MRIP Estimation Methods for Total Catch, Landings and Releases

Standing, Reef Fish, Mackerel, Ecosystem, & Socioeconomic MRIP SSC Meeting
July 8-9,2020

Presentation Outline

I. Overview of APAIS Design

II. Missing Data Situations and Imputation

III. Weighted Estimation

IV. Data and Estimate QA/QC

NOAA FISHERIES

I. Overview of APAIS Design

II. Missing Data Situations and Imputation

III. Weighted Estimation

IV. Data and Estimate QA/QC

NOAA FISHERIES

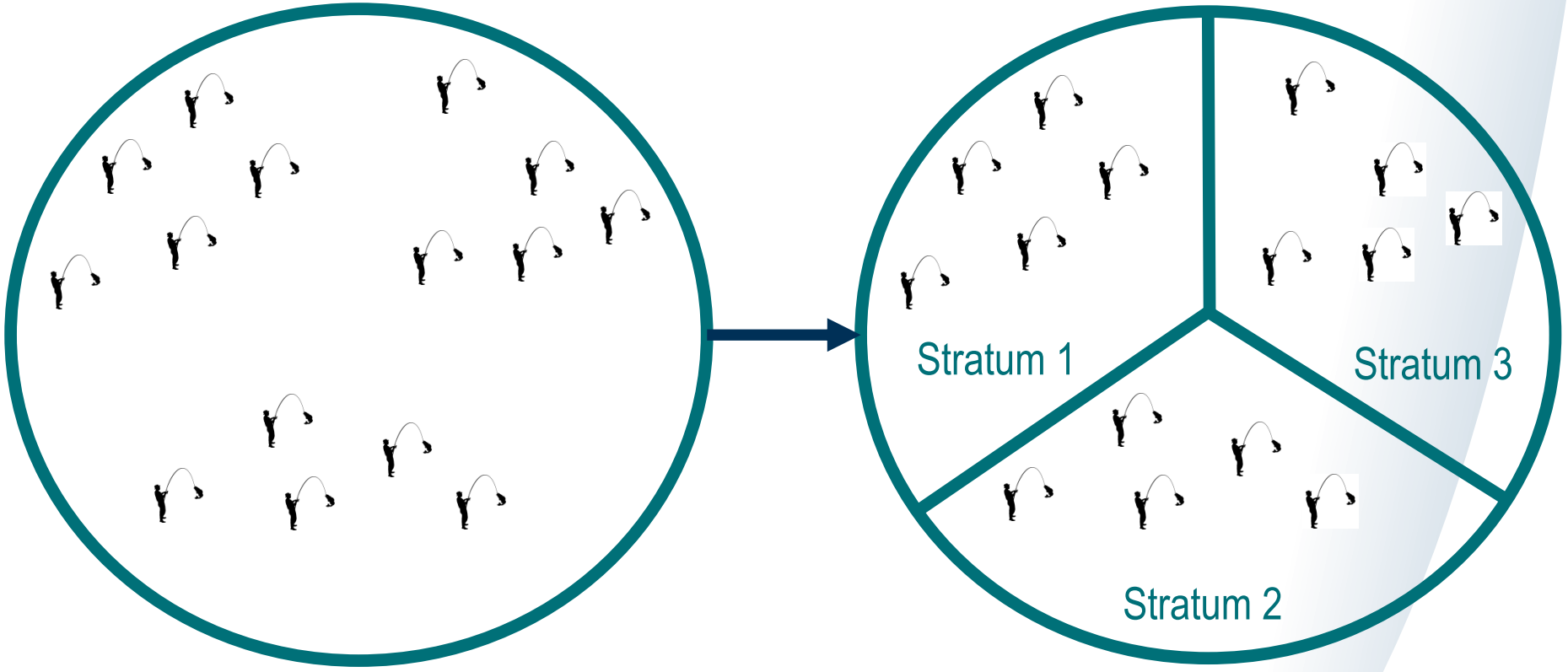
APAIS Overview

- In-person **interviews of anglers** intercepted at **public access fishing sites**
- Sample frame derived from NOAA Fisheries Public Fishing Access Site Register
- Data collected continuously, used to **estimate catch rates** and trip characteristics for two-month waves

