National Marine Fisheries Service's
Marine Recreational Information Program
Survey Design and Statistical Methods for Estimation of Recreational Fisheries Catch and Effort

Prepared by:
Katherine J. Papacostas, ECS Federal, LLC
and
John Foster, NOAA-NMFS

Acknowledgements

The following members of MRIP's Research and Evaluation Team and Survey Operations Team are gratefully thanked for their review of, and contributions to, this report: Rob Andrews, Lauren Dolinger Few, Yong-Woo Lee, Daemian Schrieber, Tom Sminkey, Lucas Johansen, Ryan Kitts-Jensen, and Dave Van Voorhees.

Table of Contents

1. Introduction 1

2. The Access Point Angler Intercept Survey (APAIS): Catch Rates 7
2.1 Sampling Design 7
2.1.1 Stratification Variables 7
2.1.2 Design Stages 9
2.1.3 Sample Frame: Public Access Fishing Site Register 10
2.1.4 Clustering Methods 10
2.1.5 Sample Selection 11
2.1.6 Headboat Sampling Design 13
2.2 Imputation Methods for Missing Values 14
2.3 APAIS Estimation Methods 15
2.3.1 Final Sample Weight 16
2.3.2 Standard Catch Rate Estimation 18
2.3.3 FES Coverage Adjustment for Out-of-Frame Angler-Trips 19
2.3.4 FHS Coverage Adjustment for Out-of-Frame Angler-Trips 20
2.3.5 Legacy CHTS Coverage Adjustment for Out-of-Frame Angler-Trips 20
2.3.6 Area Fished Proportions for Private Boat and Shore Mode Effort 21
2.4 APAIS Certification Status 22

3. The Fishing Effort Survey (FES): Shore and Private Boat Fishing Effort 23
3.1 Sampling Design 23
3.2 Data Collection Design 24
3.3 Data Editing 24
3.4 Estimation Methods 25
3.4.1 FES Weighting 25
3.4.2 Private Boat and Shore Fishing Effort Estimation 27
3.5 FES Certification Status 27

4. Legacy Coastal Household Telephone Survey (CHTS) Design: Shore and Private Boat Fishing Effort 28
4.1 Sampling Design 28
4.2 Data Collection Design 29

4.3 Imputation Methods for Missing Values *30*
4.4 Estimation Methods *30*
4.4.1 Winsorization 31
4.4.2 Private Boat and Shore Fishing Effort Estimation 31
4.5 CHTS Certification Status *31*

5. The For-Hire Survey (FHS): For-Hire Fishing Effort 32
5.1 Sampling Design *32*
5.1.2 Stratification Variables and Sample Frame 32
5.1.3 Sample Selection 33
5.2 Data Collection Design *33*
5.3 Vessel Trip Reports (VTRs) *34*
5.4 Estimation Methods *35*
5.4.1 For-Hire Fishing Effort Estimation 35
5.4.2 Dockside Validation and Adjustment for Reporting Error 36
5.5 FHS Certification Status *41*

6. Total Catch and Effort Estimation: Combining the Estimate Components from the APAIS, FES, CHTS and FHS 42
6.1 Total Effort Estimation *42*
6.1.1 FES Total Effort 42
6.1.2 CHTS Total Effort 43
6.1.3 For-Hire Effort – Combining VTR Data with FHS Data 44
6.2. Total Catch Estimation *47*
6.2.1 Standard Total Catch Estimation 47
6.2.2 Alternative Total Catch Estimation for Small Sample Sizes 48
6.2.3 Total Catch Estimation for Public Use Datasets 49
6.3. Length-Frequencies *51*
6.4 MRIP Estimation Certification Status *52*

7. Large Pelagics Survey (LPS): Large Pelagics and HMS Catch and Fishing Effort .. 53
7.1 Large Pelagics Telephone Survey (LPTS): Large Pelagics and HMS Fishing Effort *53*
7.1.1 LPTS-Private Sampling Design 54
7.1.2 Data Collection Design 54
7.1.3 LPTS Charter Boat Sampling General Protocols (Add-on to the FHS) 54
7.1.4 LPTS Imputation 55
7.1.5 LPTS Estimation Methods 55
7.2 Large Pelagics Intercept Survey (LPIS): Large Pelagics and HMS Catch Rates *56*

7.2.1 LPIS Sampling Design 56
7.2.2 Data Collection Design 58
7.2.3 LPIS Estimation Methods 58
7.3 LPS Total Effort and Catch Estimation *59*
7.3.1 LPS Total Effort 59
7.3.2 LPS Total Catch 60
7.4 Large Pelagics Biological Survey (LPBS) *61*
7.5 LPS Certification Status *61*

8. Time-Series Calibration Methods 62
8.1 CHTS/FES Calibration *62*
8.1.1 Auxiliary Variables 63
8.1.2 Model 65
8.2 MRFSS Intercept/APAIS Calibration *67*
8.2.1 Initial Pseudo-Weighting Approach used from 2004 - wave 1, 2013 68
8.2.2 Calibration Approach for the 2004 - 2013 Data 69
8.2.3 Calibration Approach for the 1981 - 2003 Data 72

9. Historical and Anticipated MRIP Design Changes 74
9.1 APAIS Changes *74*
9.1.1 Historical Changes to the APAIS 74
9.1.2 Future Improvements 75
9.2 FES Changes *76*
9.2.2 Future Improvements 76
9.3 FHS Changes *77*
9.3.1 Historical Changes to the FHS 77
9.3.2 Future Improvements 77
9.4 LPS Changes *78*
9.4.1 Historical Changes to the LPS 78
9.4.2 Future Improvements 78

10. References 79

11. Acronym Glossary 84

1. Introduction

NOAA's National Marine Fisheries Service (NMFS) began collecting data on recreational fisheries in 1981 (Gray *et al.*, 1994), shortly after the Magnuson-Stevens Fishery Conservation and Management Act (P.L. 94-265 as amended by P.L. 109-479) mandated a national program for the management of U.S. fishery resources, both commercial and recreational. This large-scale data collection effort, known as the Marine Recreational Fisheries Statistics Survey (MRFSS), consisted of two primary surveys: a fishing effort telephone survey and an intercept survey where anglers were interviewed in-person at fishing access sites to obtain information about recreational catch. Following a peer-review by the National Research Council (National Research Council 2006), the MRFSS was replaced by the Marine Recreational Information Program (MRIP) in conjunction with the Magnuson-Stevens Fishery Conservation and Management Reauthorization Act of 2006. This new program revised the MRFSS survey methods to address issues identified by the review. Catch and effort estimation methods were also significantly updated to account for the complex survey designs used in data collection, implement more advanced statistical analyses, increase the accuracy of estimates, and reduce potential for bias.

Accurate catch and effort statistics are essential in determining the effects of fishing on fish stocks and for the development of sound, evidence-based management strategies and policies. Continuous monitoring of catch and effort is needed to assess trends, evaluate the impacts of management regulations, and project how different management scenarios will influence a fishery. Recreational fisheries data are therefore a necessity for NMFS, Regional Fishery Management Councils, Interstate Fisheries Commissions, state resource management agencies, recreational fishing industries, and other stakeholder groups involved in marine fisheries.

In transitioning from legacy MRFSS to new and improved MRIP survey methodologies, it is critical to be able to convert the historical time series to new survey 'units' for the purposes of fisheries management and stock assessments. Stock assessments rely on a continuous time series of recreational, and commercial, catch as a component of estimating stock status, which subsequently informs management decisions. This conversion, known as 'calibration,' allows for the construction of a time-series of comparable estimates from the two different surveys. Calibrated estimates therefore approximate what would have been produced had the new survey designs been implemented throughout the time series.

This document describes the MRIP surveys designed and administered by the National Marine Fisheries Service (NMFS) Office of Science and Technology. Its purpose is to provide information to data-users, survey statisticians, fisheries scientists, and other parties interested in the technical details of these surveys, and in understanding how estimates are produced. The document is organized into nine sections, including descriptions of each MRIP survey and associated statistical methods. Legacy survey designs and time series calibration methodologies

are also included. The final section summarizes historical changes and anticipated improvements to be made to each survey. As the program evolves, periodic updates of this report will be completed to reflect further changes in survey designs or estimation methods.

The MRIP currently consists of the following suite of complementary surveys:

- The Access Point Angler Intercept Survey (APAIS)
- The Fishing Effort Survey (FES)
- The For-Hire Survey (FHS)
- The Large Pelagics Survey (LPS) which consists of three component surveys:
    - The Large Pelagics Intercept Survey (LPIS)
    - The Large Pelagics Telephone Survey (LPTS)
    - The Large Pelagics Biological Survey (LPBS)

These four surveys, primarily implemented along the U.S. Atlantic and Gulf of Mexico coasts, collect information on marine recreational fishing catch and effort. Data collection occurs in either one or two month reference periods, or 'waves' for each survey. At the end of each wave, estimates from component surveys are combined to estimate total effort and catch by species and recreational fishing mode (i.e. shore fishing, private boat fishing, and for-hire fishing, Figure 1). The APAIS, FES, and FHS collect complementary data that contribute to wave and annual estimates. The LPS is a separate, specialized survey for large pelagic species and highly migratory species (HMS) and, in general, does not produce estimate components to be combined with the other three surveys (Figure 2).

Section 2 describes the APAIS, which collects information about recreational catch and fishing trip characteristics. In summary, the survey is conducted at public marine fishing access sites (e.g. boat ramps, piers, beaches, marinas, etc.) where anglers are intercepted and interviewed at the completion of their fishing trips. The data collected include individual angler-trip information about catch (i.e. species identification, quantity of each species caught, and fish length and weight measurements) and other trip characteristics including mode of fishing (shore, private boat, etc.), duration, and general area of fishing. APAIS data are used to estimate mean catch per angler-trip, as well as coverage correction factors for the FES and the FHS.

Section 3 describes the FES, which is a mail survey that uses address-based sampling to collect effort data from households in coastal states. Address samples are augmented with fishing license information to allow for targeted sampling of known angler households. The FES is a self-administered mail survey; the survey is mailed out one week prior to the end of each two-month wave, and data are collected for a period of 13 weeks. FES data are used to estimate the numbers of private boat and shore fishing trips taken by anglers residing in the covered coastal states.

Section 4 describes the Coastal Household Telephone Survey (CHTS), a legacy survey which was replaced by the FES beginning in January 2018. The CHTS, which was used to produce effort estimates from 1981-2017, collected information about recreational fishing effort via telephone interviews. The survey coverage was limited to households with landline telephones, located in counties within 25-50 miles of the coastline. Sampling was conducted year-round in six, two-month waves. At the end of each wave, the CHTS data were used to

estimate the number of private boat and shore fishing trips taken by anglers covered by the survey. This transition (which occurred over three years where the FES and the CHTS were conducted side-by-side in order to compare differences in the estimates, and for calibration purposes) was made in response to declining response rates, limited coverage, and reduced efficiency of the CHTS given technological and societal changes that had occurred since the survey was first initiated.

Section 5 provides a description of the FHS, a telephone-based survey for collecting data on for-hire recreational fishing (i.e. where anglers pay boat captains/crews to take them on fishing trips) effort. The FHS samples from comprehensive lists of for-hire vessels operating in each state. Since location information is available for these vessels, a dockside validation study is conducted alongside the telephone-based data collection in order to account for any reporting errors from the survey respondents. The FHS sampling is conducted weekly with captains reporting on vessel trips taken during a one-week reference period. Two estimate components are derived from the survey; an effort estimate from the telephone survey data, and a reporting error correction factor from the validation study. The FHS is further supplemented by the Vessel Trip Reporting (VTR) Program, a census of permitted for-hire vessels in the Northeast Atlantic that is administered by the Greater Atlantic Regional Fisheries Office (GARFO).

Section 6 describes the estimation methods used to produce total recreational catch and effort estimates, which are calculated by combining the various estimate components from the APAIS, FES (the CHTS in previous years), and FHS. Wave estimates are calculated every two months throughout the year in an effort to provide preliminary estimates that can be used during the current year. Final annual estimates are calculated at the end of each year of sampling. Both wave and annual estimates are produced by species, recreational fishing mode, geographic area (e.g. management region and state), and fishing area (e.g. inshore, nearshore, and offshore waters). While the CHTS methods for total estimation are included in this section, it is important to remember that only the FES effort estimates have been used to estimate total catch for shore and private boat recreational fishing since the 2018 transition.

Section 7 details the LPS, which is a separate suite of surveys designed to monitor fishing activity that targets large pelagics and HMS. The LPS was initiated because large pelagic species and HMS are targeted much less frequently than other species, and a specialized survey approach is better suited to producing more reasonably precise estimates for these species. The LPS consists of three complementary surveys: the LPIS, the LPTS, and the LPBS. The LPIS, conducted monthly, is a dockside intercept survey for collecting detailed catch and trip characteristic data from private and charter boat captains who have completed fishing trips for large pelagics and HMS. Two key estimate components are produced from the LPIS: catch rates (individual large pelagic or HMS catch per vessel-trip) and a coverage adjustment for the LPTS. The LPTS, also conducted monthly, is a telephone survey of HMS permit holders that is used to produce estimates of effort for large pelagics and HMS in numbers of vessel-trips. The components derived from the LPIS and LPTS are combined to produce total effort and catch estimates at the end of each one-month wave, and at the end of each year. The LPBS collects measurements and biological samples from large pelagics and HMS that are designated ‘high priority’ by NMFS. In particular, biometric length and weight data, otoliths, muscle tissue, first dorsal spine, and gonads are collected from high priority species (e.g. bluefin tuna) for use in

assessing various life history parameters. While collected as part of the MRIP, the LPBS data and collected specimens do not contribute to the recreational catch estimates; the samples are processed, stored and used instead by the NOAA-Southeast Fisheries Science Center. The LPS is conducted from Maine to Virginia for only part of the year when the majority of the large pelagic and HMS fishing activity occurs (typically June-October).

Sections 2-7 also include a summary of the certification status of each MRIP survey. To achieve nationwide consistency of recreational catch and effort data, MRIP established a rigorous certification process for assuring that survey and estimation methods are scientifically sound. In general, MRIP only supports surveys that apply methods that have been MRIP certified. In select cases, MRIP supports continuing use of legacy survey methods that are not certified, provided:

1) The data produced by such surveys has been utilized in peer-reviewed applications (e.g. federal stock assessments);
2) An MRIP Regional Implementation Plan identifies the need to continue such surveys; and/or
3) A plan to certify those survey methods is in place and is being followed.

Section 8 provides an overview of the methodologies used to calibrate the catch and effort estimates derived from the APAIS and the FES with the MRFSS time series (based on the MRFSS Intercept Survey and the CHTS). The APAIS calibration uses a sample weight adjustment method known as raking, or iterative proportional fitting, to produce revised sample weights for data collected under the old MRFSS Intercept Survey design. The FES effort calibration implements an established small-area estimation method that uses linear mixed modeling to approximate the differences between the CHTS and the FES estimates back through time and produce revised historical estimates. The model essentially includes fixed effects variables that connect direct estimates to auxiliary variables (i.e. variables that aid in making estimates on incomplete data) as well as a random effects variable to capture any remaining variation not explained by the auxiliary variables.

