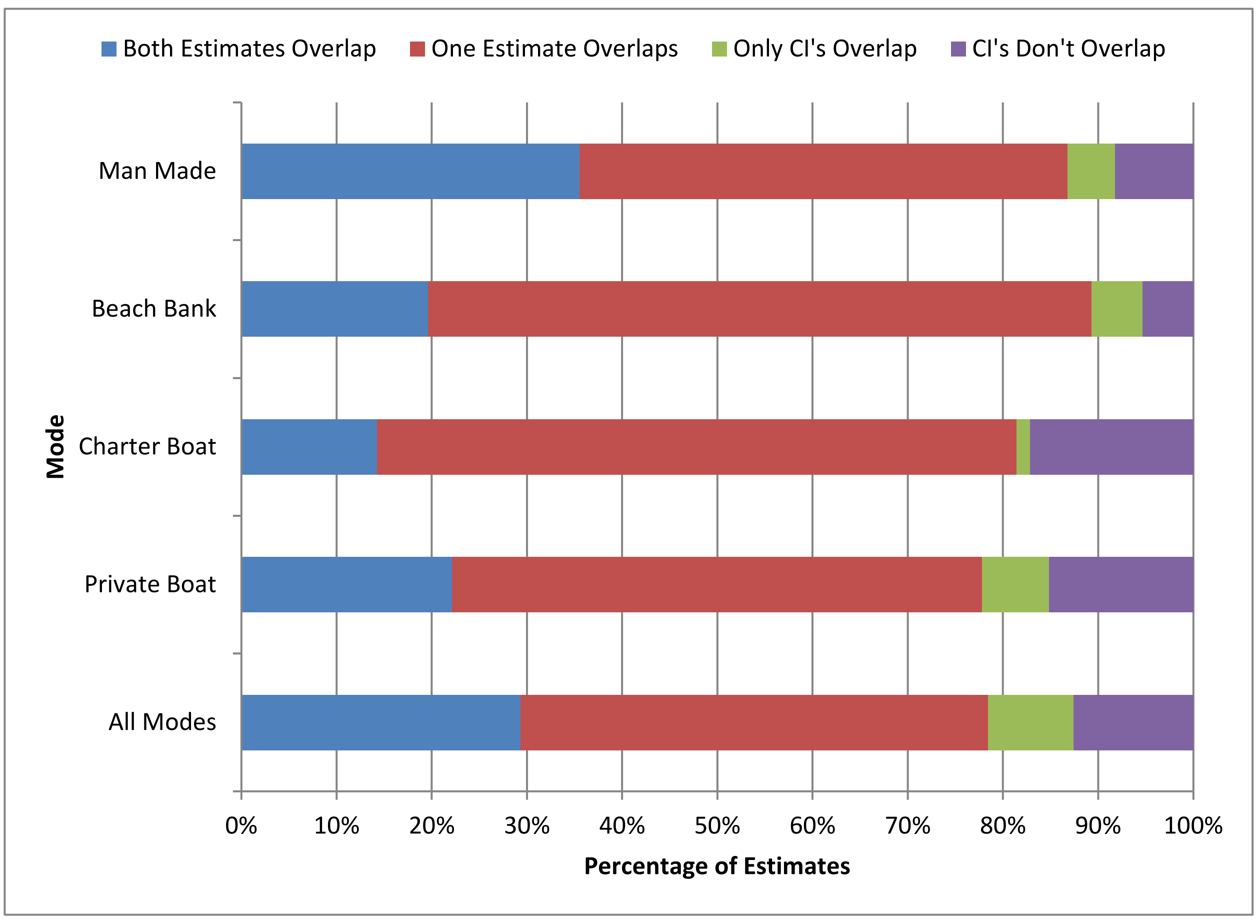


**Figure 7. Difference in 2010 recreational landings estimates, expressed as NC Pilot minus (weighted) MRFSS, for wave level comparisons (with all modes combined) with non-overlapping confidence intervals.**

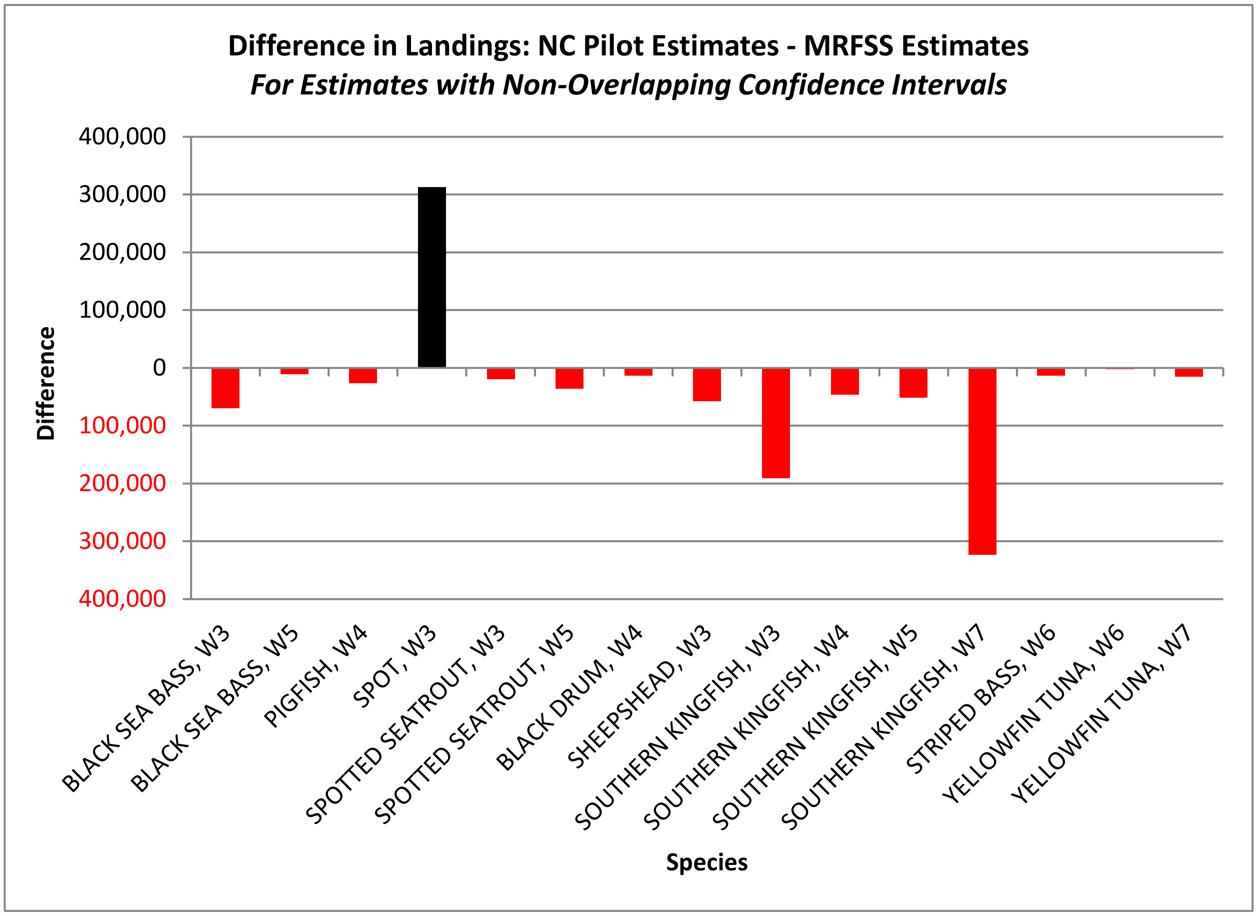
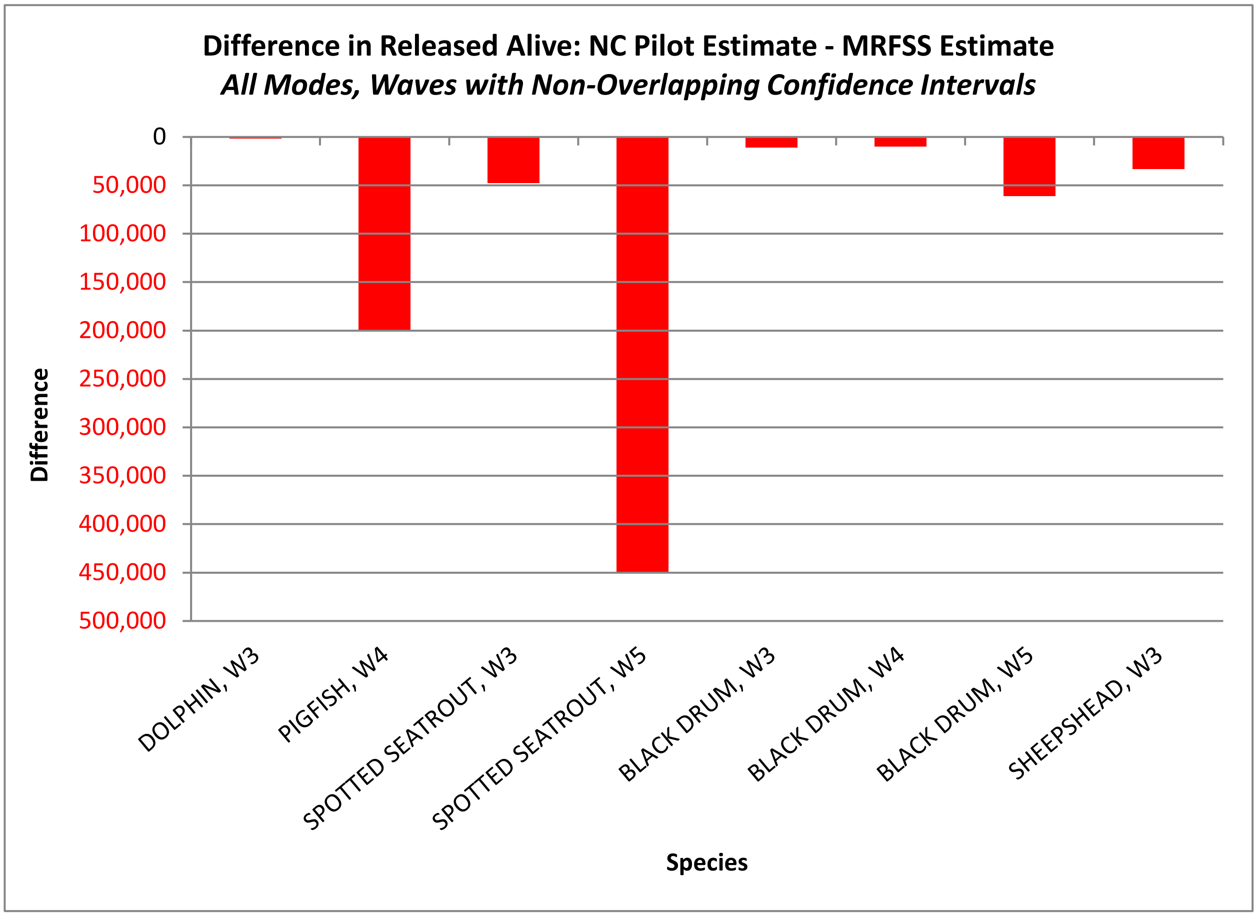


**Figure 8. Difference in 2010 recreational landings estimates, expressed as NC Pilot minus (weighted) MRFSS, for wave level comparisons (with all modes combined) with non-overlapping confidence intervals.**




While the results suggest that annual level Pilot and MRFSS point estimates across all modes were reasonably close, there were a few particular mode/wave strata level comparisons where absolute differences were rather large, regardless of whether or not confidence intervals overlapped. In some of these cases, the MRFSS estimate was considerably greater than the Pilot and in others the Pilot estimate was considerably greater than the MRFSS. Strata level catch estimates with very large differences were examined more closely. Results of this analysis are shown in Appendix E.

Proportional Standard Errors (PSEs) were consistently higher for pilot catch estimates than they were for MRFSS catch estimates due mainly to smaller sample sizes in the Pilot (Figures9 and 10).

**Figure 9. 2010 NC Pilot and (weighted) MRFSS landings Proportional Standard Errors (PSEs) with all waves and modes combined for 15 important management species.**

