Documentation and Information

APAIS Calibration and Review

The collection of documents contained in this pdf are only a small part of the resources available on the APAIS calibration and review workshop website. That website also includes recorded webinar presentations given during the review workshop for the calibration models. That information can be found at the address below.

https://www.fisheries.noaa.gov/event/access-point-angler-intercept-survey-calibration-workshop

APAIS data calibration methodology report

John Foster
National Marine Fisheries Service

F. Jay Breidt
Colorado State University

Jean D. Opsomer
Colorado State University

March 11, 2018

1 Background

In 2013, new design and estimation procedures were implemented for the Access Point Angler Intercept Survey (APAIS) conducted by the National Marine Fisheries Service (NMFS). The new procedures were introduced as part of on-going efforts to improve the statistical validity and reliability of the recreational marine fisheries estimates produced by NMFS, which followed the recommendations of a National Academies of Sciences panel review (Sullivan et al. 2006). The most important design changes include improved protocols for interview assignments in terms of interview sites and times of day, and changes to the randomization of assignments so that they better covered the target population, again in terms of sites and times of day. Associated with those design changes were changes to the estimation methods, which are now fully weighted to reflect the unequal probability sampling design. APAIS data collected since March (wave 2) 2013 follow the new design and estimation procedures.

APAIS data have been collected since 1981, and NMFS staff clearly recognize the importance of preserving the integrity of the time series of catch estimates despite these design and estimation changes. Because of this, an adjustment procedure was developed to create “pseudo-weights” for APAIS data collected between January 2004 and March 2013. These weights were constructed based on a combination of site pressures and empirical site visit frequencies, and on estimated expected fractions of trips that took place during the time the interviewer was on site relative to the daily total number of trips. Weighting the observed trips on a given site-day assignment by the inverse of this estimated fraction was meant to correct for differential representation of sampled trips within site-days. The fractions of trips were predicted by a small area estimation model fitted to data from the

Coastal Household Telephone Survey (CHTS), see Hernandez-Stumpfhauser et al. (2016). The combination of modeled site-day selection probabilities and within-day probabilities resulted in weights that better reflect the population of trips during the period 2004-2013, with respect to its overall size and distribution across states, waves and modes.

However, an implicit assumption underlying the validity of this approach is that the trips occurring during the time period the interviewer is on site are representative of those that take place during the full day. This is satisfied if either the time on site is randomly selected within the day, or the trip characteristics are not related to the time of day. The first condition was definitely not satisfied, because the large majority of site visits were made at what was considered the busiest time of the day and were also subject to a degree of interviewer discretion. The second condition appears not to be satisfied either, according to analysis reported in a technical report (see MRIP Staff 2014). Hence, there is a need to supplement the weighting procedure that is based solely on fraction of daily trips within selected site-days by a procedure that accounts for differences in trip characteristics between those that were observed during the site visit intervals and those outside of it.

There is also a desire to adjust the time series for the period prior to 2004. For that earlier period, not only are the selection probabilities within site-days unknown as above, but information allowing the construction of site-day visit probabilities is incomplete or missing, with the required design information becoming progressively more limited going back in time. Further complicating matters, the sampling procedures, including site selection and sampling intensity, underwent changes during that period, and documentation for these changes is no longer available. Hence, separate procedures are needed to calibrate estimates prior to 2004.

Correcting time series of survey data following changes in design, data collection and/or estimation methods is a challenging statistical issue. The “gold standard” approach involves conducting side-by-side measurements under the old and new methods, fitting a suitable calibration model relating estimates under both methods, and developing and applying adjustment factors based on the model results. This approach is currently being implemented by NMFS to calibrate the trip estimates obtained under CHTS and its replacement survey, the Fishing Effort Survey (FES). See NMFS Staff (2015) for more details on the CHTS and FES surveys and the transition between them.

While explicit statistical calibration would in principle be attractive for the APAIS time series as well, there are a number of reasons why that is not possible. First and most critically, there is no overlap period between the old and new designs, so that the data needed for fitting a calibration model are not available. Second, unlike CHTS and FES, which primarily involve estimating the total number of trips for a given region and time period, APAIS is used to produce numerous different estimates, covering a wide range of trip characteristics and detailed catch by species, location and type. Hence, even if an overlap period were available, it is not clear that statistical calibration would be feasible, since multiple models would likely be required for the different types of variables.

For these reasons, the proposed APAIS calibration will rely on a weight adjustment approach, which is conceptually similar to the pseudo-weight approach described above for the 2004-2013 data. By adjusting weights rather than modeling the estimates themselves, the data collected prior to 2013 are preserved but their weights are suitably modified so that the distribution of trips better reflects the actual population distribution. By incorporating the calibration adjustments into the survey weights, the historical data can continue to be made available as survey public-use (micro) datasets, greatly facilitating their acceptance among the current data users.

2 Adjustment approach for 2004-2013 data

We first consider the adjustment of the weights for the period 2004-2013 (wave 1). Because there is no overlap period between the old and new designs and the CHTS contains only limited information on trip characteristics, no direct comparison distribution is available on which to calibrate. Instead, calibration will be performed using the trip distribution for the period 2013 (wave 2)-2016 as the target distribution. This is reasonable if the mix of trip characteristics has remained constant over time, at least over the periods being considered. The validity of this assumption cannot directly be assessed, because differences in observed trip characteristics before and after 2013 can be explained by both the design and estimation changes as well as by actual changes in the fishery. However, we will modify the proposed method in situations in which we observed a significant "drift" in important trip characteristics over time, see Section 3 below. For now, assume that it is reasonable to work under the assumption that differences in trip characteristic distributions between the periods 2004-2013 and 2013-2016 are likely primarily due to the design and estimation method changes. Hence, the weight adjustment method will calibrate the weights for trips in 2004-2013 (wave 1) to the weight distribution for 2013 (wave 2)-2016.

The key decision in the proposed method is which trip characteristics to adjust for. Following the analysis results shown in MRIP Staff (2014), the following trip-level variables were identified as both important trip characteristics and ones for which the distribution in the data collected prior to 2013 deviated from those under the new methods:

- state and sub-state region (if applicable)
- year and wave

- mode
- area fished
- coastal/non-coastal household
- for-hire boat frame membership.

The values for each of these variables defines categories of trips. For instance, there are 4 modes (shore, private boat, headboat, charter) and each trip belongs to one of those modes. Taken in combination, the values for these variables define large numbers of trip domains. Under the new design and estimation approach, by summing up the weights of the trips corresponding to a given set of values for these variables, we obtain an estimate of the number of trips of that type.

As an example to explain the adjustment procedure, let $U_{D,2014}$ present the domain of all trips that occurred, say, in a particular substate region in Florida during wave 2 on a private boat by a coastal household in state waters, in 2014. The true total number of such trips that took place is equal to $N_{D,2014}$. It is unknown but it can be estimated based on APAIS intercepted trips under the new design and estimation methods, by $\widehat{N}_{D,2014} = \sum_s w_i I_{\{i \in U_{D,2014}\}} = \sum_{s_{D,2014}} w_i$. This can be repeated for any combination of values of the classification variables. However, while statistically valid, these estimates are likely to be quite variable for some of these domains because they contain only small numbers of observed trips.

Likewise, we can compute estimates for the same domains for years prior to the design change, e.g. 2012: $\widehat{N}_{D,2012} = \sum_{s_{D,2012}} w_i$. This estimate might not be valid, however, because of the recognized shortcomings of the design and estimation methods in effect at that time. If $N_{D,2012}$ were known, we might therefore decide to adjust the weights so that they sum up to $N_{D,2012}$. This is readily accomplished by replacing all w_i for $i \in s_{D,2012}$ by

$$w_i^* = \frac{N_{D,2012}}{\sum_{j \in s_{D,2012}} w_j} w_i. \tag{1}$$

The weights w_i^* in sample domain $s_{D,2012}$ now sum to the new control total $N_{D,2012}$, and can be applied to any variable y_i collected in the survey. This type of calibration to known control totals is commonly applied in surveys, to improve the precision of estimators.

Since we do not know $N_{D,2012}$, implementing this ratio-type adjustment requires that it be replaced by a sample-based quantity. As noted above, we propose to use estimates based on the data collected under the new design since 2013. In order to reduce the variability of the control total estimates and also because individual years are not meaningful targets (i.e. we are not interested in adjusting 2012 weights to match the 2014 totals, but rather, adjust pre-2013 years to post-2013 years), both the control targets and the adjustment ratios are averaged across years. Hence, the unfeasible adjustment in (1) is replaced by

$$w_i^* = \frac{\widehat{N}_{D,\text{new}}}{\widehat{N}_{D,\text{old}}}\, w_i, \tag{2}$$

where $\widehat{N}_{D,\text{new}}$ is the average of $\widehat{N}_{D,2013}$, $\widehat{N}_{D,2014}$, $\widehat{N}_{D,2015}$ and $\widehat{N}_{D,2016}$ (with the first of these omitted if the domain is in wave 1) and $\widehat{N}_{D,\text{old}}$ is the average of $\widehat{N}_{D,2004}, \ldots, \widehat{N}_{D,2012}$ (and $\widehat{N}_{D,2013}$, only if the domain is in wave 1). Unlike the (unfeasible) adjusted weights in (1), the weights w_i^* do not sum to a control total for a particular year. Instead, they correct for the overall under- or over-representation of trips in domain U_D under the old design and estimation methods relative to the new methods implemented since 2013, which is expected to lead to improved estimates for variables of interest that are related to the domain that is being adjusted.

While averaging the adjustment ratios across years as in (2) reduces their variability, the fine definition of the domains (as intersections of numerous control variables) is still expected to lead to unreliable adjustments in many domains. Therefore, the full ratio adjustment in (2) is replaced by a *raking ratio* adjustment, originally proposed in Deming and Stephan (1940) and widely used in survey calibration. The motivation for this procedure is that instead of adjusting at the finest domain level, adjustments are made iteratively on a set of a coarser domains. These coarser domains are determined by a subset of the variables mentioned above. For each of them, it is possible to compute the averages of the annual estimates as described for $\widehat{N}_{D,\text{new}}$ above. We denote the ones we use in our adjustment procedure as follows:

- AF (state, wave, mode and area fished): $\widehat{N}_{D,\text{new,AF}}$
- HS (state, wave, mode and coastal/non-coastal household status): $\widehat{N}_{D,\text{new,HS}}$
- FH (state, wave, mode and for-hire boat frame status): $\widehat{N}_{D,\text{new,FH}}$
- RE (state, wave, mode and substate region): $\widehat{N}_{D,\text{new,RE}}$

While not explicit in this notation, for each of these domains, the averages are for each of the categories of these variables. So for instance, $\widehat{N}_{D,\text{new,AF}}$ are averages of estimates for each state-wave-mode-area fished combination, and so on for the other domain definitions above.