NOAA FISHERIES

APAIS Design

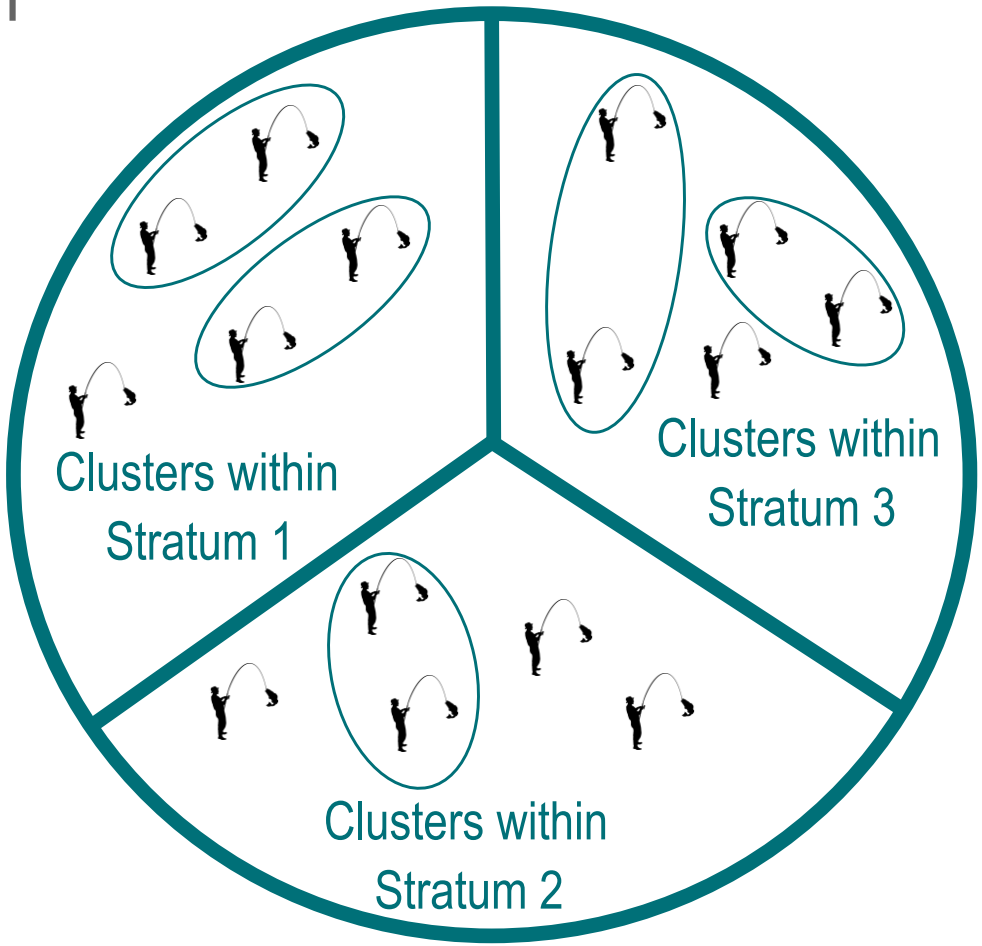
Stratified, clustered multi-stage design



NOAA FISHERIES

APAIS Design

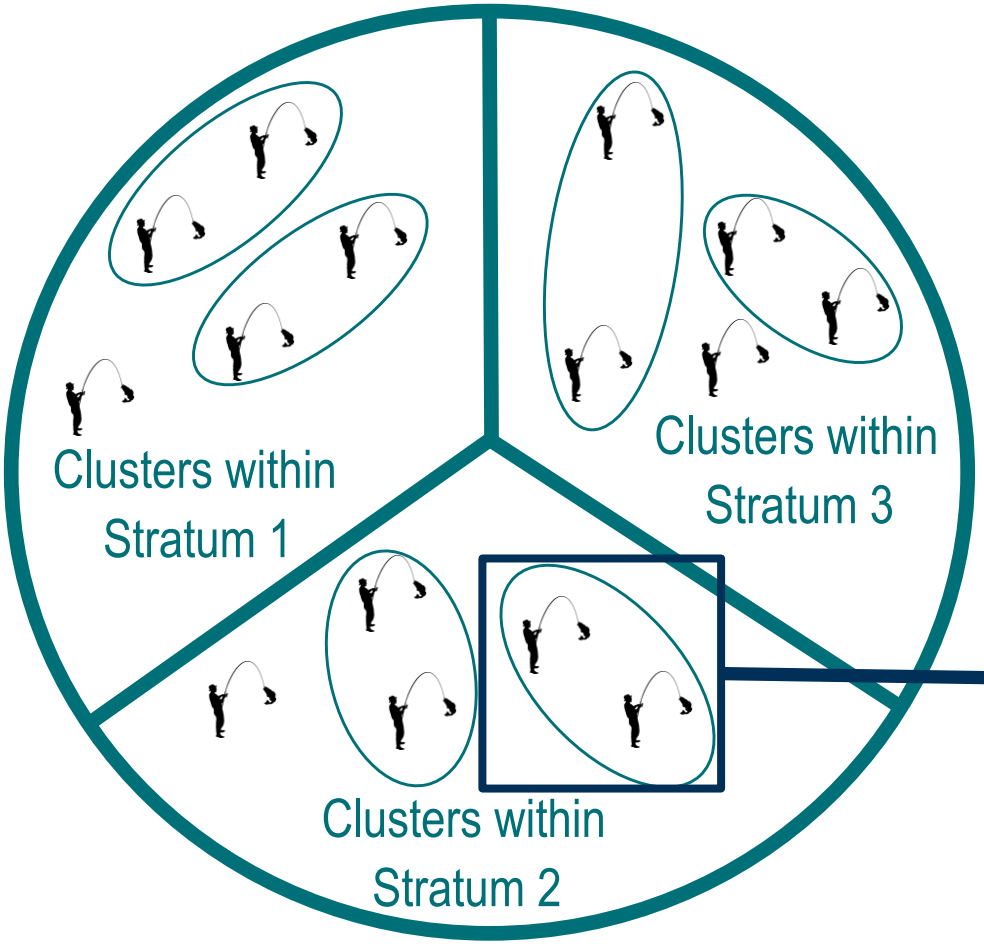
Stratified, **clustered** multi-stage design