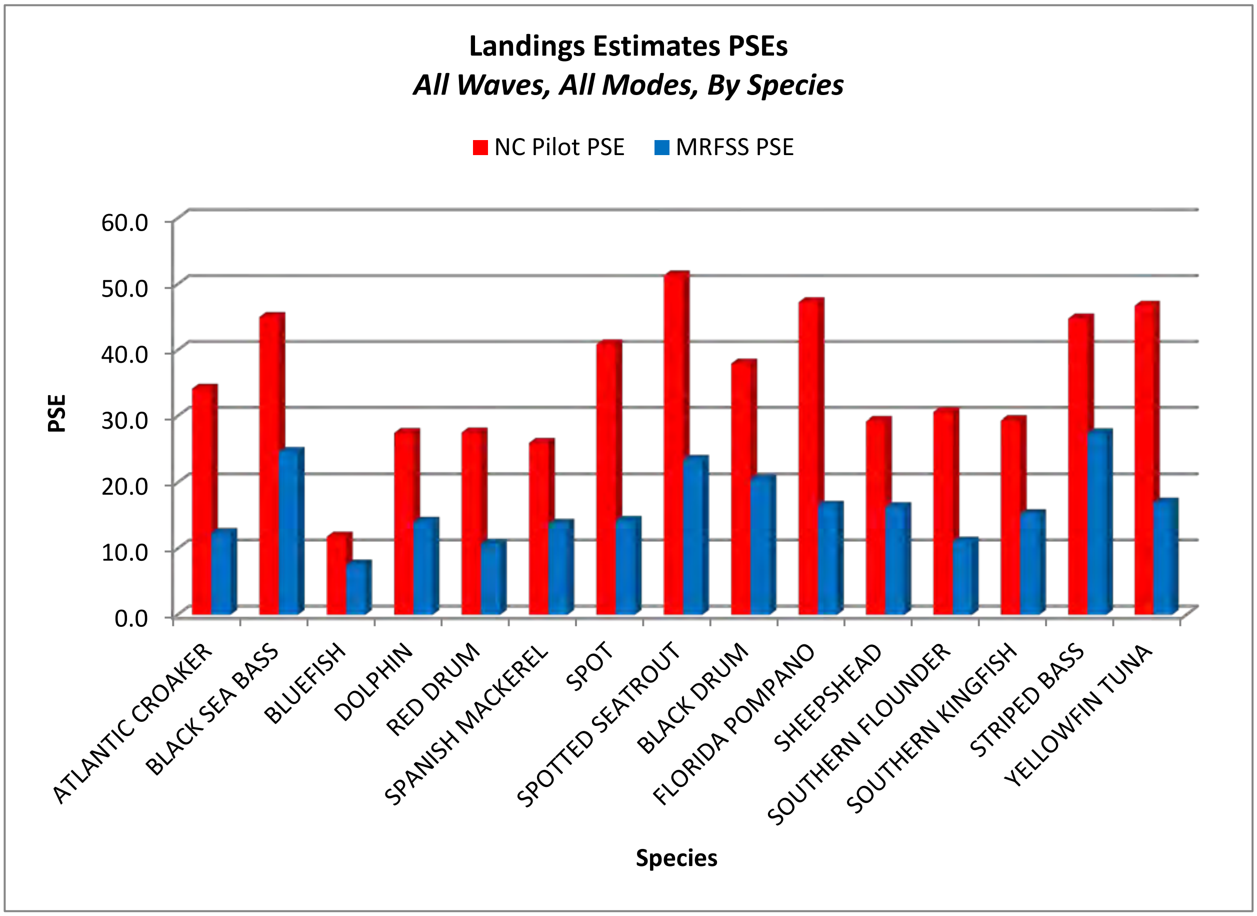
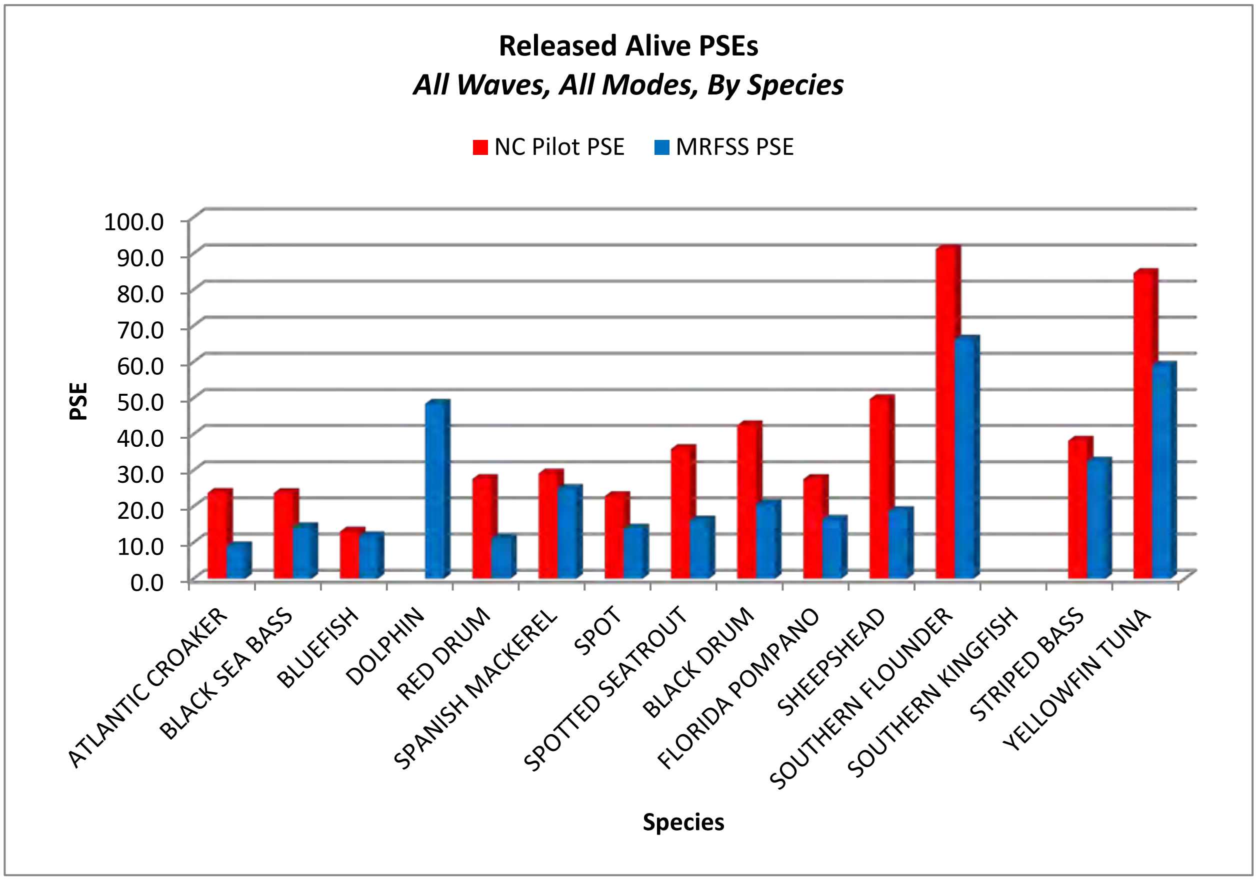


**Figure 10. 2010 NC Pilot and (weighted) MRFSS fish released alive Proportional Standard Errors (PSEs) with all waves and modes combined for 15 important management species.**




## 5. Discussion and Recommendations

This section of the report is divided into the following subsections:

1. Discussion of the differences between the MRFSS sampling design and the new Pilot sampling design as revealed in the Pilot Study results.
2. Specific recommendations for immediate implementation.
3. Recommendations for further study.

## 5.1 Discussion of Differences

**Coverage and stratification of the spatiotemporal frame**

The stratification of days into four six-hour time blocks in the Pilot design provides more representative coverage of fishing times, and, in particular, ensures a better representation in the sample of nighttime and off-peak daytime fishing trips than the MRFSS design provides. This stratification assured that angler trips ending at night, early morning or during off-peak daytime hours have a non-zero probability of being included in the sample. This eliminates possible bias in catch rate estimators that would occur if nighttime, early morning or off-peak period fishing trips differ in mean catch rates from peak period fishing trips, which are the main target of the MRFSS. The Pilot succeeded in obtaining angler intercepts in all time intervals for each mode and wave for which non-zero pressure was expected.

Furthermore, the six-hour duration for each time block stratum provided a consistent time frame for sampling that is lacking in the MRFSS design. Six-hour intervals worked well because they allowed up to two hours for samplers to travel to and from the assigned set of sites, as well as some additional time for editing of forms within an eight-hour standard work day. It was not necessary to require interviewers to regularly work overtime (more than an eight-hour day). The choice of time intervals also worked well for North Carolina. Activity peaks in the Pilot data tended to occur near the middle of the most active daytime six-hour time blocks rather than near the boundaries between them. The use of two samplers for nighttime assignments was deemed to be good idea for safety reasons, and night sampling was not problematic; no safety related issues were reported during this study.

The MRFSS design does not stratify fishing sites by subregion within a state. The stratification of sites into three geographic state subregions for the Pilot allowed for more representative coverage of different management areas and also made it easier to manage staffing of the interviewing assignments. The area north of Cape Hatteras is characterized by an assemblage of fish stocks that differs somewhat from the area south of Hatteras. In particular, two different stocks of black sea bass are identified to be separated by the Hatteras boundary. The northern area was established as a single sampling stratum for this study. The area south of Hatteras was split into two geographic strata of relatively equal stretches of coastline that could be easily covered by a staff of samplers without requiring large travel distances from a home office. There can be both statistical and management advantages to geographic stratification of sites/clusters by subregion within a state, particularly for a state like North Carolina that has both a considerable amount of coastline and regional variability in the stock composition of recreational catch. Overall precision may improve as a result of stratification if catch rates are more similar within state subregions than across state subregions. Stratification within a state can be done by dividing the site register using county boundaries (as was done for the Pilot) or well-defined geographic or natural boundaries (e.g. enclosed bay versus ocean).

**Change in definition of the primary sampling unit**

Formalization of a probability-based approach for the selection of all site assignments allows for more accurate determination of correct PSUs which facilitates the calculation of sampling weights to be used in the estimation stage. MRFSS procedures allowed samplers to leave the assigned site (PSU in the MFRSS) and visit up to two alternate sites on a given assignment. Because the Pilot design eliminated the on-site decision-making by samplers regarding the selection and sampling of alternate sites, it was now possible to calculate the correct PSU sampling weights to be included in the estimation process.

The clustering of medium and low activity sites to produce 3-site and 2-site PSUs that could be combined with high-activity 1-site PSUs maintained the ability to specify their inclusion probabilities through a formal probability sampling method, while reducing the likelihood of assignments without interviews. The sampling of predefined sites and site clusters also eliminated potential for bias in the MRFSS design that could result from samplers making unpredictable choices of alternate sites.

The Pilot design effectively eliminated sampler discretion to choose both the start time and the duration of interviewing for a given assignment. Since the temporal dimension of each PSU in the Pilot design was a specified six-hour interval, the variability among samplers in the time intervals chosen for data collection under the MRFSS design was eliminated. Under the MRFSS design, if different samplers consistently started collecting data at different times and consistently stayed on site for shorter or longer time periods than other samplers, then a spatial and temporal bias could have been introduced if catch rates varied in some consistent way with time of day and site. The potential for such a bias is eliminated with the new sampling design.

The new sampling approach allowed for more straightforward directions to be given to interviewers, thus eliminating a good deal of confusion or inconsistency regarding decisions about when and where to collect data. The pre-determined order of site visits and times for arrival and departure at each site eliminated any possible bias resulting from the variability among samplers in choices made regarding the order or duration of visits to individual sites selected in the PSU sampling approach. For the Pilot, samplers were instructed to stay a maximum of two hours on-site for all multi-site cluster assignments. For two-site clusters, this meant that samplers spent two hours at the first site, two hours at the second site, and then returned to the first site to finish out the six-hour time interval. These on-site procedural changes also assured that each site in the cluster had an opportunity to be sampled during different two-hour time blocks within a six-hour interval. If this decision were left to sampler discretion the same site may always be visited first (or last), which may introduce selection bias.

The use of ArcGIS for determining appropriate site clusters in this study is a novel approach that allows considerable flexibility in the way individual sites are sampled from wave to wave. This procedure worked very well to minimize driving time between sites, thereby maximizing the actual time period for data collection within the assigned time intervals. The accompanying computer algorithm assured that the number of sites in a PSU was determined by a cumulative measure of expected fishing pressure, resulting in less variability in the inclusion probabilities of individual PSUs. For this reason, the clustering of sites also effectively decreased the probability that any one intercepted angler trip would get an unusually high weight in the design-based estimation process.

The fixed time interval for interviewing assignments in the Pilot design also assured that angler fishing trips ending at different times within a given time block stratum would have relatively equal inclusion probabilities. MRFSS assignments had varying start times and durations that were set by decisions made by individual interviewers. The Pilot sampling design eliminates this variability and reduces the potential for bias that can result from differential sampling of time intervals when there are significant catch rate differences among angler fishing trips ending at different times.

**Sampling of interviewing locations in space and time**

In general, the clustering of lower pressure sites into multi-site PSUs in the Pilot design increased their inclusion probabilities relative to the higher pressure sites. Higher activity sites still had higher inclusion probabilities than lower activity sites in the new sampling design, but there was generally less variability among sites in their probabilities and a greater chance that the sample was spread more evenly among sites of similar pressure. Under MRFSS, sites of equal pressure could wind up having different inclusion probabilities due to differences in their proximity to other sites. If a site was located close to several lower pressure sites rather than just one or two, then it was more likely to be selected as an alternate site.

The Pilot design's elimination of "alternate site" visits made at the discretion of samplers is a very important improvement. All sites and times for sampling are fixed in the formal draw of the PSUs, and the inclusion probabilities can be easily calculated for all site clusters, sites within those clusters, and angler fishing trips encountered within selected sites and time intervals. The MRFSS design specifies when alternate sites can be visited and how they should be selected. If all samplers followed the specified procedures in the same manner, it would theoretically be possible to determine the inclusion probabilities for sites as alternate sites in the MRFSS design. This would likely require complex modeling techniques that would employ contingent probabilities and distances to neighboring sites. However, it is not clear that all samplers have interpreted and executed the prescribed MRFSS procedures in the same way. Therefore, modeling of the inclusion probabilities for sites as "alternate sites" in the MRFSS design is not straightforward. Any biases that could possibly have been introduced by interviewer errors in the execution of alternate site protocols were essentially eliminated by the new design.

The Pilot design did not allow opportunistic sampling of newly discovered sites. New sites could be identified and added to the frame for sampling in the next month or wave, but they were not included in the same month or wave that they were identified. The MRFSS sampling design allowed "new" sites to be used by samplers as possible alternate sites. The value of adding new sites opportunistically to increase coverage would be outweighed by the difficulty of determining an appropriate weight for any data that was collected at the site.

The Pilot design's emphasis on completing a certain number of assignments, rather than a certain number of angler intercepts led to a considerable reduction in the level of unobserved PSUs in any given formal sample draw. This greatly reduced the possibility of a nonresponse bias that could result from the inability to obtain observations from some of the selected PSUs (i.e., selected site-cluster-days). If observed and unobserved PSUs in the sample differ with respect to the mean catch rates of angler trips, then a high rate of non-observation in the primary sampling stage could lead to a significant bias in the catch rate estimators. Because the Pilot design places great emphasis on getting observations for all selected PSUs, it greatly reduced the potential for such non-sampling errors in the survey estimates.

In the Pilot Study, the goal of completing 100% of all the assignments that were drawn was nearly achieved. This is important for eliminating any possible bias that could result from preferentially completing some site-cluster assignments over others or from re-scheduling selected dates to match sampler requests or availability. The MRFSS design allows too much discretion in the completion of drawn site assignments and the scheduling of assignments. Consequently, many drawn assignments were either rescheduled or not completed. Changes in the pre-selected dates for some sample units and complete omissions of others could cause estimation biases. Rescheduling assignments can have unintended consequences on the sample design and could result in a distribution of assignments that is not representative of fishing activity or catch rates. Rescheduling is particularly problematic for the new estimation design because it complicates the assignment of sampling probabilities for weighting and estimation purposes. The Pilot procedure of not allowing assignments to be rescheduled removed sampler discretion in terms of which days they complete assignments and preserved the initial selection probabilities of the assignments. Whereas MRFSS assignments that are "weathered out" are rescheduled for another day, "weathered out" assignments in the Pilot were considered to be "completed" with the assumption of zero catch and effort within the cluster for that day.

The MRFSS emphasis on getting a certain target number of angler intercepts necessitates drawing many more assignments than can actually be completed with the existing staff. Therefore, many of the formally drawn assignments cannot be matched to an available interviewer. This opens the door to a possible preferential selection of some drawn PSUs over others, although the MRFSS has had strict procedures in place to try to avoid this possibility.