The raking ratio algorithm, also sometimes called iteratively proportional fitting, then proceeds as follows:

1. Initialize: set $t = 0$, set the adjusted weights $w_i^{(t)}$ equal to the initial weights w_i for the period 2004–2013 (wave 1), and compute the $\widehat{N}_{D,\text{new,AF}}$, $\widehat{N}_{D,\text{new,HS}}$, $\widehat{N}_{D,\text{new,FH}}$ and $\widehat{N}_{D,\text{new,RE}}$.

2. Let $\widehat{N}^{(t)}_{D,\text{old,AF}}$ be the averages of the estimated AF domain totals for the period 2004–2012 (include 2013 for wave 1) using weights $w_i^{(t)}$, compute the ratios $R^{(t)}_{\text{AF}} = \widehat{N}_{D,\text{new,AF}}/\widehat{N}^{(t)}_{D,\text{old,AF}}$, and set $w^{(t)}_{i,\text{AF}} = R^{(t)}_{\text{AF}} w_i^{(t)}$.

3. Starting from the weights $w^{(t)}_{i,\text{AF}}$, do the same as in 2 for the HS domains, resulting in ratios $R^{(t)}_{\text{HS}}$ and weights $w^{(t)}_{i,\text{HS}}$.

4. Starting from the weights $w^{(t)}_{i,\text{HS}}$, do the same as in 2 for the FH domains, resulting in ratios $R^{(t)}_{\text{FH}}$ and weights $w^{(t)}_{i,\text{FH}}$.

5. Starting from the weights $w^{(t)}_{i,\text{FH}}$, do the same as in 2 for the RE domains, resulting in ratios $R^{(t)}_{\text{RE}}$ and weights $w^{(t)}_{i,\text{RE}}$.

6. Set $w_i^{(t+1)} = w^{(t)}_{i,\text{RE}}$.

7. Repeat steps 2–6 until convergence, which is evaluated by measuring the change in the ratios $R^{(t)}_{\text{AF}}, R^{(t)}_{\text{HS}}, R^{(t)}_{\text{FH}}, R^{(t)}_{\text{RE}}$ for different t. Set the final adjusted weights w_i^* equal to the iterated weights $w_i^{(t)}$.

This raking ratio procedure ensures that the weights w_i^* are adjusted to match each of the "marginal" raking variables (AF, HS, FH, RE), but not the fine domains defined by the combinations of these raking variables. This prevents adjusting to overly small domains, with associated overfitting and weight instability issues.

3 Modification for temporal changes in fishery characteristics

We now return to the assumption of constant trip characteristics over time. As noted, the raking procedure described in Section 2 is based on the assumption that the estimated trip distribution since 2013 is a reasonable target for the trip distributions prior to 2013. However, if the trip characteristics in the fishery have changed over that time period, observed differences between the pre-2013 and post-2013 periods are likely to be due to a combination of the design-estimation changes and actual fishery changes. Raking as in Section 2 in this situation will result in a weight adjustment that is too large, because it will remove both the design-induced change and the actual fishery change. We therefore implemented a two-step procedure to decrease the risk of over-adjusting the weights, described in this section.

Consider a single set of control domains first, say AF above. Prior to raking, for a given state, mode and area fished, we create a dataset containing the estimated domain totals for each year and wave combination between 2004 and 2013 (wave 1), resulting in a time series of 145 data points. There are multiple such time series, for each combination of state, move and area fished. We perform a simple linear regression of the totals against a time index for each time series, and test whether the slope is significantly different from zero at the 97.5% confidence level. If the null hypothesis cannot be rejected for a given time series, we maintain the raking adjustment described in Section 2 for the AF domains. If the null hypothesis is rejected for a time series, step 2 in the raking algorithm is modified for that particular control domain, so that only the years 2010–2013 (wave 1) are used in the computation of $\widehat{N}^{(t)}_{D,\text{old},\text{AF}}$. Hence, the AF ratio adjustment $R^{(t)}_{\text{AF}} = \widehat{N}_{D,\text{new},\text{AF}}/\widehat{N}^{(t)}_{D,\text{old},\text{AF}}$ is only based on the most recent years instead of the full time period, in those domains for which a significant time trend is detected.

The same testing and modifications are applied to the remaining three control domains (HS, FH, RE). The full adjustment procedure that accounts for temporal trends therefore consists of the linear regression tests followed by shortening of the time period used for computing the ratio adjustments in any of the control domains for which a non-zero slope is detected, following by the raking algorithm.

4 Adjustments for prior periods: 1993-2003

Weight adjustments for data collected prior to 2004 were performed following the computation of the adjusted weights for 2004-2013 (wave 1). The major difficulty for the earlier periods was that unlike for 2004-2012, it is not possible to construct meaningful initial sample weights for the APAIS data. As such, the weight adjustment method described in sections 2 and 3 could not be applied directly and needed to be extended to address effects of the 2013 APAIS design change as well as any effects associated with initial weighting of the 2004-2012 APAIS data. Using 1 as the initial base weight for intercepted angler-trips was not adequate as the sample sizes, in terms of sampled site-days, were known to vary considerably over time. Unfortunately, the exact sample sizes were unavailable for these earlier years.

It was decided to divide this period in two pieces overall. This provides a hedge against incorrect time trend adjustments masking actual changes in the fishery, as well as unaccounted-for changes in design. This is similar to the argument in Section 3, but was applied globally prior to any further adjustments. Hence, we performed the adjustments for 1993-2003 and 1981-1992 separately.

Considering first 1993-2003, we investigated two approaches for creating initial weights. In a first approach, these weights were calculated by using the MRFSS effort estimates as counts of angler-trips and dividing this by the number of intercepted angler-trips. This calculation was performed in cells defined by state, year, wave, mode, area fished and sub-state region. However, while these initial weights account for the overall magnitude of the fishing effort in a cell, they completely miss relative changes in the number and distribution of site-day assignments that occurred during this period. This lead to stability issues in the development of final weights.

Hence, a second approach was developed using counts of site-days with intercepts to account for changes in site-day assignments. For this approach, counts of site-days with intercepts were tallied in cells defined by state, year, wave and mode. While the exact sampling design was unknown, these counts are a useful proxy for it, in the sense that changes in the number of site-days in these cell over time very likely correspond to changes in the underlying sampling design.

In order to incorporate the design changes, the maximum count was identified within each unique combinations of state, mode, and wave across years. Initial weights at the angler-trip level in a state-year-wave-mode cell were calculated as the count of site-days with intercepts in that cell, divided by the maximum count for that state-wave-mode combination. Hence, for cells corresponding to the year with the maximum count, the angler-trip weight is set equal 1, and for any other cell, the weight is greater than 1.

Under this approach, the initial weights will not be correct for the total number of trips, since they only account for relative changes in the design over time. This is justified by the fact that the overall "scale" of the weights, accounting for the volume of angler-trips, is not of interest in APAIS estimation, in which only rates are estimated.

Starting from these initial weights, a raking algorithm was again implemented to create updated weights. As a further adjustment for unobserved design effects, several raking control domains were added to those used for the 2004-2013 period:

- KOD (state, wave, mode and kind-of-day)
- MG (state, wave, mode and month groups)
- AC (state, wave, mode and site activity class).

The first of these corresponds to the usual weekday-weekend/holiday classification of angler-trips, but the other two require further explanation. For the MG domains, raking was attempted using individual month cells, but there were cases that would not converge. Months were therefore grouped into three classes: (1) January, March, October, December; (2) May, June, July, August; and (3) February, April, September, November. Class 1 represents the traditionally lower activity month during transition periods (month 1 in waves 1 and 2, month 2 in waves 5 and 6). Class 2 represents the peak activity period when sample sizes are generally similar or equally allocated among months within waves 3 and 4. Class 3 represents traditionally higher activity month during transition periods (month 2 in waves 1 and 2, month 1 in waves 5 and 6). For the AC domains, sites are divided into two groups, high activity and low activity, based on annual counts of intercepts by fishing mode. Sites with counts above the annual mean within cells defined by state, mode, year and sub-state region were classified as high; sites at or below the mean were classified as low.

The raking algorithm described in Section 2 was applied including these additional control variables, with the adjusted estimates for period 2004-2013 (wave 1) as the "new" estimates and those obtained with the initial weights described above for the period 1993-2003 as the "old" estimates. The linear regression testing for trend described in Section 3 was also performed, but with the modification that it was applied for both the new and the old periods. For any domains where a trend was detected in the old period, the adjustment ratio was computed on the years 2001-2003 instead of on the full period. Similarly, for domains where a trend was detected in the new period, the adjustment ratio was computed using 2004-2006 instead of the full period.

5 Adjustments for prior periods: 1981-1992

The adjustment procedure for 1981-1992 follows the same procedure as that for 1993-2003. The initial weights are again created based on relative counts of site-day assignments, and the raking procedure uses the additional control domains described in Section 4. "New" estimates are those obtained with the adjusted weights for 1993-2003 and "old" estimates are those for 1981-1992. Significant trends resulted in shortening of the period used for the raking ratios to 1990-1992 for the old period and to 1993-1995 for the new period.

References

Deming, W. E. and F. F. Stephan (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics 11*, 427–444.

Hernandez-Stumpfhauser, D., F. J. Breidt, and J. D. Opsomer (2016). Hierarchical Bayesian small area estimation for circular data. *Canadian Journal of Statistics 44*, 416–430.

MRIP Staff (2014, August 20). A descriptive analysis of the Access Point Angler Intercept Survey 2013 design change. Internal report, National Marine Fisheries Service.

NMFS Staff (2015, May 5). Transition plan for the Fishing Effort Survey. Internal report, National Marine Fisheries Service.