Section 9 summarizes the MRIP design changes that have occurred since the program's initiation in 2008, as well as updates that will be implemented in future years and incorporated into subsequent versions of this report. The largest design changes to-date include redesigning the MRFSS Intercept Survey into the APAIS, and the development of the FES to replace the CHTS in 2018. Recently, calibration methods were developed to support the transition to these two new surveys; FES calibration methods were peer-reviewed in June of 2017 and APAIS calibration methods were peer-reviewed in March of 2018. Both methods were used to produce fully calibrated time series of catch and effort estimates, which were released in July 2018. The LPS is undergoing a redesign that will be completed with a target date of 2022.

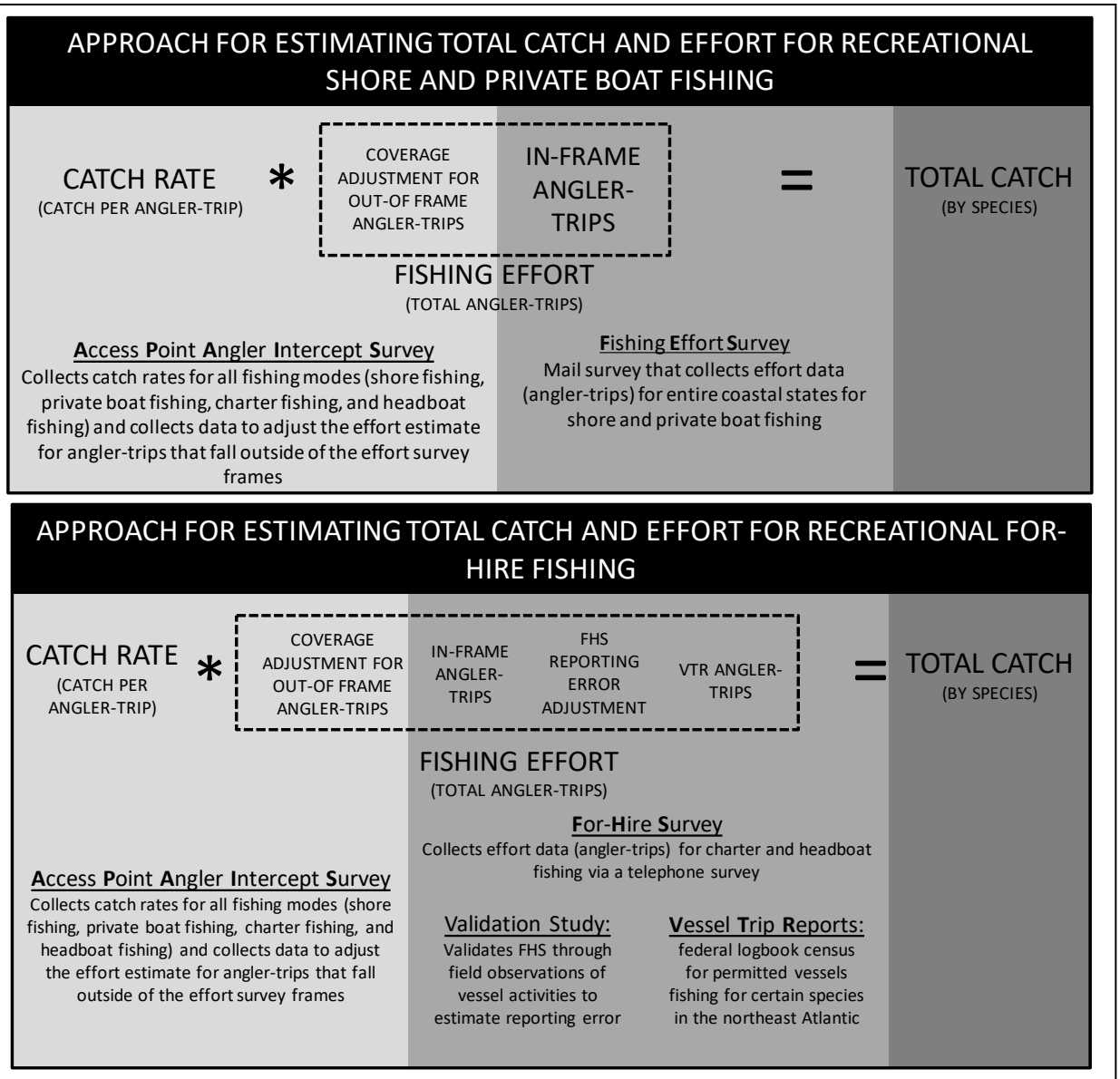

Figure 1: The MRIP approaches for estimating recreational shore, private boat, and for-hire total catch and effort. The Access Point Angler Intercept Survey (APAIS) primarily collects catch data, where anglers are intercepted and interviewed at fishing access sites (docks, beaches, etc.) The Fishing Effort Survey (FES) and the For-Hire Survey (FHS) collect effort information. The FES is a mail survey administered in coastal states, and collects shore and private boat fishing effort data. The FHS collects for-hire fishing effort data, specifically targeting charter and headboat captains; to produce the for-hire effort estimate, the FHS data is further combined with data from the Greater Atlantic Fisheries Office Vessel Trip Reporting (VTR) program, a for-hire census conducted independently from MRIP. Catch rates are combined with FES and FHS effort data, along with coverage adjustments (derived from the APAIS) and in the case of FHS, a reporting error adjustment, to produce total catch by species and by recreational fishing mode.

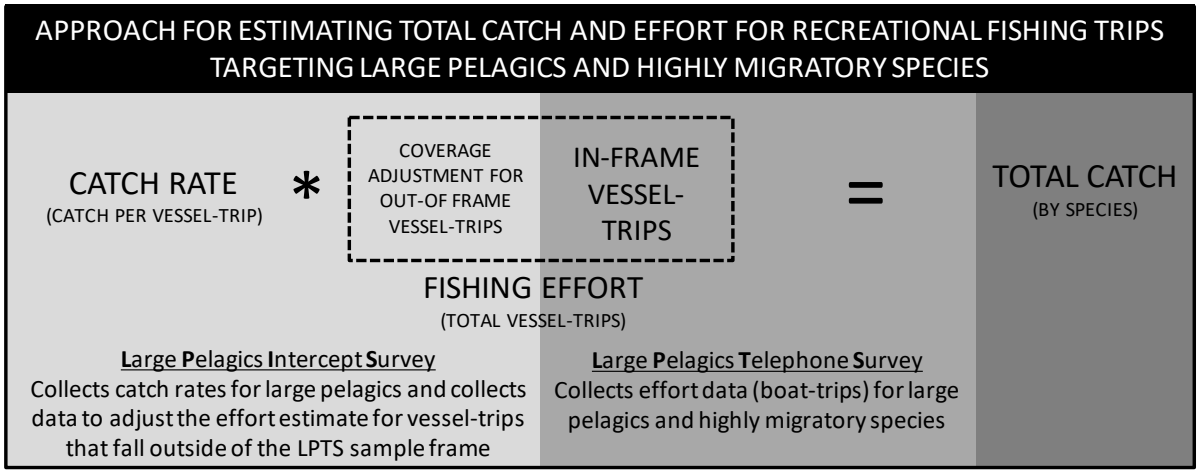

Figure 2: The MRIP approach to estimating total catch and effort for recreational fishing trips targeting large pelagics/highly migratory species. The Large Pelagics Survey (LPS) is a specialized survey to characterize the small portion of recreational angler fishing trips made annually that target large pelagics and highly migratory species (HMS). The estimates for large pelagics and HMS are separate from those of the other surveys. The LPS consists of three complementary surveys, two of which are used to produce estimates: the Large Pelagics Intercept Survey (LPIS) and the Large Pelagics Telephone Survey (LPTS) collect catch and effort information respectively. The third LPS survey, the Large Pelagics Biological Survey (LPBS) collects life history information that is used for other purposes and not taken into account in estimation.

2. The Access Point Angler Intercept Survey (APAIS): Catch Rates

The Access-Point Angler Intercept Survey (APAIS) collects individual catch data from anglers returning to public fishing access sites (e.g. boat ramps, piers, beaches, jetties, bridges or marinas). Trained interviewers administer the survey and collect data on the number and disposition (e.g. harvested or released) of each fish species caught, length and weight measurements of individual fishes, and angler-specific information about the fishing trip. Data are collected monthly and are used to calculate the mean catch per angler-trip every two months as preliminary wave estimates, and then at the end of each year as final annual estimates.

APAIS is conducted along the U.S. Atlantic and Gulf Coasts from Maine to Mississippi. A similar survey occurs in Hawaii, but the MRFSS methods (see Gray *et al.*, 1994) are currently being used there instead of the revised methods employed by MRIP. Louisiana and Texas are also excluded from the APAIS as they utilize separate, state-run recreational fishing data collection programs (e.g. Ogunyinka and Lavergne 2009; Green and Campbell 2010).

2.1 Sampling Design

The APAIS applies a time-space sampling method (i.e. sampling at pre-determined fishing access sites during specific date and time intervals), a broadly used approach in sociological studies to aid in accessing small, hard-to-reach, and/or location-based populations (e.g. Muhib *et al.*, 2001; MacKellar *et al.*, 2007; Parsons *et al.*, 2008). The APAIS further has a stratified, multi-stage cluster design which maximizes sampling efficiency and the spatiotemporal extent of the survey (Czaja, 2005). The survey's target population consists of the set of all angler-trips within a given year, month, state, and fishing mode (i.e. from shore, from a private boat or from a for-hire vessel, explained further in section 2.1.1). The sample frame for this target population consists of a list of fishing access sites, which have been clustered (by both level of fishing pressure and geographic location) and crossed with a date-time calendar so that each primary stage unit (PSU) includes both a time and space component. These PSUs, consisting of specific site-day-time combinations, are formalized through an approximated probability proportional to size without replacement approach (PPSWOR, Vijayan 1968), where site-day-times are selected in proportion to their fishing pressure.

2.1.1 Stratification Variables

APAIS sampling is stratified across time, geographically (by sub-region of the coast, state and sub-state region) and by site groups based on primary fishing mode. Temporally, there are four strata: 1) year, 2) month, 3) kind-of-day, and 4) time interval. Kind-of-day separates weekdays from weekends/holidays. Since August 2013, Friday has been considered a weekend/holiday kind-of-day because MRIP field staff observed that Friday fishing activity is more similar to that of Saturday-Sunday than Monday-Thursday. Time interval strata are used to

divide a full 24 hour sampling day into two main periods (Figure 3): a 12-hour night sampling period (Night interval: 8PM-8AM, consisting of two six-hour sub-sampling periods, interval A: 2AM-8AM and interval D: 8PM-2AM), a 12-hour day sampling period (Day interval: 8AM-8PM, consisting of two six-hour sub-sampling periods, interval B: 8AM-2PM and interval C: 2PM-8PM), and a final six-hour 'peak' period that overlaps intervals B and C (interval P: 11AM-5PM.) This final interval P, added in March 2014, was intended to improve the productivity and efficiency of the survey; interval P allows samplers to intercept anglers returning from trips during peak hours of fishing activity while minimizing disruption to the existing design.

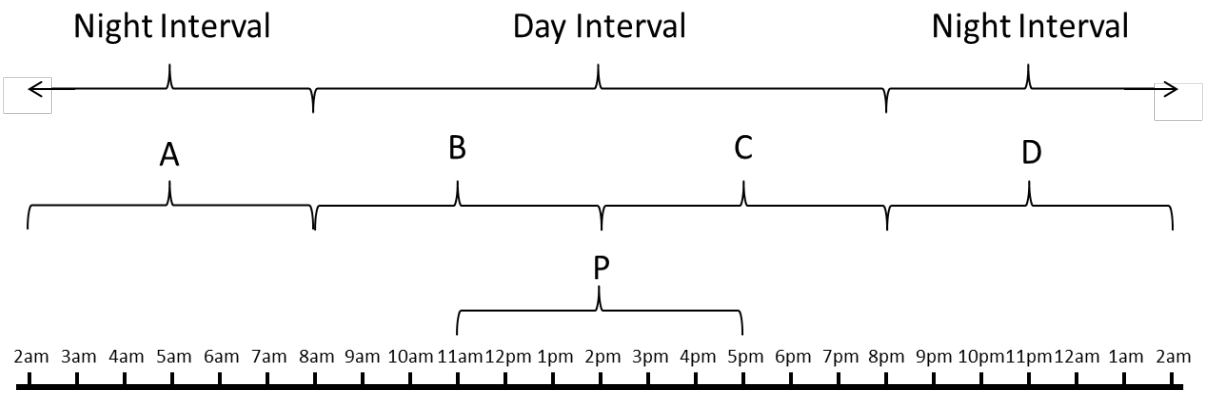

Figure 3: Time interval stratification over a 24-hour day, 2AM-2AM

Spatially, there are three strata: 1) sub-region of the coast, 2) state, and 3) sub-state region. The Atlantic and Gulf coasts are divided into 5 sub-regions: 1) the North Atlantic, which encompasses Connecticut, Rhode Island, Massachusetts, New Hampshire, and Maine, 2) the Mid-Atlantic, which encompasses Virginia, Maryland, Delaware, New Jersey, and New York, 3) the South Atlantic, which encompasses Florida's east coast, Georgia, South Carolina and North Carolina, 4) the U.S. Caribbean, which encompasses Puerto Rico, and 5) the Gulf of Mexico, which encompasses Mississippi, Alabama, and Florida's west coast. Only large states are further divided into sub-state regions (delineated along county lines) for time, cost and sampling efficiency. The number, location, and size of the sub-state regions are decided by each state. However, most of these larger states are sub-divided into two to four sub-regions with the exception of Florida, which is sub-divided into eight sub-state regions.

Site group based on primary fishing mode is the final stratification variable, of which there are four types: 1) shore mode, 2) private boat mode, 3) charter boat mode, and 4) headboat mode. Shore mode includes fishing from any natural shoreline (e.g. beaches and marsh, sand or mudflats) and fishing from artificial structures (e.g. jetties, docks, piers, breakwaters, or causeways). In North Carolina for added sampling efficiency, the shore mode is split into two modes: man-made and beach/bank. Man-made includes any artificial structure and beach/bank includes any natural shoreline. Private boat mode consists of anglers fishing from a boat they own or have rented for the trip. Charter boat mode consists of anglers fishing from a boat that includes both a small number of client anglers as well as a fishing guide or captain who directs the trip (i.e. decides when/where to fish and provides other assistance with fishing to client anglers) for a fee. Headboat mode consists of larger capacity vessels operated by a licensed

captain and crew where individuals and/or larger groups of anglers pay a fee to fish. Headboat mode is sampled separately from the other modes (discussed further in section 2.1.6).

