Review of MS Tails 'n Scales Program Proposed for MRIP Certification

Jay Breidt (Colorado State University), Mike Brick (Westat), Ginny Lesser (Oregon State University), Jean Opsomer (Colorado State University), Lynne Stokes (Southern Methodist University)

January 9, 2017

We received the following documents describing the MS Tails 'n Scales Program:

1. Red Snapper Certification doc 7.7.17 FINAL.docx
2. MDMR Certification Bullet Points Document_Final.docx

Our review will be based on these materials and refer to them as needed. Below, we address each of the terms of reference.

1. *Does the survey design follow a formal probability sampling protocol with known inclusion probabilities at all stages and/or phases of sampling?*

   The data collection consists of two components: the Tails 'n Scales trip reporting system, and the dockside intercept survey. The first consists of a required "pre-authorization" by one angler per trip, followed by reporting of trip and catch characteristics after trip completion. The reporting is greatly enhanced by enforcement mechanisms, both on-the-water interception by law enforcement and the inability to sign up for a trip unless the previous trips is closed out or canceled. From a compliance perspective, this is one of the better programs we have seen among the several self-reporting systems for recreational fisheries we have reviewed.

   The second component of Tails 'n Scales consists of a validation survey, conducted through a random sample of on-site intercepts of returning trips during the relevant fishing seasons. Data from this survey are used to construct capture-recapture type estimators for the total number of trips targeting red snapper and their associated total catch. Considered in combination, these two data collection mechanisms are a valid approach to collect statistical data in this fishery.

   Both Tails 'n Scales self-reporting and the intercept survey are well documented, and the latter has randomization protocols that appear appropriate. However, as the PPS sampling procedure is not described in full detail in the documentation, we are not able to fully assess it at this point. Specifically, it is not clear to us how the PPS sampling is performed across both sites and time slots (i.e. how are the site and time pressures combined) to select individual assignments. This should be clarified further to ensure that the weighting and variance estimation procedure correctly reflect the sampling design.

   The PPS design as currently implemented appears statistically valid, but the very large range of pressures shown in Table 3A of (1) might lead to estimates that are highly variable. One approach to reduce this problem is to first stratify the sites into pressure categories, and then select samples using PPS within these categories. The allocation to the strata does not have to be proportional to the number of sites in a stratum nor to their total pressure, but the latter is

certainly a good place to start in considering sample allocation for this program. Depending on the sampling intensity and budget, it is also possible to have a small number of "certainty sites" for the highest-pressure ones.

2. *Do the estimation methods appropriately weight the sample data to account for the sampling design and produce design-unbiased point estimates and variance estimates?*

   Following up on the previous comment, we are not able to tell from the provided documentation how the PPS design was implemented across sites and time blocks, and hence how the weights were obtained. Information on how to improve the design through stratification would be obtained from an examination of the pressures and weight distribution, We have no reason to believe it is not done correctly, but it would still be useful for documentation purposes to more fully describe it.

   On a similar topic, the documentation mentions that if a specific site/time is selected more than once, two samplers are sent to account for the expected large number of returning trips. This is certainly appropriate, but how is this assignment handled in weighting? Specifically, is it also given a double weight?

   Assuming there are no weight construction issues, the capture-recapture estimator described in the documentation is indeed appropriate. Particular strengths of the Tails 'n Scales program are the high compliance rate and the fact that the issue of matching of self-reported trips and intercepted trips is mostly avoided, since it can be done based on the trip permit number. Other estimators are possible for combined self-reporting and survey intercept estimation of effort and catch, as have been recently explored in a pilot project for the charter fishery in South Carolina. In that project, an additive adjustment was found to be a better option than the capture-recapture one implemented so far for Tails 'n Scales, so it might be worthwhile to evaluate it here as well. However, given the high quality of the record matching and the high compliance rate in Tails 'n Scales compared to those encountered in South Carolina, we expect these different estimators may lead to similar estimates. It would be of interest to conduct this comparison in order to examine the usefulness of the additive adjustment with this design.

   On p.14 of (1), an adjustment for sites with zero intercepts is mentioned but not fully described. If the number of sites with zero intercepts is small, no adjustment is likely to be needed in the general estimation procedure. If this occurs at a non-trivial number of sites, then an adjustment might indeed be warranted and using some type of larger-area average is reasonable. It might again be worthwhile to document this in more detail, so that it can be more fully assessed.

