

NOAA
FISHERIES

APAIS Calibration Methods Considered

An Overview of the Ratio Methods Considered for APAIS Calibration

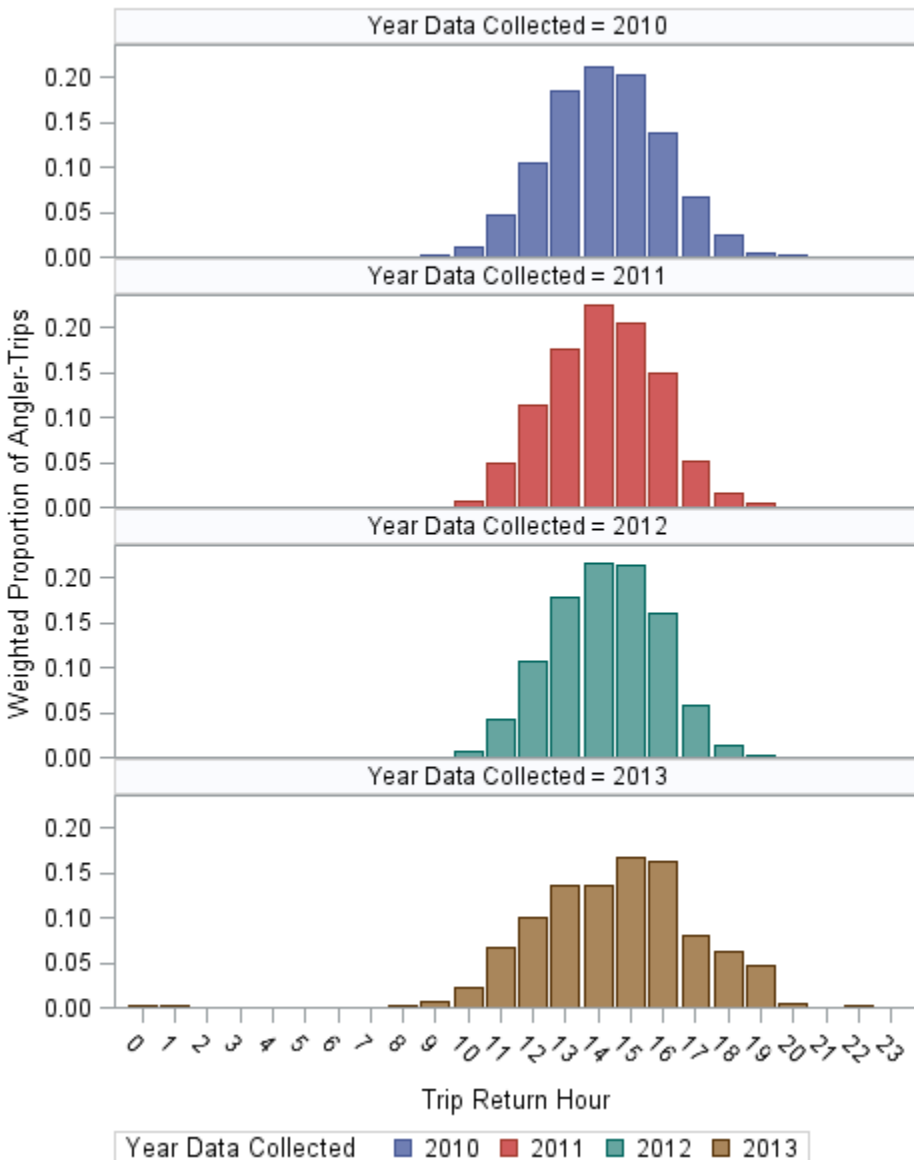
Ryan Kitts-Jensen
Office of Science
& Technology

Outline

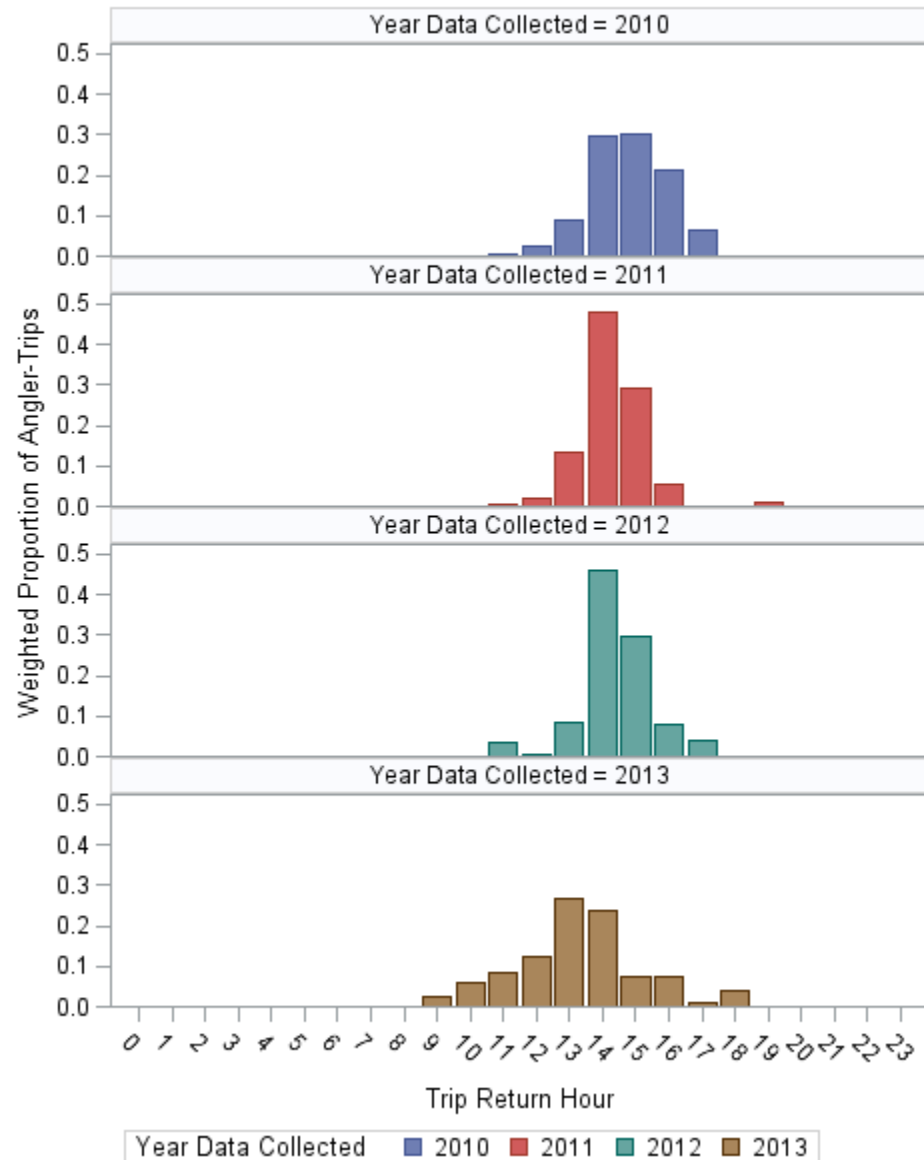
- Review of Temporal Distribution Shift in Trips
- Ratio Method 1 (Simple Ratio) Overview
- Ratio Method 2 (Complex Ratio) Overview
- Comparison of Results
- Summary of Sensitivity Analyses and Limitations

Review of Temporal Distribution Shift in Trips

North Carolina Private Mode Trip Return-Hour Distribution 2013 Represents 2013-2016 Combined

