



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

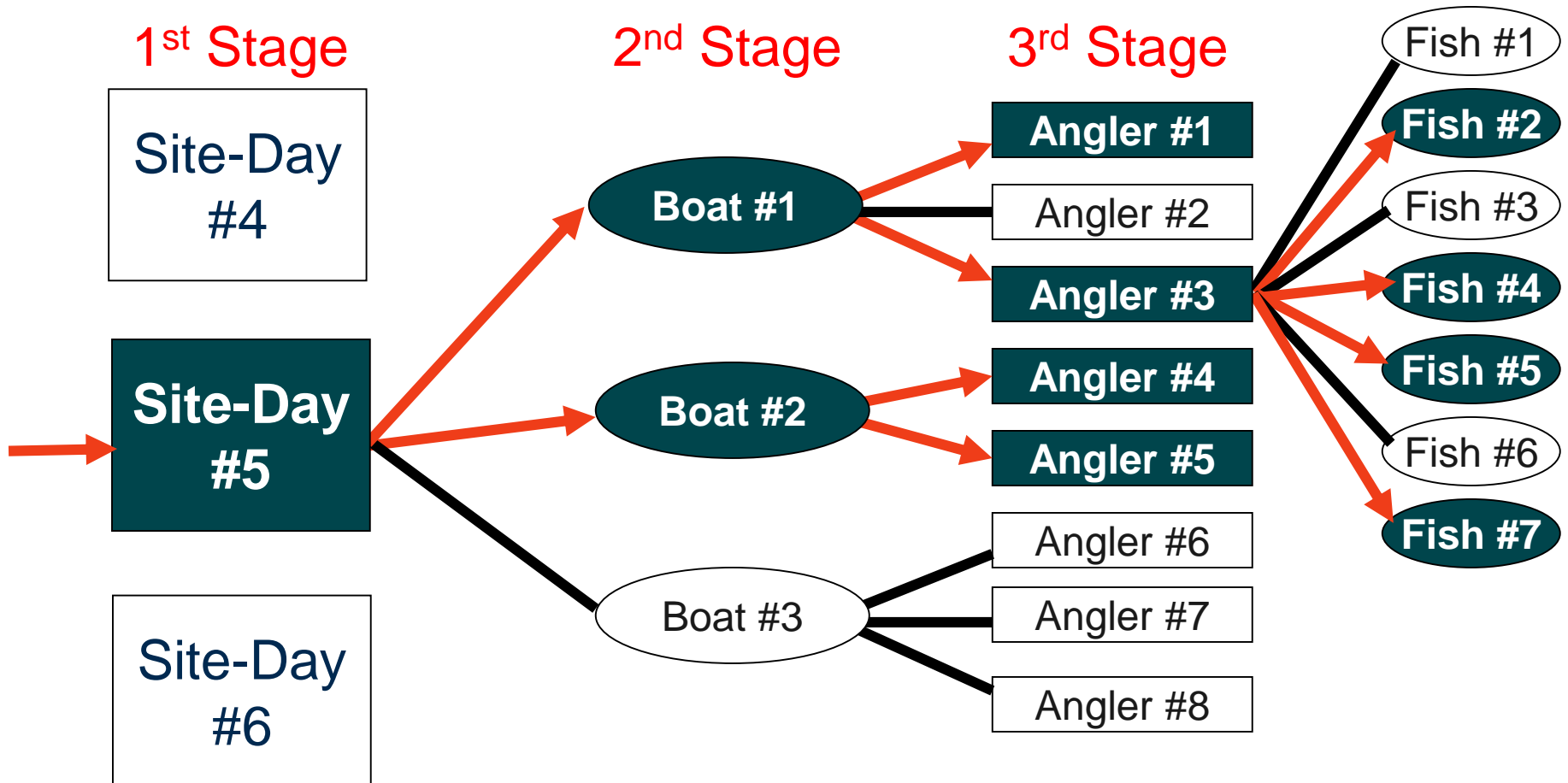
APAIS Calibration Peer Review
Silver Spring, MD
March 20, 2018

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.