No PSU assignments were rescheduled in the Pilot sampling. If an assignment could not be completed on the assigned date, it was canceled. On the other hand, many of the MRFSS PSU assignments were rescheduled in accordance with specified procedures. The rescheduling could inadvertently lead to an uneven, non-random sampling of days. This could result in either under- or over-sampling of a short-term change in catch rates for any given species, especially those known to be more or less available during brief pulse events.

The Pilot sampling resulted in a higher mean number of sites visited per PSU assignment than the MRFSS sampling, and the Pilot sampling also included more unique sites at a given level of PSU sampling. The Pilot sampling of PSUs also provided a better spread of sampling across time intervals. Although this was partly due to the temporal stratification of sampling, a comparison of the distribution of PSU sampling across one-hour intervals between 2PM and 8PM, the highest activity time block in the Pilot, showed broader coverage with the Pilot than with the MRFSS sampling design.

**Sampling of angler fishing trips**

The Pilot design effectively spread the sampling of angler trips to appropriately represent a larger temporal slice of fishing. Under the new design, samplers did not have to worry about reaching their limit too quickly. Unlike the MRFSS, the Pilot did not set an upper limit on the number of interviews allowed per assignment, instead using fixed interview time intervals. Removing the intercept limit significantly reduced any potential bias associated with sampler discretion in selection of boats (for PR and CH mode) and anglers. Under the MRFSS, samplers have been instructed to randomly select boats for sampling, and to randomly select anglers within a group, if time did not allow for interviewing all anglers. The Pilot sampler training was more straight-forward as samplers were instructed to attempt to intercept <u>all</u> eligible anglers from all boats rather than attempt to sub-sample them.

Obtaining accurate counts of completed angler trips that were missed (i.e. not intercepted) was critical to this project. These counts are incorporated into the total fishing effort for individual sites, which, under the new MRIP estimation methodology, are used to appropriately weight samples. Although MRFSS samplers have always tallied "missed eligibles" on the Assignment Summary Form, until recently this information was not used in estimation. As a result, significantly less attention had been paid to sampler procedures for counting angler trips in the past.

The greater emphasis in the Pilot to obtain accurate counts of all completed angler fishing trips while on site was very important to assure greater accuracy in the calculation of the secondary stage sampling fractions that are needed to properly weight any obtained interviews in the estimation process. The categorization of possible missed angler trips as either "confirmed" or "unconfirmed" provided a means of evaluating the relative reliability of the observed counts. In general, a very high proportion of the counted missed trips were confirmed to be recreational angler trips in the specific fishing mode of the interviewing assignment. Unconfirmed counts were more commonly recorded at high activity sites, suggesting that it is harder to get accurate counts at such sites.

Although two samplers were assigned to high activity sites in the first few waves of sampling, this was not deemed necessary in later waves. The idea was that one sampler would conduct interviews while the other was obtaining counts, and that they might alternate between counting and conducting interviews during the assignment. However, individual samplers found that they were able to get relatively accurate counts on their own even at the high activity sites. A comparison of the counts obtained in the Pilot and MRFSS sampling designs for sites in the highest pressure categories showed that the Pilot counts tended to be lower.

In the Pilot sampling design, the intercepted angler trips represented a much larger proportion of the total count of completed angler trips in the sampled time interval (6 hrs rather than 24 hrs). This meant that there was much less need to expand observed counts to estimate the total count for a sampled time period. In the MRFSS, the actual sampled time interval is a 24-hour day, but the observed counts and interviews were obtained in a much shorter time frame that could range anywhere from 2 to 8 hours. Because the observed counts in the MRFSS sampling design had to be expanded through an MRIP modeling procedure to estimate total counts for 24 hours, there was much more room for error in estimating those total counts. In the Pilot, only a minor expansion of observed counts was required to get an accurate count for the shorter time interval of 6 hours. The Pilot design sampling succeeded in getting observations from a higher percentage of the angler trips occurring within sampled PSUs. By staying on site longer, samplers executing Pilot design assignments were able to intercept a higher proportion of the trips ending during the temporal frame of the PSU. In addition, they were able to get a more representative sample because the intercepts were better distributed across the PSU time frame. MRFSS design sampling often resulted in interviewing assignments that lasted less than 6 hours, and some assignments lasted as little as 2 hours. This result is due to two factors: (1) MRFSS samplers were able to target the most active time of day at the assigned site and (2) MRFSS samplers were held to a cap of no more than 30 angler trip interviews per site within a PSU.

**Comparing estimates of catch rates**

As a result of implementing a more rigid probability sampling approach in the Pilot Study, it was possible to use available data to directly calculate representative weighting of the angler trips that were included in the survey sample without relying heavily on modeling. The inclusion probabilities for all intercepted angler trips were calculated with a design-based approach. We were able to easily calculate the sampling probabilities needed to weight the data in the estimation process, and those probabilities were less prone to possible errors than probabilities estimated through MRIP modeling procedures for the MRFSS sampling design.

## Comparing estimates of fishing effort ratios

The estimates of the proportion of fishing trips made by marine recreational anglers who could be contacted by the Coastal Household Telephone Survey of angler fishing effort were mostly similar in the two intercept surveys compared in this study. The inverse of this estimated proportion was used to adjust CHTS effort estimates to account for fishing trips made by anglers who could not be covered by CHTS sampling. Although there was some evidence that use of the Pilot sampling design resulted in an increase in this estimated proportion for the beach/bank shore mode, this study suggests that it is unlikely that the new sampling design will have significant impacts on the overall estimated APAIS effort adjustments.

## Comparing estimates of total catch

Differences in estimates of total catch by species were largely driven by differences in the estimates of mean catch per angler trip. For the large majority of management species, Pilot and MRFSS annual catch estimates (with all modes and fishing areas combined) were similar to one another. Pilot and MRFSS catch estimate confidence intervals overlapped for 13 out of 15 landings estimates comparisons and similarly for 13 out of 15 released estimates comparisons. More pronounced differences were noticed for some species as you drill down to the mode/wave/area level of estimation. In general, we expect that catch estimates based on the new Pilot design will be similar to those produced from the MRFSS design for most species. Differences observed in this study would likely have been greatly reduced if the Pilot design sampling had been conducted at the same level as the MRFSS design sampling.

For some species that are common targets for anglers ending their fishing trips during nighttime or off-peak daytime intervals, we would expect that the Pilot design estimates would be higher than the MRFSS design estimates. This may also be true for species associated with fishing tournaments because selected sites with fishing tournaments in progress (tournament weigh station sites) were not excluded under the Pilot design as they have been under the MRFSS design.

In this study, there was a suggestion that the Pilot design sampling yielded higher catch rate estimates for common night fishing targets like striped bass and red drum. On the other hand, Pilot design catch rate estimates for many of the other species tended to be somewhat lower. Although these differences were not statistically significant, their directions match what you should expect to see with the addition of nighttime and off-peak daytime sampling.

**Sample size and precision ("productivity" measures)**

In this study, the estimates generated from the MRFSS sampling design were more precise than the estimates generated from the Pilot design largely because more samplers were available to cover a greater number of sampling assignments in the MRFSS design particularly during the most active two-month periods (Waves 3-5). The number of assignments completed was consequently greater for the MRFSS sampling in those sampling waves. If the number of PSUs observed in the Pilot design had been increased to match the number of assignments completed in the MRFSS design, the variances of the Pilot mean catch rate estimates would have been lower.

The Pilot design assignments observed significantly lower mean numbers of angler trips than the MRFSS design assignments across all four fishing mode strata. Although Pilot design assignments also observed significantly lower mean numbers of caught fish weighed and measured, the Pilot design and MRFSS design assignments had similar average numbers of fish observed per angler trip. This suggests that the main difference in numbers of fish observed between the two designs was due to a difference between designs in the probability of intercepting angler trips. A larger percentage of the Pilot assignments failed to get any angler trip interviews compared to the MRFSS assignments. If both designs had completed the same number of assignments, the MRFSS design would still likely have provided estimates with greater precision due to both lower percentages of sampled PSUs without angler trip intercepts and higher mean numbers of intercepts per PSU.

The differences in the proportion of assignments with angler intercepts and the mean number of intercepted trips per assignment were greatest in the sampling for the beach/bank shore mode. This was largely because the Pilot design did not allow intercepts of incomplete angler fishing trips as has been allowed under the MRFSS design for this fishing mode. Changing the rules to eliminate "incomplete interviews" was considered to be important for eliminating the potential "length of stay" bias that results because anglers who fish longer have a greater chance of being intercepted for such interviews than those who fish for a shorter period of time. In order to be interviewed under the Pilot design, the angler must have completed their day of fishing.

This lower productivity of the Pilot design as it was implemented for this feasibility study was driven by a number of factors that could be changed in future implementation while still adhering to a strict probability sampling design. By design, MRFSS samplers visited sites much more consistently during their most active periods of fishing activity. The time-block stratification of the Pilot design sampling assured better coverage of fishing trips ending throughout a 24-hour fishing day, but the inclusion of numerous assignments directed at non-peak periods of fishing activity also resulted in both an increase in the percentage of empty assignments (i.e. no intercepts) and a decrease in the average number of angler intercepts per assignment.

Comparison of the mean number of intercepts per assignment between the MRFSS and Pilot designs for the most active 2PM-8PM interval showed a much closer match, but the MRFSS assignments still achieved slightly higher levels of non-empty assignments and mean numbers of intercepts. This can be explained at least in part by the fact that the MRFSS sampling assignments visited sites in the highest pressure categories more frequently than the 2PM-8PM Pilot design sampling assignments. This happened mostly because MRFSS samplers visited higher pressure sites more frequently than lower pressure sites as alternate sites.

## 5.2 Recommendations for Immediate Action

1. **In general, the Project Team recommends use of the new access point survey sampling design tested in this pilot study for conducting future access point surveys on the Atlantic coast and in the Gulf of Mexico.** However, we also recommend some additional changes, not implemented during the Pilot, that we have outlined in this section. The recommendations below can and should be addressed prior to implementation of the new sampling design along the Atlantic coast and Gulf of Mexico. Most of these recommendations are focused on further improving the new sampling design to increase statistical precision without increasing costs.

2. **The allocation of sampling among sampling strata should be changed as needed to maximize sampling efficiency and statistical precision.** Sampling could be allocated very differently among geographic strata, fishing mode strata, and time block strata than how it was allocated in this pilot study. Without introducing any bias, other sampling allocations will likely provide higher proportions of sampling assignments that obtain at least one interview and may also provide higher average numbers of interviews per positive assignment than were observed in the pilot study. The goal should be to find the “optimal” allocation that will provide the highest level of statistical precision for the dollar spent.

Sampling could be allocated differently among geographic strata. In this study, the sampling for the Pilot design was distributed more evenly among the three North Carolina subregions than may be desired for future implementation. By contrast, more than 60% of the MRFSS assignments were conducted in the Northern subregion, where the majority of high pressure sites are located. The distribution of Pilot design sampling could be shifted to allocate a greater proportion of it to the Northern subregion.

Sampling could also be allocated differently among the different fishing mode strata. In this study, the Pilot design sampling was spread pretty evenly among the different modes, but the MRFSS design sampling was allocated to achieve proportionately higher levels of sampling in the private boat and charter boat modes. In general, sampling in the boat modes tends to be more productive than in the shore modes. In addition, more of the key management species are caught primarily in the boat modes. Therefore, efficiency may be improved by allocating a higher proportion of the total sampling to the boat modes when implementing the new design.

Sampling could be allocated differently among the different time blocks of the Pilot design. In this study, sampling was deliberately spread across the time blocks to test the feasibility of sampling at nighttime and off-peak daytime intervals. For future implementation, the proportions of sample allocated to the nighttime and off-peak daytime blocks should probably be reduced to achieve higher levels of productivity (efficiency). As long as some sampling is allocated to all non-peak time blocks, the Pilot design will be less susceptible to possible undercoverage bias than the MRFSS design.

3. **The formal PPS sampling of sites and site clusters should be controlled to ensure all drawn assignments can be completed by existing staff.** Following the pilot study, the project team developed a “controlled selection” program for possible use in selecting PSU samples for future intercept surveys. This program is briefly described in Appendix F. It is important to clarify that the use of a controlled selection program does not imply that sampling levels would be dictated by staffing levels. Staffing levels for the access point surveys should always be set to match the sampling levels required to deliver desired levels of statistical precision on resulting estimates of mean catch per trip. Once those staffing levels are established, a controlled selection program can be used to ensure the draw of a probability sample of PSUs that can be covered by the existing staff. If staffing constraints are taken

into account, then the number of assignments drawn for any given day will not exceed the number of samplers available to work that day. Constraints on the number of assignments possible in a given day and on the possible stacking of assignments back-to-back should be built into the sample draw program such that it is possible to match all selected PSUs with an available sampler. The universe of PSU samples that can be covered by existing staff should be identified and randomly sorted prior to random selection of one of those samples. The expectation would be that all drawn site-day assignments would be completed, and none would go unobserved. This would essentially eliminate the possibility of an unobserved sample, or nonresponse, bias. With this approach the probabilities of selection and joint probabilities of selection needed for estimation purposes would also be relatively easy to calculate.

One particular constraint that should be added would be to prevent the draw of more than one assignment for the same cluster, day, and time interval, even if they are in different modes. This would be important to prevent having two samplers at the same location at the same time, which could create a perception of overall survey inefficiency. This was handled in the Pilot study by canceling some assignments to avoid such overlaps, but it would be handled better by adding a constraint to the draw program.

4. **Provide clearer instructions to samplers about how to handle the catch of charter boat captains and crew.** The MRFSS Statement of Work contains the following language regarding interviewing for-hire captains and crew: “The captain and deckhands should not be interviewed, regardless of whether or not they caught any fish during the trip…. They are not considered "recreational anglers" even though they might have fished.” Based on anecdotal information, interpretation of this procedure has been inconsistent across states and individual samplers in the MRFSS. While captain and crew should not be interviewed and are not counted as “contributors” for grouped catches, it was less clear whether or not their catch should be added to the catch of paying passengers. Excluding these fish represents a gap in the landings data whereby catch by captain and crew are not accounted for in any survey. In the Pilot design, samplers were instructed to include any catch by the captain and crew that were mixed in with the observed catch (Type A catch) recorded for a group of charter boat anglers, but they were also instructed to not count the captain and crew as contributors to the mixed group catch. This procedure should be consistently followed when recording catch at the level of the

boat trip in the future implementation of the new design. For regulatory purposes, captains may count themselves and their mates as “anglers” even if they did not fish or catch fish so the boat can keep more fish if there is a per angler bag limit. However, for survey purposes, as long as these trips are consistently not counted as “recreational” in both the intercept and effort (phone) surveys, a bias should not be introduced by including fish caught by for-hire captains and crew in group catches.

