Office of Science
and Technology
Marine
Recreational Information
Program

## MRIP Data User Seminar: Statistical Methods and Procedures

November 30, 2021 John Foster

## Overview

## Survey Design and Statistical Methods

- Surveys and Estimates Schematic
- Access Point Angler Intercept Survey (APAIS)
- Fishing Effort Survey (FES)
- For-Hire Survey (FHS)
- Catch and Effort Estimation

Component Estimates

Catch Rates (CPUE)

Effort Components:
Area Fished Proportions FES Coverage Adjustment

FHS Coverage Adjustment

Private Boat, Shore OnFrame Effort

Charter Boat, Headboat OnFrame Effort

Primary
Estimates
Total Effort

# Access Point Angler Intercept Survey (APAIS) 

## APAIS Overview



- In-person interviews of anglers intercepted at public fishing access sites
- Anglers interviewed at the end of their fishing trips

- Samplers record detailed trip characteristic and catch information including individual fish length and weight measurements


## APAIS

- Resource Links
- Design
- Sample Weighting
- Weighted Estimation Components



## APAIS Resource Links

- Survey Design and Statistical Methods


## APAIS Section 2

- Site Register
- APAIS At-a-Glance
- Outreach Information


NOAA
FISHERIES

## APAIS Design

- Complex Probability-Based Design
- Sample Frame
- Stratification
- Multi-stage Clustering
- Sample Selection using Probability Proportional to Size


## APAIS Design: Sample Frame

- List Fishing Access Sites
- Calendar
- 6-hour Time Intervals
- Primary Stage Unit
- 1 or 2 Sites (Site-Cluster)
- Date
- Time Interval



## APAIS Design: Stratification

- Space
- State, Sub-state regions
- Time
- Month
- Kind-of-Day (weekday, weekend)
- Interval (day, night)
- Fishing Access Site Group
- Grouping sites by predominant mode or other trip characteristics
- Shore, Private Boat, Charter Boat, Offshore


## APAIS Design: Multi-stage Clustering



## APAIS Design: Sample Selection

- Primary Stage Units (PSU): Site cluster-day-time interval
- PSUs selected using a probability proportional to size (PPS) approach
- Chance of being selected is related to the expected amount of fishing activity or fishing pressure

| Expected Number of <br> Angler Trips | Size Measure |
| ---: | :--- |
| 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 |

## APAIS Design: Sample Selection

- Estimates of expected fishing pressure continually updated by regional and state agency partners that conduct APAIS field sampling
- For every site, pressures provided separately for each combination of month, kind-of-day, 6-hour time interval, mode of fishing
- All fishing pressures and other site characteristics available in Public Fishing Access Site Register


## APAIS Sample Weighting

- Design aspects that impact the probability or chance of including an angler-trip in the APAIS sample must be accounted for in the sample weights and sample weights must be used in estimation
- Sample weight is the inverse (reciprocal) of the probability that a trip is included in the sample (e.g., a trip has a $10 \%$ chance of being interviewed, sample weight is $1 / 10 \%=1 / 0.10=10$ )
- APAIS has multiple stages of sampling, each stage has a separate inclusion probability and corresponding sample weight
- Final APAIS sample weight for each interviewed trip is the product of the individual weights associated with each separate stage


## APAIS Sample Weighting

1. Primary Stage Unit (PSU):

Site Cluster-Day-Time Interval

## 2. Secondary SU: Sample Duration

(time spent sampling at each site in a cluster)
3. Tertiary SU: Angler Trips (trips sampled from all trips observed)

$$
w_{1}=1 / \pi_{p s u}
$$

$$
w_{2}=\frac{6 \text { hours (total time of sample interval) }}{\begin{array}{c}
\text { sample duration } \\
\text { (time spent sampling) }
\end{array}}
$$

all trips observed (sampled + only counted)

$$
w_{3}=\frac{\text { (sampled }+ \text { only cou }}{\text { trips sampled }}
$$

$$
w_{F}=w_{1} * w_{2} * w_{3}
$$

## APAIS Weighted Estimation Components

- Catch rate, CPUE, mean catch per trip
- Area fished proportions (Ocean >3mi - EEZ, Ocean <=3mi STS, Inland)
- FES coverage adjustment - instate resident trip proportion
- FHS coverage adjustment - on-frame vessel trip proportion

$$
\hat{\wedge}=\frac{\sum w_{F i} y_{i}}{\sum w_{F i}} \quad \hat{P}_{a}=\frac{\sum w_{F i} I_{a i}}{\sum w_{F i}} \quad \hat{P}_{s}=\frac{\sum w_{F i} I_{s i}}{\sum w_{F i}} \quad \hat{P}_{f}=\frac{\sum w_{F i} I_{f i}}{\sum w_{F i}}
$$