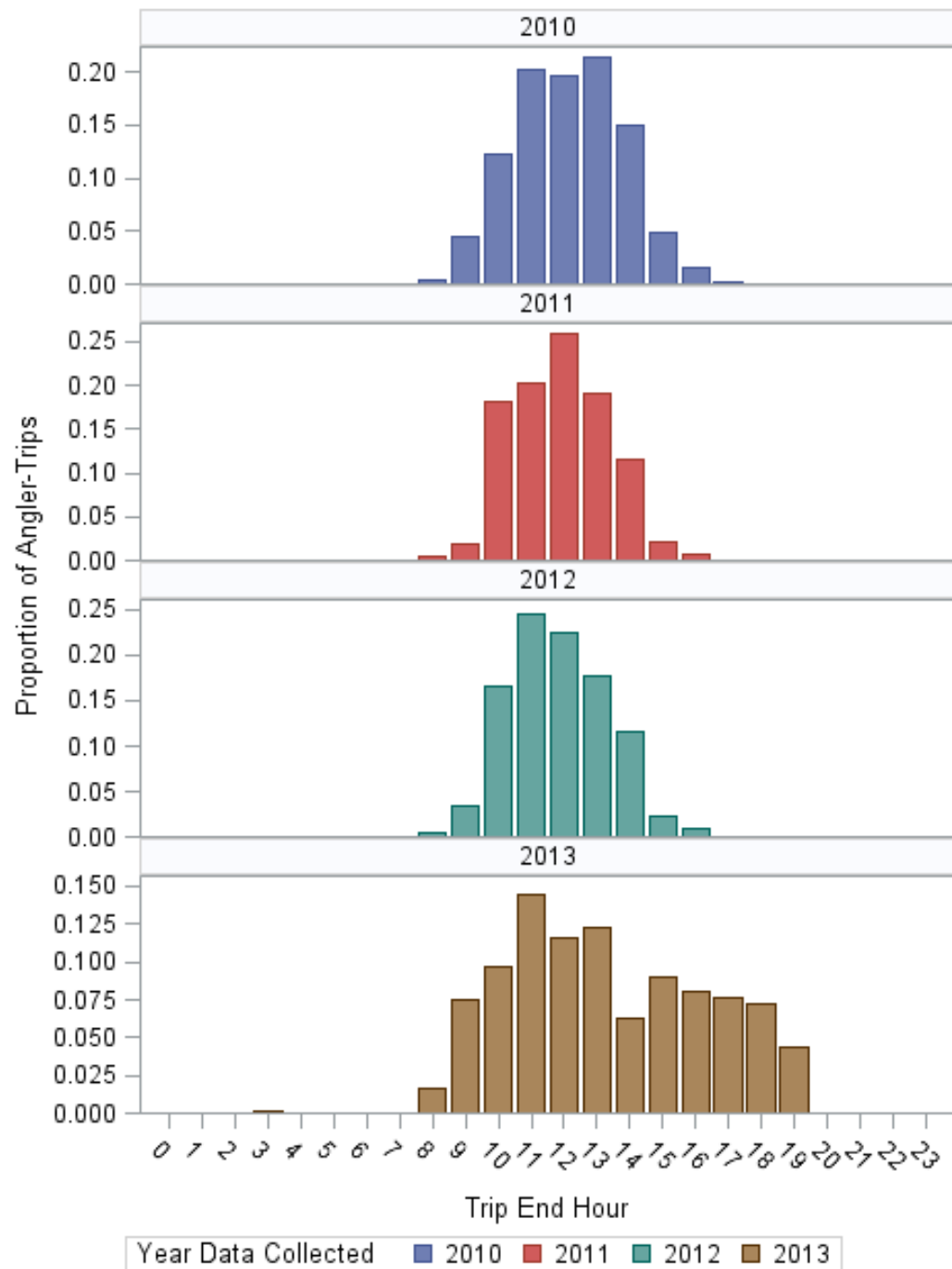


Rhode Island Charter Mode Trip Return-Hour Distribution 2013 Represents 2013-2016 Combined

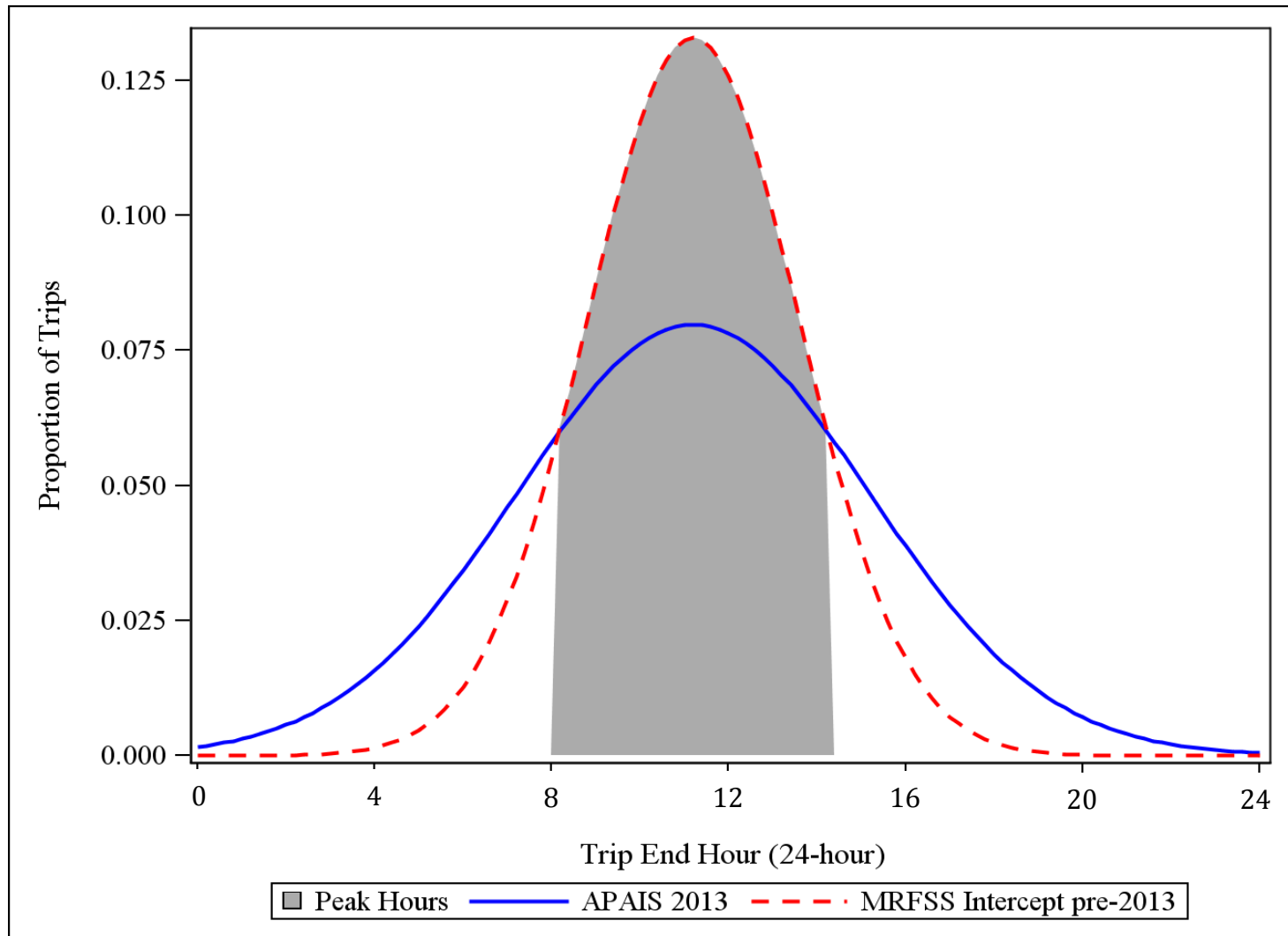


Proportions of Angler-Trips by Hour

Alabama
Private Boat
Annual
2010-2013



Hypothetical Temporal Distribution Shift



Temporal Shift – Time Block Definitions

- Temporal Coverage
 - Systematic differences possible if trips covered in 2013+ very different from trips covered in prior years
- Define time blocks for Temporal Coverage
 - **Morning:** Trips not fully covered prior to 2013
 - **Peak:** Trips fully covered prior to 2013 (assumption)
 - **Evening:** Trips not fully covered prior to 2013
 - **Total:** M+P+E

Ratio Method 1 (Simple Ratio)

Ratio Method 1 (Simple Ratio) Description

- Define a “Peak” activity range of hours (time block)
- Assume catch estimates for Peak time block are consistent and comparable between MRFSS Intercept and APAIS design years (pre-2013 MRFSS, 2013+ APAIS)
- For 2013, calculate annual Peak catch estimates, $\hat{C}_{P2013|y}$, and total (full day) catch estimates, \hat{C}_{T2013} , by sub region, state, and mode, where \hat{C}_{T2013} is the standard 2013 annual estimates by sub region, state, and mode
- Calculate the ratio of Total to Peak catch as $\hat{R}_{2013|y} = \frac{\hat{C}_{T2013}}{\hat{C}_{P2013|y}}$
- Calculate annual Peak catch estimates for prior years, \hat{C}_{Py} , by sub region, state, and mode
- Calculate revised (“calibrated”) prior year Total catch estimates by sub region, state, mode, and year (y) as $\hat{C}_{Ty}^* = \hat{R}_{2013|y} * \hat{C}_{Py}$
- Prior years $y \in \{2004, 2005, \dots, 2012\}$

Ratio Method 1 Description Continued

- Additional Notes
 - Flexible Peak definitions: the definition of Peak is allowed to vary over prior (2004-2012 MRFSS) years as well as across sub region, state, and mode
 - Allows for multiple APAIS design years (2013+) to be used together to calculate ratios
 - The ratios are calculated separately for each species and catch type (landings, releases)

Ratio Method 1 Example

Proportions of Angler-Trips by Hour, Alabama Private Boat, Annual, 2010-2013

(1) Peak time block is defined as [10:00am, 3:00pm) for AL Private Boat Mode, 2010-2012