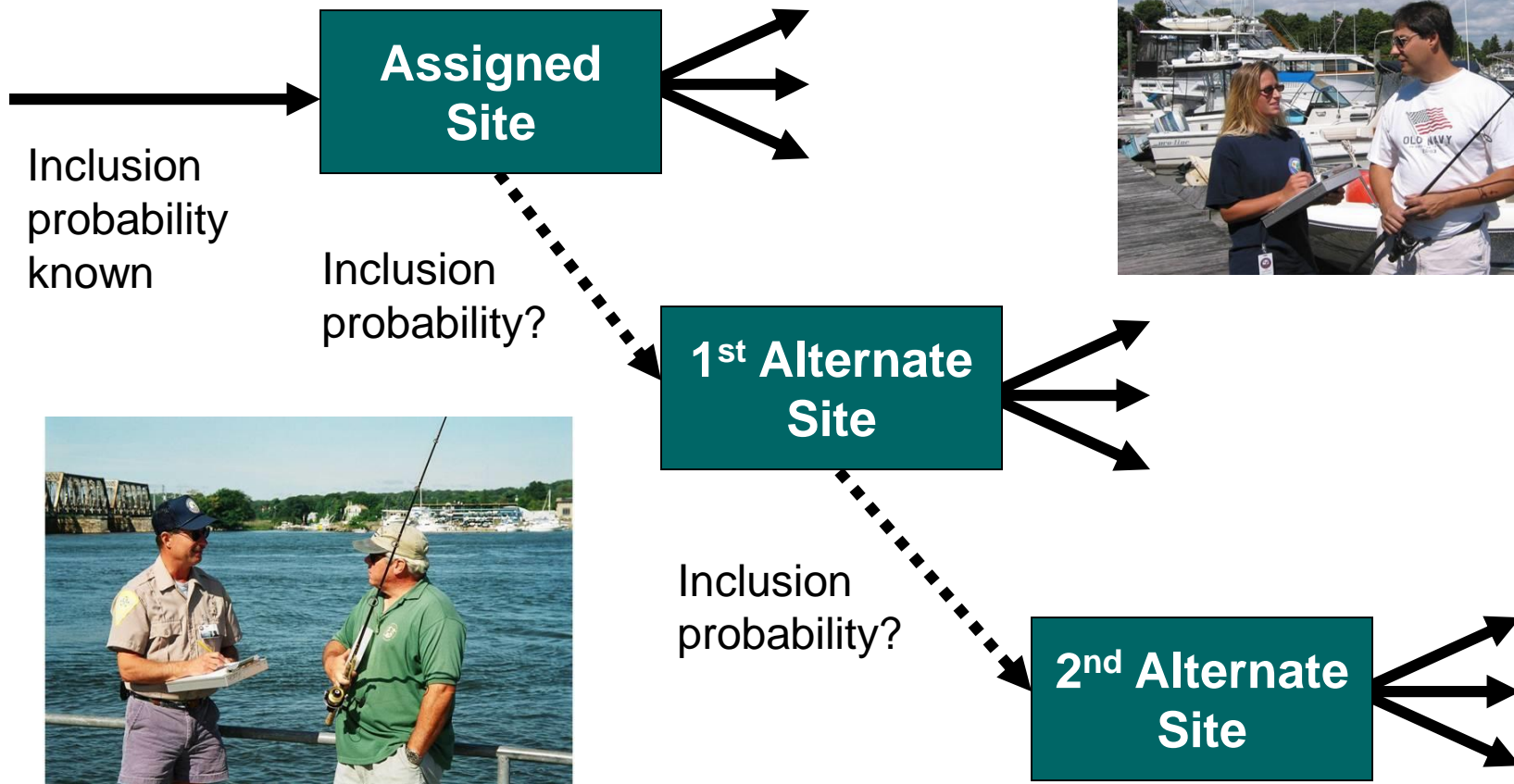
Multi-Stage Cluster Sampling

Private Boat Angler Fishing Trips



MRFSS Alternate Site Sampling

Unknown Inclusion Probabilities?



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-Stumpfhouse, 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”

$$\hat{Y} = \sum_k y_k / n$$

$$n = \sum_h^H \sum_i^{n_h} \sum_j^{b_{hi}} n_{hij}$$

MRIP Weighted Estimation

“The New Way”

$$\hat{Y} = \sum_h \frac{X_h}{X_{\bullet}} \hat{Y}_{1h}$$

Population Mean

$$= \sum_h \frac{X_h}{X_{\bullet}} \sum_i \frac{\pi_{hi}^{-1}}{\sum_i \pi_{hi}^{-1}} (X_{hi} \hat{Y}_{2hi})$$