NOAA FISHERIES

APAIS Design

Stratified, clustered **multi-stage design**

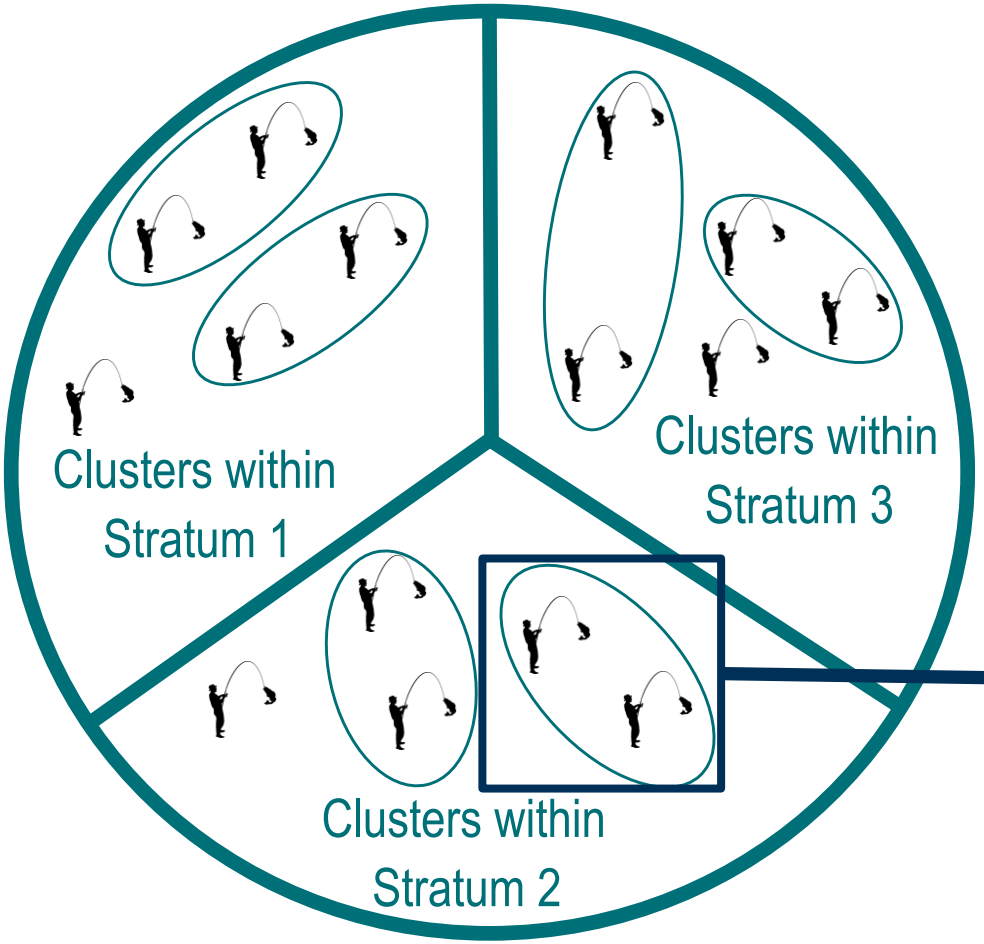


Primary Stage Unit (PSU):
Site Cluster-Day-Time Interval

NOAA FISHERIES

APAIS Design

Stratified, clustered **multi-stage design**



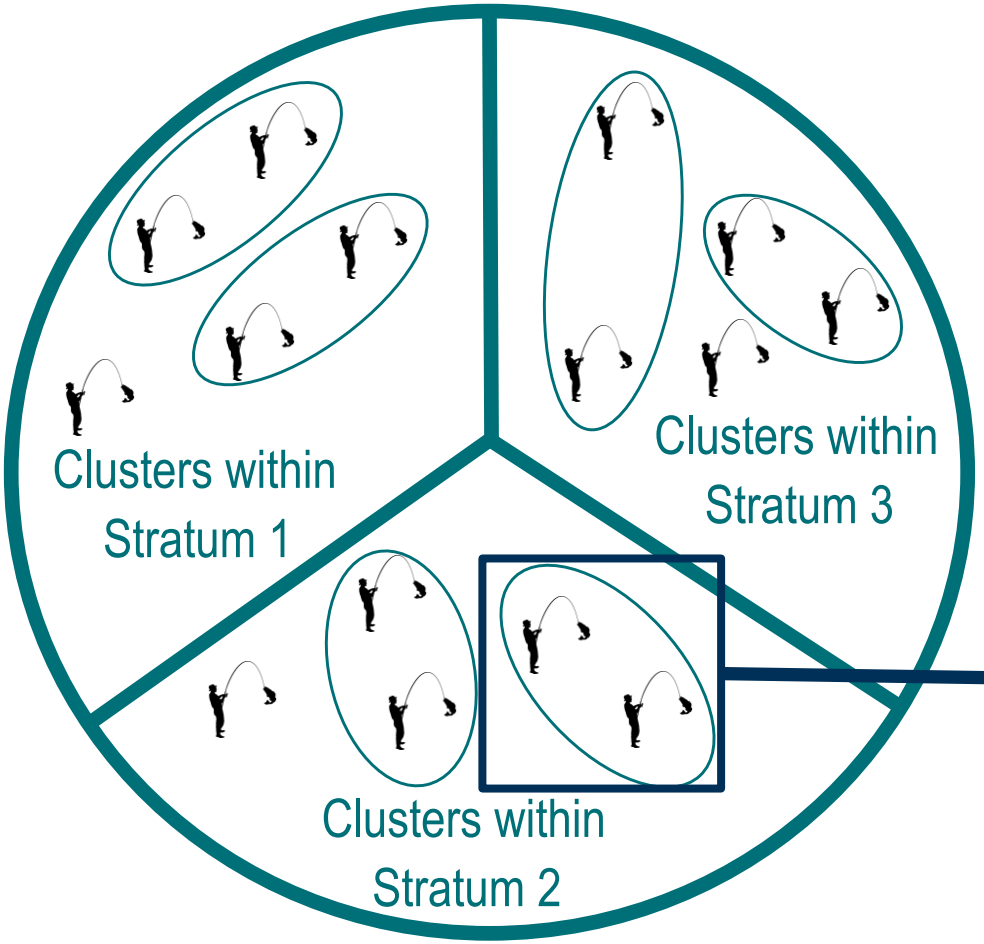
Primary Stage Unit (PSU):
Site Cluster-Day-Time Interval