Initially, sampling was done separately for every fishing mode, but methods were altered for shore, private boat and charter boat modes to increase survey productivity beginning in May 2014. In the initial MRIP method, only anglers fishing in the assigned mode were eligible for interview. For example, if shore mode was assigned, private boat and charter boat fishing trips would be excluded from both direct sampling and summary counts of unsampled fishing trips (unsampled fishing trips occur when the volume of anglers at the sampling site is too high to interview them all, or when anglers refuse to be interviewed). However, fishing mode was partially replaced with exclusive site group strata in May 2014 for private and charter boat modes and in early 2016 for the shore mode. These site groups are categorized as a) primarily shore mode, b) primarily charter boat mode and c) primarily private boat mode. With this change in stratification, samplers are allowed to randomly sample any angler-trip, regardless of mode, on the same assignment. Summary counts also include any angler-trip and are not limited to a single mode. This change was implemented because the prior design resulted in fewer interviews per assignment across modes, particularly for the charter mode, which are less frequent than other modes and difficult to intercept under such a design. For instance, in larger coastal states with many access points, guideboats (a type of charter boat) can be unpredictable and do not all consistently leave and return to the same access point. This change therefore maintains stratification related to fishing mode, but increases the overall productivity across all modes, and allows for increased success in intercepting the less predictable charter boats.

2.1.2 Design Stages

APAIS sampling occurs in up to four nested stages: 1) site cluster-day-time interval, 2) sampling duration, 3) angler-trip, and 3) catch. Site cluster-day-time interval, or a cluster of fishing access sites on a specific date during a specific time interval, is the PSU. The PSUs are sampled using a replication based draw procedure that relies on a PPSWOR approach based on expected number of angler-trips, returning during a specific time interval, for a given month and kind-of-day (see sections 2.1.4 and 2.1.5 for more detailed descriptions of the clustering methods and sample selection). While the time interval length of the PSU is six hours, APAIS samplers may spend less than the full six hours actually sampling, so sampling duration is the secondary stage unit (SSU). If samplers are able to spend the full six hours sampling at a single site cluster, then there is no subsampling at this stage. However, two site clusters always involve subsampling, since less than half of the time interval may be spent at each site when taking travel between sites into account. Additionally, on certain occasions, samplers cannot remain at a single site cluster for the full six hour window (e.g. in the case of an extreme weather event or other safety hazard or emergency situation). Angler-trip is the tertiary stage unit (TSU) nested within the SSU. The quaternary stage unit (QSU) is catch, by species, on an individual angler-trip, and is therefore nested within the TSU. Catch is subsampled for individual fish length and weight measurements. The TSUs and QSUs are sampled at random, or with equal probability without replacement. Eligible angler-trips therefore all have equal chances of being sampled from the total number of trips, and fish are sampled across all sizes from the total number of landed fish to avoid biasing the sample.

2.1.3 Sample Frame: Public Access Fishing Site Register

The public access fishing site register is a comprehensive, online database describing all the publicly accessible recreational fishing sites in the states where APAIS is conducted (https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp#) and from which the sample frame is derived. The site register has been continuously updated using historical data as well as inputs from intercept samplers, field supervisors, and state fisheries personnel since the inception of the MRFSS program. More recently, the online database has been made available to the public, giving anyone with internet access the ability to check the accuracy and coverage of the register.

Each site in the register includes a two-digit state code and a three-digit county code (using Federal Information Processing Standard codes), a unique four-digit site code, the estimated fishing pressure by fishing mode and additional information to assist samplers in completing assignments. Fishing pressure is separated into ordinal categories, with each category representing a range of angler-trips expected to return during an assignment in a given stratum (Table 1). For all sites, pressure categories are estimated for each stratum cell (i.e. each unique combination of year, sub-region, state, sub-state region, month, kind-of-day, and time interval) and mode of fishing by APAIS samplers and field supervisors. These fishing pressure categories recorded in the site register are used for clustering and in the sample selection process, described in sections 2.1.4 and 2.1.5. Other information is provided in the site register to assist samplers including contact information for the site's operator or manager, a street address, the nearest town or city, latitude/longitude coordinates, if fishing activity is affected by tide, if there is lighting at night, if the site has been deemed safe for sampling at night, etc.

Sites are never removed from the site register; when a site becomes inactive in all fishing modes, it is coded as 'retired' and is removed from the sample frame. Keeping sites in the register allows for reactivation in the future if fishing activity resumes.

Table 1: Fishing pressure categories as estimated by APAIS samplers and field staff

Pressure Category	Estimated Number of Angler-Trips
0	1 – 4
1	5 – 8
2	9 – 12
3	13 – 19
4	20 – 29
5	30 – 49
6	50 – 79
7	80+
8	0 (Mode not present at the site, or the site is inactive)

2.1.4 Clustering Methods

In the APAIS design, fishing access sites are clustered into PSU's independently within strata, and new clusters are generated monthly prior to selecting samples. Clusters consist of

either one site with high fishing activity, or two sites that are both geographically close and have less fishing activity.

A simulated annealing (e.g. Kirkpatrick, Gelatt, and Vecchi 1983) algorithm developed in a Geographic Information System (GIS) is run to create the clusters based on a set of four main constraints. First, clusters must only include sites within the same county. Second, driving times between sites within a cluster are configurable (currently set to a maximum of 60 minutes between sites for all states); driving times are imported into the GIS using the Google Distance Matrix Application Programming Interface, which provides a recommended route between a start and end point based on calculations of travel duration (Google, 2016). Third, low activity sites can be clustered with one additional site, and fourth, high activity sites are not clustered with other sites; the threshold for a single site cluster versus a two-site cluster is decided based on the distribution of fishing pressures across all sites within each given state and month. The annealing process is random, so it is possible for clusters to change between draws even if other characteristics remain the same. In addition, cluster combinations may change depending on sampling month and fishing mode since fishing pressure is not static across these strata. For single-site clusters, samplers remain at the site for the entire six-hour sampling assignment. For two-site clusters, samplers stay at the first site for three hours and then move to the second site for the remainder of the six-hour sampling assignment. The order in which samplers visit sites in two-site clusters is pre-determined as part of the sample selection. The travel time between sites is excluded from the recorded sample time.

2.1.5 Sample Selection

In the initial MRIP-APAIS pilot conducted in North Carolina in 2010 (Breidt *et al.*, 2012), the sample was selected entirely at random using a PPSWOR approach. For each draw (i.e. set of sample units), the initial inclusion probability (π_{hi}) of the i-th assignment (i.e. the PSU, or site cluster-day-time interval) out of the N_h total assignments in stratum h of the sample frame was calculated as

$$\pi_{hi} = \frac{z(a_i)}{\sum_{i=1}^{N_h} z(a_i)} n_h$$

where $z(a_i)$ is the size measure of the i-th assignment, and n_h is the number of assignments selected from stratum h. Size measures (weights) were derived from the same pressure categories used in the clustering methods, that is, the lower limit of the range of estimated angler-trips associated with the pressure category (Table 2). The only deviation from this weighting method was for the two lowest pressure categories (0 and 1) to reduce the probability of selecting very low activity PSUs; these weights were reduced to one-half of the lower limit of anglers expected to visit the site (0.5 for a pressure category of 0 and 2.5 for a pressure category of 1). While arguably ideal for generating purely random samples, this uncontrolled selection proved difficult to implement in the field because selections were sometimes made that exceeded logistical constraints (e.g. sampler availability).

A replication-based sample selection procedure was therefore introduced to control the PSU selection in June of 2013; in this procedure, standard PPSWOR (using the weights listed in

Table 2) is used to generate a large set of initial replicate draws (50,000 or greater). These replicates are then screened against the state's, or sub-state region's, constraints to create a survivor subset of replicate draws containing a minimum of 1000 replicates. From the survivors, simple random sampling is used to select one of the draws as the final selection for that state or sub-state region. This new process closely approximates standard PPSWOR methods, but allows for logistical considerations necessary for effective implementation by field staff.

Table 2: Size measures used for the probability proportional to size method used for interviewing assignments

Pressure Category: Expected Number of Angler-Trips	Size Measure (Weight)
0: 1-4 Angler-trips	0.5
1: 5-8	2.5
2: 9-12	9
3: 13-19	13
4: 20-29	20
5: 30-49	30
6: 50-79	50
7: 80+	80
8: Mode not present at site or site is inactive	0

After the draw is selected through this procedure, a new inclusion probability needs to be calculated, since the final probability is no longer the same as the initial inclusion probability. This new inclusion probability is calculated as

$$S_c : \pi_{hi} = P(hi \in A_{hi} | S_c) = \sum_{a \in A_{hi} | S_c} p(a)$$

where the inclusion probability (π_{hi}) of the i-th assignment in stratum h is the fraction of surviving draws (A_{hi}) that contain assignment i out of the survivor subset (S_c) of replicate draws. This equation is just a modification of the standard calculation for inclusion probability (π_i), where π_i is the sum of the sample probabilities $p(a)$ for the set of samples (A_i) that contain element i,

$$\pi_i = P(i \in A) = \sum_{a \in A_i} p(a)$$ (Fuller 2009).

Additionally, the night strata (8PM-8AM, consisting of the two subsampling time intervals D: 8PM-2AM and A: 2AM-8AM) and the day strata (8AM-8PM, consisting of the three subsampling time intervals B: 8AM-2PM and C: 2PM-8PM, and interval P: 11AM-5PM), are selected independently of one another. Further, interval P: 11AM-5PM, which is the more recently added sampling interval that corresponds with peak fishing activity, is selected independently from day intervals B and C with its own sample allocation. The overlapping daytime intervals B, C and P allow for the possibility of drawing the same site, 3-hour time block, and date in two intervals (B and P, 11AM-2PM or C and P, 2PM-5PM.) The inclusion probability of the sample collected during the overlapping intervals B (π_B), C (π_C) and P (π_P), require an adjustment calculated as the probability of two independent events. For the overlapping 11AM-2PM time block, the inclusion probability (π_{BP}) is

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

and, for the overlapping 2PM-5PM time block, the inclusion probability (π_{CP}) is

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

These overlapping intervals also result in a single daytime stratum for the purposes of variance estimation.

All sample selections are of fixed importance, and therefore have to be either completed or cancelled if the sampler cannot complete the assignment. For two-site clusters, the order in which the samplers visit sites is randomized prior to the assignment.

2.1.6 Headboat Sampling Design

Since 2004, a headboat at-sea sampling program has been implemented, where angler intercepts are conducted onboard headboats during fishing trips. Headboat sampling is only conducted along the Atlantic coast from Maine through Virginia. The remaining Atlantic and Gulf Coasts are covered by the separate Southeast Region Headboat Survey (Brennan, 2010) administered from the Beaufort North Carolina NOAA Laboratory. Prior to 2004, headboats and charter boats were combined into a charter/headboat fishing mode and sampled using the methods similar to those described in 2.1.5. However, analyses indicated that headboats were overrepresented compared to charter boats, and the two fishing modes were separated.

A for-hire vessel directory, called the FHS directory, is used to create the headboat sample frame. This directory is continuously updated throughout the survey year, listing unique vessel identifiers (e.g. vessel name, registration, etc.) as well as contact information for vessel representatives. The headboat sample frame consists of all possible active vessel-day combinations within each month and coastal state. A PPSWOR procedure based on fishing pressures, similar to the process detailed in section 2.1.5, is used for the headboat sample selection. The only distinction is that the headboat fishing pressures are not the categories in Tables 1 and 2, but the expected average number of vessel-trips per sub-region, state, sub-state region, month and kind-of-day based on historical data.

A few headboat sampling protocols differ from the rest of the APAIS. Firstly, sampling onboard the boats allows discarded fish to be observed, identified and measured by APAIS samplers in addition to the harvested fish, whereas with site assignments only harvested fish can be sampled directly. Weights of discards are not collected, however, due to the high potential for inaccurate weight data while on a moving vessel at sea; weights are only collected from harvested species at the dock after the trip. In addition, unlike site assignments, headboat assignments can be rescheduled if a boat cannot be sampled on the assigned date. Headboats can unpredictably cancel scheduled trips due to poor weather conditions or if there are too few customers to justify a trip, hence the need for some flexibility in assignments. In these scenarios, an alternate sampling date is scheduled within the same month and kind-of-day (weekday or weekend). Further, headboats take either full day or half-day (morning or afternoon) trips. If a

headboat schedules more than one half-day trip on an assigned vessel-day, only one of those trips is sampled at random.

2.2 Imputation Methods for Missing Values

Several methods are used to address two missing data situations associated with individual fish lengths and weights collected during APAIS angler-trip intercepts. In the first situation, either length or weight is recorded but not both. For these records, imputation is handled using a standard model for the length-weight relationship in fishes (Brodziak *et al.*, 2012). In many cases, however, both length and weight are missing. In such scenarios, a mixture of hot deck (using data from the current survey year) and cold deck (using data from prior years) imputation is used to fill in missing values. Both of these techniques are well described in the statistical literature (e.g. Lessler and Kalsbeek 1992; Andridge and Little 2017; Wang 2003).

When either length or weight data are missing, length-weight relationship models are fit to impute missing lengths (L) or weights (W), using

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

when length is unknown, and

$$W = aL^b$$

when weight is unknown. In both length and weight equations, *a* and *b* are parameters to be estimated with available length and weight data for the fish species of interest. Parameter *a* is a scaling coefficient for the weight at length of a particular fish species, while *b* is a shape parameter for the body type of a fish species. Generally, *a* is a small number between zero and one and *b* has a value around three since, in theory, the volume of a regularly shaped 3-D object is proportional to the object's length cubed. In fish that have narrow, elongated bodies, *b* is usually less than three. In wider-bodied fish, *b* is usually greater than three. Maximum likelihood estimates of *a* and *b* are calculated using a linear regression of log-transformed known length and weight data for each species of interest (see Brodziak *et al.*, 2012 for a more detailed breakdown of this calculation).

In most cases, models are fit to current year data by two-month wave and species, with data pooled across states, modes, and areas fished (inshore waters, nearshore waters, or offshore waters). If current year models fail to converge or if there are fewer than 15 complete length-weight observations for the given wave and species, then separate models are fit by species using all complete length-weight observations from the most recent 10 years. All models are fit with the SAS NLIN procedure, which fits nonlinear regression models using the least squares method (SAS Institute Inc., 2016). Very rarely, both sets of models fail to converge or produce adequate regression predicted values. These cases are passed on to imputation, described below, with both length and weight values set to missing prior to imputation.

For intercepted angler-trips with catch data but no corresponding length-weight measurements, paired length-weight observations are imputed from complete cases using hot and

cold deck imputation. Up to five imputed length-weight records are created for each species landed on a given trip. If the number of landed fish (by species) with missing data is less than five, then imputed records are only created for the number of landed fish with missing data.

Imputation proceeds from hot to cold deck through five rounds. The rounds begin with imputation cells that correspond to the most detailed APAIS estimation cells, but are aggregated to higher levels in subsequent rounds to bring in more length-weight data. Each round has an associated minimum number of complete cases required to perform imputations (Table 3).

Table 3: Imputation Rounds

Round	Imputation Cells	Minimum number of complete cases required
1	current year, wave, sub-region, state, mode, area fished, species	10
2	current year, half-year (waves 1-3, 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

The majority of imputations are made within the first three rounds. Imputations are performed by an estimation program using the SAS SURVEYIMPUTE procedure (SAS Institute Inc., 2016) in the SAS-callable version of SUDAAN 11 (RTI International, 2017). To prevent the introduction of unknown biases to the imputations, sample weights are not used.