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

NOAA FISHERIES

Office of Science and Technology

Weighted Estimation for the Access Point Angler Intercept Survey

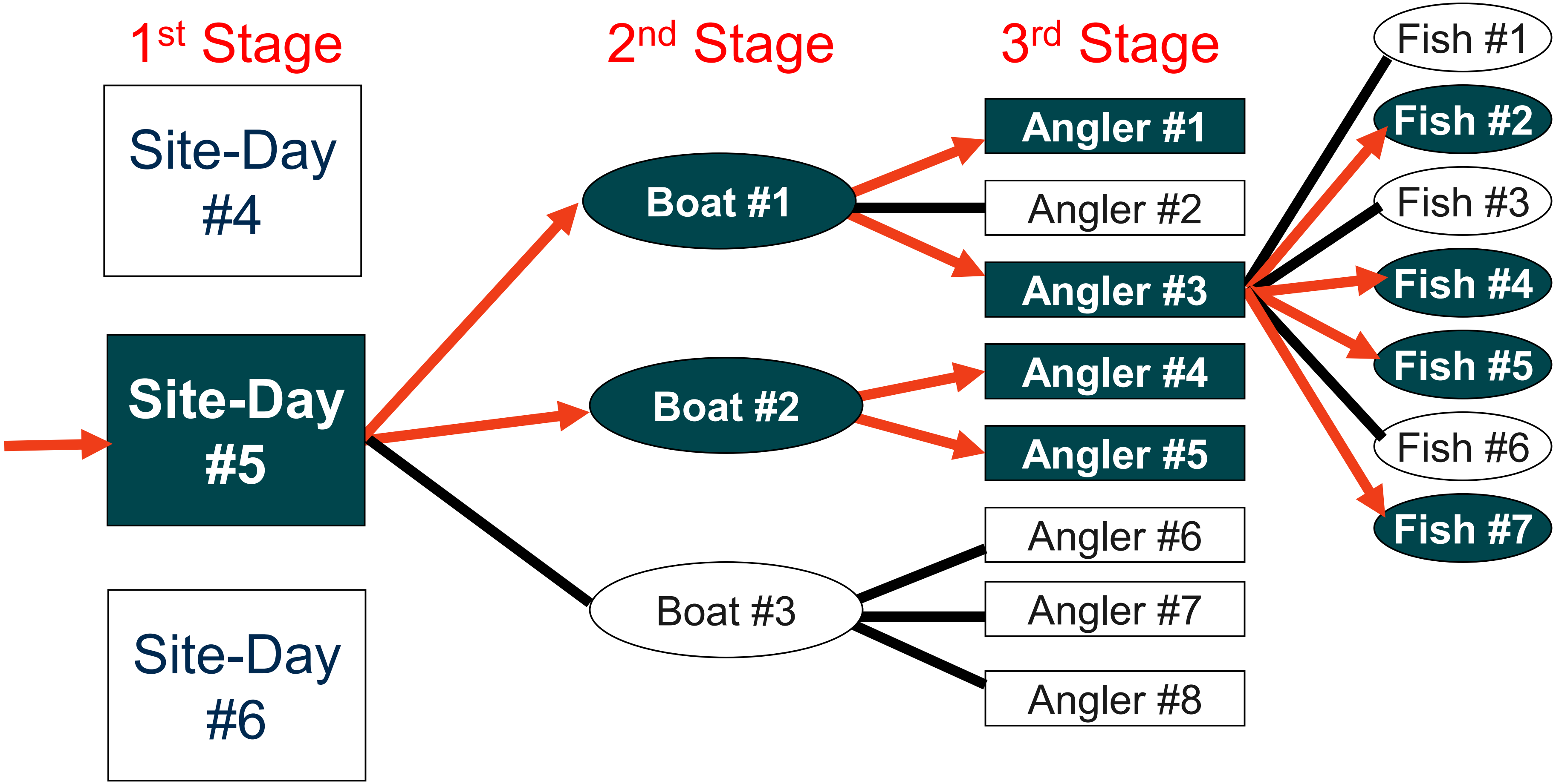
Dave Van Voorhees, F. Jay Breidt, John Foster, Han-Lin Lai, and Jean Opsomer

APAIS Calibration Peer Review
Silver Spring, MD
March 20, 2018

2006 National Academies Review

- The complex sampling design of the APAIS is not accounted for in estimation.
- APAIS point estimates and estimates of their variance are "design-biased".
- Weighted estimation needed:
  - Determine sample inclusion probabilities of intercepted angler fishing trips.
  - Use inclusion probabilities to calculate "sampling weights".
  - Apply "sampling weights" in the estimation process.

Multi-Stage Cluster Sampling
Private Boat Angler Fishing Trips



MRFSS Alternate Site Sampling

Unknown Inclusion Probabilities?

Inclusion probability known

Assigned Site

Inclusion probability?

1st Alternate Site

Inclusion probability?

2nd Alternate Site

Sample Inclusion Probabilities

1st Stage: Site-Day Sampling

- Sites selected as “assigned” sites
  - Unequal probability sampling
  - Angler trips intercepted at sites with higher probability of selection need to be “weighted down”
  - 1st Stage inclusion probabilities for “assigned” sites known
- Sites selected as “alternate” sites
  - Selection probabilities unknown
  - Need to know total probability for each site
  - Important for determining total 1st stage sampling weights

Alternate Site Sampling Weights?

- Probability of site sampled as alternate site?
  - Not known directly from a formal sample draw process
  - Contingent on:
    - Proximity to assigned site
    - Activity at assigned site
- Modeling approach used to develop pseudo-weights:
  - Historical frequencies of alternate & assigned site visits
  - Logistic modeling used to estimate alternate site inclusion probabilities

Final Site-Day Sampling Weights

- Final 1st stage sampling weights calculated to reflect inclusion probabilities of site-day being drawn directly or sampled as an alternate site
  - $w_i = (\pi_i)^{-1} = (\pi_i^D + \pi_i^A - \pi_i^D \pi_i^A)^{-1}$

where

w_i = final 1st stage weight for site-day i

π_i = combined inclusion probability for i

π_i^D = inclusion probability for i drawn directly

π_i^A = inclusion probability for i sampled as alternate

Sample Inclusion Probabilities

Later Stages: Cluster Sampling

- Intercepted trips are only a subset of the entire site-day cluster of trips
  - Must be “weighted up” to represent entire site-day cluster
  - Sampling weight is inverse of sampling fraction at site-day level
- Time spent on site is only part of the whole day
  - Time-slice sample must represent fishing trips at site over full day
  - Need count of trips for full 24 hours to calculate the right sampling fraction.

Boat and Angler Trip Cluster Sizes

- Counts of missed angler trips were made and recorded
  - Total count = intercepted trips + missed trips
- No counts were made of boat trips missed while on site
  - Counts of anglers who fished together on same boat were recorded for intercepted angler trips
    - We could estimate mean number of angler trips per boat trip
  - Total counts of boat trips could be estimated:
    - Boat trips missed = missed anglers/mean anglers per boat
    - Total boat trips = intercepted trips + estimate of trips missed

Expanding Time Slice Counts

- We obtained empirical time slice distributions of trip end times for completed angler fishing days from the Coastal Household Telephone Survey (CHTS).
- We developed a circular normal model to estimate the proportion of daily trip end times by hourly intervals.
  - Reference: Hernandez-Stumpfhouser, Breidt, and Opsomer (2016)
- We used the modeled proportions to expand counts obtained during sampling to counts for the full 24-hr day.

Alternate Mode Sampling

- Alternate mode angler trip intercepts
  - Example: shore mode angler intercepts obtained on an assignment for private boat mode sampling
  - Opportunistic sampling not based on known site selection probabilities for the assigned mode
  - Difficult to know how to weight such intercepts
  - Modeling approaches considered, but too complex
- We decided not to use alternate mode intercepts in the weighted estimation.

MRFSS Estimation
"The Old Way"

$$\hat{\bar{Y}} = \sum_{k} y_k \Big/ n$$

$$n = \sum_{h}^{H} \sum_{i}^{n_h} \sum_{j}^{b_{hi}} n_{hij}$$

MRIP Weighted Estimation "The New Way"

$$\hat{\bar{Y}} = \sum_{h}^{H} \frac{X_h}{X_\bullet} \hat{\bar{Y}}_{1h}$$

Population Mean

$$= \sum_{h}^{H} \frac{X_h}{X_\bullet} \sum_{i}^{n_h} \frac{\pi_{hi}^{-1}}{\sum_{i}^{n_h} \pi_{hi}^{-1}} (X_{hi} \hat{\bar{Y}}_{2hi})$$

Substitute $\hat{\bar{Y}}_{1h}$ by PSU mean

$$= \sum_{h}^{H} \frac{X_h}{X_\bullet} \sum_{i}^{n_h} \frac{\pi_{hi}^{-1}}{\sum_{i}^{n_h} \pi_{hi}^{-1}} \left(X_{hi} \left(\sum_{j}^{b_{hi}} \frac{X_{hij}}{X_{hi\bullet}} \hat{\bar{y}}_{3hij} \right) \right)$$

Substitute $\hat{\bar{Y}}_{2hi}$ by SSU mean

$$= \sum_{h}^{H} \frac{X_h}{X_\bullet} \sum_{i}^{n_h} \frac{\pi_{hi}^{-1}}{\sum_{i}^{n_h} \pi_{hi}^{-1}} \left(X_{hi} \left(\sum_{j}^{b_{hi}} \frac{X_{hij}}{X_{hi\bullet}} \left(\frac{\sum_k y_{hijk}}{n_{hij}} \right) \right) \right)$$

Substitute $\hat{\bar{Y}}_{3hij}$ by TSU mean

NOAA FISHERIES

Weighted Estimation in Summary

- Site-day inclusion probabilities used to weight data
  - Assigned site probabilities known (design-based weights)
  - Alternate site probabilities approximated (pseudo-weights)
- Multi-stage cluster sampling design taken into account
  - Used available data on cluster sizes at each stage
  - Expanded peak activity period counts to estimate total 24-hour counts for each sampled site-day
- Eliminated opportunistic sampling of fishing trips in alternate modes

Independent Peer Review

- Three external reviews:
  - US Census Bureau
  - 2 Reviewers selected by American Statistical Association - Survey Research Methods Section
- Response to external reviews included with final report
- Final report reviewed by MRIP Operations Team and Executive Steering Committee
- Endorsed by NOAA Fisheries AA and certified by MRIP