Secondary SU:
Sample Duration (time spent sampling each site in a cluster)

NOAA FISHERIES

APAIS Design

Stratified, clustered **multi-stage design**



Primary Stage Unit (PSU):
Site Cluster-Day-Time Interval

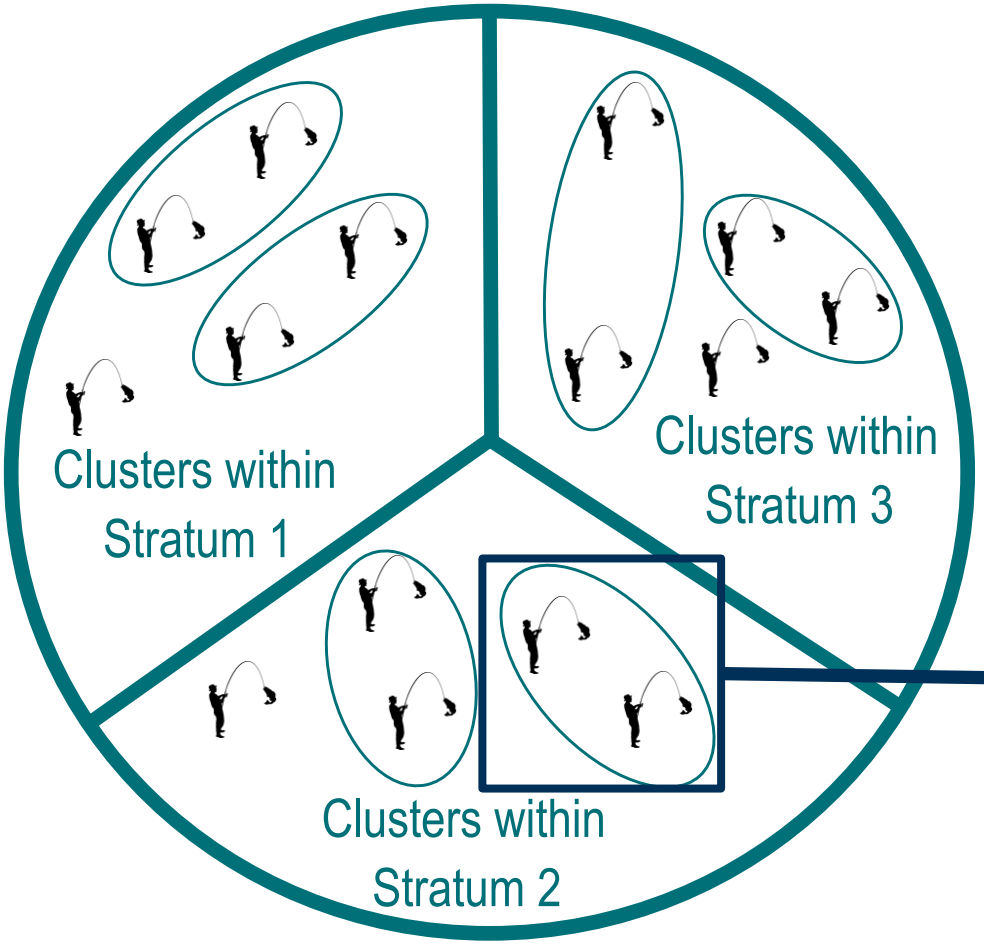
Secondary SU:
Sample Duration (time spent sampling each site in a cluster)

Tertiary SU:
Angler Trips

NOAA FISHERIES

APAIS Design

Stratified, clustered **multi-stage design**



Primary Stage Unit (PSU):
Site Cluster-Day-Time Interval

Secondary SU:
Sample Duration (time spent sampling each site in a cluster)

Tertiary SU:
Angler Trips

Quaternary SU:
Catch

NOAA FISHERIES

APAIS Sample Selection

- Probability of selecting PSU's based on **fishing pressure**
  - Higher probability of selecting high pressure sites