(2) Calculate Peak estimate for each prior year (\hat{C}_{P2010} , \hat{C}_{P2011} , \hat{C}_{P2012}) as well as Peak and Total estimates for 2013 ($\hat{C}_{P2013|2010}$, $\hat{C}_{P2013|2011}$, $\hat{C}_{P2013|2012}$, and \hat{C}_{T2013})

(3) Calculate Ratios of 2013 Total to 2013 Peak

$$\hat{R}_{2013|2010} = \frac{\hat{C}_{T2013}}{\hat{C}_{P2013|2010}}$$

$$\hat{R}_{2013|2011} = \frac{\hat{C}_{T2013}}{\hat{C}_{P2013|2011}}$$

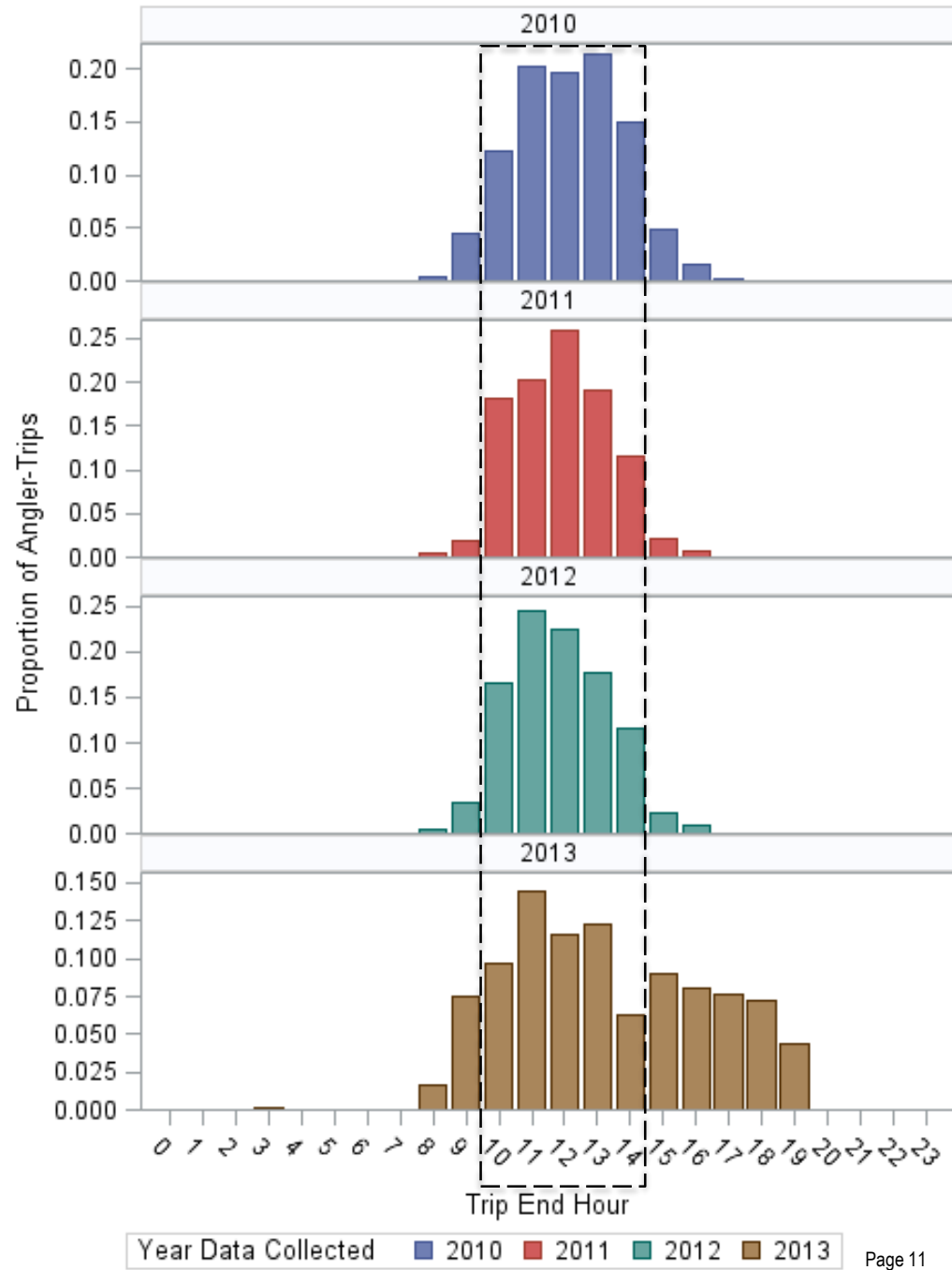
$$\hat{R}_{2013|2012} = \frac{\hat{C}_{T2013}}{\hat{C}_{P2013|2012}}$$

(4) Apply Ratios to Prior Year Peak Estimates

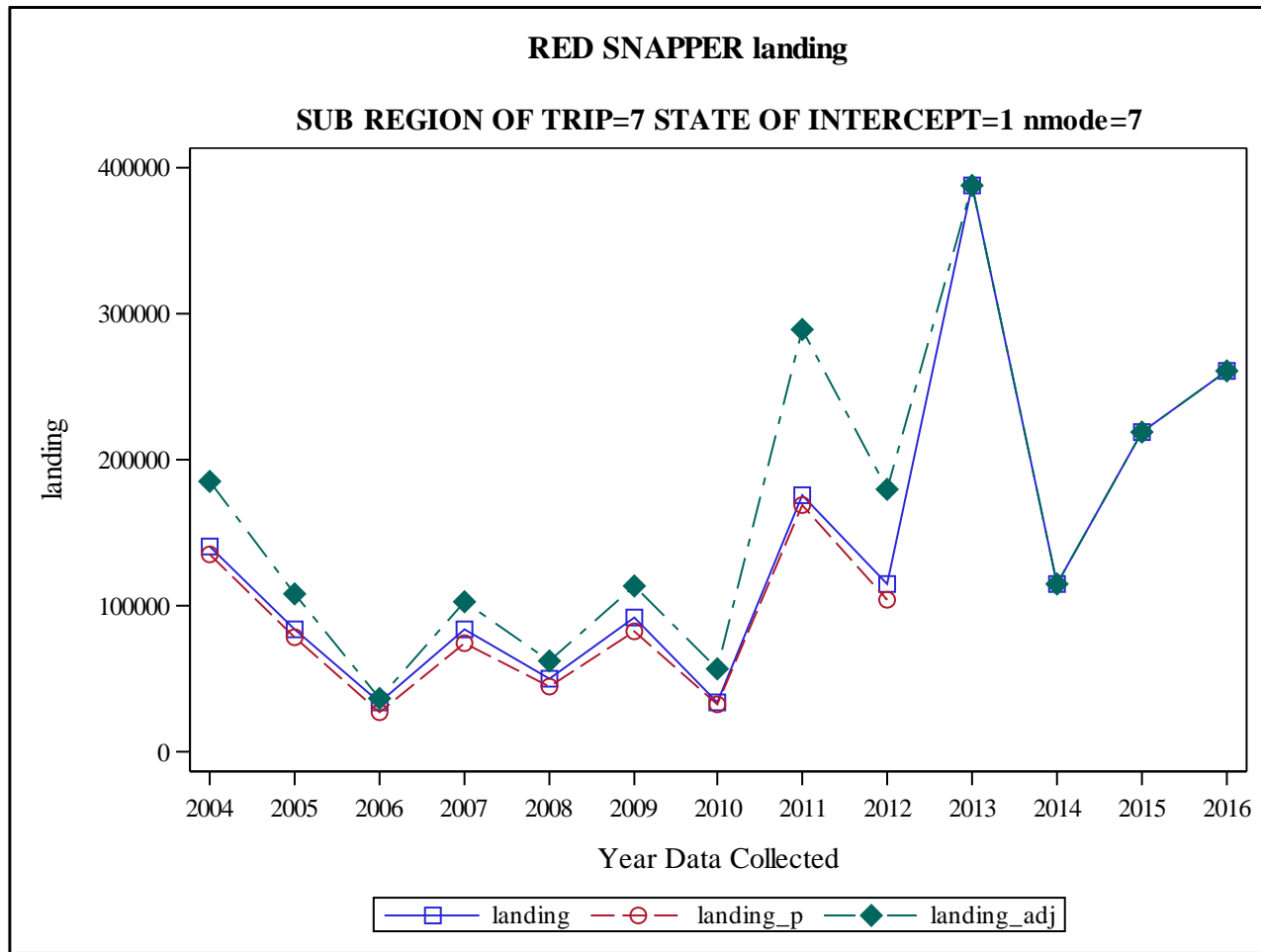
$$\hat{C}_{T2010}^* = \hat{R}_{2013|2010} * \hat{C}_{P2010}$$