A small number of missing data cases usually remains after length-weight modeling and all rounds of imputation. These cases are generally limited to records with species group codes where it was not possible to identify catch down to the species level (e.g. ‘unknown shark’, or ‘left-eye flounder’); species with almost entirely angler-reported landings (e.g. baitfish species that were unavailable for samplers to inspect, or species that were caught and released without recording measurements); and very large species (e.g. certain highly migratory species). No additional attempts are made to substitute values for these records.

2.3 APAIS Estimation Methods

In the APAIS survey, there are three categories of catch: Type A, Type B1, and Type B2. Type A catch includes landed or harvested fish that are available to be inspected by the APAIS sampler. Type B1 catch includes fish that were caught and harvested but were not available to an MRIP sampler. Type B2 catch includes fish that were caught and released alive at sea. The numbers of B1 and B2 catch in all fishing modes are reported by intercepted individual anglers. In shore mode, Type A catch is reported by individual anglers as well, but in the boat modes, Type A catch may be reported as individual angler catch or grouped catch. Grouped catch is where more than one angler contributes to the catch inspected by the APAIS sampler, but that catch cannot be separated by individual angler.

APAIS data are used to estimate catch rates, and coverage adjustments for the FES and the FHS total effort estimates described in section 6.1; previously APAIS provided coverage adjustments for the CHTS total effort, and these methods are included in section 6.1 for informational purposes only. Catch rates are estimated for individual species separately by catch category (A, B1 and B2) as catch per unit effort, where effort is expressed as angler-trips. Catch rates are estimated within domains defined by year, two-month wave, sub-region, state, mode of fishing, and area fished. As previously described, fishing modes include shore, private boat, charter boat, and headboat. Areas fished include three general saltwater fishing areas, meant to inform fishery managers at state, regional, and federal levels: 1) inland waters, 2) nearshore waters (State Territorial Seas) and 3) offshore waters (federal Exclusive Economic Zone [EEZ]). Inland waters include marine or brackish interior portions of bays, estuaries, sounds or coastal rivers. The dividing line between State Territorial Seas and the EEZ is three nautical miles in most states but 10 nautical miles off the west coast of Florida. The EEZ extends from the State Territorial Seas to 200 nautical miles from the coastline (National Marine Fisheries Service, 2016). The coverage adjustments for the FES (CHTS in prior years), and FHS are estimated using information collected by APAIS to account for out-of-frame anglers in effort estimates (i.e. anglers that are intercepted by APAIS but missed by the effort surveys).

All APAIS computations are done using the SAS SURVEYMEANS procedures (SAS Institute Inc., 2016) or the R survey package (Lumley, 2004, 2010).

2.3.1 Final Sample Weight

The final weight of each APAIS sample, needed for estimating catch rates from the collected data, is calculated as the product of three stage weights: 1) the stage I weight (or weight of the site cluster-day PSU), 2) the stage II weight (weight of the time spent sampling each site within a site cluster) and 3) the stage III weight (weight of angler-trips sampled at each individual site). Stage IV (catch) does not require weighting for the purposes of catch rate estimation, but is used in calculating length-frequencies (discussed further in section 6.3). The stage I weight (w_{hi}) is simply the inverse of the inclusion probability (π_{hi}) from the sample selection described in section 2.1.5.

$$w_{hi} = \frac{1}{\pi_{hi}}$$

The stage II weight is needed to adjust for the actual time spent sampling each site within a site cluster, relative to the full six-hour time interval assigned. This weight is primarily relevant for those sites that were sampled in two-site clusters instead of one-site clusters so that they are not underrepresented in the catch rate estimations. However, the stage II weight is also applicable to any one-site cluster where the samplers could not remain at the site for the full six-hour time interval. The weight (w_{aki}) of site a within time window k (which, for a two-site cluster, is roughly two three-hour intervals minus the travel time between sites) and site cluster-day i is calculated as

$$w_{aki} = \frac{T_i}{t_{aki}}$$

where T_i = 6, corresponding to the 6-hour time interval assigned to every site cluster-day i; and t_{aki} corresponds to the time spent sampling at site a during the ki sampling assignment.

The stage III, or angler-trip, weight is an adjustment to account for the total number of angler-trips observed at a site compared to the number of angler-trips that are intercepted and sampled (i.e. the sampling fraction). This weighting stage differs between shore and boat fishing modes, as well as among Type A and Type B1 and B2 catch in boat fishing modes.

For shore fishing, Type A, B1 and B2 catch are always reported by individual anglers, so the Stage III weight is simply the inverse of the sampling fraction of the total angler-trips at the site. Therefore, regardless of catch type, the weight (w_{akij}) of shore mode angler-trip j within site a, time window k, and site cluster i is calculated as

$$w_{akij} = \frac{M_{aki}}{m_{aki}}$$

where M_{aki} is the total number of angler-trips observed at site a during time window k; and
m_{aki} is the total number of shore mode angler-trips intercepted by the samplers at site a during time window k.

For boat fishing, the Stage III weight of an angler-trip must take into account the fraction of anglers that were intercepted compared to the total number of anglers in each observed boat fishing party. Like with shore fishing, Type B1 and B2 catch are reported by individual anglers in boat fishing; however, Type A catch on boats is often reported as grouped catch to which some, but not necessarily all, anglers in the fishing party have contributed. In these cases, expanding that sample to the total anglers at the site would lead to overestimation of catch rates. Therefore, the Stage III weight for Type A grouped catch from private boats, charter boats and headboats, is only expanded for the number of anglers contributing to grouped catch at each site. For private, charter or headboat angler-trips reporting Type B1 and B2 catch, the weight (w_{bakij}) of angler-trip j within site a, time window k, boat fishing party b and site cluster i is calculated as

$$w_{bakij} = \left(\frac{M_{aki}}{\sum_{b=1}^{B} p_{baki}}\right)\frac{p_{baki}}{m_{baki}}$$

where M_{aki} is the total number of angler-trips observed at site a during time window k;
p_{baki} is the total number of anglers in boat fishing party b (out of boat fishing parties 1 ... B) at site a during time window k; and
m_{baki} is the number of anglers intercepted by the samplers in boat fishing party b at site a during time window k.

For private boat, charter boat and headboat angler-trips reporting Type A grouped catch, the stage III weight is

$$w_{bakij}{}^{*} = \left(\frac{M_{aki}}{\sum_{b=1}^{B} p_{baki}}\right)\frac{p_{baki}}{m_{baki*}}$$

where $m_{baki}{}^{*}$ is the total number of anglers that contributed to grouped catch in boat fishing party b at site a during time window k.

The final weight is then simply the product of the stage I, stage II, and stage III weights as

$w_{hij} = w_{hi} \times w_{aki} \times w_{akij}$ (for all shore mode angler-trips),
$w_{hij} = w_{hi} \times w_{aki} \times w_{bakij}$ (for Type B1 and B2 catch of boat mode angler-trips), or
$w_{hij} = w_{hi} \times w_{aki} \times w_{bakij}{}^{*}$ (for Type A grouped catch of boat mode angler-trips).

In certain scenarios, the Stage II weight (w_{aki}) and/or the Stage III weight (w_{akij}, w_{bakij} or $w_{bakij}{}^{*}$) will equal one. For example, if a sampler spends the full six-hour time-interval at a one-site cluster, $w_{aki} = 1$. Similarly, if samplers can intercept all observed angler-trips at a site, if all anglers were intercepted *and* contributed to grouped catch, or if all Type A catch could be separated to individual intercepted anglers, $w_{akij}{}^{*} = 1$. These scenarios would more likely occur at a site with low fishing pressure and therefore fewer angler-trips for samplers to intercept. Where these conditions are met, the final sample weight is simply equal to the Stage I weight ($w_{hij} = w_{hi}$).

2.3.2 Standard Catch Rate Estimation

Mean catch per angler-trip ($\hat{\bar{y}}_d$) is calculated as a domain (d) estimate, defined by year, sub-region, state, sub-state region (county), fishing mode, wave, primary area of fishing (inland waters, nearshore waters, or offshore waters), species of fish, and catch type (Type A, B1, and B2) using the following standard weighted mean estimator (SAS Institute Inc., 2016) for a multistage stratified design:

$$\hat{\bar{y}}_d = \frac{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij} I_{d(h,i,j)} y_{hij}}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij} I_{d(h,i,j)}}$$

where $\hat{\bar{y}}_d$ is the estimated mean catch per angler-trip in domain d;
$h = 1, \ldots H$ represents the strata, each of which is defined by year, wave, month, kind-of-day, time interval (i.e. all temporal stratification variables), sub-region, state, sub-state region (i.e. all geographic stratification variables) and site-group (i.e. all fishing mode stratification variables);
$i = 1, \ldots n_h$ represents the PSUs (site-cluster-day-time intervals), sampled within stratum h;
$j = 1, \ldots m_{hi}$ represents the angler-trips sampled in PSU i;
w_{hij} is the final APAIS sample weight described in section 2.3.1;
$I_{d(h,i,j)}$ is an indicator variable, which equals 1 if (h, i, j) is in domain d or 0 if otherwise;

and

y_{hij} is the number of fish caught on angler-trip j.

The variance is estimated using a Taylor series linearization method (Dienes, 1957; SAS Institute Inc., 2016). The Taylor series obtains a linear approximation of a non-linear function, and then the variance estimate of the non-linear function is estimated by the variance of the Taylor series approximation of that function. This method of approximation depends only on the primary stage of the sampling design, so only the PSU totals within each stratum are needed (Fuller 1975; Woodruff 1971). The method estimates the variance as

$$V_h(\hat{\bar{y}}_d) = \sum_{h=1}^{H} \left(\frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\frac{\left(\sum_{j=1}^{m_{hi}} w_{hij} I_{d(h,i,j)} (y_{hij} - \hat{\bar{y}}_d)\right)}{w_{hij} I_{d(h,i,j)}} - \frac{\left(\sum_{i=1}^{n_h} \frac{\left(\sum_{j=1}^{m_{hi}} w_{hij} I_{d(h,i,j)} (y_{hij} - \hat{\bar{y}}_d)\right)}{w_{hij} I_{d(h,i,j)}} \right)}{n_h} \right)^2 \right).$$

As mentioned above, mean catch rates and associated variances are estimated using the SAS SURVEYMEANS procedures (SAS Institute Inc., 2016).

2.3.3 FES Coverage Adjustment for Out-of-Frame Angler-Trips

In addition to catch rates, APAIS data are used to calculate a coverage adjustment for the FES (detailed in section 3) that accounts for out-of-frame angler-trips. The FES samples entire coastal states for private and shore mode in-state effort, but out-of-state effort (trips made by non-residents) are out-of-frame. APAIS, however, covers these out-of-state angler-trips returning to sites on the APAIS frame, and this information is used to adjust the FES effort estimates accordingly. This coverage adjustment is made within domains defined by the sub-region, state, mode of fishing, year and wave. These values are generated using standard methods for weighted proportions (SAS Institute Inc., 2016):

$$\hat{p}_q = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} I_{hijq}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}}$$

where $\hat{p}_q$ is the estimated proportion of the in-frame angler-trips in domain *q*;

$h = 1, \ldots H$ represents the strata, each of which is defined by year, wave, month, kind-of-day, time interval, sub-region, state, sub-state region and site-group;

$i = 1, \ldots n_h$ represents the PSUs (site cluster-day-time intervals), sampled within stratum h;

$j = 1, \ldots m_{hi}$ represents the angler-trips sampled in PSU *i*;

w_{hij} is the final APAIS sample weight described in section 2.3.1; and

I_{hijq} is an indicator variable that equals 1 if angler-trip j was completed by an in-state resident and 0 if angler-trip j was completed by an out-of-state resident.

The variance is then estimated using the following Taylor series linearization (SAS Institute Inc., 2016)

$$V(\hat{p}_q) = \sum_{h=1}^{H} \left(\frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\frac{\left(\sum_{j=1}^{m_{hi}} w_{hij} (I_{hijq} - \hat{p}_q) \right)}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}} - \sum_{i=1}^{n_h} \frac{\frac{\left(\sum_{j=1}^{m_{hi}} w_{hij} (I_{hijq} - \hat{p}_q) \right)}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}}}{n_h} \right)^2 \right).$$

This adjustment is applied to the FES base effort estimate to calculate total shore and private boat effort (further details about FES base and total effort are in sections 3.4 and 6.1.1 respectively).

2.3.4 FHS Coverage Adjustment for Out-of-Frame Angler-Trips

APAIS data are also used to reduce coverage error in the effort estimates from the FHS (section 5). The FHS, which is used to estimate effort for charter and headboat modes, misses any unregistered vessels that are not listed in the for-hire vessel directory that is used to create the FHS sample frame (described in section 5.1.2). However, the APAIS intercepts many, if not all, of these trips. To determine in-frame vs out-of-frame trips, all of the for-hire angler-trips intercepted by APAIS are cross-checked with vessels in the FHS vessel directory.

The FHS coverage adjustment, which is the ratio of the total intercepted angler-trips to in-frame angler-trips ($\hat{R}_{mi}$), is estimated using the following equation:

$$\hat{R}_{mi} = \sum_{i=1}^{I} \frac{X_{mi}}{I_{mi} X_{mi}},$$

where X_{mi} is the total number of angler-trips intercepted on the i-th sampled vessel, and
I_{mi} is an indicator variable that equals 1 if the angler-trip m on vessel i is in-frame or 0 if otherwise.

The variance of $\hat{R}_{mi}$ is approximated as

$$V(\hat{R}_{mi}) = \frac{\hat{R}_{mi}(1 - \hat{R}_{mi})}{\sum_{i=1}^{I} X_{mi}}.$$

2.3.5 Legacy CHTS Coverage Adjustment for Out-of-Frame Angler-Trips

Prior to 2018, the APAIS data was also used for a CHTS coverage adjustment. The CHTS (used to estimate effort for private boat and shore fishing modes like the FES) sampled coastal counties within coastal states, and therefore, by design, missed anglers who resided in non-coastal states, or in non-coastal counties of coastal states. The CHTS coverage adjustment

for private boat mode was estimated in the exact same manner as the FHS coverage adjustment in section 2.3.4. For shore mode, however, the coverage adjustment ($\hat{R}_m$) was calculated as

$$\hat{R}_m = \sum_{i=1}^{I} \frac{X_m}{I_m X_m},$$

where X_m=1 for each shore mode angler-trip, and
I_m was an indicator variable that equaled 1 if angler-trip m was in-frame (i.e. if the angler trip was from a household in a coastal county) or 0 if otherwise.

The variance was estimated in the same manner as $V(\hat{R}_{mi})$ in section 2.3.4.

2.3.6 Area Fished Proportions for Private Boat and Shore Mode Effort

Primary area fished proportions are calculated from APAIS data and used to partition total private boat and shore mode effort estimates from the FES, and the CHTS in previous years, into three separate area-specific estimates (inland waters, State Territorial Seas, and EEZ). These proportions are estimated within domains, defined by the sub-region, state, fishing mode, year and wave. Like the FES coverage adjustment described in 2.3.3, area fished proportions are generated using standard methods for weighted proportions (SAS Institute Inc., 2016):

$$\hat{p}_a = \frac{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij} I_{hija}}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}}$$

where $\hat{p}_a$ is the estimated proportion of the total effort in fishing area *a*;
$h = 1, \ldots H$ represents the strata, each of which is defined by year, wave, month, kind-of-day, time interval, sub-region, state, sub-state region, and site-group;
$i = 1, \ldots n_h$ represents the PSUs (site cluster-day-time intervals), sampled within stratum h;
$j = 1, \ldots m_{hi}$ represents the angler-trips sampled in PSU *i*;
w_{hij} is the final APAIS sample weight described in section 2.3.1; and
I_{hija} is an indicator variable that equals 1 if angler-trip j was completed in fishing area *a* or 0 if otherwise.