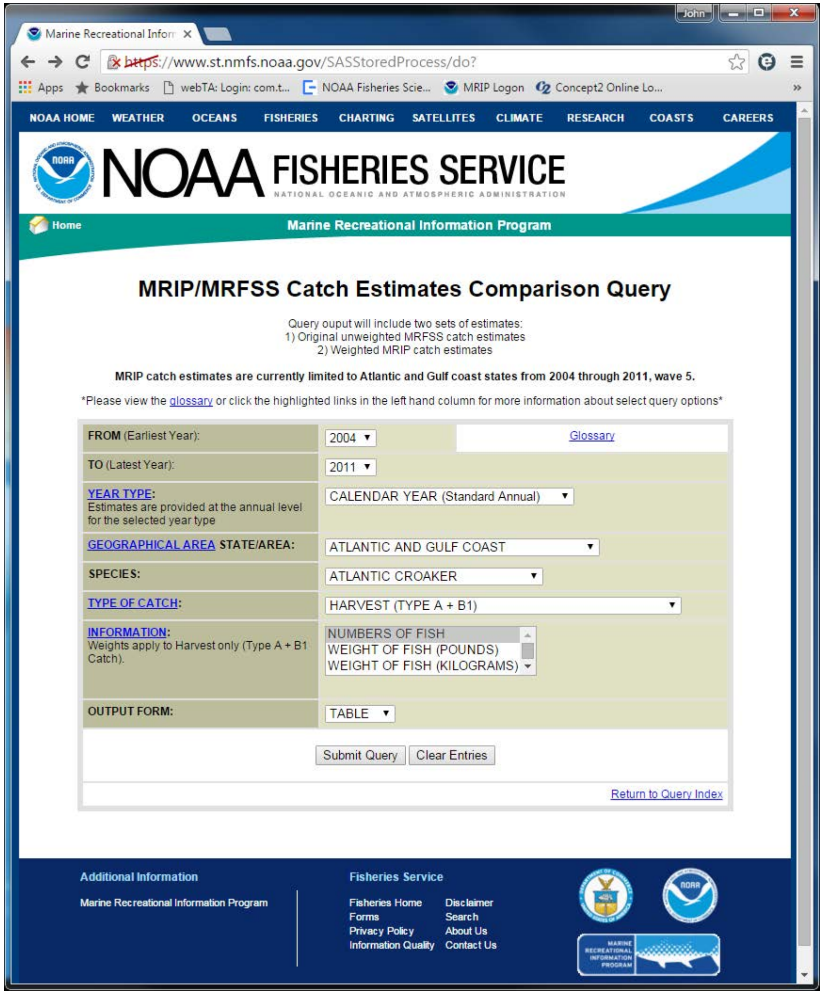
Implementation of New Method

Revision of 2004-2011 Catch Estimates

- Rigorous QC of APAIS data
- Preparation of new data structures
- Preparation and testing of new estimation programs
- Development of comparison tools:
  "New" MRIP estimates vs. "Old" MRFSS estimates
- Also used to produce 2012 weighted estimates

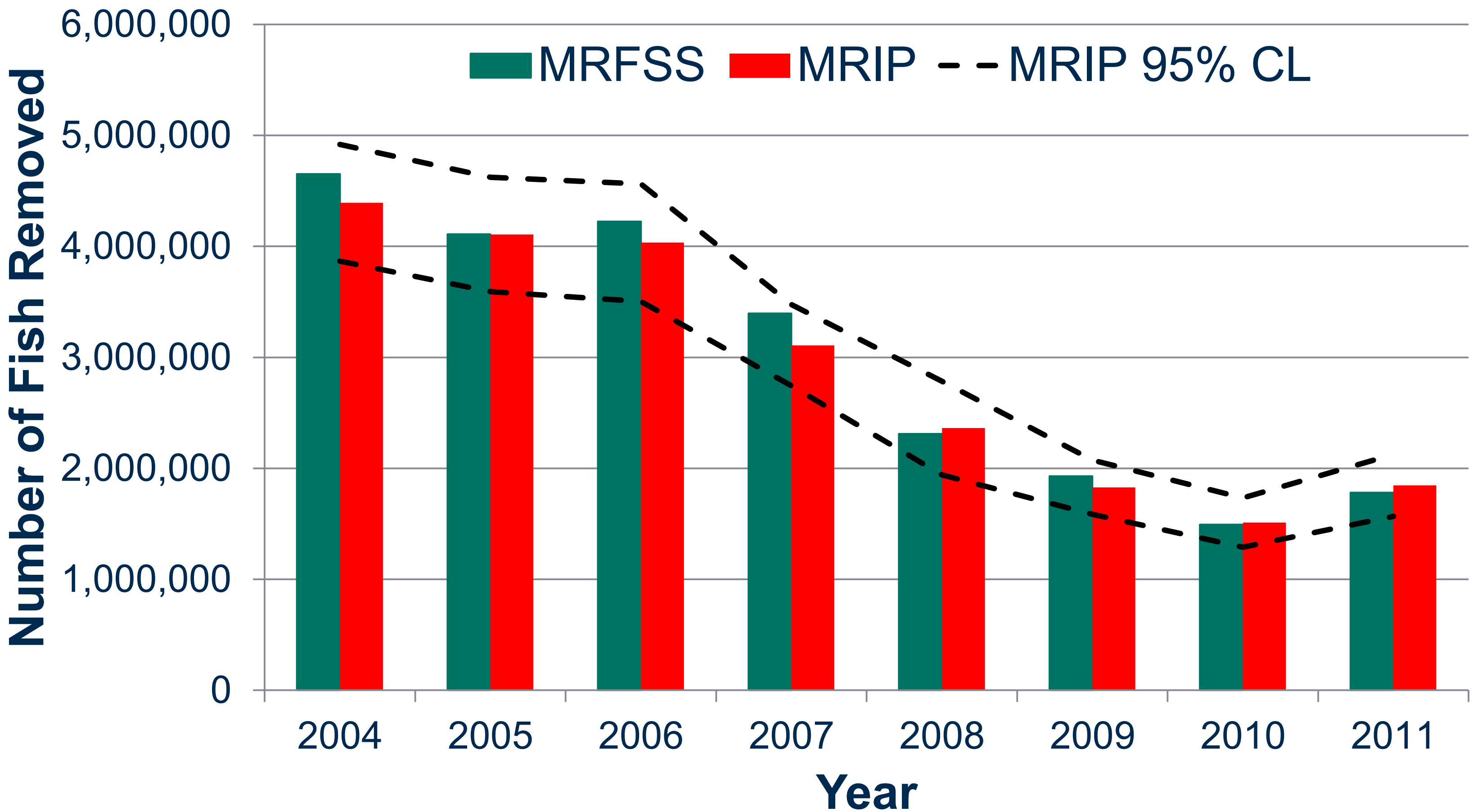
MRIP/MRFSS Comparison Tool

- Available to public through MRIP website
- Query tools for both catch and effort estimates
- Limited to annual state-level estimates
- Tabular and graphic output formats