## APAIS Weighted Estimation Components

- Estimation Domains
- Catch Rates by Species and Catch Type (e.g., landed catch, released catch)
- Sub-region, State, Year, 2-month Wave, Fishing Mode, Area Fished
- Area fished proportions
- Sub-region, State, Year, 2-month Wave, Fishing Mode (Private boat, Shore)
- FES coverage adjustment
- Sub-region, State, Year, 2-month Wave, Fishing Mode (Private boat, Shore)
- FHS coverage adjustment - on-frame vessel trip proportion
- Sub-region, State, Year, 2-month Wave, Fishing Mode (Charter boat, Headboat)


## Fishing Effort Survey (FES)

## FES Overview



- Self-administered household mail survey that includes household and individual person-level questions
- Sample frame: a comprehensive directory of residential addresses from the USPS
- Used to estimate in-state private boat and shore mode effort estimates for resident anglers


## FES

- Resource Links
- Design
- Sample Weighting
- Estimation



## FES Resource Links

- Survey Design and Statistical Methods


## FES Section 2

- Annual Reports
- FES At-a-Glance
- Outreach Information


NOAA
FISHERIES

## FES Design

- Probability-Based Design
- Sample Frame
- Stratification \& Sample Selection
- Data Collection


## FES Design: Sample Frame

- United States Postal Service Delivery Sequence File
- $\quad>95 \%$ of Residential Households
- State Saltwater Fishing License Databases
- Primary Stage Unit:
- Residential Household


## FES Design: Stratification and Sample Selection

- Space
- State
- Sub-state regions
(Coastal, Non-Coastal)
- State Saltwater Fishing License

Match Status

- Time

- Year
- 62-Month Waves
- Samples selected using equal selection probabilities within strata


## FES Design: Data Collection

- Generally follows Dillman Approach for Mail Surveys
- Mailings administered near the end of each 2-month wave



## FES Sample Weighting

- Household Sample Base Weight
- Non-response adjustment

$$
w_{B R}=w_{B} / \text { response } \text { rate }_{R}
$$

- Ratio adjustments
- Demographic Control Totals from U.S. Census Bureau
- Raking Ratio, Post-stratification

$$
w_{B}=1 / \pi_{p s u}=N_{h} / n_{h}
$$

$$
w_{B R P}=w_{B R} * \frac{C}{\hat{C}}
$$

## FES Effort Estimation

- Estimate effort as weighted sum of trips reported by sampled households

$$
\hat{T}=\sum w_{B R P} t_{i}
$$

- Estimation Domains
- State
- Year, 2-month Wave (Jan/Feb, Mar/Apr,...)
- Fishing Mode
- Private Boat

■ Shore

- State resident in-state fishing effort


## For-Hire Survey (FHS)

## FHS Overview



- List-frame telephone survey of captains and operators of for-hire vessels
- Vessels selected for weekly reporting of for-hire trips
- Used to estimate charter boat and headboat effort estimates by state, year, 2-month wave, and area fished


## FHS Resource Links

- Survey Design and Statistical Methods


## FHS Section 2

- FHS At-a-Glance
- Outreach Information



## Catch and Effort Estimation

## Catch and Effort Estimation

- Resource Links
- Catch and Effort Estimation Example
- Variance Estimation and Percent Standard Error (PSE)


## Estimation Resource Links

- Survey Design and Statistical Methods

Total Catch and Effort Estimation Section 6

$$
\widehat{Y}_{d}=\widehat{y}_{d} \times \widehat{T}_{T \ldots}
$$

- Estimation Methods Overview

$$
V\left(\widehat{Y}_{d}\right)=\widehat{\hat{y}}_{d}^{2} V\left(\widehat{T}_{T \ldots}\right)+\left(\widehat{T}_{T \ldots}\right)^{2} V\left(\hat{y}_{d}\right)-V\left(\widehat{T}_{T \ldots}\right) V\left(\hat{y}_{d}\right)
$$