$$\hat{C}_{T2011}^* = \hat{R}_{2013|2011} * \hat{C}_{P2011}$$

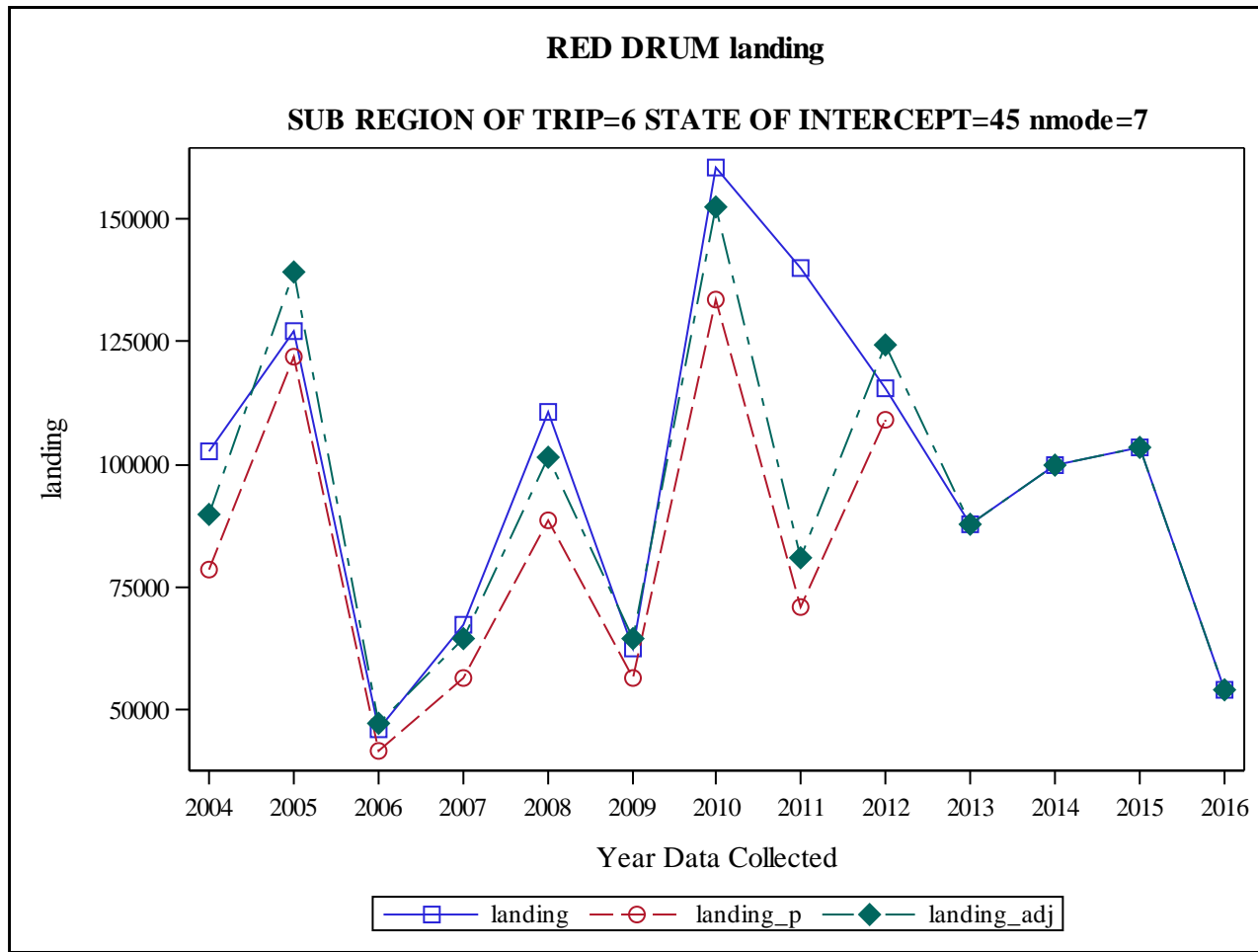
$$\hat{C}_{T2012}^* = \hat{R}_{2013|2012} * \hat{C}_{P2012}$$



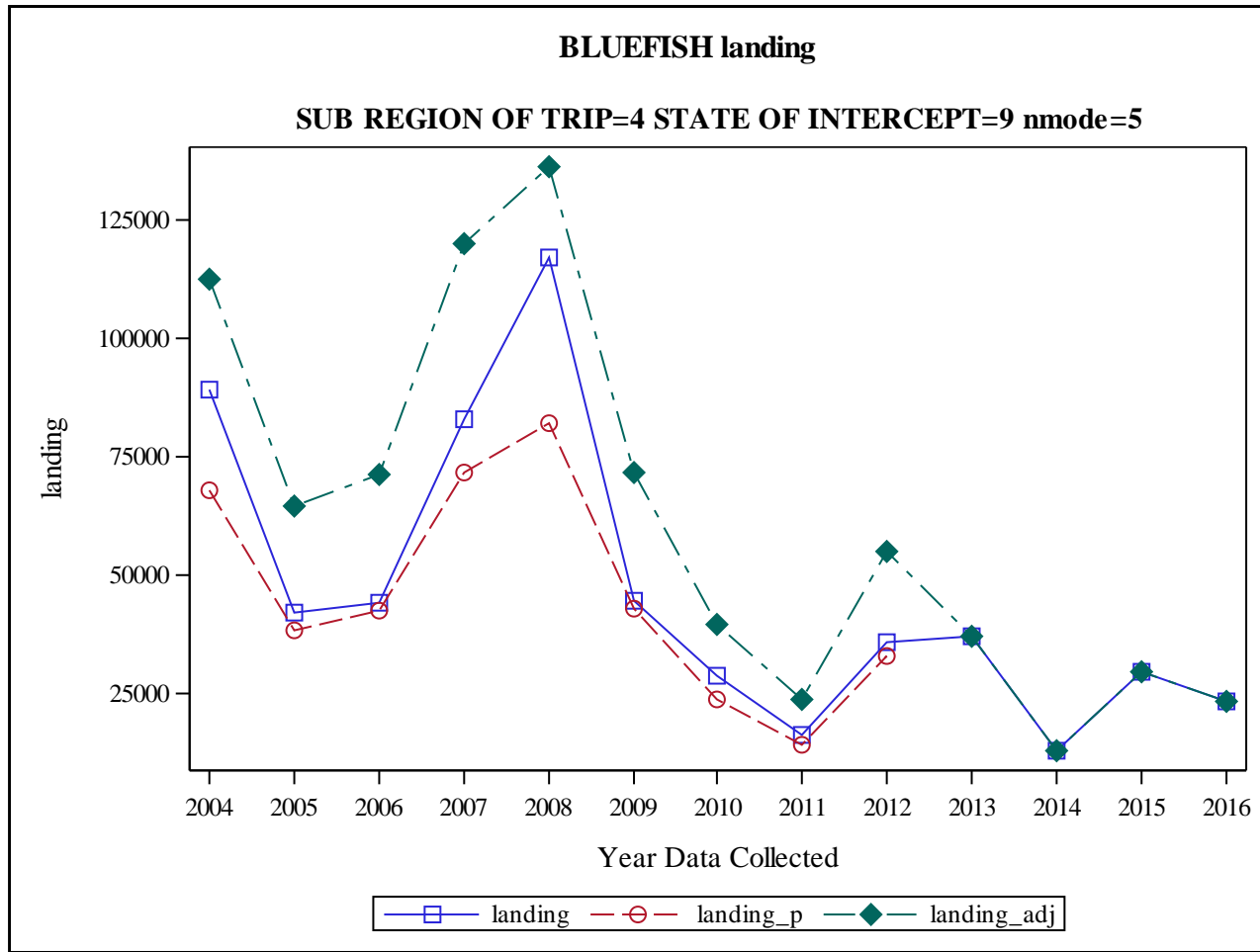
Ratio Method 1 Results – Red Snapper Landings Alabama Private Mode



Ratio Method 1 Results – Red Drum Landings South Carolina Private Mode



Ratio Method 1 Results – Bluefish Landings Connecticut Charter Mode



Ratio Method 1 Limitations

- Doesn't make use of all of the data from prior MRFSS years (2004-2012). Data collected outside of the defined Peak interval is discarded. Across all states in 2004-2012, approximately 85,000 records out of 600,000 are discarded in Private mode, and 24,000 out of 130,000 in Shore mode
- Limited data in estimation cells can cause outlier results
 - Some species by year, state, mode combinations could have zero records during the peak time block
 - Could occur in MRFSS years (2004-2012) resulting in calculated ratio being applied to zero
 - Could occur in APAIS years (2013+) resulting in undefined ratio (dividing by zero)
 - In cases where the majority of catch for a species returned in non-peak hours during APAIS years (2013+), ratios could be exceedingly large

Ratio Method 1 Limitations (continued)

- To handle outlier situations, separate ratios are calculated at higher and higher levels of aggregation and substituted when needed
 - For example, if for a given species, sub region, state, and mode, a ratio is incalculable or larger than a chosen limit (default 10), a substitute ratio is used collapsing across states (species, sub region, mode)
 - If that ratio has issues, a substitute ratio is used collapsing across species (sub region, state, mode)
 - If that ratio is still problematic, a substitute ratio is used collapsing across species and states (sub region, mode)
- Across a selection of 17 species of interest, the base ratio was replaced by a higher-level ratio 28.5% of the time for landings, and 21% of the time for releases. This is more likely to occur with less-common species where data are limited. Some cells don't need any substitutions.