5. **Collect total catch data for any intercepted angler who just completed a multi-day fishing trip.** In the pilot study, sampling under both the MRFSS and Pilot designs collected catch data for only the last day of a multi-day angler fishing trip. Angler fishing trips that span more than a single day are often referred to as over-night trips or multi-days trips. While relatively rare compared to day trips, it is still important that data from such trips are recorded consistently by samplers in a manner that will not bias catch rates or other data analyses. While there are several ways a “trip” can be defined, the project team recognized that for purposes of catch estimation this definition should ideally be consistent between the intercept survey which produces catch per trip rates and the effort (phone) survey which produces estimates of numbers of trips. Under the current MRFSS “trip” is defined as fishing during part or all of one waking day (as opposed to a calendar day) in one mode. The Coastal Household Telephone Survey asks respondents to recall the number of days fished (not number of trips) in the past two months. Using trip profile information (i.e., mode(s) fished, specific dates, and return times) it is then possible to determine the number of "trips" for estimation purposes to match the intercept survey definition. MRFSS intercept samplers are instructed to only record catch for the most recent waking day fished. Although the two survey components are consistent, under the current MRFSS intercept procedure there is no way to verify whether the catch recorded was from only the most recent waking day. In practice, anglers returning from a multi-day trip may have trouble remembering which specific fish were caught on which particular days. In addition, the most recent waking day’s catch may not be reflective of the trip as a whole since a considerable amount of time is spent in travelling back from the fishing grounds on the last day and not actively fishing.

   The NC pilot followed the same protocol as the MRFSS regarding treatment of multi-day trips. However, the project team recommends adding the following question to future Intercept forms to indicate how many fishing days the Type 3 catch represents:

*26.b. Were these fish all caught today (that is, from the time you woke up to the time you ended your fishing trip) while fishing from __________(insert mode)?

1 ☐ Yes 8 ☐ N/A

2 ☐ No

*26.c. If No, how many waking days did you spend catching these fish?

☐☐ No. of days 88 ☐ Not Applicable

This question only applies to the Type 3 (Available) portion of the catch and samplers were still instructed to obtain Type 2 (Unavailable) catch information only for the most recent waking day of fishing. Since overnight trips are possible from all modes (not just boat modes) and it is preferable to keep procedures as consistent as possible for the samplers, the team decided this additional question should be asked for all fishing modes. This additional question makes it possible to calculate an average catch per day to represent the catch for the intercepted angler's day of fishing.

6. **To increase on-site productivity and reduce driving time, instruct samplers to stay up to 3 hours (rather than only two hours) at the first site when a two-site cluster is assigned.** This may be particularly advantageous in situations where driving time between two clustered sites is long. For the Pilot Study, the project team considered increasing the maximum time spent at each site for two-site clusters (e.g. 3 hours per site) but ultimately decided to keep the two-hour limit. This decision was based on the rationale that samplers would have an easier time remembering how long to stay if the duration per site was consistent across three-site and two-site assignments. The change to three hours for the first site would make more efficient use of the on-site sampler time for the purpose of data collection.

## 5.3 Recommendations for Future Consideration

In additional to the recommendations above for immediate implementation with the new design, the project team also identified several recommendations that require additional study and evaluation. These are not presented in any specific order of priority.

1. **Consider requiring samplers to obtain counts of all boat trips on which anglers have finished fishing for the day.** The current estimation procedure develops weights within each observed site-day or site-cluster-day that are based only on the sampled fraction of the total number of angler trips counted. Given that boat angler trips are actually clustered together within different boat trips, it may be better to obtain total boat trip counts and assign counted angler trips to specific boat trips. This would allow determination of appropriate sampling fractions at both the secondary (boat level) and tertiary (angler level) stages of the multi-stage sampling design. Each boat trip represents a cluster of angler trips that fished similar locations and time periods with similar fishing gears and methods. Because these angler trips are likely to be more similar to each other than to angler trips made on other fishing boats returning to the same site within the same sampled time period, the sample inclusion probability for each boat trip could be determined and taken into account in the estimation process. The Pilot study did not obtain counts of returning boats, but a method for obtaining boat trip counts could be developed and used in future implementation of improved access point surveys of private boat or charter boat fishing. Similar to angler counts, boats counts could be divided into "confirmed" and "unconfirmed" depending on whether or not the sampler was able to screen someone on the boat regarding fishing activity.

2. **Consider collecting catch data at the boat trip level rather than at the angler trip level for the boat modes of fishing.** This would eliminate a stage of sampling, thereby reducing both sampling error and the potential for sampler errors (i.e., non-sampling errors) in the selection of boat anglers for interviews. This change would also require the development of new on-site sampling protocols. Samplers would have to conduct interviews that would obtain data on the total catch of all anglers who fished on the boat trip, as well as the location, duration, and primary fishing target of the boat fishing trip. They would also have to obtain counts of the total number of anglers who fished on the boat, as well as total counts of their observed (Type A) and unobserved (Type B) catches. It may still be necessary to interview a random sample of the anglers who fished on the boat to collect data needed to determine their potential for being contacted by an off-site telephone or mail survey of fishing effort. However, mean angler catch rates could simply be calculated by taking the total catch for the boat trip and dividing by the total count of anglers who fished.

3. **Consider including for-hire "guide boats" in the private/rental boat mode instead of the charter boat mode.** For-hire "guide boats" may have more in common with private boats than with charter boats. Guide boats tend to be smaller, more transient, use multiple access points and boat ramps, and have less predictable trip schedules compared to charter boats. They may also target species that are more likely to be targeted by private boats than by charters. As a result, guide boats may also be more likely to be intercepted at sites with private boat activity than at charter boat sites in many areas. Adding guide boats to the private boat stratum may address an undercoverage issue associated with these trips and may increase sampling efficiency by eliminating very low pressure sites guide boat sites.

4. **Evaluate options for combining boat mode trips (private/rental, guide boats, and charter boats) into a single stratum.** Sites with boat mode fishing activity often include a combination of private boats and for-hire boats. Combining these modes into a single stratum could result in more efficient sampling and fewer assignments resulting in zero intercepts obtained. If needed for management purposes, separate catch estimates could still be calculated for private boat and for-hire sectors by treating these as "domains" within the boat mode stratum.

5. **Consider implementing more rigorous protocols to ensure random sampling of observed fish for weight and length measurements.** In the pilot study, samplers selected fish for measurements in the same manner under both the Pilot and MRFSS sampling designs. However, the project team discussed ways to improve the MRFSS sub-sampling fish procedures and developed a more rigorous random sampling protocol that would be feasible for field implementation. This new procedure is described in Appendix G. We recommend testing of this protocol.

6. **Consider basing rules for clustering sites more strictly on how geographic strata are defined.** In the Pilot design, sites were only clustered together if they were within the same county. In the future it would be more appropriate to cluster sites across county boundaries if you are not stratifying the state by county. If one wants to stratify the state into geographic subregions, one just has to make sure the rules for clustering are set up so that only sites within the same geographic stratum can be clustered together.

7. **Evaluate how best to use "confirmed" and "unconfirmed" counts of trips in calculating the secondary and tertiary stage sampling fractions used to weight the**

**data.** If "unconfirmed" trips make up a small proportion of the counts, it may not be necessary to include them in the weighting of data. The number of "unconfirmed" trips could still be used to evaluate or adjust site pressures for a given time period. If this proportion is relatively large, future survey designs may want to consider an adjustment factor to account for the fact that some proportion of the "unconfirmed" trips will not actually be eligible for interviewing. It may also be interesting to compare the ratio of "confirmed" to "unconfirmed" trips across sites to determine if this ratio is relatively consistent across sites or there is a high degree of variability.

8. **Consider modifying the rules for clustering sites to use a total fishing pressure threshold as a basis for determining the number of sites in a multi-site cluster.** In the Pilot design, sites below a certain pressure threshold were clustered to form three-site clusters whenever possible. Few two-site clusters were formed, because such clusters were only formed when there were not enough lower pressure sites within close proximity to allocate to three-site clusters. However, creating more two-sit site clusters would reduce the amount of time spent driving between sites. If a selected two-site cluster exceeds an established total pressure threshold similar to the one established for stand-alone sites, then it should not be necessary to add a third site to the cluster.

9. **Evaluate the feasibility of sampling beach/bank shore mode fishing trips in all states using a strict access point survey design as tested in the pilot. In the Pilo**t study, it was assumed that all angler fishing trips ending at each identified beach/bank site could be appropriately sampled by stationing a sampler at a single access point. This may not be possible in other states where access to beach/bank fishing may be more diffuse and well-defined access points would be harder to establish. In such cases, it may be better to sample beach/bank shore angler trips through a "roving creel" sampling design that allows the collection of data for "incomplete trips". Consideration should be given to the potential disadvantages of introducing a "length of stay" bias throug**h** the use of a roving creel design. If the access point design is deemed to be appropriate, eliminating incomplete interviews will likely reduce the number of intercepts per shore mode assignment and the impact of this change will vary geographically. If the access point design is not deemed appropriate for sampling of beach/bank fishing trips, then it may be necessary to separately sample man-made shore trips and beach/bank shore trips as different strata (as was done in North Carolina).

10. **Evaluate the possible use of access point survey data to produce estimates of total fishing effort at sites included in the sampling frame.** The Project Team began to examine possible access point survey methods for effort estimation, but we recognized that further study is needed. Further study should be directed at determining whether or not on-site survey data on fishing effort could be used effectively in conjunction with off-site survey data to improve the accuracy of total fishing effort estimates. It may be very difficult to accurately identify and evaluate differences in estimates for overlap domains, because this would require some way for off-site interviews to accurately obtain information on the actual fishing sites to which anglers return from fishing. Such information could potentially be very hard to obtain and would require a substantial increase in the complexity of a telephone or mail interview. The advantage gained by doing this would have to be weighed against the possible disadvantages of increasing non-response rates.

11. **Consider splitting sites rated to have very high fishing pressure to create more total sites in the highest pressure category.** This could provide more high-pressure alternatives to assign when the number of available days for sampling is limited, such as for weekend assignments. This would provide more PSUs that are likely to be highly productive when selected. As it is now, some of the highest pressure sites get selected for all available weekend days in a month. Any increase in the selection probabilities for such sites would not increase the numbers of assignments allocated to them if all available dates are already getting saturated. However, the splitting of some of the highest pressure sites would create more high-pressure alternatives to possibly assign on the limited number of available days. Splitting these "super sites" could also have the added benefit of improving angler count data since it is more difficult to obtain accurate counts of missed eligible trips at very high pressure sites. However, the project team did note that high pressure sites should only be split if the configuration of the site allowed for a clear demarcation of angler trips returning to one site or the other and the site boundaries could easily be explained to samplers.

12. **Consider conducting separate "frame maintenance assignments" that would survey sites and provide site register updates without attempting to collect any interviews.** Such assignments could be focused on improving the quality of the site register and the accuracy of site pressure ratings. The more accurate the pressure ratings, the more efficient the sampling can become. Inaccurate site pressure

ratings would not cause any bias, as long as the inclusion probability of each site is easily known for weighting purposes. However, the proportion of assignments that obtain at least one interview should increase as the accuracy of the fishing pressures used in the PPS selection of sites and site clusters is improved. Frame maintenance assignments can also be used to identify new sites to add to the site register.

13. **Consider alternative ways to define size measures and weights for sites and site clusters in the sampling frame.** The Pilot sampling design adapted the traditional MRFSS pressure categories for use as size measures for the PSUs. The categories were translated to angler counts during each six-hour period for a site and mode. Size measures were summed over sites in a cluster when a cluster of two or three sites was used as the primary sampling unit. Depending on the clarified objectives, size measures might be based on projected catch rather than total anglers. It also appears that it may be beneficial to expand the range of fishing pressure category size measures at the high end to get more representation of the heavily fished PSUs in the sampling. This possibility should be evaluated prior to implementation of the new design in other states. It may also make better sense to simplify the measurement of expected fishing pressures across fewer size categories. Consideration should be given to the potential advantages and disadvantages of lumping (into fewer categories) versus splitting (into more categories), and decisions should be based on how reliably site pressures can be estimated and assigned to an appropriate category. If site pressures are likely to be extremely variable and hard to estimate accurately, it may be more appropriate to designate expected site pressure more simply as “high”, “medium”, or “low”. On the other hand, if site pressures are not very variable and they are easily assessed, then it may be beneficial to create more categories to more precisely match the weighting of sites and site clusters in the assignment draws with their actual activity levels.

    Pilot design sampling could also be changed in other ways to increase efficiency. More weight could be given to PSUs with higher pressure estimates in the PPS sampling. As long as lower pressure PSUs have some non-zero probability of being selected, an increase in the inclusion probabilities for higher pressure PSUs would not introduce any bias. However, too much of a shift of sampling toward the higher pressure sites would increase the variability among sites in their inclusion probabilities, thereby increasing the variability of sampling weights applied in the estimation process to the intercepts obtained. In other words, if sampling is shifted too much toward high pressure sites, the chances will be much greater that some small number of angler trip intercepts obtained within a selected low probability PSU would get an unusually high weight in the estimation process. Further study should be given to how best to balance the possible advantages of shifting PSU sampling probabilities against the possible disadvantages of creating much greater variability in the weighting of individual angler trip intercepts.

14. **Consider alternative ways to implement the desired stratification of sampling.** Some combination of "explicit" stratification and "implicit" stratification could be used. Explicit stratification creates disjoint subpopulations (in space and time), each of which is allocated a particular sample size and is sampled independently. This explicitly controls sample size within these spatio-temporal domains. Implicit strata are generally defined within explicit strata based on ordering on other dimensions; by using an ordered sampling algorithm the expected allocation to the implicit strata can be controlled, but the realized allocation may differ from expectation. To facilitate a simple sample selection scheme, define first-level explicit strata in terms of a geographic coastal area that can be covered by one team of interviewers. Order the PSUs within explicit strata by date and time of day within date. Post stratification at selected margins can be used to tune up the estimates to match known marginal distributions. An example of implicit stratification would be systematic sampling of sites within a spatiotemporal stratum after ordering by latitude. The sample size within a given latitude band would not be explicitly controlled, but there would be good representation of sites across latitudes. In particular, it would not be possible to have only southern sites within a latitude band, which could occur by chance without the implicit stratification.