Expected Number of Angler-Trips	Size Measure (Weight)
1-4 Angler-trips	0.5
5-8	2.5
9-12	9
13-19	13
20-29	20
30-49	30
50-79	50
80+	80
Mode not present at site/site is inactive	0

Probability proportional to size sampling used to select sample, with logistical field constraints (e.g. maximum number of sites visits scheduled per day) incorporated into the process

NOAA FISHERIES

I. Overview of APAIS Design

II. Missing Data Situations and Imputation

III. Weighted Estimation

IV. Data and Estimate QA/QC

NOAA FISHERIES

Catch Information Collected by APAIS

Type A
Harvest observed by APAIS sampler

Type B1
Harvest unobserved by APAIS sampler

Type B2
Releases reported by angler

Used to produce the following estimates:

- Total landings (in numbers of fish)
- Total landings (in lbs or kg)
- Mean fish weights and lengths
- Length frequencies

Used to produce the following estimates:

- Total released alive (in numbers of fish)

Total Catch (Landings + Releases in numbers of fish)

NOAA FISHERIES

Missing Data Situations

- Missing only length OR weight data
- Missing both length AND weight data

NOAA FISHERIES

Missing Length OR Weight

- Fill in length (L) or weight (W) data with standard length-weight relationship models

$$L = \left(\frac{W}{a}\right)^{\frac{1}{b}}$$

Shape parameter for the body type of the fish species

Scaling coefficient for the weight at length of the fish species

$$W = aL^{b}$$

NOAA FISHERIES

Missing Length AND Weight

- Fill in up to 5 length and weight observations for each species on a given trip using a combination of **hot and cold deck imputation**

Hot Deck

Where missing values are replaced by values from a similar unit in the **same** dataset

(e.g., species lengths/weights from the same year, wave and state)

Cold Deck

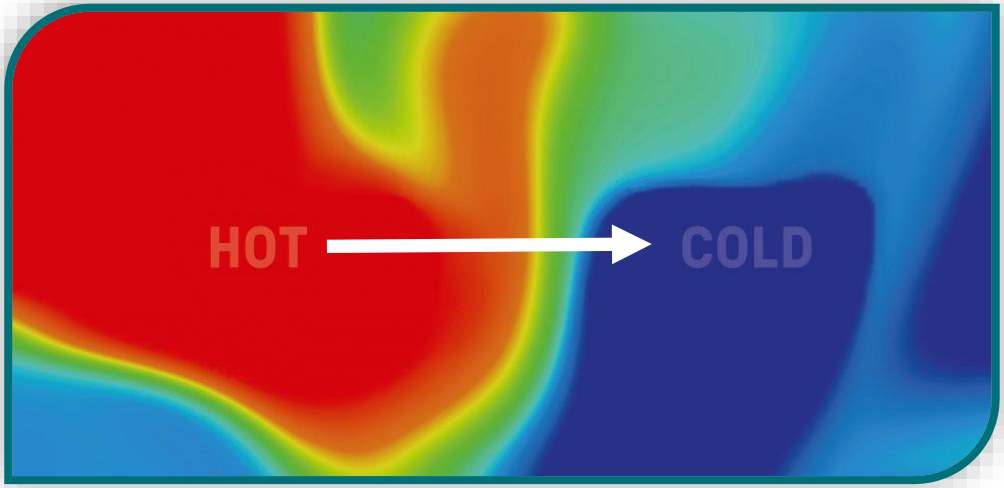
Where missing values are replaced by values from a similar unit in a **different** dataset

(e.g., species lengths/weights from the same sub-region, but from last year)

NOAA FISHERIES

Length and Weight Imputation

- 5 rounds of attempted imputation, starting with most similar data proceeding to less similar data in each round
  - The majority of imputations are completed within 3 rounds
  - Imputed data are **always from the same species and sub-region**
  - Sample weights are not factored into imputation so as to not introduce unknown biases into the data



NOAA FISHERIES

Length and Weight Imputation

Round	Imputation Cell	Minimum number of completed observations required
1	Current year, wave, sub-region, state, mode, area fished, species	10
2	Current year, half-year (waves 1-3 or 4-6), sub-region, state, mode, species	5
3	Current + most recent prior year, wave, sub-region, state, mode, area fished, species	5
4	Current + most recent prior year, sub-region, state, mode, species	5
5	Current + most recent prior year, sub-region, species	1

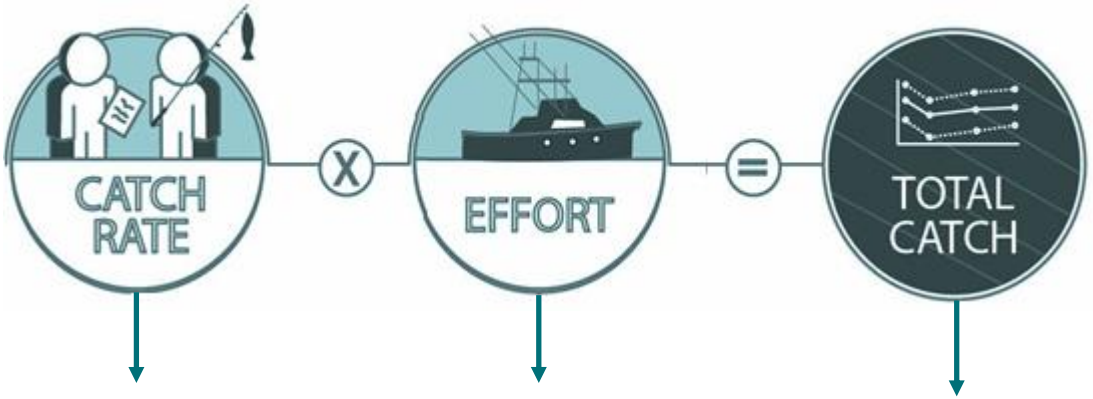
If the minimum number of observations are not met, imputation proceeds to the next round. If no imputations can occur after 5 rounds, no additional attempts are made to fill in missing data.