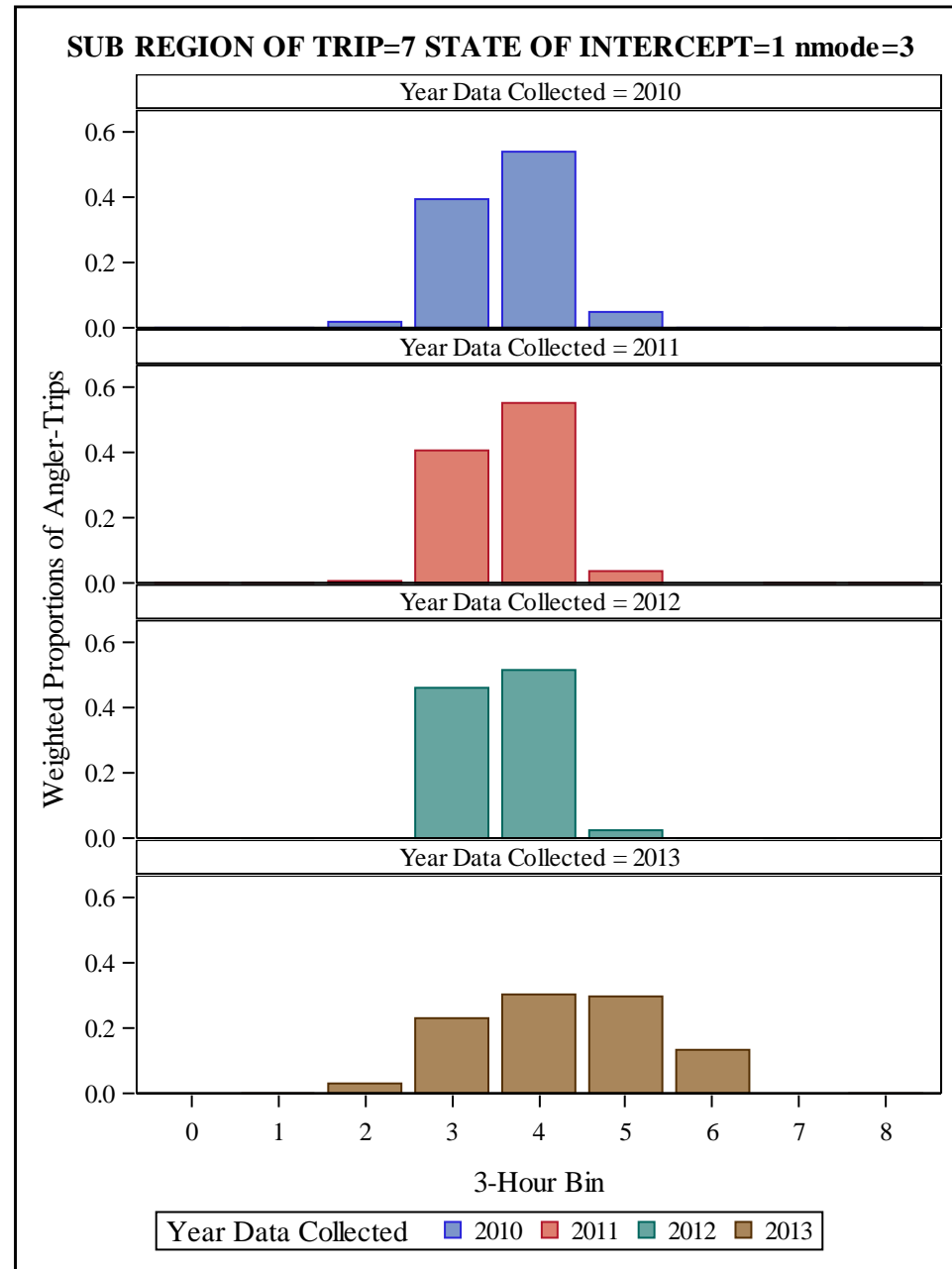
Ratio Method 1 Limitations (continued)

- Provides adjusted estimates at aggregated levels, which limits the utility of survey micro-datasets
 - Ratios calculated at year, sub region, state, mode level of aggregation, and results can be aggregated to higher-level cells, but not easily applied to lower-level cells
 - Difficult to incorporate into sample weights (species-specific adjustments)
 - Adds extra calibration step to any data users' analyses
 - Sensitive to species-specific data distributions over time in reference years
- Can be sensitive to changes in Peak time block definitions
 - Implemented ability to run sensitivity analyses by adjusting starting and ending times of Peak time block definitions by both static and randomized amounts. In many cases, differences were minor, while in some cases, an hour difference could be substantial

Ratio Method 2 (Complex Ratio)

Ratio Method 2 Intro

- **Weighted Proportions of Angler-Trips by 3-Hour Bins (Return Times)**
 - Alabama Shore Mode
 - Data from 2010-2016
 - 2013 represents combined APAIS years (2013-2016)
- Instead of only peak and non-peak designations, match a coarse APAIS (2013+) temporal distribution to a MRFSS year's coarse temporal distribution
- Example 3-Hour Time Bin Definitions
 - 3 – 9:00am-12:00pm
 - 4 – 12:00pm-3:00pm
 - 5 – 3:00pm-6:00pm



Ratio Method 2 (Complex Ratio) Description

- Calculate weighted temporal relative distributions of trips for APAIS years (2013+) and a single prior (MRFSS) year by sub region, state, and mode – trip times grouped into 3-hour bins by default for stability, but can be binned at lower levels
- Make a post-stratification adjustment to the APAIS (2013+) sample weights such that the APAIS temporal distribution matches the prior MRFSS year distribution using the adjusted APAIS sample weights
- Calculate adjusted APAIS (2013+) catch estimates, $\hat{C}_{2013|y}^*$, by sub region, state, and mode using the adjusted sample weights
- Calculate the ratio of standard MRIP 2013 annual catch estimate, \hat{C}_{2013} , to adjusted 2013 annual catch estimates by sub region, state, and mode as

$$\hat{R}_{2013|y} = \frac{\hat{C}_{2013}}{\hat{C}_{2013|y}^*}$$

- Calculate revised (“calibrated”) prior year, y , annual catch estimates by sub region, state, and mode as $\hat{C}_y^* = \hat{R}_{2013|y} * \hat{C}_y$
- Repeat for each individual prior year, $y \in \{2004, 2005, \dots, 2012\}$

Ratio Method 2 Example

Proportions of Angler-Trips by 3-Hour Bins, Alabama Shore, 2010 vs. 2013-2016

(1) Calculate a standard catch estimate for combined APAIS years \hat{C}_{2013} and prior year \hat{C}_{2010}

(2) For each 3-Hour Bin, post-stratify the weights by the ratio of prior year proportion to APAIS proportion

- Bin 2, 2013 down-weighted to match 2010
- Bin 3, 2013 up-weighted to match 2010
- Bin 4, 2013 up-weighted to match 2010
- Bin 5, 2013 down-weighted to match 2010
- Bin 6, 2013 zeroed-out to match 2010