The variance is then estimated using the following Taylor series linearization (SAS Institute Inc., 2016)

$$V(\hat{p}_a) = \sum_{h=1}^{H}\left(\frac{n_h}{n_h - 1}\sum_{i=1}^{n_h}\left(\frac{\left(\sum_{j=1}^{m_{hi}} w_{hij}(I_{hija} - \hat{p}_a)\right)}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}} - \sum_{i=1}^{n_h}\frac{\frac{\left(\sum_{j=1}^{m_{hi}} w_{hij}(I_{hija} - \hat{p}_q)\right)}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}}}{n_h}\right)^2\right).$$

The proportions for each area of fishing are multiplied by the FES total effort estimates in order to produce inland, State Territorial Seas and EEZ effort estimates. Prior to 2018, the proportions for each area were multiplied by the CHTS total effort estimates; see sections 6.1.1 and 6.1.2 for FES and CHTS total effort estimation.

2.4 APAIS Certification Status

The APAIS survey design and estimation methods were certified in December of 2012. The survey was designed between 2008 and 2010. In 2010, a year-long pilot study in North Carolina was conducted to test the feasibility of the new approach (Breidt *et al.*, 2012). The study specifically assessed the effects on various measures of survey performance and select estimates through side-by-side comparisons with the ongoing MRFSS Intercept Survey.

The pilot study was evaluated by independent peer-reviewers and the MRIP Operations Team; the Operations Team evaluated both the peer-reviews and the report, and recommended the methods be certified following a point-by-point response to the reviews. Both the MRIP Executive Steering Committee and NMFS Leadership supported the Operations Team recommendation, and the methods were officially recommended for large-scale implementation beginning in wave 2 of 2013.

3. The Fishing Effort Survey (FES): Shore and Private Boat Fishing Effort

The Fishing Effort Survey (FES) is used to estimate shore and private boat recreational fishing effort (in angler-trips) in coastal states along the Atlantic Coast and the Gulf of Mexico, as well as in Hawaii. The survey was designed as a replacement for the CHTS, to yield higher quality data by providing nearly complete coverage of coastal states, more efficient sampling, and much higher response rates than the CHTS. The FES was developed through a rigorous testing and evaluation process (Brick *et al.*, 2012, 2016; Andrews *et al.*, 2013).

3.1 Sampling Design

The FES utilizes address-based sampling (ABS). Sample are selected from the United States Postal Service Computerized Delivery Sequence File (CDS), which covers all residential addresses within the study states (https://postalpro.usps.com/address-quality/cds). The sample is stratified both geographically and by angler license status (Figure 4). Within each coastal state, sampling is stratified into coastal and non-coastal sub-state regions defined by geographic proximity to the coast. Generally, the coastal stratum consists of counties with borders within 25 miles from the shoreline, and the non-coastal stratum consists of counties beyond 25 miles from the shoreline. However, the designation of coastal counties in most states varies throughout the year based on historical fishing activity. Due to their small geographic areas, all counties in Rhode Island, Connecticut and Delaware are considered coastal. All counties in Florida are also classified as coastal due to the relatively high rate of fishing throughout the state.

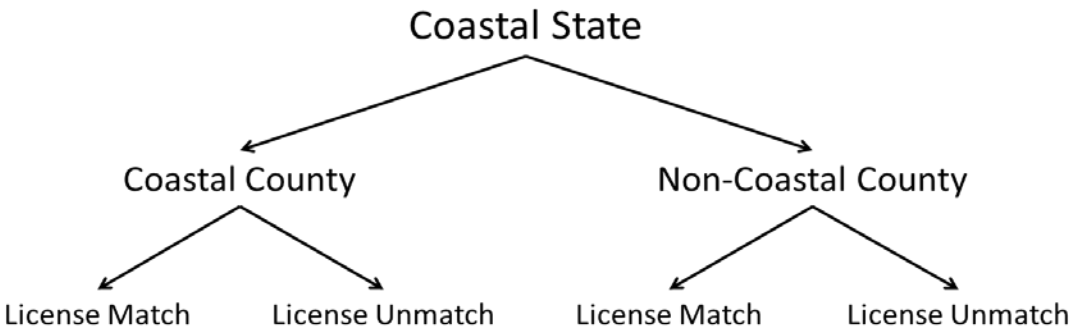

Figure 4: FES Stratification

Within the geographic strata, addresses are matched to the National Saltwater Angler Registry (NSAR), which consists of a list of licensed saltwater anglers. This creates two additional strata: license matched (households with one or more licensed anglers) and license unmatched (households that cannot be matched to NSAR). This stratification provides

additional information to optimize sampling; previous studies (e.g. Andrews *et al.*, 2010; Brick, *et al.*, 2012; Andrews *et al.*, 2013) have demonstrated that residents of households that match to license databases respond to fishing surveys at a higher rate and are more likely to have fished during the reference wave than residents of unmatched households. Addresses are selected from within each stratum using simple random sampling.

Within each state and wave, sample sizes are determined for each stratum using Neyman allocation (e.g. Wright 2014), where the sample is distributed among strata in proportion to the product of the population size and the standard deviation. The goal of Neyman allocation is to maximize precision of estimates for a fixed sample size. Standard deviations are based upon historical FES data and estimates. The sample size for each state and wave is currently targeted to produce estimates with coefficients of variation of 0.20 for each state and wave.

3.2 Data Collection Design

The FES is a self-administered mail survey, administered for six, two-month reference waves annually. Data collection starts with an initial mailing one week prior to the end of the reference wave so that materials are received right at the end of that wave. This initial mailing is delivered by regular, first class mail and includes a cover letter stating the purpose of the survey, a survey questionnaire, a post-paid return envelope, and a $2 cash incentive.

One week after the initial mailing, a follow-up, thank you and reminder postcard is mailed out via regular first class mail to all sampled addresses. For addresses that could be matched to a landline telephone number, an automated voice message is also delivered as a reminder to complete and return the questionnaire. These follow-up procedures are carried out because previous work has demonstrated that varying the delivery mechanism (e.g. switching from regular first class mail to telephone or special mail), may improve response rates in mail surveys (Brick *et al.*, 2012).

Three weeks after the initial survey mailing, a final mailing is delivered to all addresses that have not yet responded to the survey. As with prior mailings, the final mailing is delivered via first class mail. It includes a nonresponse conversion letter, a second questionnaire, and a pre-paid return envelope.

3.3 Data Editing

FES data processing includes imputation of missing values (item nonresponse) and editing of inconsistent, illogical, or out-of-range values. Data editing is a sequential process whereby: 1) survey data are compiled, 2) possible errors are identified and flagged, 3) missing data and errors are corrected via automated imputation and editing, and 4) unlikely and out-of-range values are flagged and examined via manual review.

The FES consists of household questions and person-level questions for up to five individual household members. Key items are survey responses that are used to estimate fishing

effort, including the reported number of household members and sufficient person-level information to enumerate the total number of shore and private boat fishing trips taken during the reference wave by each household member. Returned questionnaires are classified as complete, partially complete or incomplete (returned blank). Complete surveys include responses for all key survey questions, and the number of complete household member sections matches the reported number of household members. Incomplete surveys, which are returned entirely blank, are treated as unit nonresponse and addressed through a nonresponse weighting adjustment (described in section 3.4.1). Partially completed surveys include some level of item nonresponse to key survey questions. In contrast, surveys are classified as illogical if the count of complete household members is greater than the reported number.

Partially completed and illogical surveys are corrected via automated processing, which reconciles inconsistencies between the reported number of household members and the count of completed individual household member sections. Illogical responses are reconciled by either editing the number of reported household members or eliminating inconsistent, 'extra', information. For partially completed surveys, imputed values for key items are either derived from complete auxiliary information (e.g. responses to household-level questions or counts of completed person-level sections) or, when auxiliary information is not available, based upon assumptions about the likelihood of fishing activity by individual household members. Surveys that cannot be resolved through automated processes are referred to subject matter experts for manual review and editing.

3.4 Estimation Methods

The FES estimates fishing effort by residents of sampled states in number of angler-trips. The basic approach uses a Horvitz-Thompson weighted total estimator (Horvitz and Thompson, 1952) with sample weights that reflect sample inclusion probabilities, a nonresponse adjustment, and a post-stratification adjustment to known population totals. A final adjustment that accounts for non-resident (i.e. out of frame) fishing activity is applied to estimate total effort by fishing mode. This adjustment is derived from the APAIS, and its related estimation methods are described in section 2.3.3. Total effort calculations for the FES are described in section 6.1.1.

3.4.1 FES Weighting

FES weights are calculated in several steps. In the first step, base weights (w_i) for each sampled address within a given stratum are calculated as the inverse of the inclusion probabilities

$$w_i = \frac{1}{\pi_i}$$

where π_i is the probability that unit *i* is included in the sample.

In the second step, base weights are adjusted to compensate for unit nonresponse (i.e. when households fail to mail back the completed survey). The sample is partitioned into nonresponse adjustment cells, or weighting classes, by state, sub-state region (coastal or non-coastal), license match (matched or unmatched), and matching telephone status (i.e. whether or not the sampled address could be matched to a landline telephone number). The base weights of

the respondents in a given adjustment cell ($w_{ci.r}$) are divided by the response rate for that cell ($\widehat{\emptyset}_c$) to calculate the adjusted weight (w_{ci}^*)

$$w_{ci}^* = \frac{w_{ci.r}}{\widehat{\emptyset}_c}$$

where $\widehat{\emptyset}_c = \frac{\sum w_{ci.r}}{\sum w_{ci.r} + \sum w_{ci.nr}}$,

$\sum w_{ci.r}$ is the sum of the base weights of each respondent within adjustment cell c, and

$\sum w_{ci.nr}$ is the sum of the base weights of each nonrespondent within adjustment cell c.

In the third step, the nonresponse-adjusted weights are post-stratified to account for incomplete coverage of the target population. Post-stratification is commonly used to make respondent data conform to target population totals from other sources independent from the survey (Brick and Kalton 1996). Here, the most recent, reliable estimates of the number of residential households available from the American Community Survey (United States Census Bureau 2016) are used as population control totals. Nonresponse adjusted weights are post-stratified to household-level control totals within coastal and non-coastal strata (as defined at the time of sampling for each wave). The resulting post-stratified weight (w_{hi}^*) of address i in stratum h is calculated as

$$w_{hi}^* = w_{ci}^* \left(\frac{H_h}{\widehat{H}_h} \right)$$

where the adjustment factor is equal to the ratio of the control total (H_h, from the American Community Survey) to the estimated total based upon the sum of nonresponse adjusted weights ($\widehat{H}_h$).

Following these three weighting adjustments, a final weight trimming process is applied to mitigate the impacts of extreme values on the precision of survey estimates. Highly variable weights can result in large variances, so it is often desirable to minimize the frequency and size of extreme weights. There is also a tradeoff, however, between increasing precision and biasing estimates through weight trimming procedures. The Estimated Mean Square Error (MSE) Trimming procedure allows for evaluating various trimming levels to identify an optimal level that minimizes the estimated mean square error of an estimate (i.e. minimizes the sum of sampling variance and the square of the estimated bias, Potter 1990; Potter 1988, Henry and Valliant 2017). The MSE for various levels of trimming ($\widehat{MSE}(\widehat{T}_t)$) is estimated as

$$\widehat{MSE}(\widehat{T}_t) = (\widehat{T}_t - \widehat{T})^2 - V(\widehat{T}) + 2[V(\widehat{T}_t)V(\widehat{T})]^{1/2}$$

where $\widehat{T}$ is the effort estimate, using untrimmed weights,

$\widehat{T}_t$ is the effort estimate using trimmed weights, and

$V(\widehat{T})$ and $V(\widehat{T}_t)$ are the estimated variance of $\widehat{T}$ and $\widehat{T}_t$ respectively.

The automated procedure, completed in SAS 9.4 (SAS Institute Inc., 2016), is carried out by repeatedly reducing maximum weighted values by increments of 5% and redistributing excess

weights among untrimmed sample cases. The $\widehat{MSE}(\widehat{T_t})$ is estimated for each incremental adjustment until the minimum value is identified, indicating that the optimal level of trimming has been reached.

3.4.2 Private Boat and Shore Fishing Effort Estimation

For each state and wave, the FES fishing effort ($\hat{T}_r$), or the number of in-state private and shore recreational fishing trips taken by state residents, is estimated as

$$\hat{T}_r = \sum_{h=1}^{H} \sum_{i=1}^{n_h} w_{hi}^* t_{hi}$$

where w_{hi}^* is the final weight of address i in stratum h, and
t_{hi} is the reported number of recreational fishing trips for address i in stratum h.

The variance is estimated using a Taylor series linearization

$$\hat{V}(\hat{T}_r) = \sum_{h=1}^{H} \left(\frac{n_h}{n_h - 1} \left(\sum_{i=1}^{n_h} w_{hi}^* t_{hi} - \frac{1}{n_h} \sum_{i=1}^{n_h} w_{hi}^* t_{hi} \right)^2 \right).$$

These effort estimates are produced using the SAS SURVEYMEANS procedures (SAS Institute Inc., 2016).

3.5 FES Certification Status

The FES design has been certified by MRIP. The development of the FES was a longer process than that of the APAIS; four different methods were piloted and built upon prior to seeking certification in 2015. First, a telephone survey of licensed anglers (called the Angler License Directory Survey, or ALDS), was piloted in Florida, Alabama, Mississippi, and Louisiana in 2007, and in North Carolina in 2008. In 2008, a dual-frame telephone survey approach that involved integrating the CHTS and the ALDS was developed and tested to determine if undercoverage of both surveys was decreased. Then, a dual-frame mail survey (using both a licensed saltwater database as one sample frame and residential addresses as a second sample frame) was developed and tested in North Carolina in 2009 and in both North Carolina and Louisiana in 2010. Finally, a dual-frame, mixed-mode survey was tested in 2012, which had a nearly identical design to the dual-frame mail survey, but samples were randomly allocated to mail and telephone treatment groups.

Following these tests and subsequent evaluations, the FES was developed and tested in Massachusetts, New York, North Carolina, and Florida for slightly over a year (September 2012-December 2013). The FES was independently peer-reviewed and certified by NMFS Leadership

in February of 2015, following recommendations by both the MRIP Operations Team and the MRIP Executive Steering Committee.