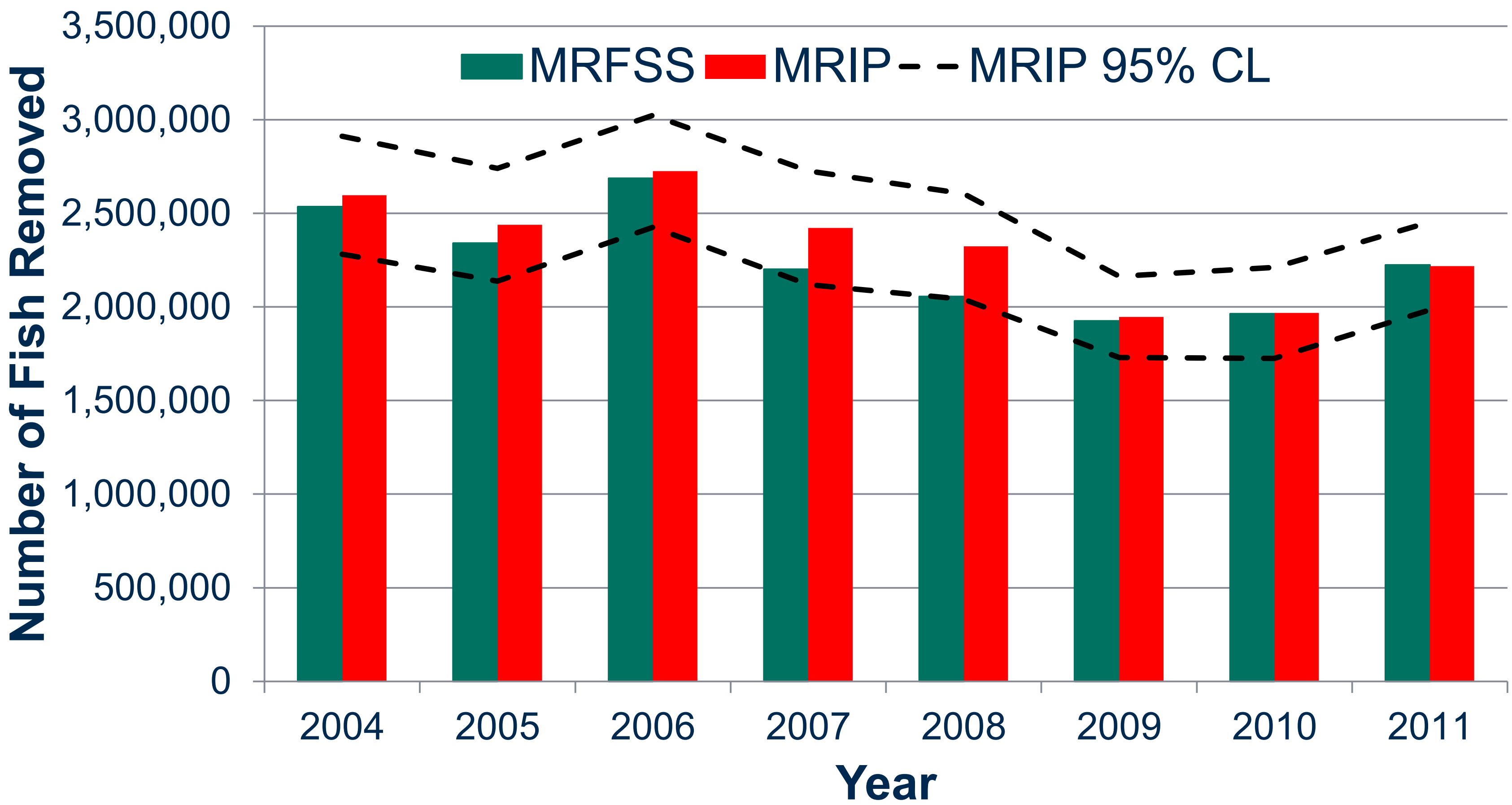
NOAA FISHERIES

Summer Flounder

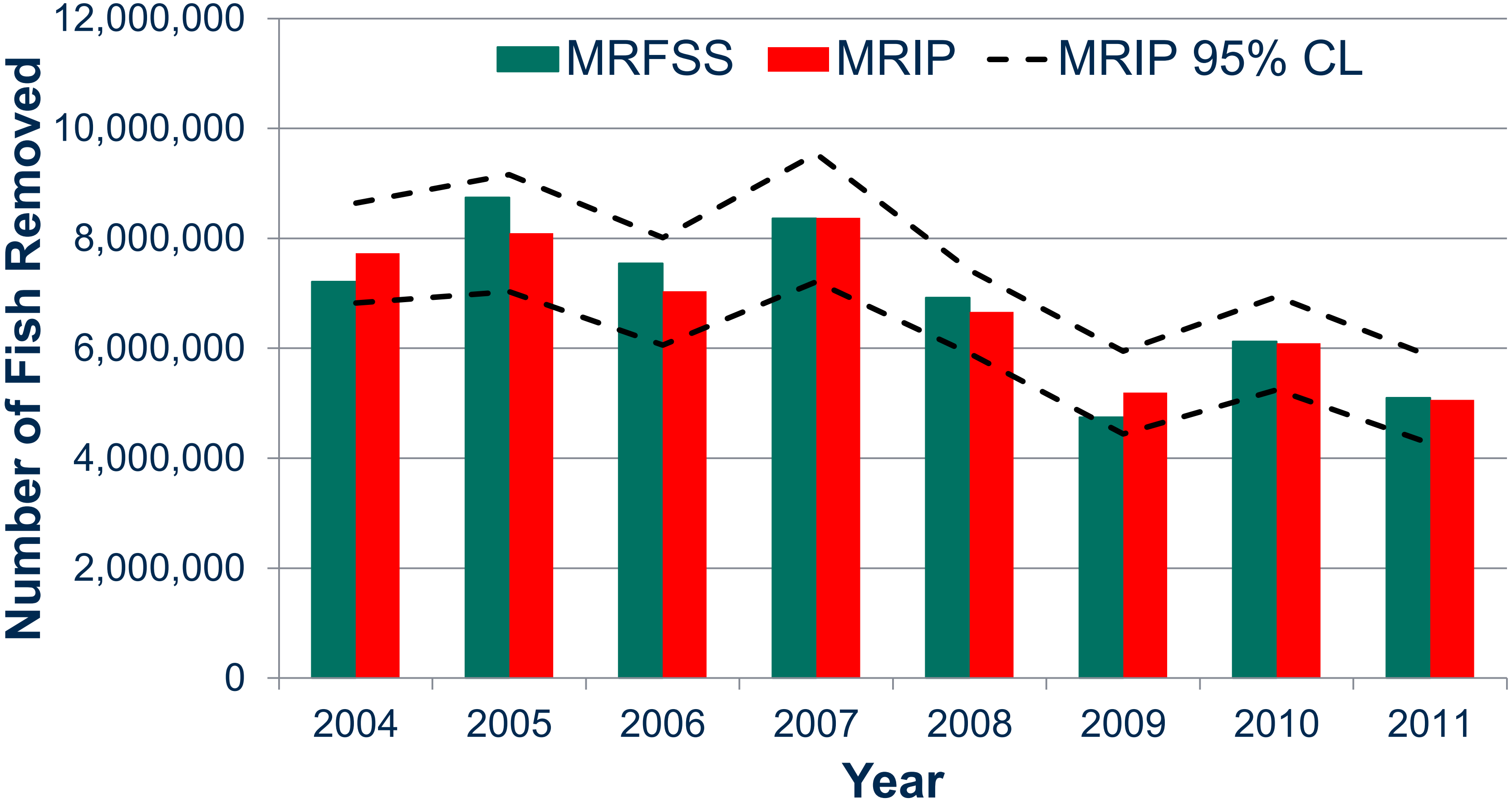


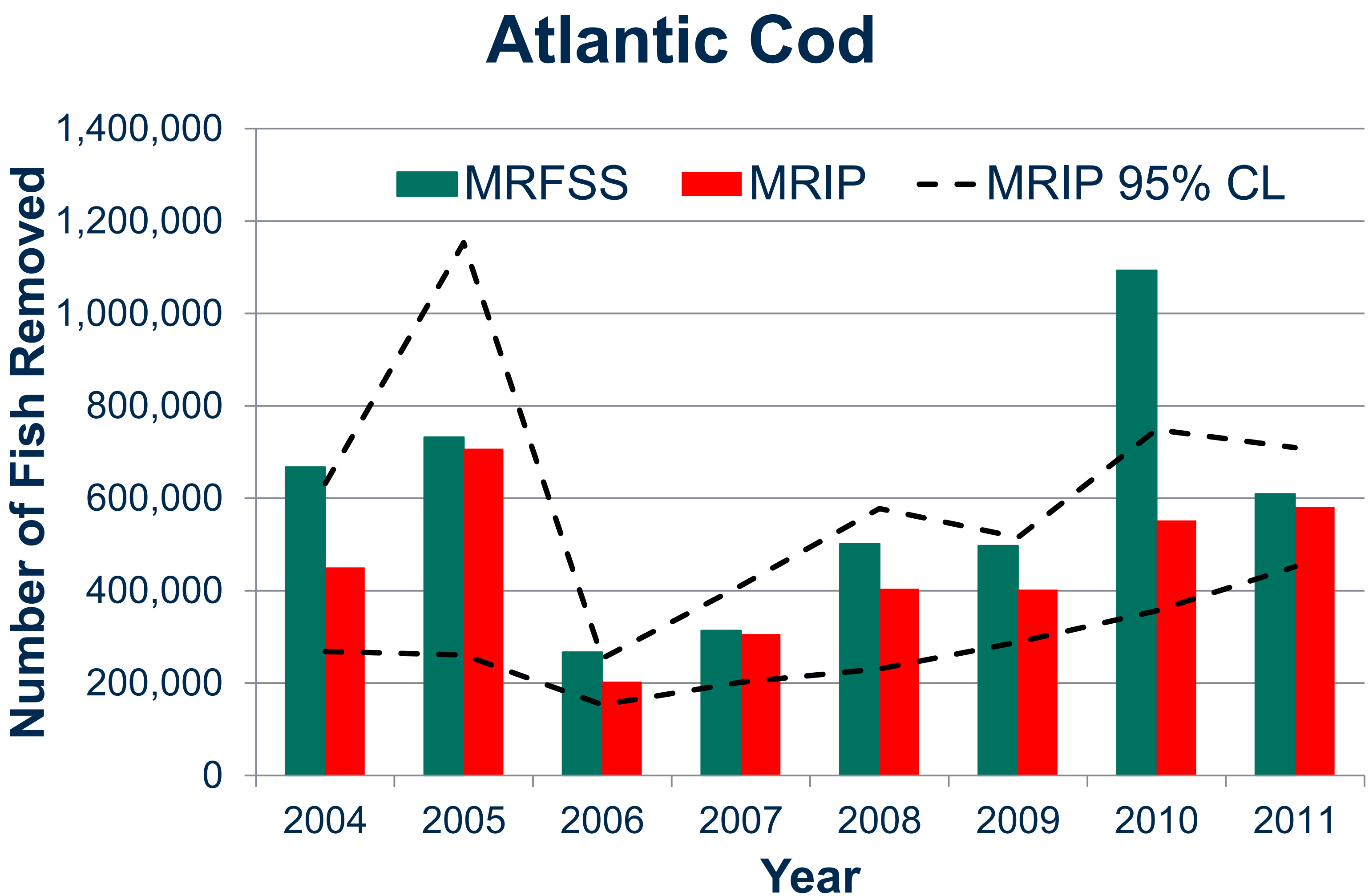
NOAA FISHERIES

Striped Bass

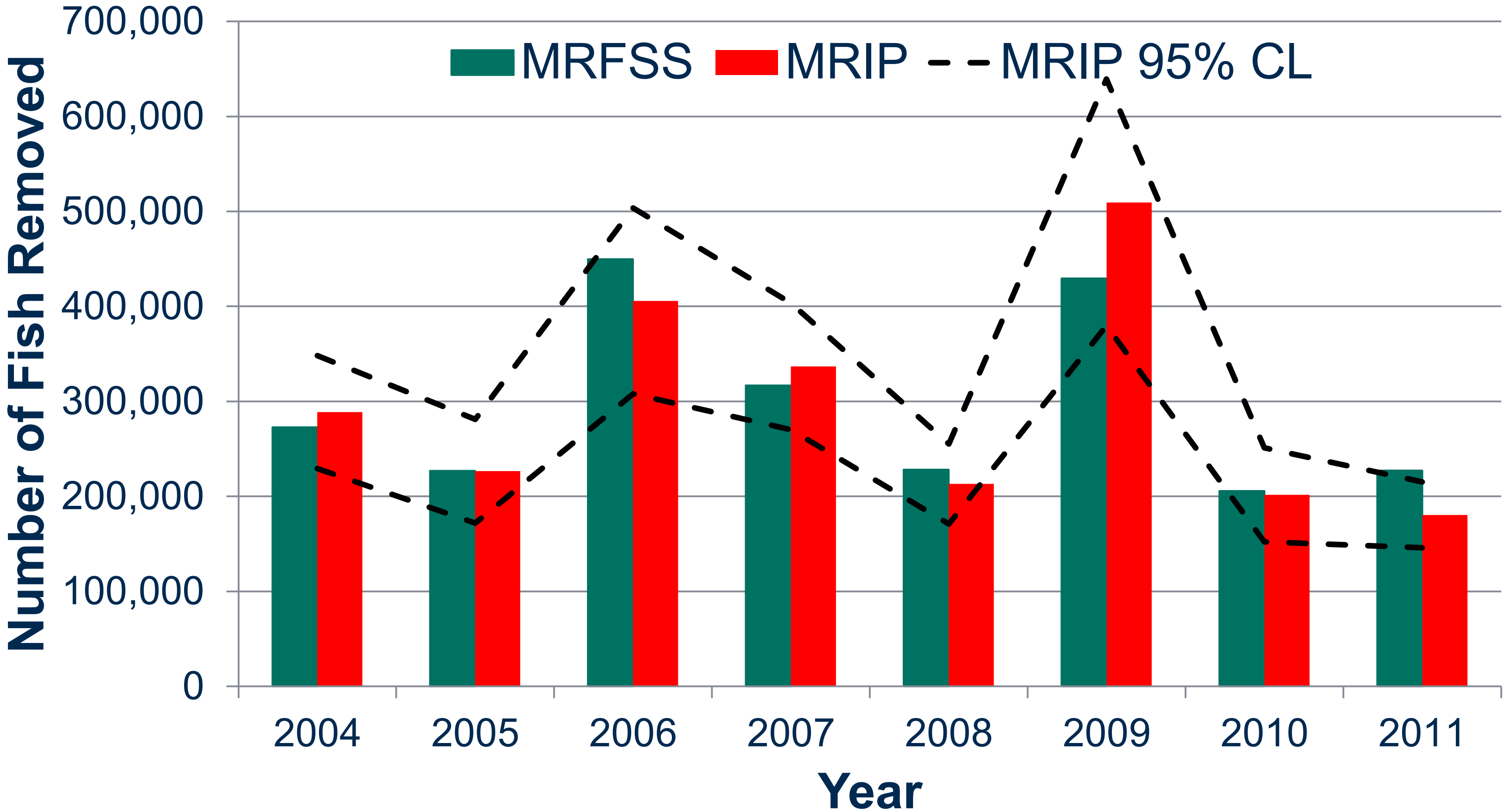


Bluefish

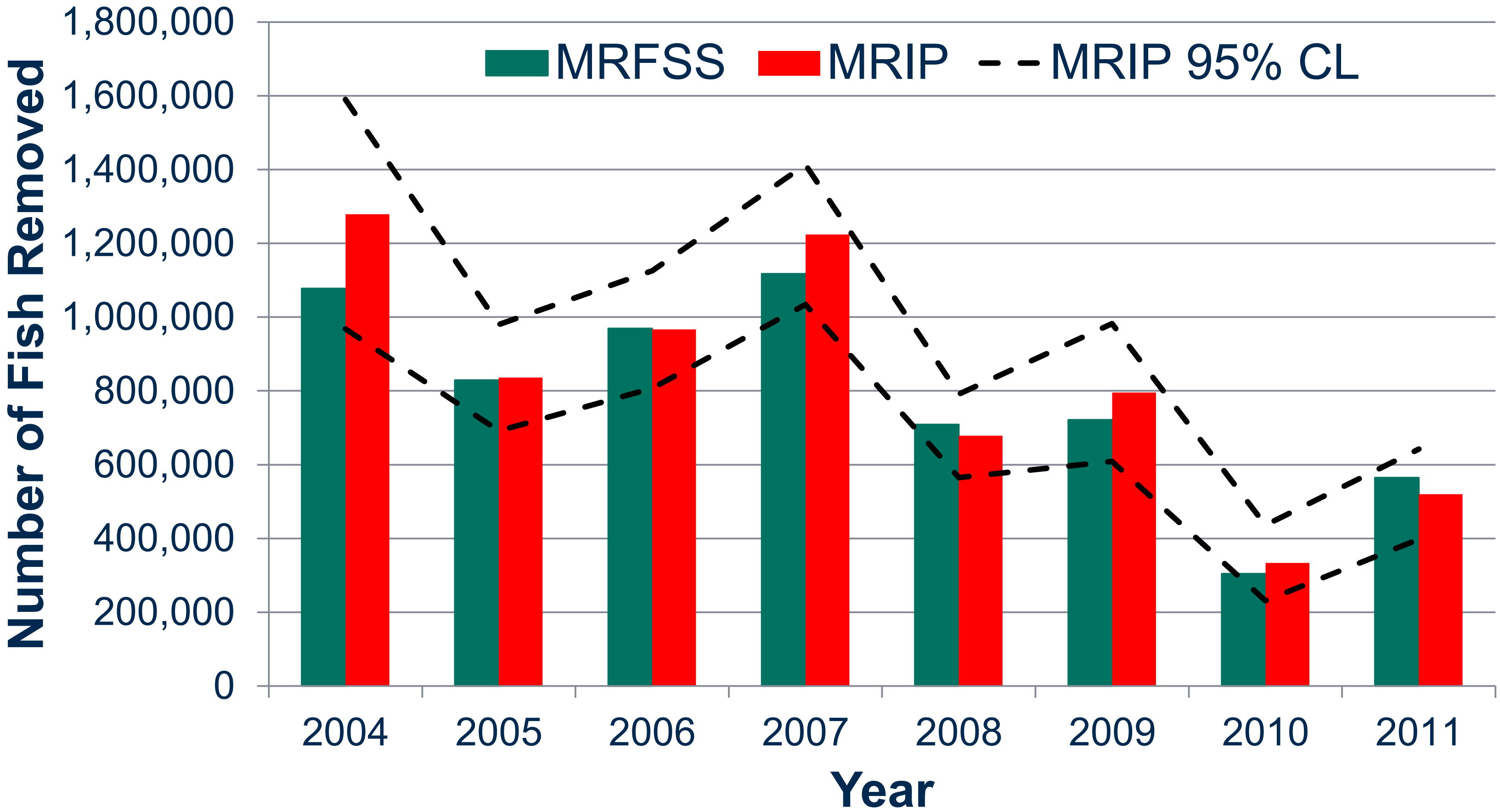

MRFSS
MRIP
MRIP 95% CL
Number of Fish Removed
12,000,000
10,000,000
8,000,000
6,000,000
4,000,000
2,000,000
0
2004
2005
2006
2007
2008
2009
2010
2011
Year


Atlantic Cod
MRFSS
MRIP
MRIP 95% CL
Number of Fish Removed
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
2004
2005
2006
2007
2008
2009
2010
2011
Year

King Mackerel - Gulf of Mexico



Red Snapper – Gulf of Mexico



Statistical Precision

- Estimates of the variance of point estimators of total catch were higher with weighted estimation.
- MRFSS unweighted variance estimates were statistically invalid.
- Explanation:
  - The variance depends mostly on number of site-days included in sample – not number of trips intercepted
  - The sample size of importance is number of site-days

APAIS Calibration #1

2012 MRIP/SEDAR Workshop

Revised estimation resulted in a split time series:

- 1981-2003 MRFSS unweighted estimation
- 2004-2011 MRFSS weighted estimation

Terms of Reference:

- Review studies comparing MRFSS methods to new MRIP methods and propose work to further facilitate calibration.
- Based on years with paired estimates, propose method for calibrating weighted to unweighted estimates, and demonstrate how calibration would be used to hind-cast earlier estimates.
- Recommend plan for implementing the resulting calibration into updated and benchmark stock assessments.

Key Recommendations

- Weighted estimates for 2004-2011 are “best available” and should be used in stock assessments
- Re-estimate catch for 1981-2003
  - Constant “ratio of means” estimators (weighted/unweighted) based on comparisons (2004-2011) should be used to hind-cast revised 1981-2003 estimates and associated variances.
  - Trended ratio estimators based on 8 years of data <u>not</u> advised.
- Variances of hind-casted estimates should incorporate both:
  a) calibrated variance of the catch estimates and
  b) variance associated with ratio estimator used for calibration.

Key Recommendations

- Until revised estimates are incorporated into a new stock assessment, unweighted APAIS data should be used to estimate catches to be compared with an ACL.