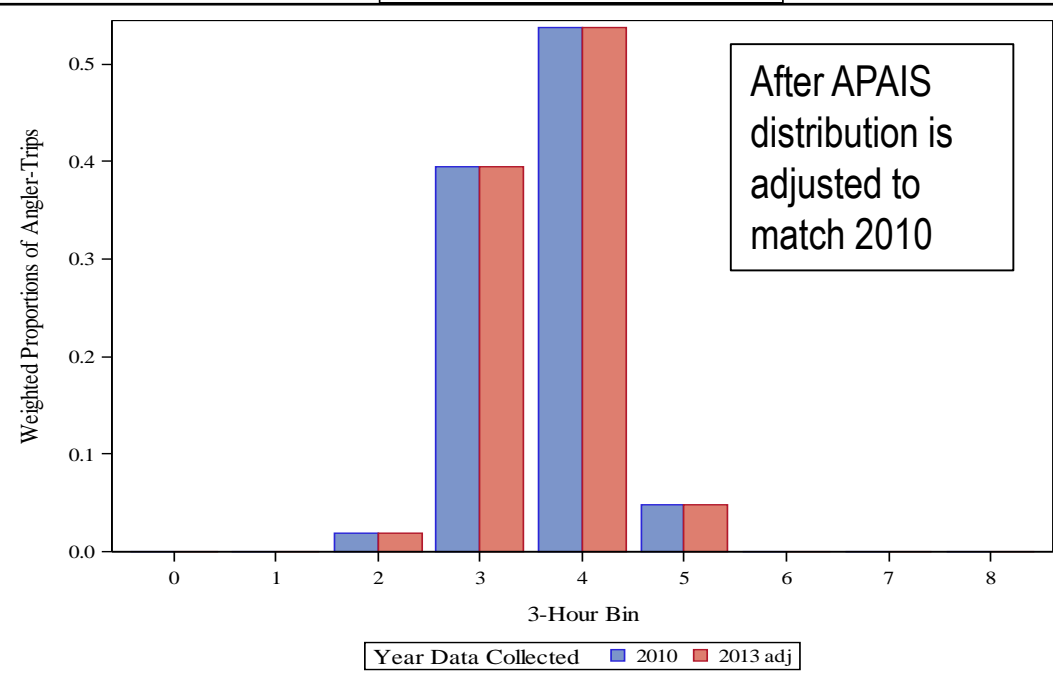
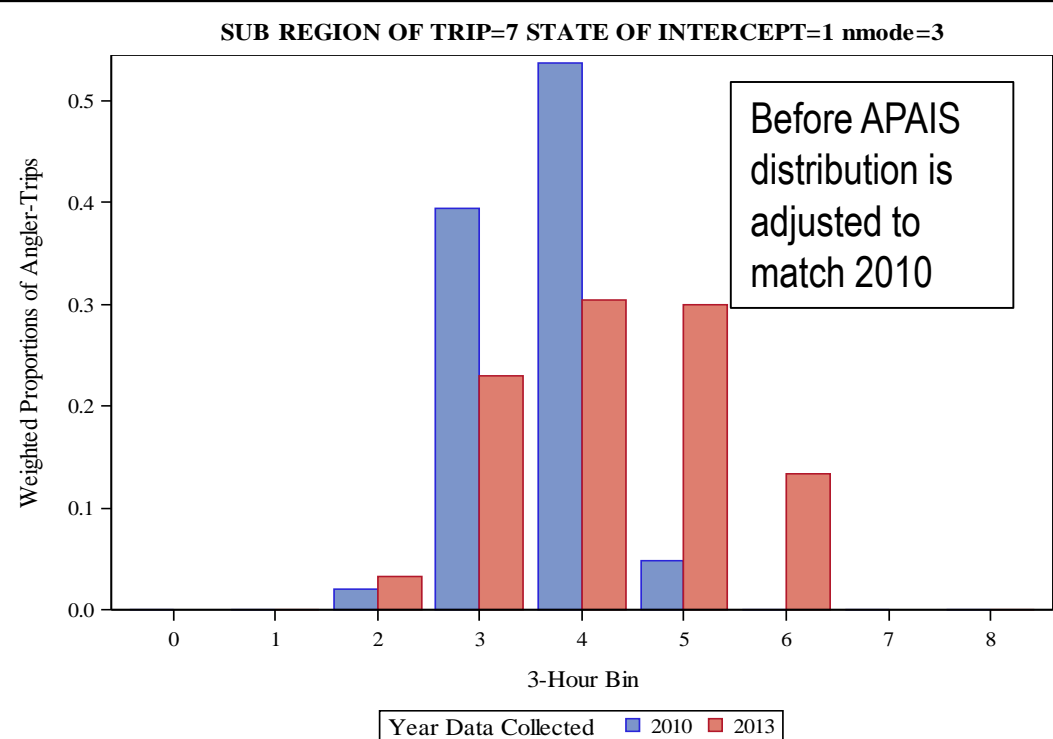
(3) Calculate an adjusted catch estimate for combined APAIS years based on new weights $\hat{C}_{2013|2010}^*$

(4) Calculate a ratio of the standard catch estimate to the adjusted catch estimate

$$\hat{R}_{2013|2010} = \frac{\hat{C}_{2013}}{\hat{C}_{2013|2010}^*}$$

(5) Apply ratio to the prior year estimate (2010) to create revised estimate $\hat{C}_{2010}^* = \hat{R}_{2013|2010} * \hat{C}_{2010}$

(6) Repeat for each prior MRFSS year (2004-2012)



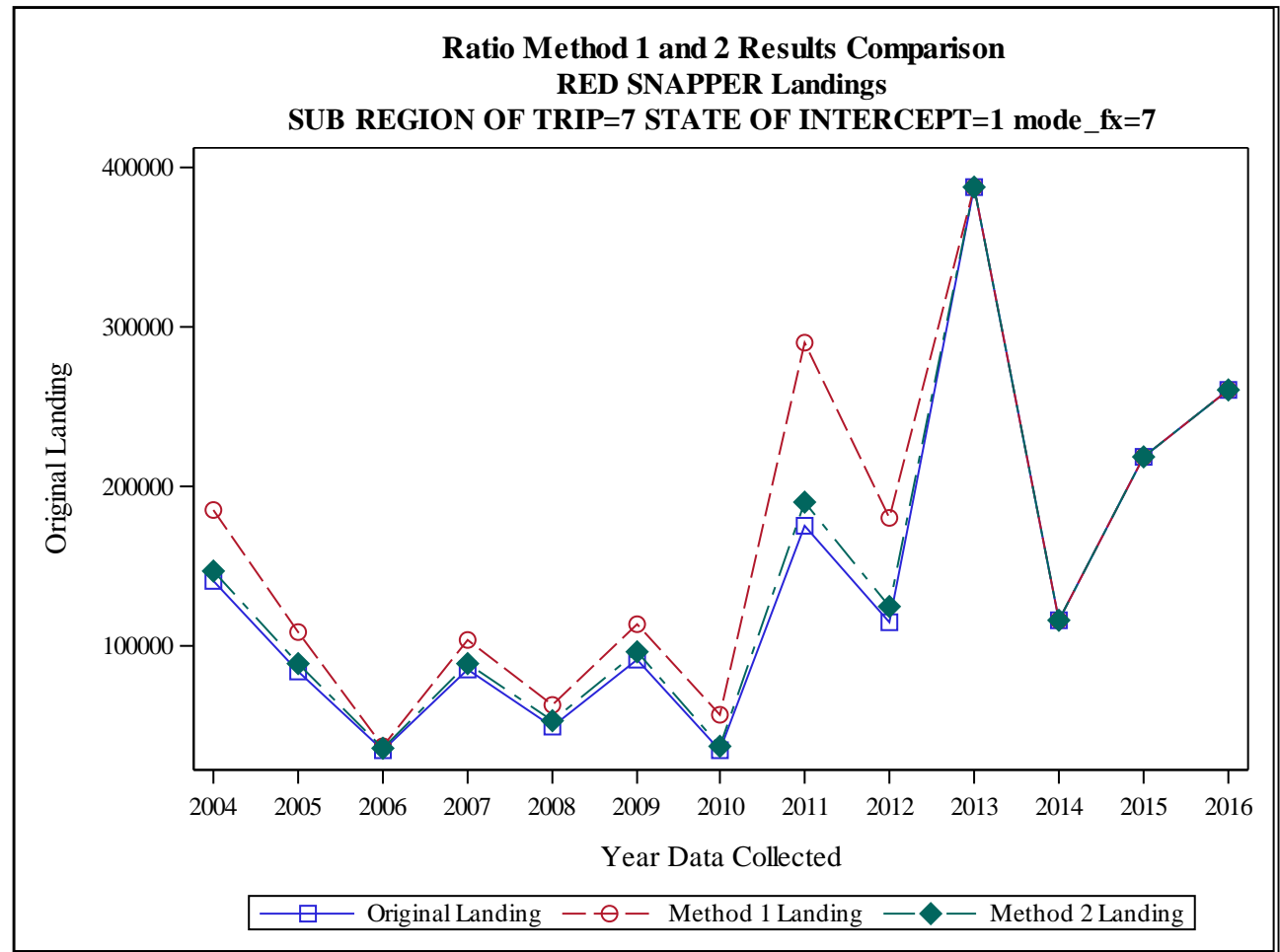
Comparison of Ratio Method 1 and 2 Results