4. Legacy Coastal Household Telephone Survey (CHTS) Design: Shore and Private Boat Fishing Effort

The CHTS estimated marine recreational shore and private boat fishing effort in states along the Atlantic Coast and the Gulf of Mexico, in Puerto Rico, and in Hawaii. The survey used a computer-assisted, random digit dialing (RDD) approach (Glasser and Metzger, 1972) to identify and interview household members that participated in saltwater fishing activities. The dialing area generally encompassed counties within 25 miles from the coastline, but was expanded to 50 miles from May through October (waves 3-5) in the South Atlantic and had been expanded to 100 miles during these months in North Carolina. The survey was limited to coastal counties because historical APAIS data indicated that 65-90% of saltwater fishing trips were taken by residents of coastal counties (Andrews 2015). Data collection for the CHTS occurred during a two-week period at the end of each two-month wave.

The CHTS for many years had been a reliable and cost effective method for collecting household information on fishing effort. However, advancements in communications technology and shifts in public perceptions and attitudes resulted in diminishing rates of coverage and response, minimizing the efficiency and effectiveness of the CHTS, and RDD surveys in general (Curtin *et al.*, 2005; Blumberg and Luke, 2015). In 2016, CHTS response rates were less than 10%, and, for the first time, most U.S. households were wireless-only (Blumberg and Luke 2017). The CHTS was completely replaced by the FES in 2018, but between 2015 and 2017, both surveys were administered simultaneously for the purposes of calibrating effort estimates (see section 8.1 for these calibration methods).

4.1 Sampling Design

A list-assisted, RDD approach was used to sample landline telephone numbers. RDD works with blocks of 100 phone numbers, each consisting of the first five digits of a seven-digit telephone number within an area code (e.g. Glasser and Metzger 1972). The CHTS sample frame included all blocks within the dialing area that contained at least one listed household number. Only full-time residential households with landline telephone service were included in the survey; institutional housing, businesses, wireless phones and pay phones were excluded.

Sampling for each two-month reference period (wave) was stratified by state and county. Within each stratum, the sample was selected via simple random sampling without replacement. Within a wave and state, the samples were distributed among counties in proportion to the square root of the county population; this allocation approach was designed to avoid having very small sample sizes in counties with small populations.

4.2 Data Collection Design

After sample selection, telephone numbers were pre-dialed to identify and eliminate non-working numbers. Then, approximately 85% of the working numbers were loaded into a Computer Assisted Telephone Interviewing (CATI) system for dialing. The remaining 15% of the numbers were reserved in the event that the sample yields fall below the desired level. Sampling productivity was monitored throughout the sampling period, and if sample yields were lower than anticipated (e.g., there was high nonresponse), the reserved 15% of the numbers were released to complete the desired number of interviews.

The CATI system automatically dialed telephone numbers and scheduled dialing attempts on unresolved records to meet minimum protocols, which were as follows:

- A minimum of five attempts were made to categorize each sampled number as an interview (partial or complete), an eligible non-interview (non-respondents), ineligible or unknown eligibility (based on the standard final disposition codes for survey research [American Association for Publich Opinion Research 2016]).
- At least one weekday attempt and three night (after 5pm) or weekend attempts needed to be made per number.
- At least one of the nighttime attempts needed to be a weekend attempt.
- No calls were attempted before 8:00AM or after 9:00PM local time for the area being called.

The CATI system also automatically queued and dialed scheduled callbacks, notifying CHTS samplers when callbacks were scheduled. The interview began by determining if any household members participated in saltwater recreational fishing during the previous two months. Each participating angler was then asked to provide details about all fishing trips that occurred during the prior 60 days, beginning with the most recent trip.

Although data was preferably collected through a direct interview with an angler, proxy data (i.e. information gathered from someone in the fishing household other than a specific angler) was collected in certain situations. When a respondent indicated that all household trips were made as a group, then those responses were duplicated without interviewing the other anglers in the household. Further, an adult could answer in place of a child if that adult knew details of the child's fishing trips. Finally, proxy interviews were permitted if an angler could not be reached after five call attempts.

The CATI system was programmed to identify unreasonable and/or inconsistent values. Outlier values (a reported number of household trips greater than the 95th percentile from the previous four years for that state, wave and fishing mode) were identified and subjected to follow-up interviews.

There were three main data elements collected by the CHTS: household information, angler information, and trip information. The household information included the total number of anglers living in fishing households. The angler information included the number of fishing trips taken by each angler within the household during the wave. The trip-specific information

included the fishing mode, date, state and county, access type (private or public access) and time for each fishing trip.

4.3 Imputation Methods for Missing Values

A hot deck imputation procedure was applied to compensate for item nonresponse. The procedure replaced missing values in the data with values randomly selected from complete observations in the current survey (e.g. Andridge and Little 2017; Lessler and Kalsbeek 1992). Generally, for nonresponding households, data was imputed in two rounds from a donor dataset comprised of households where all resident anglers were reached for interview. The first round was done by state, county and angler. If further imputation was needed, the second round was carried out by county and angler. Missing angler and trip data were also imputed from donor datasets comprised of complete angler and trip records.

Missing household data were the first level of information that was imputed. The first step of imputation was to randomly assign the households into two categories: fishing or non-fishing. Some of the households initially identified as fishing were subsequently reclassified as non-fishing on further contact. Each completely missing household was categorized as fishing or non-fishing using a simple Bernoulli trial, (a random experiment that has two possible outcomes, where the probability of each outcome remains constant each time the experiment is conducted, e.g. Uspensky 1937; Papoulis 1984) where the probability of being a fishing or non-fishing household was based on historical information.

Trip imputations were done if a missing household was categorized as fishing. Trip information was imputed by copying data from a suitable donor household that had complete data. A suitable donor was a household in the same state (or if unavailable, the same sub-region) that had the same number of anglers. The mode and the county of each fishing trip from the donor household were duplicated for the recipient household.

Angler imputations were done if a household had some anglers with complete information, but others provided incomplete information (e.g. one angler in a household indicated that other anglers lived in the household, but those remaining anglers were not reached for interview). In this scenario, missing data from the complete anglers was copied and applied to incomplete anglers, again using a simple random selection process.

If a given angler detailed some fishing trips with complete information but some trips with missing information, the imputation randomly copied information from the angler's complete trips and applied it to the incomplete trips.

Finally, if the fishing mode of the angler-trip was the only missing information then the mode was assigned from that angler's other trips, with multinomial probabilities based on the frequency of each mode for that angler.

4.4 Estimation Methods

The CHTS fishing effort (in number of angler-trips) was estimated in stages. First, outlier values were identified and truncated to minimize extreme variability in estimates. Next,

county-level fishing effort was estimated by multiplying mean shore and private boat trips per household within each county by the total number of households in the county. County estimates within each state were then summed to estimate total effort by residents of coastal counties. Finally, total effort for each state was estimated by multiplying coastal resident effort estimates by a coverage adjustment derived from the APAIS survey (section 2.3.5; see section 6.1.2 for a more detailed description of total effort estimation from the CHTS data).

4.4.1 Winsorization

A procedure known as winsorization was applied to CHTS data to reduce the impact of outlier values (Barnett and Lewis, 1994). Using data from the previous four years of the survey, the numbers of trips per household were truncated to the 95th percentile of this historical data. Truncation was done separately for each state, wave, and fishing mode (private boat or shore). This outlier reduction resulted in a negative bias but also minimized large fluctuations in estimates that could occur if outliers were retained in the data.

4.4.2 Private Boat and Shore Fishing Effort Estimation

The fishing effort estimates were produced using a Horvitz-Thompson estimator (Horvitz and Thompson, 1952). First, the stratum-level effort ($\hat{T}_h$) was estimated as

$$\hat{T}_h = N_h \frac{\sum_{i=1}^{n_h} t_{hi}}{n_h}$$

where N_h was the total number of households in stratum h;
n_h was the total sample size (i.e. the number of households completing the survey) in stratum h; and
t_{hi} was the reported number of recreational fishing trips of household i in stratum h;

The stratum-level effort was then summed over all coastal counties in a given state, two-month wave and fishing mode to produce the following total estimate of fishing trips ($\hat{T}$):

$$\hat{T} = \sum_{h=1}^{H} \hat{T}_h.$$

The variance was then estimated as

$$V(\hat{T}) = \sum_{h=1}^{H} \frac{N_h^2}{n_h} \frac{\sum_{i=1}^{n_h} (t_{hi} - \hat{\bar{t}}_h)^2}{n_h - 1}.$$

4.5 CHTS Certification Status

Certification was never sought for the CHTS because of the number of shortcomings identified by the NRC's 2006 review of the MRFSS. Being a legacy survey, however, it continued until a superior design could be developed and implemented. Once the FES was developed and certified, which was nearly an 8-year process from first pilot study to the signing

of a certification memorandum, it required a 3-year benchmarking period to compare the effort estimates derived from the CHTS and FES. The CHTS was discontinued as of January 2018.

5. The For-Hire Survey (FHS): For-Hire Fishing Effort

The For-Hire Survey (FHS) gathers data via telephone interviews for estimating for-hire fishing effort. Respondents, which are all contacts for for-hire vessels, are asked to report vessel-fishing activity for the prior week and then recount details about each trip, including area fished, the number of anglers who fished from the boat, hours spent fishing, method of fishing and target species. These fishing effort data are used in conjunction with the catch data collected from the charter and headboat intercepts of the APAIS in order to estimate total for-hire catch.

The FHS was initially developed because the CHTS did not adequately cover for-hire fishing modes; historical MRFSS data indicated that the majority of anglers that fish on for-hire boats reside either out-of-state or within a coastal state but not in a coastal county (i.e. outside of the 25-50 mile CHTS coverage area). Whereas the design of the CHTS uses an RDD approach, the FHS survey design uses a continually updated directory of known for-hire vessels to create a sample frame that exclusively targets for-hire fishing. FHS data collection occurs on a weekly basis within two-month waves, annually. The survey has been continuously implemented in the Gulf of Mexico since 1997, and along the Atlantic Coast since 2003 (Lai and Foster, 2008).

5.1 Sampling Design

The FHS has a stratified design, with for-hire vessels as sampling units. The sample frame is generated from a continually updated list of known for-hire vessels, from which samples are selected without replacement.

5.1.2 Stratification Variables and Sample Frame

Sampling is stratified by sub-region, state, vessel type (headboat or charter boat), and sample week within each two-month wave. For the purposes of the survey, the sample week is Monday through Sunday.

The sample frame is constructed from a continually updated directory of known for-hire vessels (called the FHS directory) two weeks prior to the sampling wave. Vessel records in the FHS directory contain a vessel identifier (vessel name or registration number); county and state (and site, if known) in which the vessel operates; contact information for the vessel representative (captain, owner, or proxy) including name, address and telephone number; vessel status; and the vessel's cooperation level. Vessel status is listed in the directory as active, (i.e. currently participates in for-hire fishing activities), inactive (i.e. does not currently participate in for-hire fishing activities; for example, a vessel would be considered inactive if is being repaired, or has switched to commercial fishing for a period of time) or ineligible (i.e. will no longer participate in for-hire activities). Cooperation levels are either cooperative, where the vessel

representative responds to FHS telephone interviews, or non-cooperative, where the FHS vessel representative does not respond or refuses to participate outright. The FHS directory is updated regularly based on input from APAIS samplers, state FHS coordinators and vessel representatives. The directory can also be updated with information obtained during the telephone survey (e.g. if a vessel representative reports that a vessel will be inactive for a certain period of time, such information will be added to the directory).

To be included in the sample frame, a vessel must meet three criteria. First, the vessel's status must be active. Second, there must be complete contact information, including the vessel identifier and at least one telephone number for the vessel representative. Third, the county and state in which the vessel operates must be known. If the vessel does not meet these criteria, it remains in the vessel directory but is excluded from the sample frame. Vessels that are non-cooperative are kept in the sample frame but are automatically coded as a refusal and are not actually contacted if selected for sampling.

5.1.3 Sample Selection

Prior to the sample selection, the sample frame is sorted by three additional variables, creating three additional implicit strata: business county, vessel length, and permit type. The business county variable is the county in which the vessel operates. The vessel length variable simply categorizes the vessels as small, medium and large. Permit types are Highly Migratory Species (HMS) or non-HMS; additional questions, related to the Large Pelagics Survey, are asked for the HMS-permitted vessels (see section 7.1.3). In addition to these three variables, a uniform random variable is created and used to randomly order vessels within the business county, vessel length, and permit type groups.

Sample selection is then systematically done without replacement at the stratum level (by vessel type, state, sample week, and by the implicit strata business county, vessel length, and permit type, from the sample frame sorting process). The FHS has a fixed sampling rate of 10% within strata. In addition, there is a minimum sample size requirement of three vessels from each stratum.

5.2 Data Collection Design

The sample selection is completed on the 13th of the month, before the start date of each two-month wave. All phone interviews are conducted using a CATI system, similar to the one used for the CHTS described in section 4.2, in the 7 days immediately following a sample week, giving the vessel representatives a recall period of 7-14 days. This timeframe was chosen to minimize both the potential for recall error as well as the reporting burden during the survey interview.

A minimum of 10 attempts are made to contact the selected vessel representatives during a reporting period. The first attempt is made on the first day following the sample week (i.e. Monday, because the survey runs Monday-Sunday) and the remaining attempts are spread over the rest of the following week as day and evening attempts. Day attempts are before 5PM, and evening attempts are after 5PM but before 9PM. If someone other than the selected vessel

representative answers the phone during one of the 10 initial attempts, then five additional attempts are made for 15 total call attempts.

To improve response rates, an advance letter is mailed to the representatives of all selected vessels one week before the sample week (i.e. two weeks before the phone interview). The letter details the dates of the sample week that representatives will be asked about during the interview, the contact information of the organization conducting FHS interviews, a logsheet with the questions that will be asked, as well as some alternative reporting options for completing the survey. If the vessel representative cannot take the phone interview, he or she can fill out and submit the logsheet by fax or mail, or report their trip information through an online reporting website.

The key data collected in the phone interviews are:

- the number of vessel-trips with paying passengers in the sample week;
- the date of each vessel-trip;
- the mode of each vessel-trip;
- the number of anglers on each vessel-trip;
- the state/county and site where each vessel-trip returned;
- the fishing methods used during each vessel-trip;
- the targeted species for each vessel-trip;
- the fishing area for each vessel-trip;
- the distance from shore where each vessel-trip occurred;
- the hours spent fishing for each vessel-trip; and
- the return time for each vessel-trip.

For quality control of the telephone interviews, 10% of each interviewer's work is validated either by having a supervisor listen to those phone calls, or by independently re-contacting the vessel representative at a later date to confirm the interview.

5.3 Vessel Trip Reports (VTRs)

In addition to the main telephone survey, an administrative record data source called Vessel Trip Reports (VTRs) are used to supplement the FHS data in order to calculate for-hire catch and effort. The VTR program, operated by the Greater Atlantic Regional Fisheries Office (GARFO), is a federal logbook census for permitted vessels fishing for species in the northeast region. The VTR program operates primarily from Maine to Virginia, but location is driven by species distributions rather than state-lines; North Carolina is the southernmost state where VTRs are typically submitted. There are six permit types covered by the program, including 1) Bluefish, 2) Black Sea Bass, 3) Summer Flounder, 4) Northeast Multispecies, 5) Scup, and 6) Squid/Mackerel/Butterfish. By law, federally permitted vessels are required to document their fishing effort and harvest by submitting VTRs to GARFO. Vessel representatives are required to submit reports for every fishing trip taken; if a boat did not fish during a calendar month, the vessel is still required to submit a report stating that they did not fish. The GARFO requires that VTRs be submitted monthly (i.e. received or postmarked by the 15th day of the month following the trip landing date). At the end of each reporting period, GARFO sends all of the VTR data

that they receive from vessels to MRIP for integration with the FHS for producing total wave and annual for-hire effort estimates.