15. **Consider defining different time intervals for the temporal stratification of sampling in other states.** Time intervals other than the ones used in the NC pilot study may be considered for use in other states. If so, the time interval sizes and boundaries should be chosen to both ensure reasonable sampler productivity while maintaining representative sampling. Implementation of a new intercept survey design will provide site-specific pressure information for various time intervals that could be used to fine-tune the intervals selected for this pilot. Such information may also reveal "dead" times when no intercepts are ever obtained and therefore sampler coverage is not needed (although care should be taken to confirm that this is truly the case and remains so over time). Optimal time intervals may also vary by region or state to reflect the geographic diversity that exists in recreational fisheries.

## 6. Literature Cited


Breidt, F.J., H.L. Lai, J.D. Opsomer, and D. A. Van Voorhees (2011) A Report of the MRIP Sampling and Estimation Project: Improved Estimation Methods for the Access Point Angler Intercept Survey Component of the Marine Recreational Fishery Statistics Survey.http://www.countmyfish.noaa.gov/projects/downloads/Final%20Report%20of%20New%20Estimation_Method_for_MRFSS_Data-01242012.pdf

Chromy. J.R., S.M. Holland, and R. Webster (2009) Consultant's Report: For-Hire Recreational Fisheries Surveys.http://www.countmyfish.noaa.gov/projects/downloads/MRIP_FHWG%20ForHire%20Methods%20Review%20Final.pdf

National Research Council, Committee on the Review of Recreational Fisheries Survey Methods (2006) Review of Recreational Fisheries Survey Methods. 202 pp. http://www.nap.edu/catalog/11616.html


## 7. Acknowledgements


The project team would like to thank the following people who all contributed to the success of this project: North Carolina samplers Travis Williams, Kelly Hawk, Jess Hawkins, Jesse Bissette, Jarrod Rabatin, James Woolard, Mary Alice Young, Wes Collett, and Blake Hocker for collecting all the NC Pilot data; Tim Haverland (NOAA Fisheries) for creating the site clustering program; Lauren Dolinger Few (NOAA Fisheries) for facilitating data transfers and formatting files for the clustering program; Laura Johansen (NOAA Fisheries) for performing multiple data analyses; Rebecca Ahrnsbrak (NOAA Fisheries) for producing the catch estimates comparison graphs; Tom Sminkey (NOAA Fisheries) for reviewing this report and providing valuable feedback; and John Foster (NOAA Fisheries) for assisting with development of the sample draw and estimation programs.

**Review of "A Pilot Study of a New Sampling Design for the Access Point Angler Intercept Survey," by the MRIP Design and Analysis Workgroup**

I am very positive about the work that was done on the pilot project and the changes that put the Intercept survey on a more scientific footing. I am in full agreement with the recommendation that the new design be adopted. The report is clear and thorough.

I. Changes requested

I have just one suggestion that would require any significant changes to the report, and then a few other suggestions that would require more minor changes.

The major one first:

1. The proposed new design and implementation rules undoubtedly produce estimates whose properties can be evaluated with much more precision than those of the MRFFS. I also believe that with adjustments and improvement, the new design will produce estimates that are, on average, better than those of the MRFFS. However, that message does not come through strongly enough in this report, in my opinion. The productivity measures (interviews per hour, etc.) and the standard errors reported (Figures 4,5,9 and 10) always show the pilot at a disadvantage. There are mentions of possible biases in estimates from the MRFFS, but then in other places the new estimation method developed by Breidt et. al. are described (p. 13) as approximately design unbiased.

The only statements about standard error are vague but suggest that even if sampling hours were the same, the Pilot would have been better but would still compare unfavorably to MRFFS. (p. 60 "If the number of PSU's observed in the Pilot design had been increased to match the number of assignments in the MRFFS design, the variances of the mean catch rates would have been lower." but "If both designs had completed the same number of assignments, the MRFFS design would still likely have provided estimates with greater precision") But how much better? The report doesn't make it easy for the reader to make a fair comparison. It would be useful for evaluating to have something like design effects for the two methods, since there are competing factors make the Pilot both less efficient (lower interviews per assigned hour) and more efficient (less variable weights, as mentioned on p. 54) It isn't clear (to me) how to do this precisely from the data, but even some rough measure of "variance per sampler hour" would be helpful.

Another analysis that would be useful to see would be a prediction of how much the design effect could be improved after implementation of some of the suggested improvements in allocation to the various strata. It seems that this could be calculable for at least a few major species or total catch, to give an idea of the potential value of the new method.

And of course this doesn't even consider the possible residual biases in the MRFFS data. The similarity of both sets of estimates would seem to suggest that there is little bias left in the newly weighted MRFFS, but then we get statements like this one on p. 8 :"… the potential for non-sampling errors was much larger under the MRFFS…" So does the statement on p. 8 refer to the "old MRFFS" or does it mean that even the “new MRFFS” is believed to still have substantial non-sampling errors. If the latter, it would be helpful to briefly describe what kinds of non-sampling errors the authors think still remain. Presumably, the authors feel there are smaller non-sampling errors than with just the original MRFFS estimators. Is that the interpretation of the authors? I find that surprising, although it doesn’t mean that even if it was true in NC in this year it will always be true. Does this mean then that the average catch per trip really does not vary much by time of day (night vs. day), or at alternative sites in the same vicinity (chosen by MRFFS samplers)? Or is it that the precision was so low we couldn’t tell there was a difference? More discussion of what evidence there is about this would be helpful to explain the advantage of the new design to those who might not be so convinced of the superiority of a strictly probability sampling design on theoretical grounds alone.

2. I would like to see more interpretation of results, or explanation of most likely causes, when there are differences between MRFFS and Pilot results. Here are some places where that would be informative:

a. p. 8 The summary states that the proportion of anglers reachable by the CHTS is comparable for all but beach/bank mode. Is there any intuition or speculation about why that would be true? Also, I believe it is true that an angler is not in the CHTS frame if he or she does not have a landline phone. If that is correct, it should be noted as part of the description of “the effort ratio estimator.”

b. p. 37 Do you have any speculation on why the jagged curve for the pilot? is it just because of the small sample size, or could it be because of the rules for clusters that have people driving from one place to another at certain hours?

c. p. 38 It is hard to know what to make of the differences seen in Table 6. What is the difference between reported and observed? Were these fish chosen because they are ones that are particularly common? How is the average # of fish computed? Is it average per angler overall, or average per angler that targeted or has any of that fish? Do these differences seem reasonable or explainable (e.g., are some fish caught more at night, for example, and so may be more likely to be seen in the Pilot?)

d. p. 47 The fact that the MRFFS estimates are higher when there is a difference is described as “interesting.” Is there a reason to believe this is a real difference and indicates a bias on the part of MRFFS data, due to either the noncoverage of night fishing or some other reason?

e. Table 7. Does it make sense that these estimates would differ due to noncovearge of night fishing in MRFFS, or is there some other reason? Maybe only locals fish from the beach at night?

3. On p. 7 and again on p. 38 are mentions that the telephone survey (CHTS) is used along with the intercept data to come up with total catch estimates. I believe these descriptions are too cryptic for a reader who is not already familiar with the estimation method to understand. Even a brief explanation that total catch is estimated as (total trips by coast county anglers) *(catch per trip for all anglers) *(# of all anglers/# of coastal county anglers) would be helpful, and would help elucidate some of the

comments about why bias can result if certain assumptions (e.g., avg. catch the same for coastal and noncoastal) do not hold.

II. Clarifications Requested:

1. p. 4 When the report says that the Pilot compared the performance of the new sampling design and estimates of catch per trip, etc. with the “traditional MRFFS sampling design” (p. 4), does this mean that the comparisons are with the old design but the new estimators developed by Breidt et al.? The remainder of the report seems to suggest the comparisons are always to the new analysis method for MRFFS, but the use of the word traditional made me wonder as I was reading the executive summary.

2. How does the new design obtain nearly 3 sites per assignment? (2.46, from p. 8). Does that mean that most assignments are 3-site clusters? It might be good to have a summary somewhere of how many units in the frame are of each size.

3. What does the word positive mean in this statement on p. 10 “provide higher average number of interviews per *positive* assignment.” Also on p. 61.

4. p. 11 #6. It is not clear in what way cluster formation would be modified. More two-site clusters? Why would that happen?

5. p. 11 Recommendation # 10. Apparently there was some indication that the site list had some flaws. Could these be clarified?

6. p. 12 #12. Clarify what is meant by “explicit” and “implicit”

7. p. 23 “at least one psu was selected from each interval…” Clarify that this means “per month” (I think).

8. p. 36 The x-axis scales are different on these plots. I don’t think they should be. In fact, why should n’t they all be from 0 to 24?

9. p. 39 I don’t’ understand the difference in the last two sentences of the paragraph before Table 7.They seem contradictory, but then I don’t know what “fishing area” means.

10. Figures 4 and 5 should have a label of Pilot and MRFFS to be consistent with the other descriptions, rather than NC and MRFFS.

11. p. 10 I do not understand the meaning of #1 under future recommendations.

12. p. 53 The word *inappropriate* seems wrong. It is not that samplers are making bad choices, but that they are making choices at all that is wrong with MRFFS.

13. p. 69 The need for frame maintenance assignments is mentioned a couple of times, but no information on the reason for this suggestion is mentioned. Was there some problem with the list of sites that was discovered during the pilot?

14. Probability sampling is referred to as probabilistic sampling on p. 5. Is this meant to imply something different than what we usually mean by probability sampling?

15. p. 15. Why are there quotes around the statement under item #3? I presume because it is quoting from the previous report, but it should say so.

16. pp. 52, 53,55 and several other places There are comments about the fact that one advantage of the Pilot covering all times is that it eliminates potential bias due to different catch rates at night. But it also eliminates potential bias caused by different coastal county anglers' percentage at night. I believe this should be emphasized.

17. p. 52 "It was not necessary to require samplers to regularly work overtime." Was this a problem with MRFFS? If so, then state this advantage.

III. Other comments :

1. Executive summary and p. 13 mentions that the MRFFS survey was "…not providing coverage of fishing trips ending on private property or at night." Much discussion was devoted to the issue of improving night fishing, but no other discussion of private property was included in the report. It seemed a bit like a "bait-and-switch" as I kept looking for discussion of innovations for the private property issue. Maybe this should be removed from the overview.

2. p. 12 # 11 The idea of basing pressure measures on # of fish landed seems a good one, especially if this allocation is very different than the one being used. It would be even better if the variability in fish landings could be used, though maybe mean and variance of fish landings are positively correlated.

4. p. 68 Suggestion #8. I like this idea a great deal. In fact, the potential for cost savings if the entire effort estimation were to be moved to the intercept survey is large. With the new design, my guess would be that number of trips could be more accurately estimated from the intercept survey than from telephone or mail, due to the large measurement error and nonresponse of the effort survey. The only problem I see is the private access sites. How much of the total effort is on private access sites? Maybe a telephone survey would still be necessary but only to estimate this ratio rather than the current coastal county ratio estimated from the intercept. Why not just compare total public access estimates from telephone and the intercept survey rather than needing to estimate by site (thought this is perhaps not a discussion for this report)? If the intercept survey were to be used for both effort and catch, one might consider counting both arriving and departing anglers, since the spatial-temporal time frame would be appropriate for either. The two estimators would be correlated but not identical, so should provide a little extra information basically for free.

**Review of Report on a Pilot Study of a New Sampling Design for the NOAA Fisheries Access Point Angler Intercept Survey**

**Reviewer 1**

**11-8-12**

## 1. Introduction

A pilot study was conducted in 2010 in North Carolina to test the feasibility of implementing a new sampling design for the Access Point Angler Intercept Survey, a component of the Marine Recreational Fisheries Statistics Survey (MRFSS) of the NOAA Fisheries Service. The new design was compared to the traditional MRFSS sampling design. This review discusses the quality and completeness of the report on the pilot study.

## 2. Key ideas in sample surveys

In order for results from sample surveys to be representative of a population, the survey must be designed and implemented following certain principles. First, it is important to identify the target population and how the population can be enumerated or listed, either explicitly or implicitly. Second, it is important to know the probabilities that units in the population are included in the sample. These probabilities can be influenced by the structure of the population and the sampling scheme. Stratification can be used to increase precision of estimators as well as to ensure adequate coverage of subpopulations. Cluster sampling often reflects the structure in the population and can be utilized to reduce costs and for practical considerations, but often at the cost of decreasing precision of estimators. Third, adequate plans need to be in place to deal with likely problems in survey data collection, including refusal to participate, difficulty in answering questions, resource limitations, and other challenges. When the target population is well defined and probabilities of sample inclusion for members of the population are known, the sampling strategy is an example of probability sampling and can be the basis for scientific statements about the population.

## 3. Overall comments on the redesign plans

The changes proposed and implemented for the Access Point Angler Intercept Survey are consistent with professional scientific survey practice. These changes, if implemented in an optional way, are expected to yield estimates that are improved in terms of less bias and reduced variance in this survey.

The report (Breidt et al. 2012) provides a thorough and professional evaluation of the pilot study. Numerical and graphical presentations are sufficient for comparing results. Unfortunately, results using the new design are not clearly superior to the results using the previous design with weight adjustment. In particular, sample yield is lower using the new design. In part this reflects the requirement that interviewers not substitute sites or fishing modes as they see fit. This is an important change for reducing selection bias. As in most surveys, it is not really possible to measure the amount of bias, because it could only be compared to a much more rigorous and large effort.

Despite the lower sample yield, the recommendation that seems most reasonable is to continue with comparisons of the new design with the original design. The authors have given several suggestions for improvement. These should be considered carefully before proceeding to other areas of application. Importantly, one should consider changes in stratification, cluster definition, and allocation to improve performance of estimates. Additional analysis of the pilot data could provide some guidance about adjusting the specific implementation of the new design. The new design has the potential to produce much better survey estimates. For that promise to be realized, additional work will be necessary in refining the approach.

## 4. Specific comments on the report

### *a. Time of day stratification*

Time of day stratification versus sampling at peak intervals: This suggestion is quite important to eliminate interviewer judgment about when to interview. Without control in time of day stratification, there could be substantial bias in estimating mean catch and number of fishing trips.

Is interviewing assigned at night? It would seem prudent to not sample in the middle of the night. Or to sample with a low frequency in the middle of the night. Perhaps the four time frames are not enough. What if you excluded 12-4am and divided the remaining 20 hours into 5 slots each of 4 hours? You could still have two low intensity sites in a cluster.

### *b. Geographic stratification*

Geographic stratification versus sampling across entire state: This suggestion is important for ensuring coverage of the whole state and allocating interviewers to areas.

Effective and creative selection of strata could lead to efficiency and adequate information for estimation in some sub-state areas. Stratification should be implemented separately in each state. It could reflect NOAA Fisheries Service divisional units in addition to major state-level geographic features.

Heavy stratification into small strata could lead to gains in efficiency. One would want to track and estimate interviewer effects if few interviewers are collecting information in a single strata or across strata.