NOAA FISHERIES

I. Overview of APAIS Design
II. Missing Data Situations and Imputation
III. Weighted Estimation
IV. Data and Estimate QA/QC

NOAA FISHERIES

Catch Estimation – Basic



Estimated number of angler trips

Estimated number of fish caught per angler trip

Estimated total number of fish caught

NOAA FISHERIES

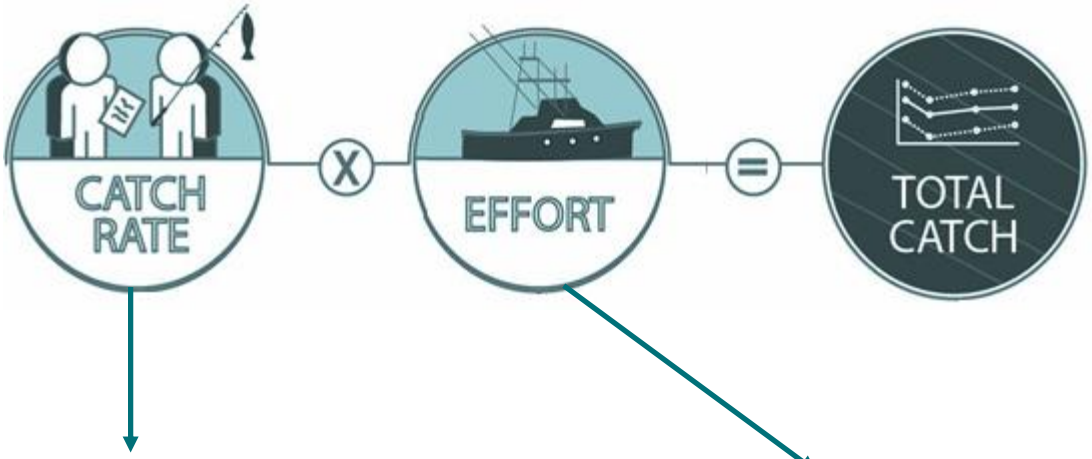
Catch Estimation – Broken Down

Weighted APAIS catch rate

- includes 3 sample weighting components
- calculated using standard weighted mean estimator

NOAA FISHERIES

Catch Estimation – Broken Down



Weighted APAIS catch rate

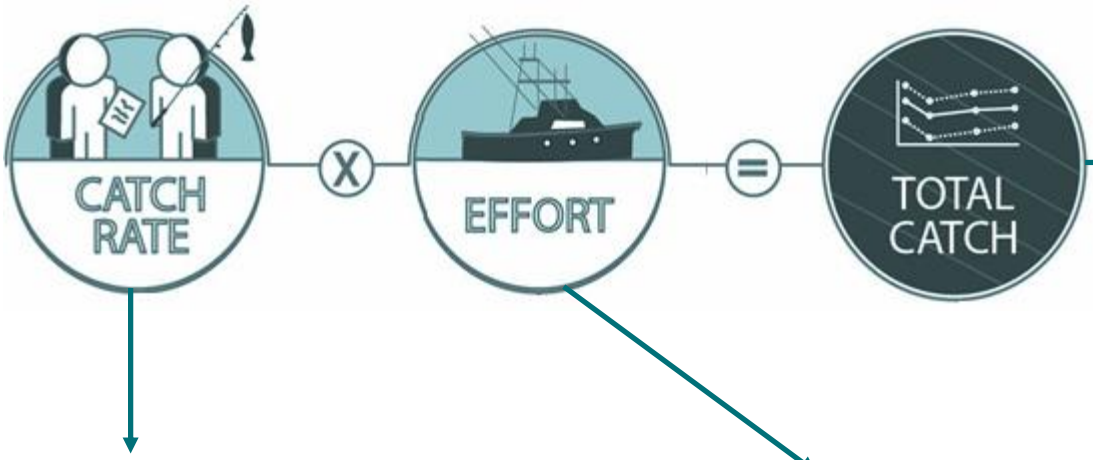
- includes 3 sample weighting components
- calculated using standard weighted mean estimator

Weighted FES Effort

- includes 3 sample weighting components
- calculated using standard weighted total estimator
- **From APAIS**: an adjustment factor to account for out-of-state angler trips
- **From APAIS**: partitioned by area fished (inland, nearshore, offshore)

NOAA FISHERIES

Catch Estimation – Broken Down



- Observed Landings (A)
- Unobserved Landings (B1)
- Releases (B2)
- Total Landings (A+B1)
- Landings + Releases (A+B1+B2)