Because participation is mandatory, the VTR program provides a census of effort and harvest for federally permitted for-hire vessels, however it is used as a supplement rather than a replacement for the FHS for several reasons. First, the VTR program does not cover non-permitted for-hire vessels, and therefore misses some vessels that are then captured through the FHS. In addition, VTR data are self-reported and not subject to validation through independent observation like the FHS. Integrating both datasets therefore helps overcome the shortcomings of both approaches.

As some vessels are covered by both the FHS and VTR, the datasets are merged to identify vessels in both frames. For vessels that are in both frames, only the VTR data are used and the FHS data are excluded for those vessels. The decision to exclude FHS data where it overlaps with VTR data was made because the VTR data, being a census, is assumed to be more complete and has a variance of zero, which produces a more precise estimate. Estimates are then calculated using all of the VTR data as well as all remaining FHS data from non-VTR vessels (see section 6.1.3 for total for-hire effort estimation methods).

5.4 Estimation Methods

For-hire fishing effort is estimated in numbers of angler-trips per sub-region, state, two-month wave, vessel type, and fishing area (inshore, nearshore, offshore). To get a total effort estimate, the effort estimate component is corrected by two other estimate components – the coverage adjustment calculated from the APAIS (section 2.3.4) and a reporting error from a validation study conducted in conjunction with the FHS (section 5.4.2). Here the FHS effort estimate component is described along with the reporting error calculation. Total for-hire wave and annual effort estimation is described in more detail in section 6.1.3.

5.4.1 For-Hire Fishing Effort Estimation

To obtain a base estimate of fishing effort in a given wave, vessel type, and fishing area, the number of angler-trips in each fishing area per sample week is estimated, and then estimates are summed over the weeks in the two-month wave within vessel type and fishing area domains. The number of angler-trips per vessel type and sample week ($\hat{T}_h$) is estimated as

$$\hat{T}_h = \frac{N_h}{n_h} \sum_{i=1}^{n_h} t_{hi} = N_h \hat{\bar{t}}_h$$

where t_{hi} is the number of angler-trips aboard vessel i (out of vessels 1 ... n_h) in stratum h;
N_h is the total number of vessels in stratum h (vessel type by sample week); and
n_h is the sample size within stratum h.

The associated variance is estimated as

$$V(\hat{T}_h) = {N_h}^2 \frac{(1 - \frac{n_h}{N_h})}{n_h(n_h - 1)} \sum_{i}^{n_h} (t_{hi} - \hat{\bar{t}}_h)^2.$$

The base estimate of the number of angler-trips by vessel type and area ($\hat{T}$) and its variance in a wave are then estimated as

$$\hat{T} = \sum_{h} \hat{T}_h$$

and

$$V(\hat{T}) = \sum_{h} V(\hat{T}_h).$$

5.4.2 Dockside Validation and Adjustment for Reporting Error

The FHS collects self-reported data (activity reported by vessel captains) and consequently may be subject to recall error. A dockside validation study is therefore implemented on for-hire vessels in the FHS sample frame, during the sample week (i.e. the week prior to the actual telephone interviews). The study is carried out by APAIS samplers during their assignments at fishing access sites with known for-hire activity. Samplers determine the status (i.e. docked or away) of the for-hire vessels in the FHS sample frame. If the vessel is away from the dock, an attempt is made to identify the reason it is gone (e.g. out fishing, other). At least one validation attempt is made for all vessels drawn for sampling by the FHS during the sample week, but multiple validations per week are encouraged and often made. The validation sites are visited at a time of day when the vessel is likely to be away from the dock on a for-hire fishing trip. Multiple vessel-trips per vessel per day are recorded.

The validation data are classified into seven categories (see Table 4). When estimating reporting error, validation categories 0 and 3 are excluded since no validation attempts were made for those records. Category 6 is also excluded; because the validation survey is carried out by APAIS samplers when they happen to be at sites where FHS vessels are located, validations of the FHS vessel status sometimes occur at a time different from what is reported by the vessel representative during the following week's telephone interview. In these scenarios, it is not possible to validate whether or not the trip occurred as reported. In addition, because validation samples may not be sufficient to handle within-week variability, the reporting error correction factors are calculated within domains defined by year, wave, sub-region, state, vessel type or fishing mode (charter, headboat).

Table 4: Validation Categories for Combining FHS and Validation Study Information

Validation Category	Information Collected Through FHS	Information Collected from Validation Study
0	No trip reported	No validation attempt made
1	No trip reported	Validated in-dock
2	No trip reported	Validated out for fishing
3	Reported trip	No validation attempt made

4	Reported trip	Validated in dock
5	Reported trip	Validated out for fishing
6	Reported trip	Validated (either fishing or in dock) at the wrong time such that it is unknown whether fishing occurred at reported time

The mean reporting error is estimated as the 'change' from what is reported by the survey respondent to what is observed from the dockside validations. The calculations are done in a sequence of steps where the outputs of the previous calculations feed into the subsequent calculations (Figure 5). The reporting error is required to adjust both the preliminary wave estimates as well as the final annual estimates with different ranges of data being used for the preliminary and final reporting errors. To calculate the preliminary wave-level reporting error, validation data are used from the preceding five to six waves to produce a robust and data-rich error estimate for that wave; the mean estimated error is first calculated individually for the current wave and then pooled with the error of the previous five to six waves. At the end of the year, the final reporting error is calculated only using the current year's data to produce a single correction factor. This final reporting error is used to recalculate each of the wave effort estimates, within the year, which are then summed to produce an annual estimate. These wave and annual total effort estimates are described in more detail in section 6.1.3.

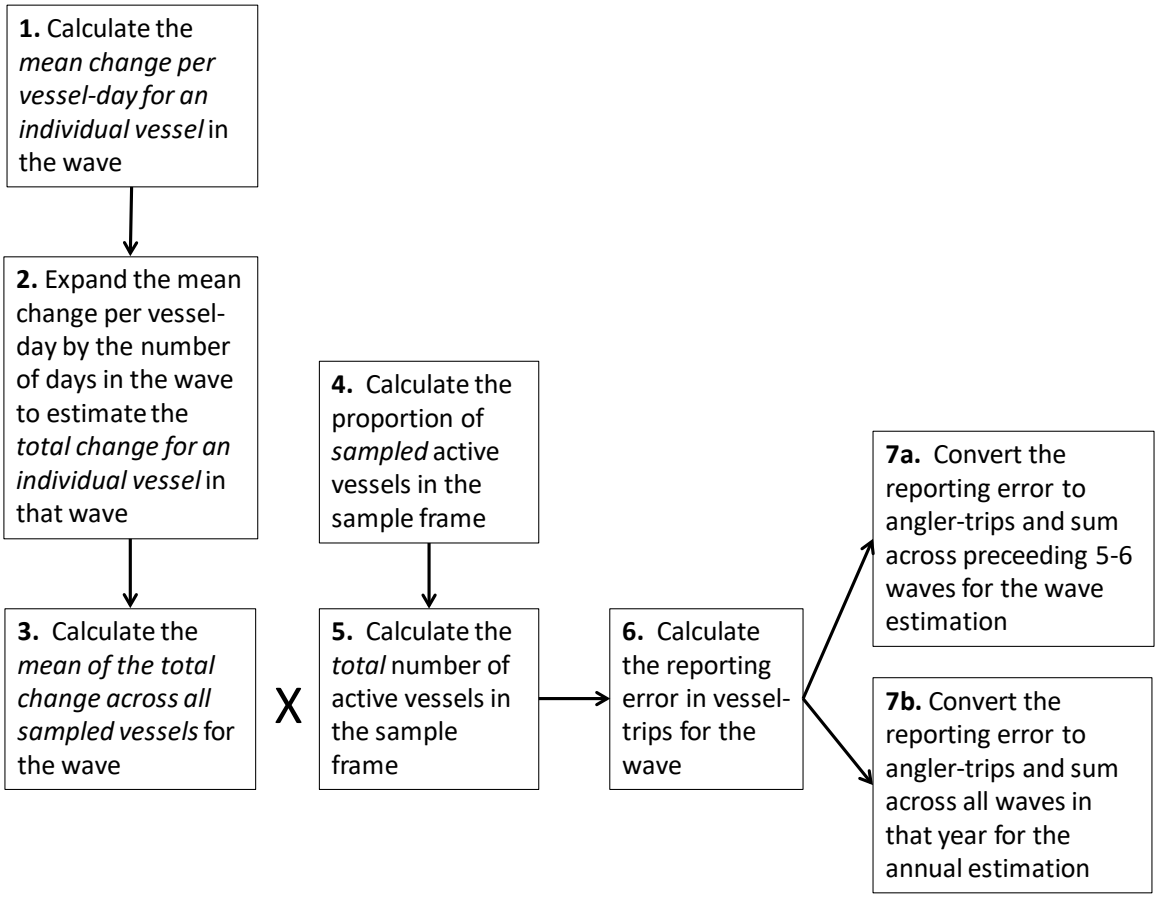

Figure 5: Steps in calculating FHS reporting error

Note that the error is calculated using data from the current and preceding five to six waves, and applied to the base effort estimate for the current wave; first the error is calculated individually for a wave and then pooled across the waves to be applied. The reporting error is recalculated at the end of the year, using the data from all waves in that year, to adjust the annual estimate of fishing effort.

Step 1: Calculate the mean change per vessel-day for an individual vessel in the wave

To calculate the reporting error, first the mean change per vessel-day of the i-th sampled vessel ($\hat{\bar{x}}_i$) in a given wave and its variance are estimated as

$$\hat{\bar{x}}_i = \sum_{j=1}^{m_i} \frac{x_{ij}}{m_i}$$

and

$$V(\hat{\bar{x}}_i) = \sum_{j=1}^{m_i} \frac{(x_{ij} - \hat{\bar{x}}_i)^2}{m_i(m_i - 1)}$$

where x_{ij} is the change of the i-th sampled vessel (for $i = 1, \ldots n$ vessels in validation survey) in

the j-th vessel-day (for j = 1,…,m_i vessel-days sampled); if the validation category is 1 or 5, $x_{ij} = 0$, if category is 2, $x_{ij} = 1$ and if category is 4, $x_{ij} = -1$; and

m_i is the number of vessel-days sampled in the given wave.

Step 2: Expand the mean change per vessel-day by the number of days in the wave to estimate the total change for an individual vessel in that wave

The mean change per vessel-day for the i-th sampled vessel is then expanded by the total number of days in the given wave to estimate the total change per wave ($\hat{z}_i$). This estimated total change per wave and its variance are then calculated as

$$\hat{z}_i = M_i \hat{\bar{x}}_i$$

and

$$V(\hat{z}_i) = M_i{}^2 V(\hat{\bar{x}}_i)$$

where M_i is the total number of days in the wave (i.e. approximately 60 days, since each wave is two months in length).

Step 3: Calculate the mean of the total change across all sampled vessels for the wave

The mean is then taken of the total change across all sampled vessels in the wave to obtain the overall mean change per vessel-day and its variance using the following calculations:

$$\hat{\bar{z}} = \frac{\sum_{j=1}^{m_i} \hat{z}_i}{n}$$

$$V(\hat{\bar{z}}) = \sum_{i=1}^{n} \frac{(\hat{z}_i - \hat{\bar{z}})^2}{n(n-1)}$$

where n is the total number of vessels in the wave.

Step 4: Calculate the proportion of sampled active vessels in the sample frame

The overall mean change per vessel-day is needed to estimate the number of vessel-trips with which to adjust the wave estimate. Prior to that step, two separate calculations are needed to estimate the proportion of active vessels from total vessels that are sampled, and then the total number of active vessels within the wave. The proportion ($\hat{p}$) of active vessels sampled within the targeted population and its associated variance are estimated as

$$\hat{p} = b/B$$

and

$$V(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{(B-1)}$$

where b is the number of sampled active vessels in the validation survey, and
B is the number of sampled vessels (active vessels + inactive vessels) in the validation survey.

Step 5: Calculate the total number of active vessels in the sample frame

The total number of active vessels in the sample frame for the wave ($\widetilde{N}$) can then be estimated by multiplying the proportion ($\hat{p}$) of sampled vessels that were active by the total number of vessels in the sample frame (N) for that wave

$$\widetilde{N} = \hat{p}N$$

with the variance of $\widetilde{N}$ estimated as

$$V(\widetilde{N}) = N^2 V(\hat{p}).$$

Step 6: Calculate the reporting error in vessel-trips for the wave

From the above calculations for the number of active vessels in the sample frame ($\widetilde{N}$) and overall mean change per vessel-day in a wave ($\hat{\bar{z}}$), the estimated number of vessel-trips with which to adjust the effort estimate in a given wave (Δ) and its variance are

$$\Delta = \widetilde{N}\hat{\bar{z}}$$

and

$$V(\Delta) = \hat{\bar{z}}^2 V(\widetilde{N}) + \widetilde{N}^2 V(\hat{\bar{z}}) - V(\widetilde{N})V(\hat{\bar{z}}).$$

Step 7a: Convert the reporting error to angler-trips and sum across the preceding 5-6 waves for the wave estimation

Next, since the validation study is conducted by observing vessel-trips but the FHS effort estimate is in angler-trips, the reporting error needs to be expanded from vessel-trips to angler-trips. In this step, the reporting error is also pooled across the preceding five to six waves, which again, is done to produces a more data-rich estimate. This final reporting error (ϕ_w) is calculated as

$$\phi_w = 1 + \frac{\sum_{w=1}^{W} \Delta}{\sum_{w=1}^{W} \hat{B}_a}$$

where $\hat{B}_a$ is the unadjusted vessel-trip estimate for a given wave, which is summed across the preceeding five to six waves, which is estimated as

$$\hat{B}_a = \sum_h b_{ha}$$

with a variance of

$$V(\hat{B}_a) = \sum_h V(\hat{b}_{ha})$$

where $\hat{b}_{ha} = \frac{N_h}{n_h}\sum_{i=1}^{I} b_{hai} = N_h \hat{\bar{b}}_{ha}$ and $V(\hat{b}_{ha}) = {N_h}^2 \frac{(1-\frac{n_h}{N_h})}{n_h(n_h-1)}\sum_i (b_{hai} - \hat{\bar{b}}_{ha})^2$;

b_{hai} is number of vessel-trips taken by vessel i, in fishing area a and in sampling period h;

N_h is the total number of vessels in the sample frame within sampling period h; and

n_h is the sample size within sampling period h.