### *c. Clustering sites*

In the old design sites were not clustered together, and interviewers were allowed to visit other sites. Allowing interviewers to choose sites invalidates the assumptions of probability sampling and would be likely to lead to bias in estimation. Not clustering sites together could lead to high variability in probabilities of selection if selection is based on activity level. Although not related to bias, high variability in probabilities can be disadvantageous due to associated high uncertainty in estimators. Grouping low activity sites together and visiting all grouped sites should stabilize probabilities of selection, but still allow high rates of sampling among high activity sites. This seems like a good compromise. Control of routes and time spent by interviewers is a critical addition to the design.

Some high activity sites might be included with certainty every survey year due to their importance to state-level fishing. Other sites might appear every few years in surveys. It would be a good idea to review information used to create probabilities of selection every year.

Given the availability of GIS and other technologies for recording time and location, it probably would be worthwhile to design measures of performance of interviewing assignments. That is, if you can collect some process information about travel and ease of collecting information by site, then it might be possible to improve resource planning in the future.

d. *Sampling Frame and Probabilistic Sampling; Elimination of Opportunistic Sampling; Angler Trip Counts*

Specifying a formal sampling frame and implementing a probability sampling scheme is a major advance over the former method. Statistical models can be useful for improving efficiency and removing some amount of bias, but it is best to start with a probability design. Use of models to improve efficiency can still be considered, but one is relying on models to a much smaller degree if the starting point is a probability sampling design.

Eliminating opportunistic switching of fishing modes removes a potential source of bias.

Getting accurate numbers of angler trips ending in a given interval is important as stated in the report for assessing secondary probabilities of selection. It will be important to assess whether adequate resources are available for both counting anglers finishing trips as well as interviewing in sites. It might be quite difficult for a single individual to interview effectively and count over a broad area at the same time. According to Appendix A, page 28 (page 33 of PDF)

> At sites with low activity you should be able to both count and conduct interviews at the same time. At sites with moderate-to-high activity you will alternate between counting and conducting interviews by the hour. Be sure to record the start and stop time for the time spent counting and the time spent interviewing as two separate sampling periods, even if you do not switch sites."

It is recommended that an experiment be conducted to compare the current new plan (simultaneous counting and interviewing) with dedicated counting plus interviewing. The latter likely will require a second person.

e. *Issuing and Completing Assignments:*

The new design requires issued assignments to be completed without rescheduling. The former design allowed interviewers to reschedule. As stated in the report, "Eliminating assignment rescheduling greatly reduces the possibility of a nonresponse bias that could result from a failure to obtain observations from some of the selected assignments. It also eliminates possible temporal undercoverage biases that could result from the rescheduling of assignments." It is the opinion of the reviewer that these statements most certainly are true.

*f. Interviewing limits:*

The previous design had a cap on interviews. Instead, the pilot design has interviewers utilize the full time frame for sampling. There are a couple of possibilities motivating this change. First, for a given number of clusters, a larger sample size within each cluster is better. Second, interviews of a higher quality provide better information. The change in policy obviously should increase the number of interviews in some clusters. It might also positively impact interview quality as long as the interview period is not too long and tiring. If one knows that one can stop interviewing after a given number of interviews, then one might do the required number of interviews quickly in order to be done with the assignment. Requiring interviewing to continue for a given period of time eliminates the incentive to finish quickly.

It is recommended that available technology (GIS, computer time/date stamps, etc.) be used to monitor interviewers if there is any question about the legitimacy of data collection reporting.

*g. Eligibility for Interviews:*

As I understand the report, the new design allows interviewing children under age 5 and individuals who are returning from a contest. These considerations are beyond my expertise to evaluate. Of course, interviewing children usually requires parent/guardian permission and attention to what the children might or might not be able to reasonably answer. I could see one defining the ultimate sampling unit either as an individual angler or as a group of anglers together (in a boat, in a group such as a family with small children). There could be advantages and disadvantages to both arrangements in terms of ease of interviewing, clarify of definition of unit, and value of information in estimation.

The inclusion of tournaments is specified on page 30 (PDF page 35) of Appendix A, but I see no mention of ages of interviewees. Given that children under age 5 were not included before, it is likely important to add a section stating that they should be interviewed. Procedures for interviewing children (e.g., parental/guardian permission, presence of parent/guardian) should be included.

*h. Complete vs. Incomplete Beach/Bank Interviews*

The inclusion of incomplete angling trip under the old design probably was problematic for estimation. Even if anglers were asked what percent of their trip was completed it would have had some degree of speculation. It is likely as stated in the report that those fishing longer would have a higher probability of being selected for an incomplete trip and also have a higher number of fish, thereby producing bias. Although it makes finding someone to interview harder (they have to be done fishing), this change likely removes a source of bias.

*i. Questionnaires and Data Forms*

These seem to have been adapted appropriately for the new design. Are any thoughts being given to electronic data capture on, for example, a smart phone, GIS position recording, or time/date stamp to automate the recording of some information?

*j. Estimation methodology*

A prime advantage of a probability sampling design is that design-based estimation should produce unbiased (or nearly unbiased in the care of ratio estimation) estimates of desired quantities. It is appropriate to switch to a design-based estimation method.

One could consider the use of statistical models in order to utilize auxiliary information available at all sites. It is recommended that one study correlations among survey outcome variables and information available at each site. Some of this information could be quantitative and other information categorical or dichotomous. Environmental: area of lakes, length of fishing bank, etc. Population: population size within 10, 25, and 50 miles of site, percent of population nearby below poverty and below 200% of poverty, etc. Administrative: is it a state park? Is a state permit required? Is hunting allowed at the site? Is there a swimming beach? It is uncertain what will be predictive, but thinking broadly about potential relationships could guide useful data collection. Even if not used in a statistical model to improve estimation, it still might be interesting to examine associations of auxiliary variables with outcomes.

## 5. Comments on some survey questions

The introduction to the potential respondent could be rephrased. Instead of, Hi, I'm representing NCDMF, it would be better to not use an acronym in the opening contact. See page 33 (PDF page 38) of the Appendix A (NOAA 2011).

The screening question (recreation) is as follows (page 33, PDF page 38 of the Appendix A, NOAA 2011):

> Was the primary purpose of your trip today for recreation, that is, for fun and relaxation, or was it to provide income either from the sale of fish or from the sale of the fishing opportunity?

The screening criterion might be better stated, or additional instructions could be provided to interviewers. How would someone be classified who is fishing to have something to eat? Presumably some people fish regularly in order to supplement their diet. I would assume as recreation. What if they trade some fish to someone for vegetables or wild game? Does this qualify as recreation or income?

## 6. Comments on results and recommendations

In the pilot study, the original design resulted in more interviews per assignment and per hour. Estimates of mean catch rates were not very different under the two methods. If measured simply by estimated sampling variance, the results using the new method are not superior. As Breidt et al. (2012; page 9) notes,

> The estimates generated from the MRFSS sampling design were more precise than the estimates generated from the Pilot design largely because a greater number of sampling assignments were completed under the MRFSS design.

One can speculate, however, that the new design actually could be better in terms of selection bias and other problems. As in most surveys, it is not really possible to measure the amount of bias, because it could only be compared to a much more rigorous and large effort.

> It should be noted, however, that the potential for non-sampling errors was much larger under the MRFSS than under the new design, which is not reflected in these precision comparisons. (Breidt et al 2012; page 9)

One also can speculate that the performance of the new design can be improved.

> Although sampling under the new design in this study yielded a much larger percentage of completed assignments with no angler trip interviews and a much smaller number of interviews per positive assignment, changes in the allocation of sampling across sampling strata could greatly reduce these differences. (Breidt et al 2012; page 9)

Importantly, one could try to analyze the existing pilot survey in order to better inform choices about the new sample design.

*Allocation of sample*

The report gives suggestions on how to improve. The 2nd and 3rd recommendations are critical. One should study when, where, and for which mode the MRFSS survey gained such large amounts of sampled units. Time stratification and whether to exclude the middle of the night are topics that need consideration.

One should also identify resources before allocating sample. It sounds like resources were woefully inadequate in some areas. One can implement legitimate probability sampling schemes with controls connected to geography and resources, and more effort apparently is needed in this direction.

*Splitting high intensity sites*

Among issues to study in the future, splitting high intensity sites so that size is not so large among the largest probably is a good idea. Handling a mix of very small and very large sites is difficult when a limited amount of time is available for each and there are substantial travel times to the next site. It is recommended that stratification become finer and clusters be formed to be more uniform in size. Recommendations for future research #9, 12, and 13 likely should be moved up in priority. One aspect of this is time stratification. Surely one can omit 12-4am and do a better job. One could consider splitting the remaining twenty hours into five 4-hour blocks.

One then could still visit two low intensity sites (up to 2 hours each) in a single cluster if they are geographically close enough.

*Adjusting allocation to mode, or allowing mode switching*

In Breidt et al. (2012), it is noted that

> For the Pilot, assignments were allocated evenly across the four modes in each state subregion: Man-made (MM), Beach Bank (BB), Private/Rental (PR), and Charter (CH). Allocation of mode-specific assignments within each state subregion and day type (i.e. kind of day) was determined monthly.

It seems to this reviewer that the actual modes used would not be close to even in every site. Being able to switch modes would have given the original survey a big advantage. Perhaps one can learn from the actual data collected (pilot versus original design) how mode restrictions impacted data collection.

**Summary**

The changes proposed and implemented for the Access Point Angler Intercept Survey are consistent with professional scientific survey practice. These changes, if implemented in an optional way, are expected to yield estimates that are improved in terms of less bias and reduced variance in this survey.

The report (Breidt et al. 2012) provides a thorough and professional evaluation of the pilot study. Numerical and graphical presentations are sufficient for comparing results. Results using the new design are not clearly superior to the results using the previous design with weight adjustment. The recommendation, however, is to continue with comparisons of the new design with the original design. The authors have given several suggestions for improvement. These should be considered carefully before proceeding to other areas of application. Importantly, one should consider changes in stratification, cluster definition, and allocation to improve performance of estimates. The new design has the potential to produce much better survey estimates. For that promise to be realized, additional work will be necessary in refining the approach.

It is recommended that one study correlations among survey outcome variables and information available at each site. Some of this information could be quantitative and other information categorical or dichotomous. Thinking broadly about potential relationships could guide useful data collection. Even if not used in a statistical model to improve estimation, it still might be interesting to examine associations of auxiliary variables with outcomes.

**References**

Breidt, F.J., et al. (2012). A Pilot Study of a New Sampling Design for the Access Point Angler Intercept Survey. Report to the Marine Recreational Information Program's (MRIP) Design and Analysis Work Group (DAWG).

NOAA. (2011). Appendix A. Field Procedures Manual. *Procedures Manual: Development and Testing of Alternative Sampling Design for the MRFSS Intercept Survey.* North Carolina Division of Marine Fisheries. National Oceanic and Atmospheric Administration National Marine Fisheries Service (NOAA Fisheries). February of 2011.

**Response to Reviewers Comments**
**Report on a Pilot Study of a New Sampling Design**
**for the NOAA Fisheries Access Point Angler Intercept Survey**

We appreciate the careful review and insightful comments provided by the two reviewers. We have added clarifications to the report in some sections following reviewers' advice. Listed below are the reviewer's comments and our responses.

**1. *Time of day stratification:***

Time of day stratification versus sampling at peak intervals: This suggestion is quite important to eliminate interviewer judgment about when to interview. Without control in time of day stratification, there could be substantial bias in estimating mean catch and number of fishing trips. Is interviewing assigned at night? It would seem prudent to not sample in the middle of the night. Or to sample with a low frequency in the middle of the night. Perhaps the four time frames are not enough. What if you excluded 12-4am and divided the remaining 20 hours into 5 slots each of 4 hours?

**Response**. Data from the pilot study showed that for certain modes (e.g. shore mode) fishing activities targeting specific species do occur at night in some states and seasons. Therefore the inclusion of the night fishing corrects the temporal and spatial undercoverage of night fishing in the current MRFSS. We recognize the importance of having an appropriate sample allocation across all time intervals to accurately represent the expected effort at each time interval and anticipate a low sampling intensity in intervals A (2:00 am - 8:00 am) and D (8:00 pm - 2:00 am).

As explained in the report we selected the six-hour sampling interval because it would allow for a standard eight-hour workday when travel time to and from the sampling cluster is included.

**2. *Geographic stratification:***

Geographic stratification versus sampling across entire state: This suggestion is important for ensuring coverage of the whole state and allocating interviewers to areas.

Effective and creative selection of strata could lead to efficiency and adequate information for estimation in some sub-state areas. Stratification should be implemented separately in each state. It could reflect NOAA Fisheries Service divisional units in addition to major state-level geographic features. Heavy stratification into small strata could lead to gains in efficiency.

**Response**: We have considered geographic stratification as a means to ensure adequate spatial distribution of the study and to improve precision of the estimates at different levels. We agree with the reviewer that stratification could be implemented at different levels (state and regional) so data collected can be used to inform decision makers and stakeholders. As we move on with the implementation of the sampling design, we will be considering stratification to improve precision estimates.

**3.** ***Clustering sites:***

In the old design sites were not clustered together, and interviewers were allowed to visit other sites. Allowing interviewers to choose sites invalidates the assumptions of probability sampling and would be likely to lead to bias in estimation. Not clustering sites together could lead to high variability in probabilities of selection if selection is based on activity level. Although not related to bias, high variability in probabilities can be disadvantageous due to associated high uncertainty in estimators. Grouping low activity sites together and visiting all grouped sites should stabilize probabilities of selection, but still allow high rates of sampling among high activity sites. This seems like a good compromise. Control of routes and time spent by interviewers is a critical addition to the design.

Some high activity sites might be included with certainty every survey year due to their importance to state-level fishing. Other sites might appear every few years in surveys. It would be a good idea to review information used to create probabilities of selection every year. Given the availability of GIS and other technologies for recording time and location, it probably would be worthwhile to design measures of performance of interviewing assignments. That is, if you can collect some process information about travel and ease of collecting information by site, then it might be possible to improve resource planning in the future.

**Response**. When clustering the sites, we used a GIS algorithm that clusters sites in the sampling frame by minimizing the driving distance between sites while maintaining similar size measures (effort) across clusters. This algorithm allows for high activity sites to be placed in a single site cluster and to combine nearby sites so the total effort across clusters are similar. This clustering of the sites in the sampling frame will be performed for each sampling wave (two month period) and year, therefore there is no certainty that two sites will be in the same cluster during the same period. Also, a site that has high activity in a given mode in one wave not necessarily will have the same activity level the next wave. Furthermore, the sampling frame size measures are updated for sampling purposes before the beginning of each wave to reflect the changing trends in fishing practices across time and space. We will revise the clustering method section (section 3.1.1.3) in the report to better explain the clustering algorithm. Also, we will revise current examples and will add few more to cover the comments provided by the reviewer and other relevant instances that can clarify the clustering process to the reader.