Weighted APAIS catch rate

- includes 3 sample weighting components
- calculated using standard weighted mean estimator

Weighted FES Effort

- includes 3 sample weighting components
- calculated using standard weighted total estimator
- **From APAIS**: an adjustment factor to account for out-of-state angler trips
- **From APAIS**: partitioned by area fished (inland, nearshore, offshore)

NOAA FISHERIES

APAIS Sample Weights

PSU – Stage I Sample Weight

$$w_1 = \frac{1}{inclusion\ probability}$$

APAIS Sample Weights

PSU – Stage I Sample Weight

$$w_1 = \frac{1}{inclusion\ probability}$$

Sample Duration – Stage II Sample Weight

$$w_2 = \frac{length\ of\ assigned\ time\ interval}{time\ spent\ subsampling\ at\ a\ site\ within\ a\ cluster}$$

APAIS Sample Weights

PSU – Stage I Sample Weight

$$w_1 = \frac{1}{inclusion\ probability}$$

Sample Duration – Stage II Sample Weight

$$w_2 = \frac{length\ of\ assigned\ time\ interval}{time\ spent\ subsampling\ at\ a\ site\ within\ a\ cluster}$$

Angler-Trip – Stage III Sample Weight

$$w_3 = \frac{Total\ angler\ trips\ observed\ at\ a\ site}{Total\ angler\ trips\ intercepted\ at\ a\ site}$$

APAIS Sample Weights

PSU – Stage I Sample Weight

$$w_1 = \frac{1}{inclusion\ probability}$$

Sample Duration – Stage II Sample Weight

$$w_2 = \frac{length\ of\ assigned\ time\ interval}{time\ spent\ subsampling\ at\ a\ site\ within\ a\ cluster}$$

Angler-Trip – Stage III Sample Weight

$$w_3 = \frac{Total\ angler\ trips\ observed\ at\ a\ site}{Total\ angler\ trips\ intercepted\ at\ a\ site}$$

Catch – Stage IV Sample Weight

$$w_4 = \frac{Total\ number\ of\ fish\ harvested\ by\ an\ angler\ trip}{Number\ of\ fish\ sampled\ from\ that\ angler\ trip}$$

APAIS Catch Rate Estimates

Mean catch per angler trip is calculated as a **domain estimate**, defined by year, wave, region, state, fishing mode, area fished (inland, nearshore, offshore), species and catch type.

Final sample weight ($w_1 * w_2 * w_3$ if estimating numbers of fish, or $w_1 * w_2 * w_3 * w_4$ if estimating total landed weight) for angler trip in domain d

Number of individuals OR weight (lbs or kg) of fish caught on angler trip in domain d

Catch rate in domain d

$$\hat{\bar{y}}_d = \frac{\sum w_d y_d}{\sum w_d}$$

This is a standard weighted mean estimator used in survey statistics (e.g. SAS Institute Inc, 2016)

NOAA FISHERIES

FES Sample Weights

1. **Base design weight** Of household i in stratum h	$w_{hi} = \frac{1}{inclusion\ probability}$

NOAA FISHERIES

FES Sample Weights

1. Base design weight Of household i in stratum h	$w_{hi} = \frac{1}{inclusion\ probability}$
2. Nonresponse adjustment Done to minimize response bias	Sample partitioned into nonresponse adjustment cells, and adjusted by response rates ($w_{hi}{}^{*}$)

NOAA FISHERIES

FES Sample Weights

1. Base design weight Of household i in stratum h	$w_{hi} = \frac{1}{inclusion\ probability}$
2. Nonresponse adjustment Done to minimize response bias	Sample partitioned into nonresponse adjustment cells, and adjusted by response rates (w_{hi}^{*})
3. Post-stratification adjustment Done to improve representativeness of sample (common technique used to conform population totals to an independent survey)	Sample matched to demographic controls from the U.S. Census Bureau's American Community Survey residential household estimates (w_{hi}^{**})