Ratio Method 1 and 2 Results Comparison

Red Snapper Landings, Alabama Private Mode

Method 1 Net Effect:
36% increase

Method 2 Net Effect:
6% increase

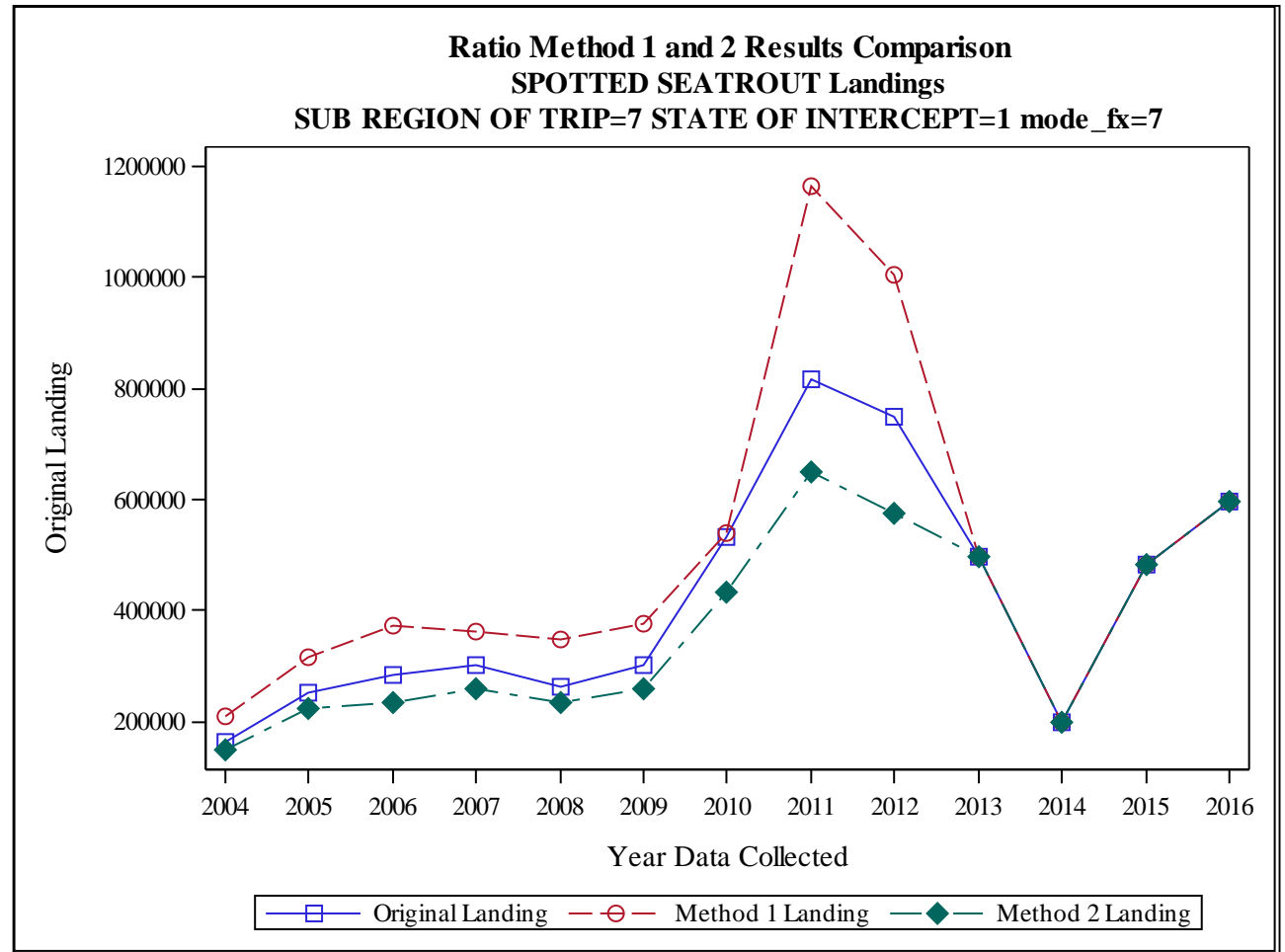


Ratio Method 1 and 2 Results Comparison

Spotted Seatrout Landings, Alabama Private Mode

Method 1 Net Effect:
27% increase

Method 2 Net Effect:
16% decrease

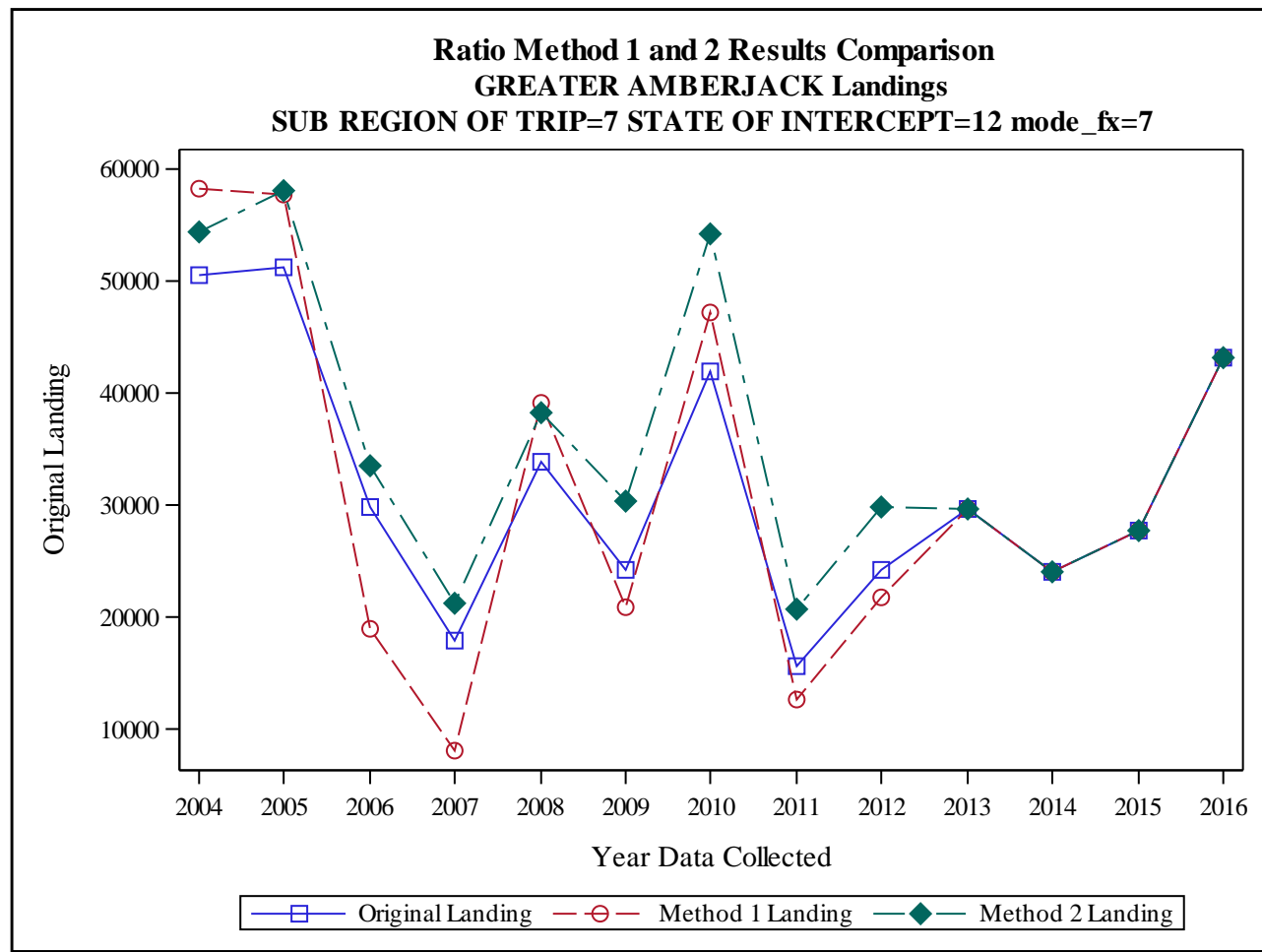


Ratio Method 1 and 2 Results Comparison

Greater Amberjack Landings, West Florida Private Mode

Method 1 Net Effect:
9% decrease

Method 2 Net Effect:
19% increase

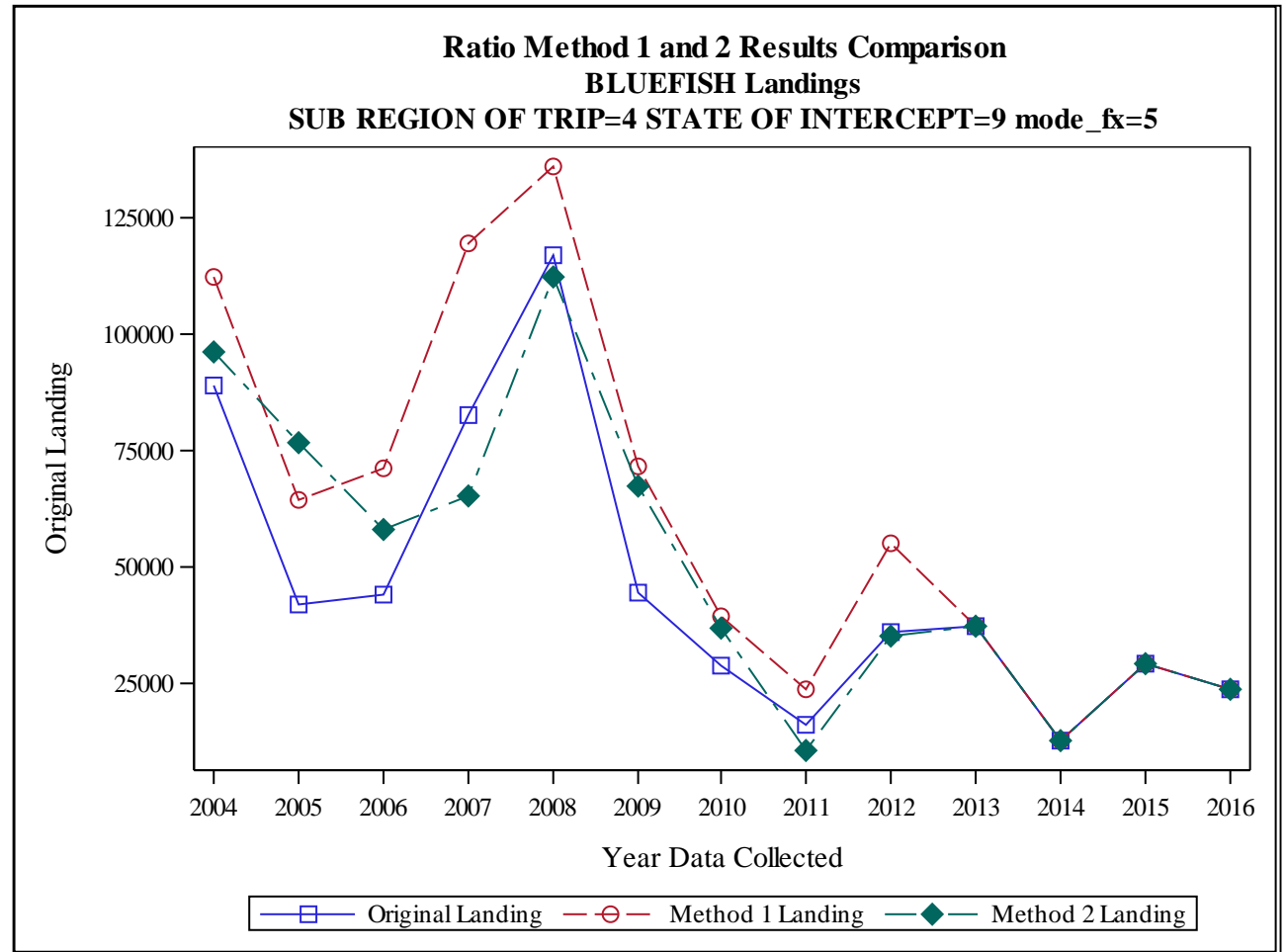


Ratio Method 1 and 2 Results Comparison

Bluefish Landings, Connecticut Charter Mode

Method 1 Net Effect:
45% increase

Method 2 Net Effect:
16% increase



Summary of Sensitivity Analyses and Ratio Method Limitations

Summary of Sensitivity Analyses

- Ratio Method 1
 - Peak time block definitions
 - Fixed peak versus flexible peak by year
- Ratio Method 2
 - Different hour bin lengths for temporal distribution match (1-hour bins, 2-hour bins, 3-hour bins)
- Both Ratio Methods
 - Using additional years of APAIS data for calibration versus a single year (2013)

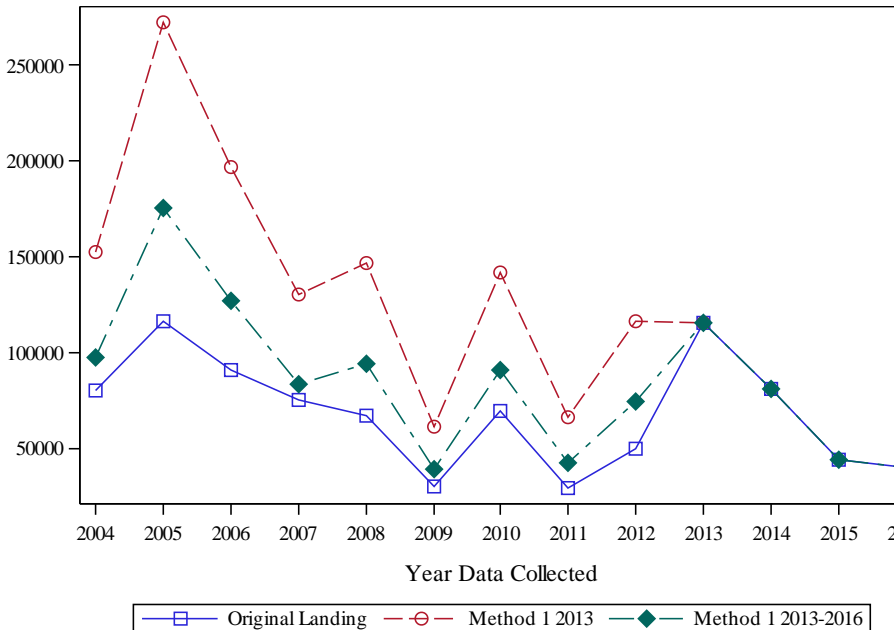
Sensitivity Example: 2013-data vs. 2013-2016