3. *Are appropriate methods in place to measure and/or correct for potential biases due to undercoverage, nonresponse, or response errors?*

By construction, the combination of mandatory self-reporting and randomized intercepts is designed to correct for undercoverage and self-reporting errors. Especially with high participation in the self-reporting component, this should lead to high quality estimates for the MS recreational red snapper fishery. Other issues, such as angler nonresponse in the intercept survey, possible differences between private and public sites, are present in most other MRIP surveys, so acknowledging them is sufficient at this stage.

As noted in the documentation, there are a number of additional auxiliary data sources that are unique to MS and, while not necessary as part of the basic estimation procedure, can provide further insights in some of the sources of non-sampling errors. These include the home visits to a random sample of anglers returning to private sites, the on-the-water law enforcement intercepts, and the flight counts of angler vessels.

4. *How sensitive is the accuracy of the survey to assumptions made about segments of the target population that are not covered by the survey frame? What can be done to reduce or limit that sensitivity?*

   The undercoverage due to vessels returning to private sites is inherent in this intercept survey, as it is in the APAIS, even though at an estimated 30% of the trips, it might represent a larger fraction in MS than in many other states. If these trips are different in either their catch characteristics or in their compliance behavior, then this might indeed lead to bias. However, the combination of mandatory pre-approval for trips and on-the-water enforcement makes it likely that the latter factor will have at most a minor impact. Regarding possible differential catch reporting, the home visits mentioned above will provide some information on this issue even if it is unlikely to allow for estimation of the magnitude of biasing effects.

5. *How sensitive is the accuracy of the survey to other potential sources of nonsampling error? What can be done to reduce or limit that sensitivity?*

   Because trip matching will be done based on individually issued authorization numbers, matching errors, often a major source of non-sampling errors, is not present in MS. Requiring that these numbers be issued prior to going on a trip also greatly reduces the potential dependence between the “capture” (self-reporting) and the “recapture” (intercept) events. Overall, the Tails ‘n Scales program appears to be a very good way to avoid several of the key complications present in implementing survey-based capture-recapture surveys.

6. *How sensitive is the survey design to potential errors in implementation? What can be done to evaluate, reduce or limit that sensitivity?*

In most survey programs, careful and accurate implementation of the stated procedures is an essential underpinning of the quality of the resulting estimates and associated measures of precision. Because this program includes two separate components that require matching at the individual trip level, this is certainly the case here. But the mandatory pre-authorization, the high level of enforcement and issuing of penalties, the public relations campaigns are all aspects of the program that greatly improve its overall quality. As already noted, trip matching by authorization numbers bypasses one of the major hurdles in implementing capture-recapture surveys. In addition, the high level of compliance ensures that the resulting estimators are expected to be efficient.

The main aspect we were not able to fully evaluate is whether the PPS design and its associated weighting procedures are properly implemented, because the documentation was incomplete in this area.

7. *How does the survey design compare to the survey design it would replace or supplement? Is it more statistically sound and efficient, or is it at least comparable in its statistical validity and efficiency? What design features are most important in supporting this assessment?*

   The Tails 'n Scales Program is intended to provide more data on red snapper catch and to do so in a more timely manner than the general-purpose APAIS and CHTS (or FES). All indications are that this program can indeed achieve these purposes. The statistical methodology underlying the combined estimators is not in doubt, because it can be explained using standard design-based theory. It might be useful to evaluate alternative forms of the combined estimator, but this can be done at a later time and does not change the data collection and survey design aspects of this program.

8. *How does the survey design compare with other survey designs previously certified by MRIP for estimating fishing effort and/or catch for the same fishing mode(s)? Is it more statistically sound and efficient, or is it at least comparable in its statistical validity and efficiency? What design features are most important in supporting this assessment?*

   The intercept component of the Tails 'n Scales is very similar to the APAIS, which has been used as the standard MRIP catch data collection approach. The self-reporting component is not yet part of any certified programs, although a similar program in AL is currently undergoing review. Matching and estimation procedures have been pilot-tested in South Carolina and are also being implemented in AL. We recommend that these various programs coordinate future efforts in implementation of their statistical methods and try to harmonize their approaches, to facilitate data integration and comparisons across states and regions.

9. *Is the survey collecting data and producing information products that will meet the needs of the primary customers (stock assessment scientists and fishery managers)? [To be addressed by NMFS staff.]*