- A full benchmark assessment should not be required if changes are small, recreational catches don't dominate overall catch, and changes in age composition are minor.
- Implementation of the revisions should not be delayed to wait for possible future revisions to effort estimates.
- Stock assessment scientists should conduct sensitivity analyses of the hind-casted catch estimates and length frequencies.

Questions?

NOAA FISHERIES

Office of Science and Technology

MRIP: A New Design of the Access Point Angler Intercept Survey

Tom Sminkey, Lauren Dolinger Few, John Foster, Dave Van Voorhees,
NOAA Fisheries, ST1

What is an Access Point Intercept Survey?

- **On-site survey** to collect catch data (access point)
- **Sampling** of completed angler fishing trips (intercept)
- **Spatiotemporal sampling frame:** matrix of fishing access sites and time intervals
- **Multi-stage cluster sampling** (survey)

National Research Council Review (2006)

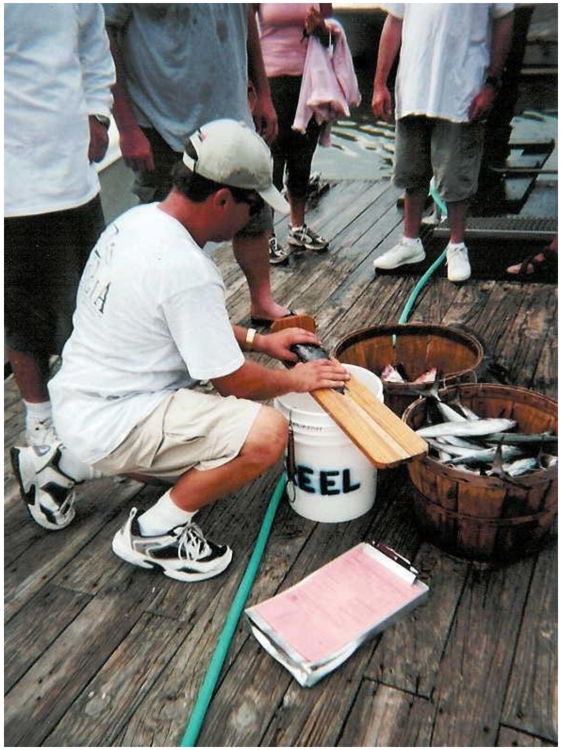
Recommendations for Improvements and Revisions to Access Point Intercept Survey

Need to eliminate "alternate" sites –
unknown and inconsistent selection probabilities

Need to get accurate counts of all completed trips on site –
needed for sample weighting

Should consider approach to cover trips throughout the day –
peak fishing period has been focus, need to cover all time periods

New APAIS Sampling Design



Project Team started in 2009

- Develop new intercept design

2010 North Carolina pilot study:

- Conducted side-by-side with old design (MRFSS)
- Final Report (Breidt, et al., 2012):
  - Recommended coast-wide implementation
  - Recommended possible further enhancements
- Independent peer reviews endorsed implementation

APAIS: MRFSS vs MRIP Design

MRFSS (pre-Mar., 2013)

- Single site sampled
- Alternate sites allowed; use and sites field selected
- "Peak" sampling time selected by sampler
- "Peak" time sampling mandated
- Fishing mode assigned; alternate mode allowed
- Early years did not tally all eligible anglers
- Sampling probability unknown

MRIP (Mar., 2013 – present)

- Site-cluster sampled
- Cluster sites predetermined; mandatory visits
- Sampling time of day and length ot stay predetermined
- All periods of day sampled
- Initial: Single mode stratified interviewing; modified to allow all available modes
- All eligible anglers tallied; supports sample weight computation
- Sampling probability known

What's Different in the New Design?