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

$$= \sum_h \frac{X_h}{X_{\bullet}} \sum_i \frac{\pi_{hi}^{-1}}{\sum_i \pi_{hi}^{-1}} \left(X_{hi} \left(\sum_j \frac{X_{hij}}{X_{hi\bullet}} \hat{y}_{3hij} \right) \right)$$

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

$$= \sum_h \frac{X_h}{X_{\bullet}} \sum_i \frac{\pi_{hi}^{-1}}{\sum_i \pi_{hi}^{-1}} \left(X_{hi} \left(\sum_j \frac{X_{hij}}{X_{hi\bullet}} \left(\frac{\sum_k y_{hijk}}{n_{hij}} \right) \right) \right)$$

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

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

The screenshot shows a web browser window displaying the NOAA Fisheries Service website. The page title is "MRIP/MRFSS Catch Estimates Comparison Query". The URL is "https://www.st.nmfs.noaa.gov/SASStoredProcess/do?". The page features a navigation menu with links for NOAA HOME, WEATHER, OCEANS, FISHERIES, CHARTING, SATELLITES, CLIMATE, RESEARCH, COASTS, and CAREERS. The main content area includes a "Home" link and the "Marine Recreational Information Program" logo. The query tool is titled "MRIP/MRFSS Catch Estimates Comparison Query" and provides instructions on the query output, which includes two sets of estimates: 1) Original unweighted MRFSS catch estimates and 2) Weighted MRIP catch estimates. It also notes that MRIP catch estimates are currently limited to Atlantic and Gulf coast states from 2004 through 2011, wave 5. A link to the "Glossary" is provided for more information. The query form includes the following fields:

FROM (Earliest Year):	2004	Glossary
TO (Latest Year):	2011	
YEAR TYPE: Estimates are provided at the annual level for the selected year type	CALENDAR YEAR (Standard Annual)	
GEOGRAPHICAL AREA STATE/AREA:	ATLANTIC AND GULF COAST	
SPECIES:	ATLANTIC CROAKER	
TYPE OF CATCH:	HARVEST (TYPE A + B1)	
INFORMATION: Weights apply to Harvest only (Type A + B1 Catch).	NUMBERS OF FISH WEIGHT OF FISH (POUNDS) WEIGHT OF FISH (KILOGRAMS)	
OUTPUT FORM:	TABLE	

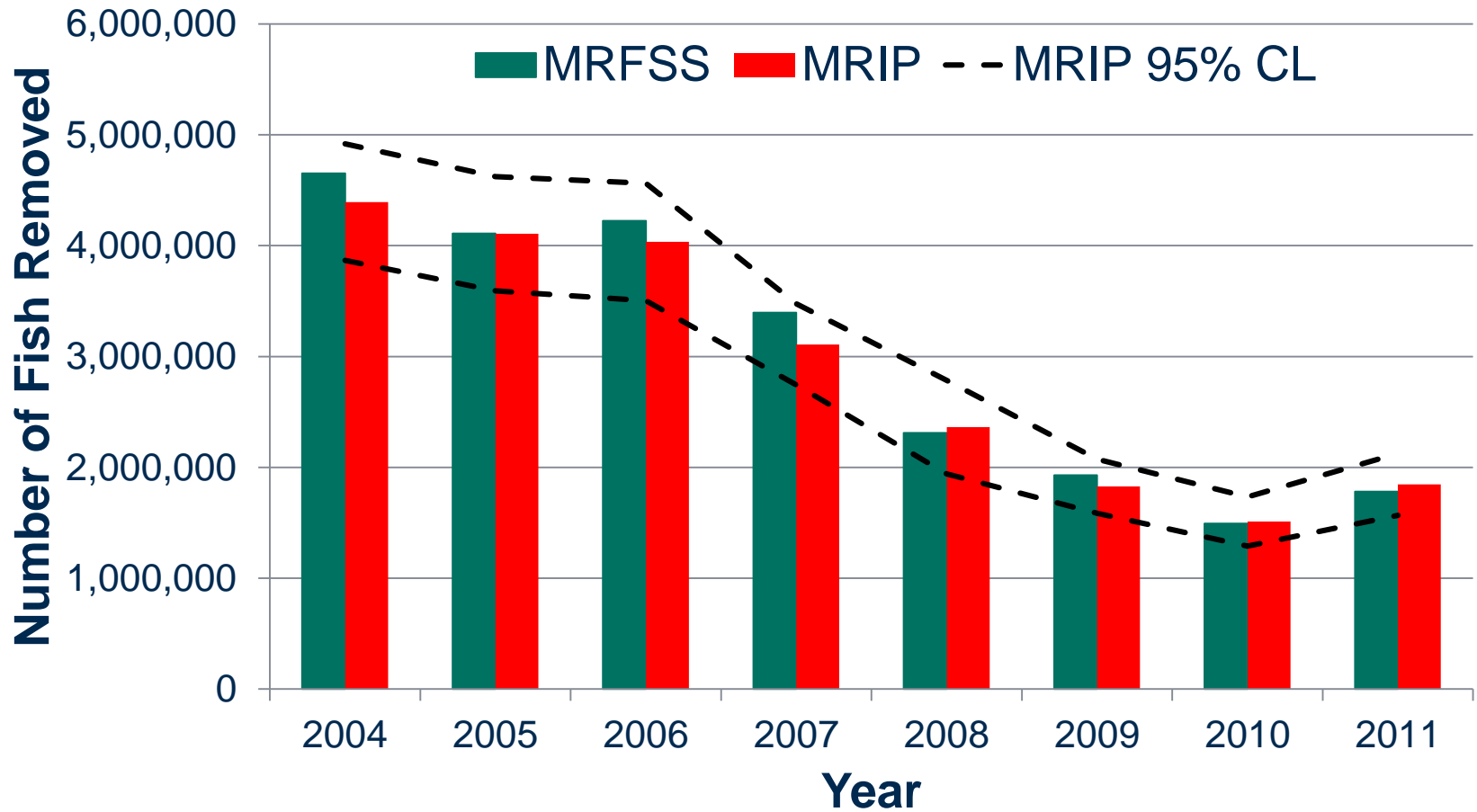
Buttons for "Submit Query" and "Clear Entries" are located below the form. A "Return to Query Index" link is also present.