NOAA FISHERIES

FES Effort Estimates

Estimated number of angler trips by year, wave, region, state, fishing mode

$$\hat{T} = \sum w_{hi}^{**} t_{hi}$$

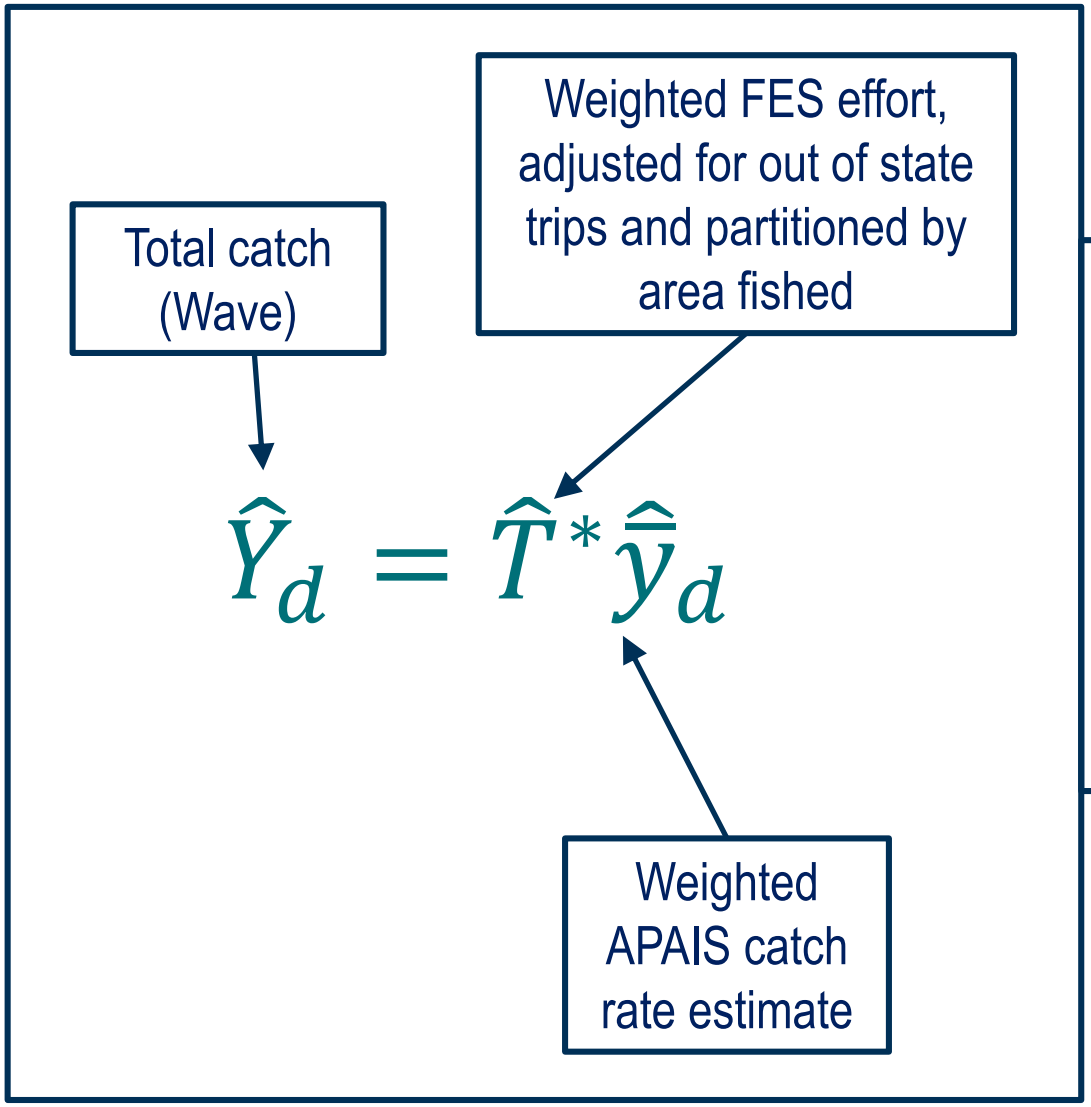
Trips taken by household *i* in stratum *h*

Final sample weight of household *i* in stratum h (comprised of base weight, nonresponse weight and post-stratification adjustment)

This is a Horvitz-Thompson total estimator, a standard method for estimating the total of a stratified sample.

NOAA FISHERIES

Total Catch Estimates



Total catch (Annual)

$$\hat{Y}_D = \sum \hat{Y}_d$$

NOAA FISHERIES

I. Overview of APAIS Design
II. Missing Data Situations and Imputation
III. Weighted Estimation
IV. Data and Estimate QA/QC

NOAA FISHERIES

APAIS QA/QC

Step 1
- Each state's lead biologist checks and edits raw fish records for unusual data elements

Step 2
- States runs automated checks to identify 1) potential errors and 2) produce warnings when data elements seem unusual (e.g. high fish counts).

Step 3
- Data Delivered to NMFS

Step 4
- NMFS runs the same automated checks to identify 1) potential errors and 2) produce warnings when data elements seem unusual (e.g. high fish counts).
  - Done to double check for any raw data issues that may need to be addressed prior to using the data in estimation.

NMFS and partners follow standard protocols to flag unusual data and minimize data processing errors.

NOAA FISHERIES

FES QA/QC

Step 1
- Verify survey deliverables

Step 2
- Identify item nonresponse and illogical responses and make minor, logic-based edits if possible.

Step 3
- Check for contradictory, nonsensical, and unlikely or extreme values

Step 4
- Adjust weight(s) of non-representative* values

*Non-representative values = those that contribute a disproportionate amount of effort to the total effort estimate within a domain for the current wave OR represent outliers within the time series

NMFS and partners follow standard protocols to flag unusual data and minimize data processing errors.

NOAA FISHERIES

Estimate Review

- Estimates are reviewed each wave prior to public release by:
  - NMFS OST Fisheries Statistics Division
  - Representatives from NOAA Regional Offices and Science Centers
- Any additional data errors that are detected at this stage are corrected by NMFS staff or state partners
- Any outliers are investigated by NMFS staff to determine appropriate action

Summary

- MRIP weighted estimation is reflective of complex survey designs, and the techniques used are standard in the survey statistics field.
- Minimal imputation is used to fill in missing data.
- Data are systematically and repeatedly reviewed by multiple parties to minimize processing errors.
- Outlier estimates are reviewed on a case-by-case basis.

NOAA FISHERIES