Maximize number of site-days observed

- Not the number of angler interviews!
- Precision of multi-stage survey estimators depends almost exclusively on number of primary sampling units (site-days) observed

Improved sample frame:

- Spatial component consists of single-site and multi-site clusters
- Increased temporal stratification: 6-hour time intervals
- Increased geographic stratification: state sub-regions

What's Different in the New Design?

Fully formalized probability sampling:

- Probability-proportional-to-size sampling of site-time units (PSUs)
- Attempt to intercept all completed angler trips on site

Samplers do not decide when/where to conduct interviews

- Fixed time interval for each site assignment
- Fixed order of sites for multi-site assignments
- Alternate mode sampling eliminated

No limit on number of interviews per assignment

What's Different in the New Design?

Accurate counting of all trips within sampled site/time unit

- Sampling fractions at each stage known
- Important for proper weighting of data

Emphasis on completing all assignments

- "Controlled selection"
  - Draws thousands of possible sets of assignments
  - Eliminates sets that don't match constraints
  - Selects one of remaining sets at random
- No canceling or re-scheduling of assignments

2013 Design Overview

Complex Stratified Multi-stage with Clustering	
Strata	• Sub region, State, Mode, Month, Kind-of-Day, Interval
Primary Stage Units	• Site cluster-day-interval • A:2am-8am, B:8am-2pm, C:2pm-8pm, D:8pm-2am
PPS Selection	• Estimated Measure-Of-Size defined as expected fishing activity or “pressure” (counts of angler-trips per time period)
SRS at lower stages	

2013 Design Adjustments

Goals	How
• Accommodate field staff constraints • Improve interviewing productivity • Improve spatial and temporal sample distribution • Maintain same temporal and spatial coverage	• Adjustments to site/cluster pressures & clustering rules • Adjustments to sampling strata and allocation of sample to strata • Addition of temporal and spatial sorting variables to assignment draw

New replication-based draw program

Generate large set (S_u) of replicate sample draws using uncontrolled (base) design

Filter S_u replicates through constraints to create survivor subset of replicates (S_c)

Select one replicate (a) from S_c using simple random sampling

Replicate a is official sample draw for intercept survey

Standard definition of inclusion probability

$$\pi_i = P(i \in A) = \sum_{a \in A_{(i)}} p(a)$$

(Fuller, 2009)

Modify definition to condition on survivor subset S_c

$$\pi_i = P(i \in A|\boldsymbol{S_c}) = \sum_{a \in A_{(i)}|\boldsymbol{S_c}} p(a)$$

$\boldsymbol{\pi_i}$ is proportion of survivor draws that contain $\boldsymbol{i}$

2013 Design Adjustments - Conclusions

Effectiveness of 2013 Changes

- Substantive improvements in interviewing productivity
- Improvements to Charter mode not satisfactory

Additional changes warranted in 2014

- How can we better target sample to productive times of day but still maintain full temporal coverage?
- What else can be done for Charter mode?

APAIS 2014 Design

2 Primary changes

- **Peak interval** – Period of day with highest fishing activity
- **Mixed boat sampling** – both Private/Rental and Charter Boat modes sampled on each assigned day/site-cluster/time

APAIS 2014 Design: Peak Interval

Create a new sample interval

that more closely corresponds to peak fishing activity

Minimize disruption to existing design

P:11AM-5PM

- 6-hour interval
- Covers 2nd half of B interval and 1st half of C interval

B:8AM-11AM	B:11AM-2PM	C:2PM-5PM	C:5PM-8PM
	P:11AM-2PM	P:2PM-5PM	

APAIS 2014 Design: Peak Interval

Keep existing B and C intervals

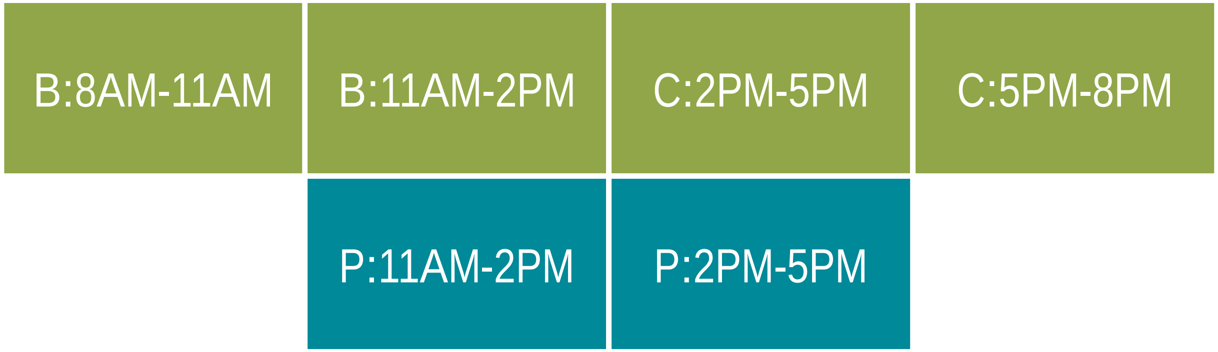
- Maintain full coverage 8AM-8PM

Overlapping Intervals

- Possible to draw same site/3-hr time block/date in two intervals
- Requires adjustments to inclusion probabilities, strata definitions

$$\pi_{BP} = \pi_B + \pi_P - (\pi_B * \pi_P)$$
$$\pi_{CP} = \pi_C + \pi_P - (\pi_C * \pi_P)$$

- Requires special field procedures



NOAA FISHERIES

APAIS 2014 Design: Mixed Boat Sampling

Improve Charter mode
efficiency and productivity

Allow samplers to interview
both Private boat and
Charter boat anglers
during the same assignment

Treat mode of fishing
as domain variable
instead of stratification variable

Replace mode
with site group stratification
in sample frame

APAIS 2014: Mixed Boat Sampling

Existing PR and CH mode strata replaced with site groups

- **Site groups are exclusive** – a site can only belong to one group
- **Site groups are still related to mode**
  - CH sites (only CH, primarily CH, or high CH activity)
  - PR sites (only PR, all other sites not in CH site group)

Site groups have separate clustering, sample allocation, draws

Led to improved productivity particularly for smaller guide boats in Charter sector

APAIS 2016: All Mode Mixed Sampling

Shore Mode added to Mixed Boat mode sampling

- **Site groups are still related to mode**
- **Sites exclusively assigned to a group - hierarchical**
  - CH sites (only CH, primarily CH, or high CH activity)
  - PR sites (only PR, PR relatively high)
  - SH sites (only SH, SH high, low PR and/or CH, what's left)

APAIS 2016: Offshore Stratum (PR mode)

- **Sites with PR activity assigned to a new group**
- **Historical site-intercept data suggests relatively high proportion of PR trips returning fished in offshore (Federal) waters**
- **Improves sampling of trips with rarer occurrence fishes**
- **Strategically used by state (FL, AL, NC?) – assists state surveys**

Thank you!

Questions?

Calibration Workshop II

John Carmichael

APAIS Calibration Model Peer Review

March 20-22, 2018

Calibration Workshop Overview

When & Where

- **September 8-10, 2014 in Charleston SC**

Goals

- **Consider if APAIS changes impacted catch estimates**
- **Evaluate how to adjust estimates to maintain the time series**
- **Provide guidance on addressing future changes**

APPROACH

- Panel of technical representatives from the Northeast through the Gulf of Mexico
  - Councils, States, Commissions, NMFS, University
- Plenary sessions for presentations and general discussion
- Breakout groups to develop recommendations
  - Addressing survey design changes
  - Evaluating 3 methods for this calibration

TERMS OF REFERENCE

- Review Calibration Workshop I (2012) approaches
- Review evaluations of 2013 APAIS changes
- Evaluate the feasibility of separating sampling and fishery changes
- Recommend calibration approaches for pre-2013 estimates.
- Discuss key factors calibration approaches should consider and how future data may affect calibration approaches

Calibration Recommendations

Calibration is required

- Continuity is necessary
- It is not appropriate to compare estimates based on the new survey design to management parameters such as Annual Catch Limits (ACL) based on the old design.
- The appropriate long-term solution is to calibrate existing estimates to the new survey method estimates.
- Interim methods are needed for management and assessment

Calibration Recommendations

<u>Consider 3 calibration approaches & thoroughly evaluate before selecting a final approach</u>

- Two ratio methods
  - Complex and Simple
  - Can be applied in short term, serve as the interim approach
- One model based method
  - More time and effort, benefit from future data
- Address temporal changes in survey coverage
- Regional assistance is needed to develop and evaluate
  - Calibration Workgroup Ongoing since Workshop

Direct Catch Ratio

- Simpler of the 2 ratio methods
- Total Catch / Peak Catch
- Assumes distribution of catch throughout the day is unchanged
- Does not use info from non-peak times
- Recommended as the interim approach

Complex Ratio

- Incorporates relative effort distributions and trip sampling weights
- More use of non-peak info
- Assumptions to meet for an unbiased estimator were unknown