Additional Information: Marine Recreational Information Program

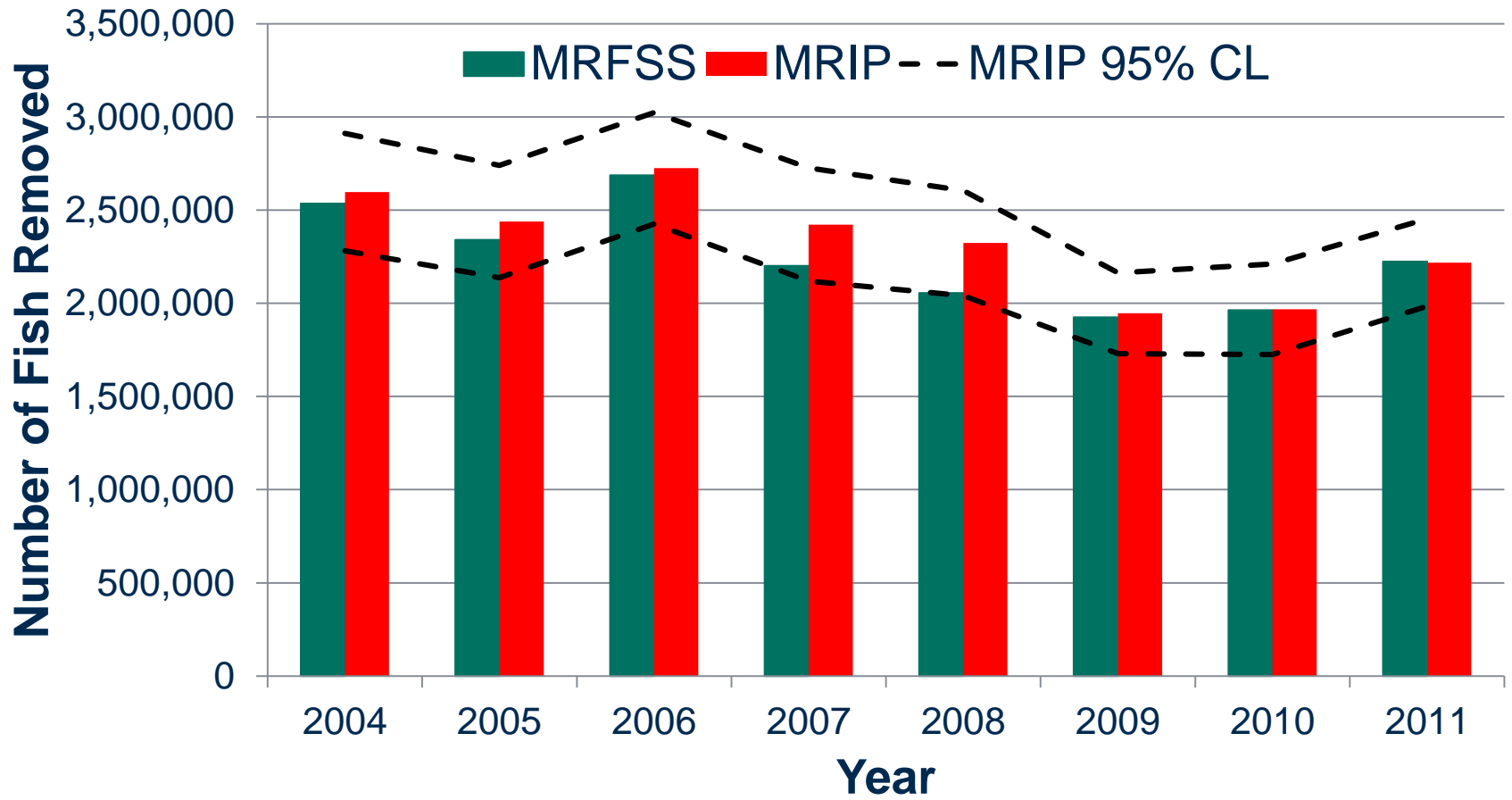
Fisheries Service: Fisheries Home, Disclaimer, Forms, Search, Privacy Policy, About Us, Information Quality, Contact Us