Currently field interviewers collect information on time and length of interview, site location, and site visit arrival and departure times. This information is used to evaluate productivity and performance measures and to inform the allocation of resources for data collection purposes. We also perform a quality check on interviewer performance by selecting a random sample of their intercepts and verifying among other topics interviewer time and location as well as interviewing technique.

**4.** ***Sampling frame and probabilistic sampling; elimination of opportunistic sampling; angler trip counts:***

Specifying a formal sampling frame and implementing a probability sampling scheme is a major advance over the former method. Statistical models can be useful for improving efficiency and removing some amount of bias, but it is best to start with a probability design. Use of models to improve efficiency can still be considered, but one is relying on models to a much smaller degree if the starting point is a probability sampling design.

Eliminating opportunistic switching of fishing modes removes a potential source of bias. Getting accurate numbers of angler trips ending in a given interval is important as stated in the report for assessing secondary probabilities of selection. It will be important to assess whether adequate resources are available for both counting anglers finishing trips as well as interviewing in sites. It might be quite difficult for a single individual to interview effectively and count over a broad area at the same time. According to Appendix A, page 28 (page 33 of PDF)

> "At sites with low activity you should be able to both count and conduct interviews at the same time. At sites with moderate-to-high activity you will alternate between counting and conducting interviews by the hour. Be sure to record the start and stop time for the time spent counting and the time spent interviewing as two separate sampling periods, even if you do not switch sites."

It is recommended that an experiment be conducted to compare the current new plan (simultaneous counting and interviewing) with dedicated counting plus interviewing. The latter likely will require a second person.

**Response:** We thank the reviewer for the review and positive comments about the sampling design. Obtaining accurate number of angler trips is very important since these are the base for the selection probabilities and estimation processes. During the pilot study we did not carry out a formal experiment to compare the efficiency of 'simultaneous counting and interviewing' (a single interviewer) versus 'dedicated counting plus interviewing' (two interviewers), however, both modes of counting and interviewing were implemented at some time during the pilot study. Feedback from the field interviewers suggested that simultaneous counting and interviewing was feasible and didn't affect their productivity. Sites with high activity will be in a single site cluster and the field interviewer will count and interview in that site for 6 hours.

**5. *Issuing and completing assignments:***

The new design requires issued assignments to be completed without rescheduling. The former design allowed interviewers to reschedule. As stated in the report, "Eliminating assignment rescheduling greatly reduces the possibility of a nonresponse bias that could result from a failure to obtain observations from some of the selected assignments. It also eliminates possible temporal undercoverage biases that could result from the rescheduling of assignments." It is the opinion of the reviewer that these statements most certainly are true.

**Response:** We appreciate the positive and encouraging comments from the reviewer regarding issuing and completing assignments.

**6.** ***Interviewing limits:***

The previous design had a cap on interviews. Instead, the pilot design has interviewers utilize the full time frame for sampling. There are a couple of possibilities motivating this change. First, for a given number of clusters, a larger sample size within each cluster is better. Second, interviews of a higher quality provide better information. <u>The change in policy obviously should increase the number of interviews in some clusters. It might also positively impact interview quality as long as the interview period is not too long and tiring.</u> If one knows that one can stop interviewing after a given number of interviews, then one might do the required number of interviews quickly in order to be done with the assignment. <u>Requiring interviewing to continue for a given period of time eliminates the incentive to finish quickly.</u>

It is recommended that available technology (GIS, computer time/date stamps, etc.) be used to monitor interviewers if there is any question about the legitimacy of data collection reporting.

**Response:** We appreciate the positive and encouraging comments from the reviewer regarding interviewing limits. See bullet (c.) for response to monitoring interviewers' issues.

**7.** ***Eligibility for interviews:***

As I understand the report, the new design allows interviewing children under age 5 and individuals who are returning from a contest. These considerations are beyond my expertise to evaluate. <u>Of course, interviewing children usually requires parent/guardian permission and attention to what the children might or might not be able to reasonably answer. I could see one defining the ultimate sampling unit either as an individual angler or as a group of anglers together (in a boat, in a group such as a family with small children).</u> There could be advantages and disadvantages to both arrangements in terms of ease of interviewing, clarify of definition of unit, and value of information in estimation.

<u>The inclusion of tournaments is specified on page 30 (PDF page 35) of Appendix A, but I see no mention of ages of interviewees. Given that children under age 5 were not included before, it is likely important to add a section stating that they should be interviewed. Procedures for interviewing children (e.g., parental/guardian permission, presence of parent/guardian) should be included.</u>

**Response:** The ultimate sampling unit is the angler. The current procedures manual instructs the interviewer on obtaining parental permission, or permission from a legal guardian to interview anglers of age 5 to 18 years. The interviewer is instructed to interview the young angler or child only after the permission is granted by the parents or legal guardian. We plan to include more detail instructions in the procedure manual for carrying out proxy interviews (an adult answers the questions for the children) when the child is very young or unable to provide the answers. These instructions will be valid for both the tournament intercepts and regular intercepts.

**8.** ***Complete vs. incomplete beach/bank interviews:***

The inclusion of incomplete angling trip under the old design probably was problematic for estimation. Even if anglers were asked what percent of their trip was completed it would have had some degree of speculation. It is likely as stated in the report that those fishing longer would have a higher probability of being selected for an incomplete trip and also have a higher number of fish, thereby producing bias. Although it makes finding someone to interview harder (they have to be done fishing), this change likely removes a source of bias.

**Response:** We appreciate the positive and encouraging comments from the reviewer regarding complete vs incomplete beach bank interviews.

**9.** ***Questionnaires and data forms:***

These seem to have been adapted appropriately for the new design. <u>Are any thoughts being given to electronic data capture on, for example, a smart phone, GIS position recording, or time/date stamp to automate the recording of some information?</u>

**Response:** We appreciate the positive and encouraging comments from the reviewer. We are considering implementing electronic data collection procedures in the future. We agree that this will be the next step to improve data collection, and that such will not only facilitate the recording and secure transmission of all information but also will improve data quality and facilitate data management as well.

**10.** ***Estimation methodology:***

<u>A prime advantage of a probability sampling design is that design-based estimation should produce unbiased (or nearly unbiased in the care of ratio estimation) estimates of desired quantities. It is appropriate to switch to a design-based estimation method.</u>

One could consider the use of statistical models in order to utilize auxiliary information available at all sites. It is recommended that one study correlations among survey outcome variables and information available at each site. Some of this information could be quantitative and other information categorical or dichotomous. Environmental: area of lakes, length of fishing bank, etc. Population: population size within 10, 25, and 50 miles of site, percent of population nearby below poverty and below 200% of poverty, etc. Administrative: is it a state park? Is a state permit required? Is hunting allowed at the site? Is there a swimming beach? It is uncertain what will be predictive, but thinking broadly about potential relationships could guide useful data collection. Even if not used in a statistical model to improve estimation, it still might be interesting to examine associations of auxiliary variables with outcomes.

**Response:** We appreciate the positive and encouraging comments from the reviewer. We recently developed new design-based estimation methods to produce catch and effort estimates by region, state, mode, species, etc. to inform policy makers for managing the east coast fish resources. As we move forward with the survey we will consider model based approaches to improve estimates or to compensate for non-response (e.g. imputation models).

**11. *Comments on some survey questions:***

The introduction to the potential respondent could be rephrased. Instead of, Hi, I'm representing NCDMF, it would be better to not use an acronym in the opening contact. See page 33 (PDF page 38) of the Appendix A (NOAA 2011).

The screening question (recreation) is as follows (page 33, PDF page 38 of the Appendix A, NOAA 2011):

> "Was the primary purpose of your trip today for recreation, that is, for fun and relaxation, or was it to provide income either from the sale of fish or from the sale of the fishing opportunity?"

The screening criterion might be better stated, or additional instructions could be provided to interviewers. How would someone be classified who is fishing to have something to eat? Presumably some people fish regularly in order to supplement their diet. I would assume as recreation. What if they trade some fish to someone for vegetables or wild game? Does this qualify as recreation or income?

**Response:** We thank the reviewer for the feedback of the survey forms. Our current procedure manual instructs the interviewer to use the name of the institution and not the acronym. In addition, it instructs the interviewer to include non-commercial (not to provide income) "subsistence fishing" trips as "recreational" trips that are eligible for an interview.

**12. *Comments on results and recommendations:***

In the pilot study, the original design resulted in more interviews per assignment and per hour. Estimates of mean catch rates were not very different under the two methods. If measured simply by estimated sampling variance, the results using the new method are not superior. As Breidt et al. (2012; page 9) notes,

> "The estimates generated from the MRFSS sampling design were more precise than the estimates generated from the Pilot design largely because a greater number of sampling assignments were completed under the MRFSS design."

One can speculate, however, that the new design actually could be better in terms of selection bias and other problems. As in most surveys, it is not really possible to measure the amount of bias, because it could only be compared to a much more rigorous and large effort.

> "It should be noted, however, that the potential for non-sampling errors was much larger under the MRFSS than under the new design, which is not reflected in these precision comparisons." (Breidt et al 2012; page 9)

One also can speculate that the performance of the new design can be improved.
Although sampling under the new design in this study yielded a much larger percentage of completed assignments with no angler trip interviews and a much smaller number of interviews

per positive assignment, changes in the allocation of sampling across sampling strata could greatly reduce these differences.

Importantly, one could try to analyze the existing pilot survey in order to better inform choices about the new sample design.

**Response:** We thank the reviewer for the insightful comments. This is indeed the intent expressed in the report.

**13.** ***Allocation of sample:***

The report gives suggestions on how to improve. The 2nd and 3rd recommendations are critical. One should study when, where, and for which mode the MRFSS survey gained such large amounts of sampled units. Time stratification and whether to exclude the middle of the night are topics that need consideration.

One should also identify resources before allocating sample. It sounds like resources were woefully inadequate in some areas. One can implement legitimate probability sampling schemes with controls connected to geography and resources, and more effort apparently is needed in this direction.

**Response:** We thank the reviewer for the insightful comments. As discussed before (bullet a.) we are working on obtaining accurate size measures (effort) for the sites for all modes and time intervals. These size measures will inform the sampling design and the sample allocation across kind of day and time interval. We anticipate that the observed fishing activity at night for some modes will result in few sites selected at intervals A and D.

We agree that identification of resources is essential in the success of a sampling design as the one proposed here. For example, a small number of interviewers would be able to successfully cover a large area and a diverse schedule that combine day and night sampling. We anticipate that some of the limitations encountered in the pilot study will not be an issue when implementing the new sampling design.

**14.** ***Splitting high intensity sites:***

Among issues to study in the future, splitting high intensity sites so that size is not so large among the largest probably is a good idea. Handling a mix of very small and very large sites is difficult when a limited amount of time is available for each and there are substantial travel times to the next site. It is recommended that stratification become finer and clusters be formed to be more uniform in size. Recommendations for future research #9, 12, and 13 likely should be moved up in priority. One aspect of this is time stratification. Surely one can omit 12-4am and do a better job. One could consider splitting the remaining twenty hours into five 4-hour blocks. One then could still visit two low intensity sites (up to 2 hours each) in a single cluster if they are geographically close enough.

**Response:** As discussed in bullet (c.) above the current clustering algorithm produces clusters by minimizing the distance between them while maintaining an almost homogeneous size measure across clusters. The NOAA staff in charge of running this algorithm works with the states to split large sites into smaller and manageable sizes sites that can be surveyed in 6 hour intervals.

**15.** ***Adjusting allocation to mode, or allowing mode switching:***

In Breidt et al. (2012), it is noted that

> "For the Pilot, assignments were allocated evenly across the four modes in each state subregion: Man-made (MM), Beach Bank (BB), Private/Rental (PR), and Charter (CH). Allocation of mode-specific assignments within each state subregion and day type (i.e. kind of day) was determined monthly."

It seems to this reviewer that the actual modes used would not be close to even in every site. Being able to switch modes would have given the original survey a big advantage. Perhaps one can learn from the actual data collected (pilot versus original design) how mode restrictions impacted data collection.

**Response:** For the implementation of the proposed sampling design we are not proposing equal allocation across modes. In the pilot study we used equal allocation across modes to ensure enough sample to produce some estimates at different levels and to be able to make some comparisons between modes and across surveys. We are proposing to allocate the sample based on current and historical data as it is currently done in the current MRFSS.

**16.** ***Performance of the proposed new design:***

The proposed new design and implementation rules undoubtedly produce estimates whose properties can be evaluated with much more precision than those of the MRFSS. I also believe that with adjustments and improvement, the new design will produce estimates that are, on average, better than those of the MRFFS. However, that message does not come through strongly enough in this report, in my opinion. The productivity measures (interviews per hour, etc.) and the standard errors reported (Figures 4,5,9 and 10) always show the pilot at a disadvantage. There are mentions of possible biases in estimates from the MRFFS, but then in other places the new estimation method developed by Breidt et. al. are described (p. 13) as approximately design unbiased.

The only statements about standard error are vague but suggest that even if sampling hours were the same, the Pilot would have been better but would still compare unfavorably to MRFSS (p. 60):

> "If the number of PSU's observed in the Pilot design had been increased to match the number of assignments in the MRFFS design, the variances of the mean catch rates would have been lower."

but then:

> "If both designs had completed the same number of assignments, the MRFSS design would still likely have provided estimates with greater precision"

But how much better? The report doesn't make it easy for the reader to make a fair comparison. It would be useful for evaluating to have something like design effects for the two methods, since there are competing factors make the Pilot both less efficient (lower interviews per assigned hour) and more efficient (less variable weights, as mentioned on p. 54). It isn't clear (to me) how to do this precisely from the data, but even some rough measure of "variance per sampler hour" would be helpful.

Another analysis that would be useful to see would be a prediction of how much the design effect could be improved after implementation of some of the suggested improvements in allocation to the various strata. It seems that this could be calculable for at least a few major species or total catch, to give an idea of the potential value of the new method.