- Survey Statistics Overview

$$
\widehat{Y}_{D}=\sum_{w=1} \widehat{Y}_{d l}
$$

- Applied Survey Data Analysis (Textbook)

$$
V\left(\widehat{Y}_{D}\right)=\sum_{w=1} V\left(\widehat{Y}_{d l}\right)
$$

- SAS ${ }^{\circledR}$ PROC Surveymeans


## Catch and Effort Estimation Example

- New Jersey Wave 4 (Jul/Aug) 2019
- Private Boat (PR)
- PR Summer Flounder Landings (No.)
- $\mathrm{n}=857$ APAIS PR Intercepts

| Area Fished | PR Effort (No. Angler Trips) | PR Summer Flounder Landings (No. Fish) |
| :---: | :---: | :---: |
| All | 1,590,161 | 609,019 |
| Ocean - STS | 287,957 | 129,534 |
| Ocean - EEZ | 562,778 | 273,132 |
| Inland | 739,426 | 206,354 |

Primary
Estimates

$$
\begin{aligned}
& \text { Total PR Landings } \\
& \text { by Area }
\end{aligned}
$$

## FES

## Private Boat

On-Frame
Landings
Rates by Area (LPUE)

Effort


Private Boat Total Effort


## APAIS Components

- New Jersey Wave 4 (Jul/Aug), 2019, Private Boat Mode
- $n=857$ APAIS PR Intercepts

| Component | Value | Raw <br> Count | Raw <br> Proportion | Weighted <br> Count | Weighted <br> Proportion |
| :---: | ---: | ---: | ---: | ---: | ---: |
| FES <br> Coverage <br> Adjustment | NJ Resident Angler | 660 | 0.770128 | 54,338 | $\mathbf{0 . 7 0 8 0 6 1}$ |
|  | Out-of-State Angler | 197 | 0.229872 | 22,404 | 0.291939 |
| Area Fished <br> Proportions | Ocean STS (<=3mi) | 189 | 0.220537 | 13,897 | $\mathbf{0 . 1 8 1 0 8 7}$ |
|  | Ocean EEZ (>3mi) | 320 | 0.373396 | 27,160 | $\mathbf{0 . 3 5 3 9 1 2}$ |
|  | Inland | 348 | 0.406068 | 35,685 | $\mathbf{0 . 4 6 5 0 0 1}$ |

## APAIS Components

- New Jersey Wave 4 (Jul/Aug), 2019, Private Boat Mode
- $\mathrm{n}=857$ APAIS PR Intercepts

| Component | Area | Raw Landings Count | Raw <br> Trip <br> Count | Raw <br> LPUE | Weighted Landings Count | Weighted Trip Count | Weighted LPUE |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| LPUE by Area | O-STS (<=3mi) | 86 | 189 | 0.455 | 6251.4242 | 13,897 | 0.44984 |
|  | O-EEZ (>3mi) | 123 | 320 | 0.384 | 13182 | 27,160 | 0.4853 |
|  | Inland | 77 | 348 | 0.221 | 9958.8215 | 35,685 | 0.27907 |

## Variance Estimation

- Sampling error - measure of uncertainty about a point estimate related to variability in the population characteristic being estimated, sample size and other design factors
- Variances for MRIP estimation components estimated using Linearization (Taylor Series approximation) - a standard approach for complex survey designs
- Variances for MRIP catch and effort estimates generally estimated using Goodman's Formula for the Variance of Products


## Percent Standard Error (PSE)

- Coefficient of Variation on the percent scale

> Square Root of Variance (aka Standard Error)

- $\operatorname{PSE}=100^{*}$
Point Estimate
- Relative measure of uncertainty, useful for comparing precision of estimates with very different magnitudes
- 30\%, 50\%
- MRIP Survey and Data Standards (Standard 7)

Office of Science
and Technology
Marine
Recreational Information
Program

## MRIP Data User Seminar: Statistical Methods and Procedures

November 30, 2021 John Foster