Logos for NOAA and the Marine Recreational Information Program are displayed at the bottom right.

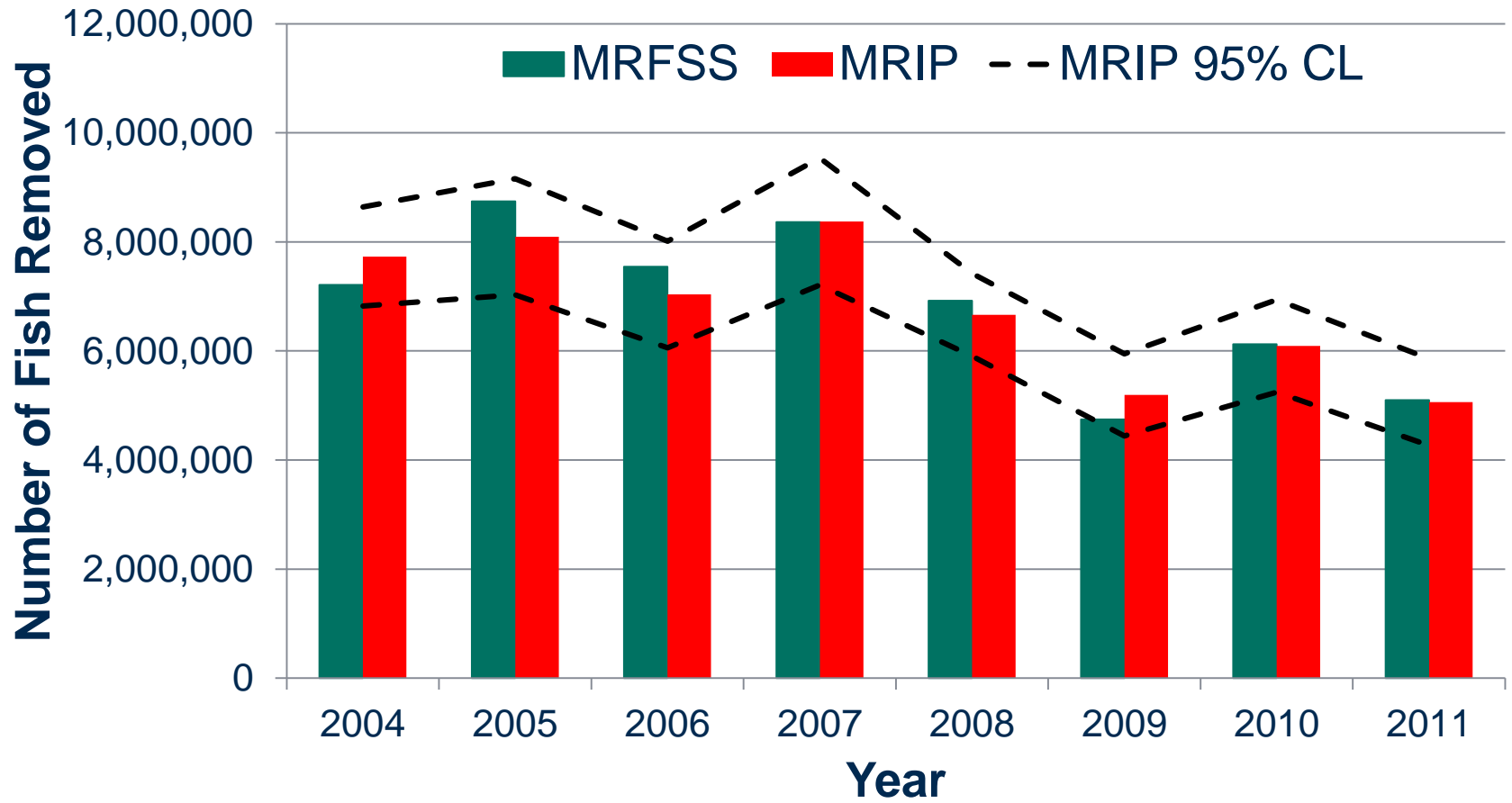
Summer Flounder



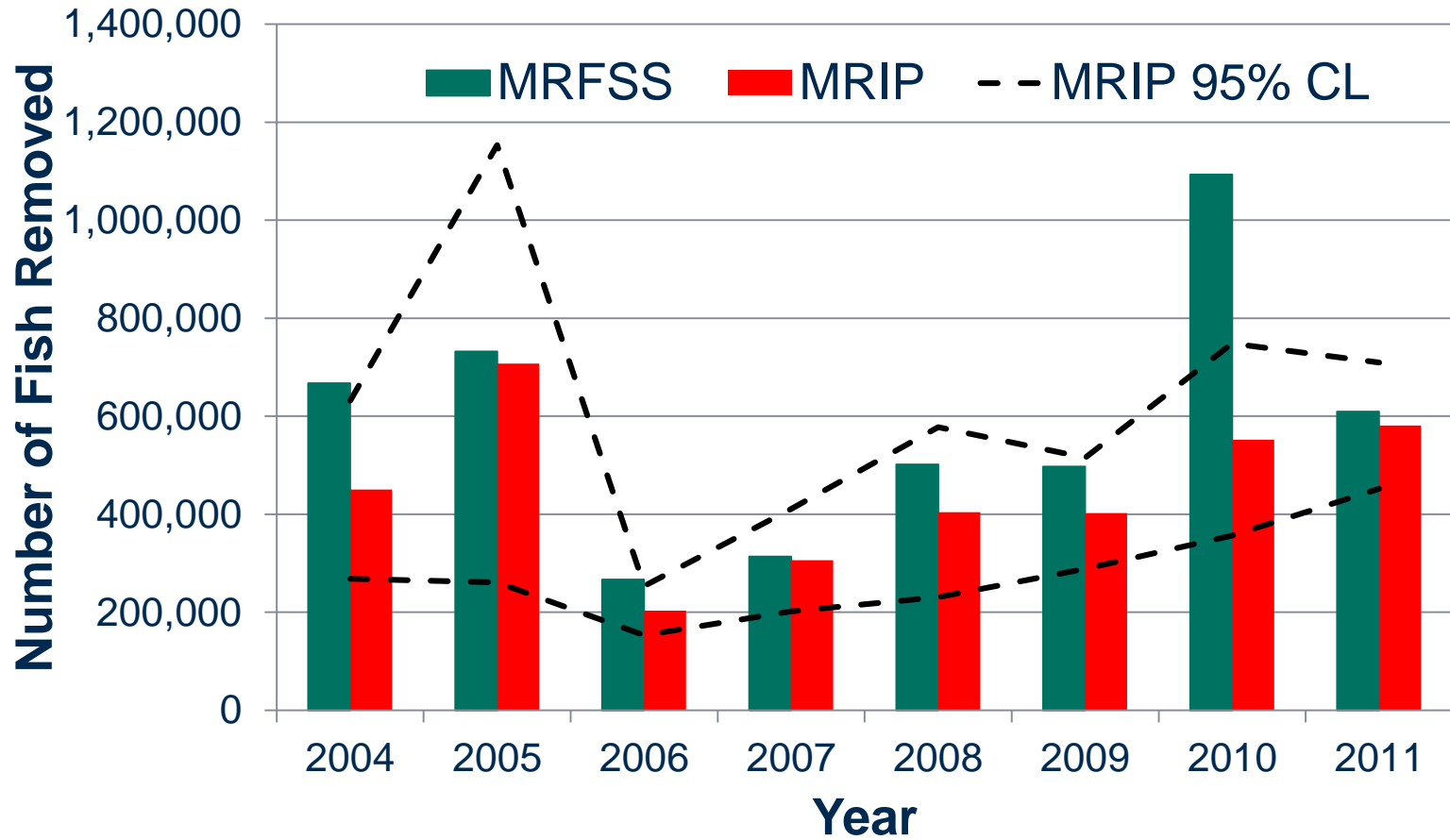
Striped Bass