Model Based

- Regression model to classify trips as catch periods (morning, peak, evening)
  - Not actual time, rather what the trip resembles
- Adjust prior years so trip ratios match 2013 and create adjusted trip weightings
- Assumes model will predict periods, and time of day periods capture characteristics
- Uses more explanatory variables & improves with more data

Future Survey Change Recommendations

- Consider calibration during initial design
  - Side by side testing, Avoid "calibrating calibrations"
- Outreach and education are critical
- Existing estimates are needed until management and assessments are updated
- Peer Review calibration methods
- Revise time series

APAIS Calibration Methodology 1

Jean Opsomer
Westat and Colorado State University

March 20, 2018

Outline

1. Calibration background
2. Background for APAIS adjustment approach
3. 2004-2013
4. 1993-2003
5. 1981-1992

1. Background: Recreational Angler Surveys 3

- Time series of catch estimates are crucial input in stock assessment models
  – consistency is clearly critical
- Estimates are obtained through two surveys
  – CHTS → FES
  – APAIS (old) → new
- New surveys are significantly improved but have undergone major methodology changes, leading to time series discontinuities

Many official surveys have implemented changes 4

- Current Population Survey (1994, questionnaire redesign)
- National Household Education Survey (2009, RDD to mail)
- National Crime Victimization Survey (2013, dual frame)
- National Survey of Fishing, Hunting & Wildlife-Associated Recreation (multiple times, mode-questionnaire changes)
- Survey of Graduate Students and Postdoctorates in Science and Engineering (postdoc definition)
- National Resources Inventory (1997, manual to automated photo-interpretation of land cover/use)

How do they deal with survey changes? 5

- Possible options:
  - do nothing
  - add disclaimers
  - calibrate

Do Nothing

- Appropriate in many cases, even for longitudinal surveys
  - small adjustments to methodology, with immaterial effects (common)
  - larger adjustments, but statistical comparison of results reveals no significant effects, e.g. NHES (less common)
- Some repeated surveys make no claims about longitudinal validity of estimates, e.g. FHWAR, Survey of Doctoral Recipients

Add Disclaimers 7

- When significant changes occur, survey agency alerts data users and provides information about change
- Data users can still perform valid time series analyses, by incorporating changes in models
- This is most commonly implemented approach
  - survey agency does not have to model their data, which is both easier and does not open them up to criticism
  - data users are "free" to choose best way to account for changes

Calibrate

- Develop approaches to preserve integrity of time series, by “matching” estimates before and after change
- Statistical calibration requires overlap sample: side-by-side data of old and new measurements
- CHTS – FES: developed calibration model between modes based on overlap data between both surveys, and incorporating changes in composition of CHTS sample over time (previously reviewed)
- APAIS?

2. Background for APAIS adjustment approach 9

- New APAIS design and estimation procedures implemented in 2013 (wave 2), fully replacing previous methods
- APAIS “pseudo-weights” developed for 2004 – 2013 (wave 1), accounting for
  - selection of site-days as implemented in the field, including alternate sites
  - fraction of days interviewer on site, as fraction of full 24-hour period
- 1981-2003: no weights available, limited/no design information (less for earlier years)

APAIS calibration?

- No overlap period available to fit calibration model
- Very large number of estimates (catch by species by type of trip)

→ Calibrate by adjusting/creating angler-trip weights, preserving micro-data

→ Replace exact calibration by reduction/removal of observed temporal discrepancy in 2013

3. 2004-2013 adjustment 11

- Issue: characteristics of trips before 2013 and after 2013 differ more than expected from “typical” angler behavior changes
  - pseudo-weights account for relative frequency of trips by types of sites, waves, modes, and kind-of-day
  - differences still apparent in other characteristics, e.g. area fished, coastal/non-coastal household
- Can we modify weights of pre-2013 trips to correct for trip characteristic discrepancies?

Ratio calibration

- Consider domain *D* consisting of set of trip characteristics
- Let $U_{D,\ 2012}$ = set of trips with those characteristics in 2012, of size $N_{D,\ 2012}$
- We know that under valid sampling design settings,

$$\widehat{N}_{D,2012} = \sum_{i \in s} w_i \, I_{\{i \in U_{D,2012}\}} = \sum_{i \in s_{D,2012}} w_i$$

is unbiased for $N_{D,\ 2012}$, but that is not true here

- Unfeasible calibrated weights

$$w_i^* = \frac{N_{D,2012}}{\widehat{N}_{D,2012}} w_i$$

Ratio calibration (2) 13

- $N_{D,\,2012}$ unknown, so replace by sample-based quantity obtained from post-2013 design, e.g. $\widehat{N}_{D,\,2014}$
- But: not interested in specific years, so replace by less variable multi-year adjustment

$$w_i^* = \frac{\widehat{N}_{D,\mathrm{new}}}{\widehat{N}_{D,\mathrm{old}}} w_i$$

with $\widehat{N}_{D,\,\mathrm{new}}$ = average of annual estimates for domain D under new design (2013, 2014, 2015, 2016) and $\widehat{N}_{D,\,\mathrm{old}}$ = same under old design (2004, 2005,…, 2013)

Ratio calibration (3)

- Would like to apply ratio corrections to correct for discrepancies in trip distributions by:
  - state and sub-state region
  - year and wave
  - mode
  - area fished
  - coastal/non-coastal household
  - for-hire boat frame membership
- Too many small domains if we consider all possible combinations

Raking calibration

- Consider less detailed domains only, and sequentially ratio adjust on each until convergence
- Raking control domains:
  - AF (state, wave, mode, area fished): $\widehat{N}_{D,\text{ new },\text{AF}}$
  - HS (state, wave, mode, coastal/non-coastal household status): $\widehat{N}_{D,\text{ new },\text{HS}}$
  - FH (state, wave, mode, for-hire boat frame status): $\widehat{N}_{D,\text{ new },\text{FH}}$
  - RE (state, wave, mode, substate region): $\widehat{N}_{D,\text{ new},\text{ RE}}$

Raking algorithm 16

1. Initialize: set t = 0, $w_i^{(t)} = w_i$ for 2004-2013 (wave 1), compute $\widehat{N}_{D,\text{ new},\text{ AF}}$, $\widehat{N}_{D,\text{ new},\text{ HS}}$, $\widehat{N}_{D,\text{ new},\text{ FH}}$, $\widehat{N}_{D,\text{ new},\text{ RE}}$
2. Let $\widehat{N}_{D,old,AF}^{(t)}$ = averages of estimated AF domain totals for 2004-2012 (include 2013 for wave 1) using weights $w_i^{(t)}$, compute ratios $R_{AF}^{(t)} = \widehat{N}_{D,new,AF}/\widehat{N}_{D,old,AF}^{(t)}$, and set $w_{\text{i},AF}^{(t)} = R_{AF}^{(t)} w_i^{(t)}$
3. Starting from $w_{\text{i},AF}^{(t)}$, repeat for HS domains, resulting in ratios $R_{HS}^{(t)}$ and weights $w_{\text{i},HS}^{(t)}$.

Raking algorithm (2) 17

4. Starting from $w_{\mathrm{i},HS}^{(t)}$, repeat for FH domains, resulting in ratios $R_{FH}^{(t)}$ and weights $w_{\mathrm{i},FH}^{(t)}$

5. Starting from $w_{\mathrm{i},FH}^{(t)}$, repeat for RE domains, resulting in ratios $R_{RE}^{(t)}$ and weights $w_{\mathrm{i},RE}^{(t)}$

6. Set $w_i^{(t+1)} = w_{\mathrm{i},RE}^{(t)}$

7. Repeat steps 2-6 until convergence (measured by change in $R_{AF}^{(t)}$, $R_{HS}^{(t)}$, $R_{FH}^{(t)}$, $R_{RE}^{(t)}$), and set $w_i^* = w_i^{(t)}$

What happens if fishery changes over time? 18

- Previous procedure adjusts for design changes if underlying fishery characteristics do not change
- However, what if there are both fishery changes and design changes?

– lack of data collection overlap under both methods makes confounding unavoidable

– as long as fishery changes are gradual over time, they can be detected in historical time series

Modification for temporal changes 19

- Create time series datasets of total trip estimates for each raking control variable for 2004-2013 (wave 1):
    - AF: 145 year-wave estimated totals for each state, mode, area fished
    - (same for HS, FH, RE)
- Fit linear regression and test for significance of slope
    - for categories where slope is not significant: no temporal trend, apply raking as before
    - for categories where slope is significant: temporal trend, replace raking ratio by one computed using 2010-2013 (wave 1) only

Modification for temporal changes (2) 20

- Computing raking ratios on most recent years avoids removing (most of) time trend in fishery characteristics
- But: increases variability of adjustment

1993-2003 adjustment 21

- We would like to apply same procedure, but starting weights not available
- First step: create initial angler-trip weights

  – naïve attempt: use CHTS total trip estimates divided by number of intercepted trips

  – better attempt: need to account for (unknown) APAIS design changes, so develop proxy for them and include in initial weights

Initial weights

- Count number of site-days with intercepts in state-wave-mode-year domains: $C_{D,1993},\ldots,C_{D,2003}$
- Maximum count across years = $C_{D,\max}$
- Initial weight for angler trip in domain *D* is

$$w_i = \frac{C_{D,\max}}{C_{D,(\text{year})}}$$

→ not calibrated for absolute number of trips in domain, but captures changes in site-day sampling intensity over time

Raking variables and algorithm

- Same as before, plus
  - KOD (state, wave, mode and kind-of-day)
  - MG (state, wave, mode and month groups)
  - AC (state, wave, mode and site activity classes)

→ Account for design effects (already included in initial weights for 2004-2013)

- Raking algorithm as before, with
  - "new" = 2004-2013 (wave 1)
  - "old" = 1993-2003

Modification for temporal changes

- New: create time series datasets of total trip estimates for each raking control variable for 2004-2013 (wave 1)
- Old: create time series datasets of total trip estimates for each raking control variable for 1993-2003
- Fit linear regression and test for significance of slope in old and new time series
  - slope is not significant: apply raking as before
  - slope is significant in old time series: replace raking ratio by one computed using 2001-2003
  - slope is significant in new time series: replace raking ratio by one computed using 2004-2006

1981-1992 adjustment 25

- Same procedure as for 1993-2003
  - New period: 1993-2003
  - Old period: 1981-1992
- Temporal trend detection: use 1990-1992 (old) and 1993-1995 (new) if detected