Striped Bass Landings, Connecticut Private Mode

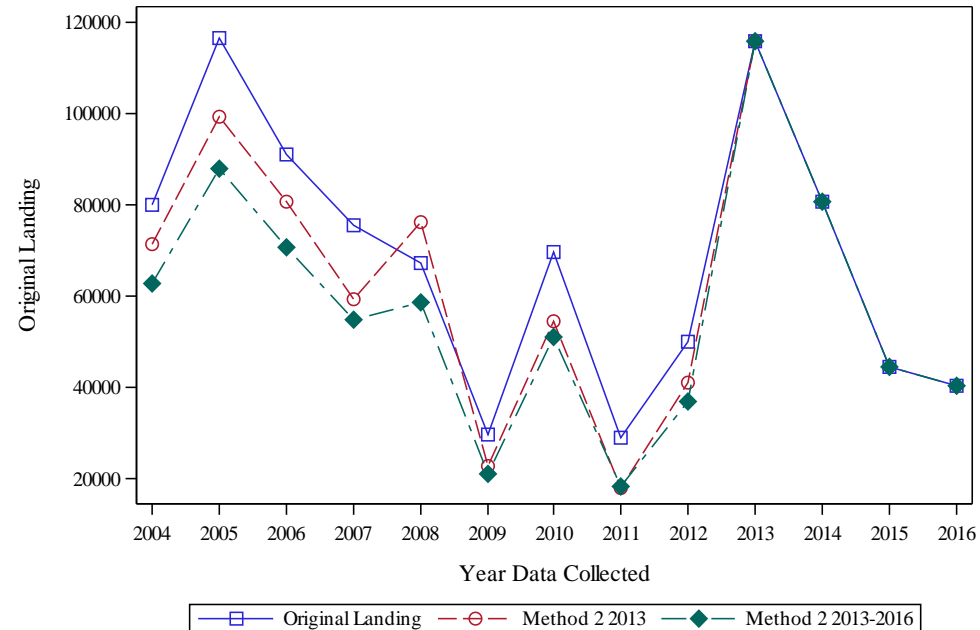
Method 1

Method 2

Ratio Method 1 2013 Data vs. 2013-2016 Data
STRIPED BASS Landings
SUB REGION OF TRIP=4 STATE OF INTERCEPT=9 mode_fx=7



Ratio Method 2 2013 Data vs. 2013-2016 Data
STRIPED BASS Landings
SUB REGION OF TRIP=4 STATE OF INTERCEPT=9 mode_fx=7



Summary of Limitations of Ratio Methods

- Ratio method 1 excludes non-peak data from prior years (~150,000 records discarded across all modes in 2004-2012)
- Ratio methods do not adjust fishing effort in prior years
- Ratio methods negatively impact survey micro-datasets
 - Difficult to incorporate into sample weights
 - Adds extra calibration step to any analysis
 - Ratio methods sensitive to species-specific data distributions over time in reference years
- Limited data in estimation cells can cause outlier results
 - Both Ratio Methods coded to substitute higher-level ratios when needed
- Ratio methods focus on temporal coverage change, while a weight adjustment approach is more general and could account for other effects



NOAA
FISHERIES

APAIS Calibration Methods Considered

An Overview of the Ratio Methods Considered for APAIS Calibration

Ryan Kitts-Jensen
Office of Science
& Technology