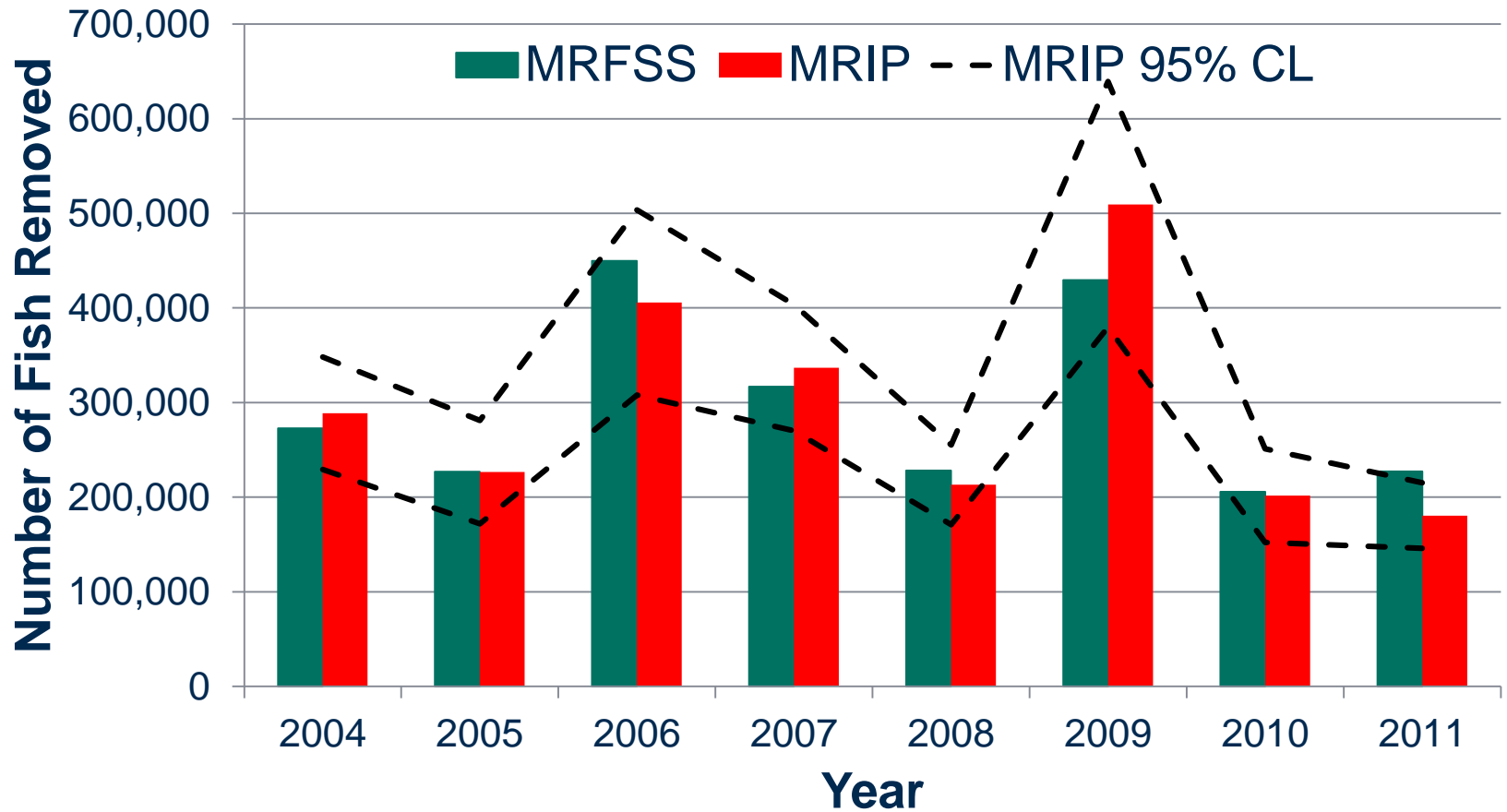
Bluefish



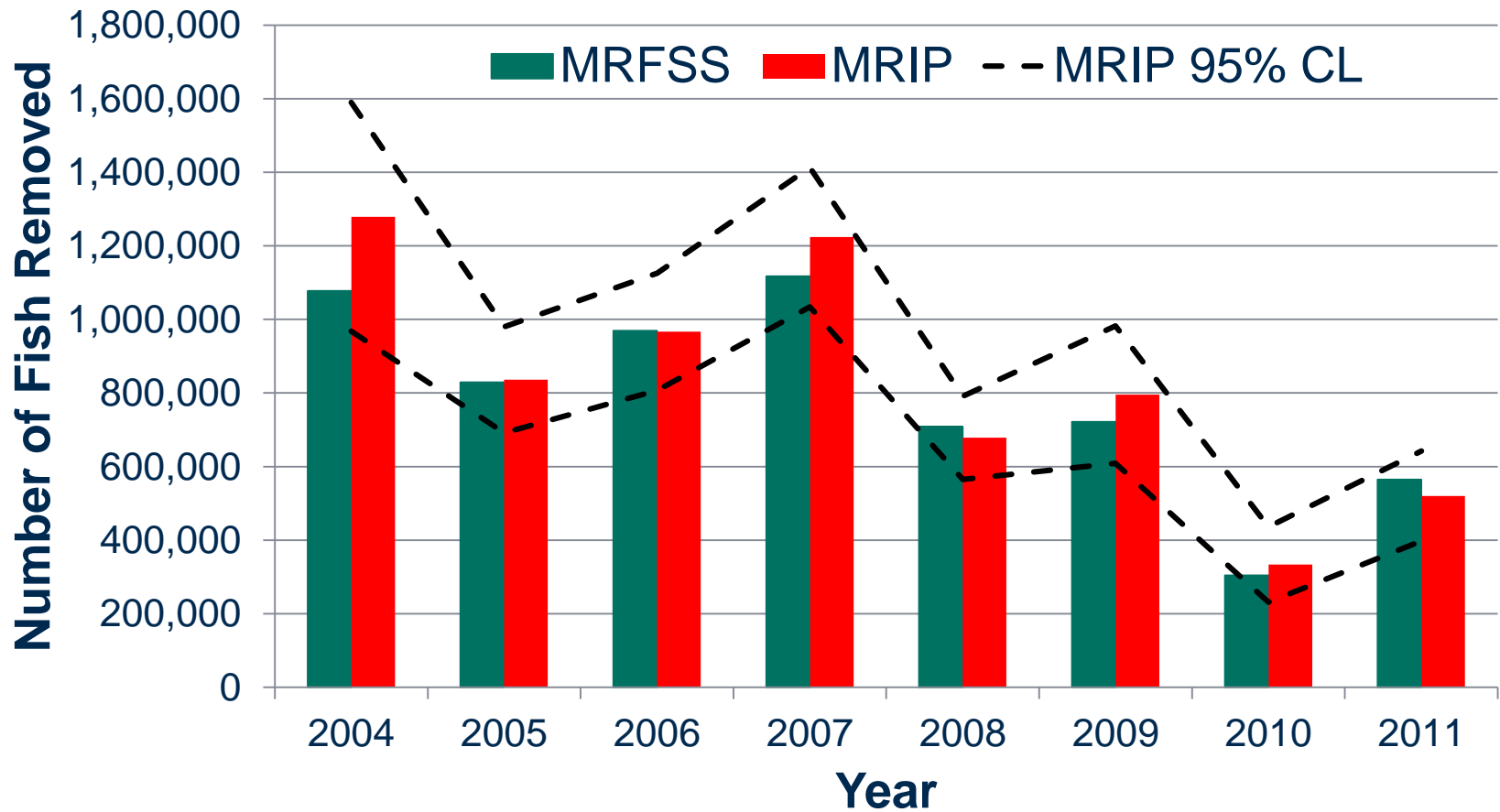
Atlantic Cod



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 not 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?