And of course this doesn't even consider the possible residual biases in the MRFSS data. The similarity of both sets of estimates would seem to suggest that there is little bias left in the newly weighted MRFSS, but then we get statements like this one on p. 8 :"… the potential for non-sampling errors was much larger under the MRFSS…" So does the statement on p. 8 refer to the "old MRFSS" or does it mean that even the "new MRFSS" is believed to still have substantial non-sampling errors. If the latter, it would be helpful to briefly describe what kinds of non-sampling errors the authors think still remain. Presumably, the authors feel there are smaller non-sampling errors than with just the original MRFSS estimators. Is that the interpretation of the authors? I find that surprising, although it doesn't mean that even if it was true in NC in this year it will always be true. Does this mean then that the average catch per trip really does not vary much by time of day (night vs. day), or at alternative sites in the same vicinity (chosen by MRFSS samplers)? Or is it that the precision was so low we couldn't tell there was a difference? More discussion of what evidence there is about this would be helpful to explain the advantage of the new design to those who might not be so convinced of the superiority of a strictly probability sampling design on theoretical grounds alone.

**Response:** We appreciate the careful review and insightful comments provided by the reviewer. We agree with the reviewer that plots and figures need better interpretation to convey the message we tried to communicate to the reader. We will revise some of these paragraphs as suggested by the reviewer. We will also revise the reference to the MRFSS and "old MRFS" to be consistent through the report.

The MRFSS sampling design suffers from an inherent space and time undercoverage resulting from the discretion the interviewers have in the MRFSS to select the interviewing time and alternate sites. This undercoverage was removed in the new sampling design by incorporating stratification by mode, kind of day and time interval and clustering of sites.

The goal of the comparisons discussed in the report was to learn specific aspects to be improved in a future implementation of the new sampling design. For example, we need adequate sample allocation across modes and time intervals. Given the difference in number of assignments (sample size) between the MRFSS in the pilot study, we expected the precision of the estimates of the MRFSS to be sometimes better than those obtained from the pilot. An implementation of the new sampling design will result in larger number of assignments for each state which will likely result in more precise estimates than those obtained in the pilot study. We will add some paragraphs to clarify this goal in the report.

**17.** ***Interpretation of results:***

I would like to see more interpretation of results, or explanation of most likely causes, when there are differences between MRFFS and Pilot results. Here are some places where that would be informative:

a. ***Effort ratio estimator (p. 8)***:

The summary states that the proportion of anglers reachable by the CHTS is comparable for all but beach/bank mode. Is there any intuition or speculation about why that would be true? Also, I believe it is true that an angler is not in the CHTS frame if he or she does not have a landline phone. If that is correct, it should be noted as part of the description of "the effort ratio estimator."

**Response:** There is not enough information to fully explain the observed differences in the beach/bank mode discussed in page 8. Possible causes are the removal of the incomplete trip interviews and the inclusion of nighttime sampling. Local residents may fish for longer durations and may be more likely to fish at night than the out of state residents or non-coastal state residents. We will add text pointing out these possible explanations and suggest that further study is warranted.

b. ***Average number of interviews obtained per two-hour interval (p. 37):***

Do you have any speculation on why the jagged curve for the pilot? Is it just because of the small sample size, or could it be because of the rules for clusters that have people driving from one place to another at certain hours?

**Response:** We thank the reviewer for bringing to our attention that we overlook explaining the
jagged curve depicting the frequency of intercept per hour (Figure 3). We agree with the reviewer that the up and down pattern of the curve for the pilot might correspond to the driving between sites. We will add a sentence to the corresponding paragraph commenting on this.

c. ***Average numbers of fish reported and observed (p. 38):***

It is hard to know what to make of the differences seen in Table 6. What is the difference between reported and observed? Were these fish chosen because they are ones that are particularly common? How is the average # of fish computed? Is it average per angler overall, or average per angler that targeted or has any of that fish? Do these differences seem reasonable or explainable (e.g., are some fish caught more at night, for example, and so may be more likely to be seen in the Pilot?)

**Response:** We agree with the reviewer that adding a sentence explaining the definition of the '% Difference' column shown in Table 6 (page 38) will help with the interpretation of the table results. We will add few sentences in the paragraph before the table to clarify that 1) the species showing in the table were selected because these are the most common species in North Carolina, and 2) the average was calculated as the average number of fish among anglers who caught that species.

***d. Differences in catch estimates (p. 47):***

The fact that the MRFFS estimates are higher when there is a difference is described as "interesting." Is there a reason to believe this is a real difference and indicates a bias on the part of MRFFS data, due to either the noncoverage of night fishing or some other reason?

**Response:** We will add a sentence in the report to discuss some of the reasons that might have resulted in the observed differences.

***e. Percent of beach/bank trips by coastal residents (Table 7):***

Does it make sense that these estimates would differ due to noncoverage of night fishing in MRFFS, or is there some other reason? Maybe only locals fish from the beach at night?

**Response:** We appreciate the comment from the reviewer. Although non coverage of night fishing in the MRFSS is a possible cause for the observed difference, the data we obtained in the study did not show a significant difference in the proportion of coastal residents for sampled nighttime and daytime beach/bank trips.

***f. Estimation of total catch:***

On p. 7 and again on p. 38 are mentions that the telephone survey (CHTS) is used along with the intercept data to come up with total catch estimates. I believe these descriptions are too cryptic for a reader who is not already familiar with the estimation method to understand. Even a brief explanation that total catch is estimated as (total trips by coast county anglers) *(catch per trip for all anglers) *(# of all anglers/# of coastal county anglers) would be helpful, and would help elucidate some of the comments about why bias can result if certain assumptions (e.g., avg. catch the same for coastal and noncoastal) do not hold.

**Response:** We thank the reviewer for this comment and will add some sentences on page 38 (and will make a reference to it in page 7) to explain the calculation of total catch.

**18.** ***Clarifications Requested:***

***a. Which version of MRFSS? (p. 4):***

When the report says that the Pilot compared the performance of the new sampling design and estimates of catch per trip, etc. with the "traditional MRFSS sampling design" (p. 4), does this mean that the comparisons are with the old design but the new estimators developed by Breidt et al.? The remainder of the report seems to suggest the comparisons are always to the new analysis method for MRFFS, but the use of the word traditional made me wonder as I was reading the executive summary.

**Response:** We agree with the reviewer that we need to be consistent through the report when mentioning the "traditional MRFSS sampling design" which refers to the current MRFSS. The estimates for the traditional MRFSS used in the comparisons were obtained by applying the new estimation procedures to the data collected using the old design. We will revise the report accordingly.

***b. Number of sites per assignment (p. 8):***

How does the new design obtain nearly 3 sites per assignment (2.46, from p. 8)? Does that mean that most assignments are 3-site clusters? It might be good to have a summary somewhere of how many units in the frame are of each size.

**Response:** We will add few sentences describing the frequency of clusters with 1, 2 and 3 sites visited during the pilot study.

***c. What is a "positive" assignment? (p.10 and p.61):***

What does the word positive mean in this statement on p. 10 "provide higher average number of interviews per positive assignment." Also on p. 61.

**Response:** The term positive assignment refers to an interviewing assignment that obtained at least one interview. We will revise the corresponding sentences in the report.

***d. Recommendation on site clustering rules (p. 11, recommendation #6):***

It is not clear in what way cluster formation would be modified. More two-site clusters? Why would that happen?

**Response:** When clustering the sites, we used a GIS algorithm that clusters sites in the sampling frame by minimizing the driving distance between sites while maintaining similar size measures (effort) across clusters. This algorithm allows for high activity sites to be placed in a single site cluster and to combine nearby sites so the total effort across clusters are similar. In the pilot we didn't enforce the constraint of achieving a total pressure across all sites, so sites were clustered based on proximity and having a pressure lower than a given threshold. If a threshold level of pressure was used to determine whether or not a given site could be clustered with one or two additional sites, more two-site clusters would be allowed. This could reduce the average driving time and increase average interviewing time per assignment. We will revise the explanation of this recommendation to clarify.

***e. Site list maintenance (p. 11, recommendation # 10):***

Apparently there was some indication that the site list had some flaws. Could these be clarified?

**Response:** The maintenance of the site registry is an ongoing task that needs to capture trends in space and time. Size measures for some sites might not be reflect the actual fishing activity for given modes or site status (e.g. inactive and active) may not be updated. We are currently working on obtaining accurate size measures (effort) for the sites for all modes and time intervals; including new sites (e.g. new piers, private piers, etc) and making sites inactive based on the accessibility or closings due to maintenance or for ecological reasons (e.g. bird nestings). We will add few sentences explaining the nature of these maintenance tasks.

***f. Explicit vs. implicit stratification (p. 12 Recommendation #12):***

Clarify what is meant by “explicit” and “implicit”

**Response:** We will add few sentences and examples to illustrate the concept of explicit and implicit strata.

***g. One psu per month? (p. 23):***

[The text reads] “at least one psu was selected from each interval…” Clarify that this means “per month”

**Response:** We thank the reviewer for detecting the missing “per month” in page 23. We will add these two words to the corresponding sentence.

***h. X-axis in Figure 2 (p. 36):***

The x-axis scales are different on these plots. I don’t think they should be. In fact, why should n’t they all be from 0 to 24? Redraw graphs using 0-24 x-axis.

**Response:** We will redraw the plots and use the same 0-24 range in the x-axis.

***i. Meaning of “fishing area”(p. 39):***

I don’t’ understand the difference in the last two sentences of the paragraph before Table 7.They seem contradictory, but then I don’t know what “fishing area” means.

**Response:** Fishing area refers to location of the fishing site (ocean within 3 miles, ocean outside of 3 miles, and inland) as described in the last sentence of the first paragraph on page 39. We will add few sentences to clarify this definition and will review the two sentences referred by the reviewer to better convey the message to the readers.

***j. Figure labeling (Figures 4 and 5):***

Figures 4 and 5 should have a label of Pilot and MRFFS to be consistent with the other descriptions, rather than NC and MRFFS.

**Response:** We thank the reviewer for the detailed review of the report. We will change the word “NC” for “Pilot” in both figures to be consistent across the report.

***k. Future recommendation #1 (p. 10):***

I do not understand the meaning of #1 under future recommendations.

**Response:** We will add few sentences in the repot to clarify recommendation #1 to consider obtaining total counts of boat trips in addition to total counts of angler trips.

***l. “Inappropriate” sampler choices (p. 53):***

The word *inappropriate* seems wrong. It is not that samplers are making bad choices, but that they are making choices at all that is wrong with MRFSS.

**Response:** We agree with the reviewer that the word “inappropriate” seems wrong in the sentence and paragraph context. We will rephrase the sentence to make emphasis that is the discretion the interviewers have to select alternate sites in the MRFSS is the main issue affecting the MRFSS.

***m. Frame maintenance assignments (p. 69):***

The need for frame maintenance assignments is mentioned a couple of times, but no information on the reason for this suggestion is mentioned. Was there some problem with the list of sites that was discovered during the pilot?

**Response:** The site registry (frame of the MFRSS and new sampling design) must be updated to reflect the space and temporal trends on fishing activity across the different modes. New sites must be added, size measures (effort) must be updated, contact information, site closures and security issues must be updated to better inform the sample selection and the site clustering exercise. We will provide information on the rationale for this recommendation. We did not discover any specific problems with the list of sites used in the pilot study. We just think that the completeness of the list and the accuracy of the site pressure estimates in it would be improved by allocating more resources to frame maintenance tasks.

***n. “Probabilistic” sampling? (p. 5):***

Probability sampling is referred to as probabilistic sampling on p. 5. Is this meant to imply something different than what we usually mean by probability sampling?

**Response:** No. we were referring to probability sampling. We will make the changes in the report.

*o. Quotes? (p. 15):*

Why are there quotes around the statement under item #3? I presume because it is quoting from the previous report, but it should say so.

**Response:** We will add the corresponding reference to item #3.

*p. Night sampling and the percentage of coastal county anglers (pp. 52, 53,55 and several other places):*

There are comments about the fact that one advantage of the Pilot covering all times is that it eliminates potential bias due to different catch rates at night. But it also eliminates potential bias caused by different coastal county anglers' percentage at night. I believe this should be emphasized.

**Response:** The reviewer is correct that we need to emphasize that with the new sampling design we will also eliminate potential bias by different coastal county anglers' percentage at night. We will add a sentence in the corresponding paragraphs.

*q. Overtime for samplers? (p. 52):*

"It was not necessary to require samplers to regularly work overtime." Was this a problem with MRFFS? If so, then state this advantage.

**Response:** This was not a problem with MRFSS sampling. We were simply pointing out that it was not necessary in the pilot study to require any samplers to work more than a 40-hour week to complete the assignments drawn under the new sampling design.

**19.** ***Other comments:***

*a. Fishing trips ending on private property:*

Executive summary and p. 13 mentions that the MRFFS survey was "…not providing coverage of fishing trips ending on private property or at night." Much discussion was devoted to the issue of improving night fishing, but no other discussion of private property was included in the report. It seemed a bit like a "bait-and-switch" as I kept looking for discussion of innovations for the private property issue. Maybe this should be removed from the overview.

**Response:** We thank the reviewer for the careful review. The site registry which is the sampling frame for the MRFSS and the new sampling design is affected by possible under coverage of private landing sites (e.g. piers, marinas, etc). We will add a statement in both the summary and page 13 clarifying this issue.

***b. Recommendation #11 for future consideration (p. 12):***

The idea of basing pressure measures on # of fish landed seems a good one, especially if this allocation is very different than the one being used. It would be even better if the variability in fish landings could be used, though maybe mean and variance of fish landings are positively correlated.

**Response:** We appreciate the reviewer's comment and suggestion on our intent to explore new ways to obtain better size measures. As we move on with the new sampling design we will explore the impact of different size measures (e.g. # of fish landed) or using the variability in fish landing (as suggested by the reviewer) on sample size selection and sample allocation.

***c. Recommendation #8 for future consideration (p. 68):***

I like this idea a great deal. In fact, the potential for cost savings if the entire effort estimation were to be moved to the intercept survey is large. With the new design, my guess would be that number of trips could be more accurately estimated from the intercept survey than from telephone or mail, due to the large measurement error and nonresponse of the effort survey. The only problem I see is the private access sites. How much of the total effort is on private access sites? Maybe a telephone survey would still be necessary but only to estimate this ratio rather than the current coastal county ratio estimated from the intercept. Why not just compare total public access estimates from telephone and the intercept survey rather than needing to estimate by site (thought this is perhaps not a discussion for this report)? If the intercept survey were to be used for both effort and catch, one might consider counting both arriving and departing anglers, since the spatial-temporal time frame would be appropriate for either. The two estimators would be correlated but not identical, so should provide a little extra information basically for free.

**Response:** We agree with the reviewer that in order to get information from the private access sites, other means of data collection (such as a phone or mail survey) in addition to the intercept survey is needed. As we move on with the sampling design we will explore the reviewer's suggestions and other approaches to obtain better estimates of effort.