The variance of ϕ_w is then estimated as

$$V(\phi_w) = \frac{(\sum_{w=1}^{W}\Delta)^2 \sum_{w=1}^{W} V(\hat{B}_a)}{\left(\sum_{w=1}^{W}\hat{B}_a\right)^4} + \frac{\sum_{w=1}^{W} V(\Delta)}{\left(\sum_{w=1}^{W}\hat{B}_a\right)^2} - \frac{2\rho(\sum_{w=1}^{W}\Delta)\sqrt{\left(\sum_{w=1}^{W} V(\hat{B}_a)\right)\left(\sum_{w=1}^{W} V(\Delta)\right)}}{\left(\sum_{w=1}^{W}\hat{B}_a\right)^3}$$

where $\rho = \frac{\left(\sum_{w=1}^{W}\hat{B}_a\Delta - \left(\sum_{w=1}^{W}\hat{B}_a\right)\left(\sum_{w=1}^{W}\Delta\right)\right)\Big/\sum_{w=1}^{W} n}{\sum_{w=1}^{W} n}$; and

n is the number of vessels sampled in a given wave.

This wave reporting error is applied to the base effort estimate, along with the coverage adjustment calculated from the APAIS (section 2.3.4), to obtain the wave estimates.

Step 7b: Convert the reporting error to angler-trips and sum across all waves in that year for annual estimation

The annual reporting error (ϕ_A) is estimated in the same manner as the wave reporting error (ϕ_w), with the exception that $\hat{B}_a$ is estimated by summing the unadjusted effort across the six waves comprising the current year (data from previous years are excluded.)

This reporting error for the year is applied to the base effort, along with the coverage adjustment calculated from the APAIS to produce annual effort estimates.

5.5 FHS Certification Status

The FHS survey design is in the early stages of certification and MRIP team members are organizing an independent peer-review to begin in 2019.

6. Total Catch and Effort Estimation: Combining the Estimate Components from the APAIS, FES, CHTS and FHS

MRIP produces both wave and annual estimates of total catch and effort by combining estimate components from the APAIS with those from the FES and FHS. Wave-level total effort is estimated by adjusting the base effort estimates from both effort surveys (section 6.1.1 and 6.1.3) by coverage adjustments derived from the APAIS. The FES private and shore mode effort is further partitioned by area fished. The methods for total CHTS effort estimation, while no longer used, are included below as well for informational purposes (6.1.2). Annual total effort estimates are calculated as a simple sum of the wave estimates.

There are several methods for estimating total catch: the standard method (section 6.2.1), an alternative method for small sample sizes (section 6.2.2), and the public-use method (section 6.2.3). Standard wave-level total catch is estimated as the product of the total effort and catch rates, and then total catch at the annual level is just the sum of the wave estimates. The alternative methods for small sample sizes (which are, on rare occasion, caused by low APAIS sampling productivity for the charter boat mode in select states) involve using collapsed annual domains instead of wave domains, so this method only produces annual estimates. A third total catch estimation method, used at the wave-level, is just a simplified version of the standard method designed for use with MRIP's public-use datasets; it is made available to data users in the form of template R and SAS programs.

6.1 Total Effort Estimation

6.1.1 FES Total Effort

For the FES, total effort is estimated for each coastal state, two-month wave, and fishing mode (private boat or shore), and is partitioned by area fished (inland waters, State Territorial Seas, EEZ). For these wave estimates, total effort ($\hat{T}_{I,FES}$) is first estimated by dividing the FES base effort estimates ($\hat{T}_r$, described in section 3.4.2) by coverage adjustments derived from the APAIS for out-of-frame angler-trips ($\hat{p}_q$, detailed in section 2.3.3) as

$$\hat{T}_{I,FES} = \frac{\hat{T}_r}{\hat{p}_q}.$$

The variance $V(\hat{Y}_{I,FES})$ is estimated using the delta method (e.g. Author, Bieler, and Williams 1993) as

$$V(\hat{T}_{I,FES}) = \frac{V(\hat{T}_r)}{V(\hat{p}_q)} = \frac{1}{\hat{p}_q{}^2} V(\hat{T}_r) + \frac{\hat{T}_r{}^2}{\hat{p}_q{}^4} V(\hat{p}_r).$$

Then, the adjusted FES effort estimate is multiplied by the area fished proportions ($\hat{p}_a$, point and variance estimation shown in section 2.3.6) to estimate effort by area fished($\hat{T}_{T,FES}$) as

$$\hat{T}_{T,FES} = \hat{T}_{I,FES}\hat{p}_a.$$

The variance $V(\hat{T}_{T,FES})$ is estimated using Goodman's formula (Goodman 1960)

$$V(\hat{T}_{T,FES}) = \hat{T}_{I,FES}{}^2 V(\hat{p}_a) + \hat{p}_a{}^2 V(\hat{T}_{I,FES}) - V(\hat{T}_{I,FES})V(\hat{p}_a).$$

Annual effort and variance are estimated as

$$\hat{T}_{Annual,FES} = \sum_w \hat{T}_{T,FES}$$

and

$$V(\hat{T}_{Annual,FES}) = \sum_w V(\hat{T}_{T,FES}).$$

6.1.2 CHTS Total Effort

CHTS total fishing effort for each two-month wave was estimated by sub-region, state, fishing mode (private boat or shore mode) and partitioned by area fished (inland water, State Territorial Seas, and EEZ). The estimate was produced in two steps. First, the CHTS base effort estimate ($\hat{T}_h$, point and variance estimation shown in section 4.4.2) was expanded by coverage adjustments derived from the APAIS for out-of-frame angler-trips ($\hat{R}_m$, point and variance estimation shown in section 2.3.5). Total effort ($\hat{T}_{T,CHTS}$) and variance were estimated as

$$\hat{T}_{I,CHTS} = \hat{T}_h \hat{R}_m$$

$$V(\hat{T}_{I,CHTS}) = \hat{T}_h{}^2 V(\hat{R}_m) + \hat{R}_m{}^2 V(\hat{T}_h) - V(\hat{T}_h)V(\hat{R}_m)$$

Total effort estimates by area fished($\hat{T}_{T,FES}$) were produced by multiplying total effort by APAIS estimates of the relative distribution of effort by area fished ($\hat{p}_a$, point and variance estimation shown in section 2.3.6)

$$\hat{T}_{T,CHTS} = \hat{T}_{I,CHTS}\hat{p}_a$$

$$V(\hat{T}_{T,CHTS}) = \hat{T}_{I,CHTS}{}^2 V(\hat{p}_a) + \hat{p}_a{}^2 V(\hat{T}_{I,CHTS}) - V(\hat{T}_{I,CHTS})V(\hat{p}_a).$$

Wave estimates were summed to estimate annual effort

$$\hat{T}_{Annual,CHTS} = \sum_w \hat{T}_{T,CHTS}$$

with a variance of

$$V(\hat{T}_{Annual,CHTS}) = \sum_{w} V(\hat{T}_{T,CHTS}).$$

6.1.3 For-Hire Effort – Combining VTR Data with FHS Data

For for-hire fishing modes, VTR data and FHS data are integrated to produce both wave and annual effort estimates. To do this, the FHS sample frame and the VTR permit frame are merged in order to identify vessels in one of three categories: 1) non-federally permitted vessels listed in the FHS frame only, 2) federally permitted vessels that are in both the FHS and VTR frames, and 3) federally permitted boats that are listed in the VTR frame only. For those vessels in category 2, the FHS data is excluded and only the data from the VTR frame is used in estimation. As the VTR data is self-reported and subject to reporting error like the FHS, the VTR data are corrected for reporting error using the methodology described for the FHS in section 5.4.2: VTR data for vessels on the FHS frame are verified by comparing reported effort to dockside observations from the FHS validation study. These methods for estimating for-hire effort are described below.

6.1.3.1 Wave Estimates

FHS total effort estimates are calculated by sub-region, state, two-month wave, and fishing mode (charter boat or headboat). These estimates are produced in two steps. The FHS base effort estimate ($\hat{T}$ described in section 5.4.1) is first multiplied by the coverage adjustment for out-of-frame angler-trips ($\hat{R}_{mi}$ calculated from the APAIS data in section 2.3.4) and then by the adjustment for reporting error (ϕ_w, calculated from the dockside validation component of the FHS, detailed in section 5.4.2). The first calculation to produce an initial estimate of for-hire effort ($\hat{T}_{I,FHS}$) is therefore

$$\hat{T}_{I,FHS} = \hat{T}\hat{R}_{mi}.$$

and the variance is estimated using Goodman's formula as

$$V(\hat{T}_{(I,FHS)}) = \hat{T}_{I,FHS}{}^2 V(\hat{R}_m) + \hat{R}_m{}^2 V(\hat{T}_{I,FHS}) - V(\hat{T}_{I,FHS})V(\hat{R}_m).$$

Then, the reporting error (ϕ_w) is applied to this initial estimate of for-hire effort to obtain the total wave estimate ($\hat{T}_{T,FHS}$) as

$$\hat{T}_{T,FHS} = \hat{T}_{I,FHS}\phi_w.$$

The variance is calculated, again using Goodman's formula, as

$$V(\hat{T}_{T,FHS}) = \hat{T}_{I,FHS}{}^2 V(\phi_w) + \phi_w{}^2 V(\hat{T}_{I,FHS}) - V(\hat{T}_{I,FHS}) V(\phi_w).$$

Because the VTR represents a census of vessel activity, the number of weekly angler-trips for each state is simply the sum of the angler-trips for individual permitted vessels, and the variance of the estimate is zero. However, VTR census counts are also subject to reporting error like the FHS. The VTR data for vessels on the FHS frame are therefore verified by comparing the reported trips to dockside observations from the FHS validation study. The correction factors for this domain are then calculated by the FHS wave reporting error methodology (ϕ_w, described in section 5.4.2). VTR vessels that are not duplicated by the FHS frame are assumed to be similar to those VTR vessels that are on the FHS frame, so the reporting error is applied to all VTR vessel trips. The VTR total effort ($\hat{T}_{T,VTR}$) for a given wave is therefore estimated as

$$\hat{T}_{T,VTR} = \phi_w \sum_{i=1}^{n_h} t_{hai}$$

where t_{hai} denotes the number of angler-trips aboard vessel i, in fishing area a and in sampling period h.

The variance for $\hat{T}_{T,VTR}$ is calculated using Goodman's formula

$$V(\hat{T}_{T,VTR}) = \phi_w{}^2 V\left(\sum_{i=1}^{n_h} t_{hai}\right) + \left(\sum_{i=1}^{n_h} t_{hai}\right)^2 V(\phi_w) - V(\phi_w) V\left(\sum_{i=1}^{n_h} t_{hai}\right),$$

which can be simplified to

$$V(\hat{T}_{T,VTR}) = \phi_w{}^2 * 0 + \left(\sum_{i=1}^{n_h} t_{hai}\right)^2 * V(\phi_w) - V(\phi_w) * 0 = \left(\sum_{i=1}^{n_h} t_{hai}\right)^2 V(\phi_w).$$

Then, the total for-hire effort ($\hat{T}_{T,FHS+VTR}$) for a wave is simply

$$\hat{T}_{T,FHS+VTR} = \hat{T}_{T,FHS} + \hat{T}_{T,VTR}$$

and the variance is

$$V(\hat{T}_{T,FHS+VTR}) = V(\hat{T}_{T,FHS}) + V(\hat{T}_{T,VTR}).$$

6.1.3.2 Annual Estimates

The annual for-hire effort estimation methods are analogous to the steps in the wave estimation where the FHS and VTR estimates are produced and then summed. The only key difference in the calculations is that the FHS and VTR estimates are adjusted using the annual

reporting error correction factor (ϕ_A, described in section 5.4.2) instead of the wave reporting error (ϕ_w, described in section 5.4.2).

First, the FHS annual total effort estimate is calculated as the sum of the wave estimates. However, these wave estimates used to produce annual estimates are slightly different from those described above in section 6.1.3.1, because the annual reporting error (ϕ_A) is used to re-estimate them prior to producing the annual FHS effort estimate. These wave estimates ($\hat{T}_{T,FHS^*}$) are produced by first re-estimating the initial for-hire wave effort ($\hat{T}_{I,FHS^*}$) using the annual reporting error correction factor (ϕ_A)

$$\hat{T}_{I,FHS^*} = \hat{T}\phi_A.$$

The variance of this re-estimated, initial wave effort estimate is estimated using Goodman's formula

$$V(\hat{T}_{I,FHS^*}) = \hat{T}^2 V(\phi_A) + \phi_A{}^2 V(\hat{T}) - V(\hat{T})V(\phi_A).$$

Then the coverage adjustment ($\hat{R}_m$) is applied to the re-estimated initial wave estimate of for-hire effort to obtain the re-estimated wave total effort estimates ($\hat{T}_{T,FHS^*}$)

$$\hat{T}_{T,FHS^*} = \hat{T}_{I,FHS^*}\hat{R}_{mi}.$$

The variance of the re-estimated wave total effort estimates are again estimated using Goodman's formula

$$V(\hat{T}_{T,FHS^*}) = \hat{T}_{I,FHS^*}{}^2 V(\hat{R}_m) + \hat{R}_m{}^2 V(\hat{T}_{I,FHS^*}) - V(\hat{T}_{I,FHS^*})V(\hat{R}_m).$$

The FHS annual total effort is then simply estimated as the sum of the re-estimated wave total effort

$$\hat{T}_{Annual,FHS} = \sum_{w=1} \hat{T}_{T,FHS^*},$$

and its variance is estimated as

$$V(\hat{T}_{Annual,FHS}) = \sum_{w=1} V(\hat{T}_{T,FHS^*}).$$

The annual VTR total effort ($\hat{T}_{Annual,VTR}$) is produced in the same manner as the VTR wave total effort estimate, where the number of angler-trips for each state in a given year is simply the sum of the angler-trips for individual permitted vessels. Again, the only difference in producing this estimate is that the wave reporting error correction factor (ϕ_w) is replaced by the annual reporting error correction factor (ϕ_A)

$$\hat{T}_{Annual,VTR} = \phi_A \sum_{i=1}^{n_h} t_{hai}$$

where t_{hai} denotes the number of angler-trips aboard vessel i, in fishing area a and in sampling period h.

The variance for $\hat{T}_{Annual,VTR}$ is estimated using Goodman's formula

$$V(\hat{T}_{Annual,VTR}) = {\phi_A}^2 V\left(\sum_{i=1}^{n_h} t_{hai}\right) + \left(\sum_{i=1}^{n_h} t_{hai}\right)^2 V(\phi_A) - V(\phi_A)V\left(\sum_{i=1}^{n_h} t_{hai}\right),$$

which can be simplified to

$$V(\hat{T}_{Annual,VTR}) = {\phi_A}^2 * 0 + \left(\sum_{i=1}^{n_h} t_{hai}\right)^2 * V(\phi_A) - V(\phi_A) * 0 = \left(\sum_{i=1}^{n_h} t_{hai}\right)^2 V(\phi_A).$$

Then, the total annual for-hire effort ($\hat{T}_{Annual,FHS+VTR}$, combining the FHS and VTR estimates) is simply

$$\hat{T}_{Annual,FHS+VTR} = \hat{T}_{Annual,FHS} + \hat{T}_{Annual,VTR},$$

and the variance is

$$V(\hat{T}_{Annual,FHS+VTR}) = V(\hat{T}_{Annual,FHS}) + V(\hat{T}_{Annual,VTR}).$$

6.2. Total Catch Estimation

6.2.1 Standard Total Catch Estimation

Total catch per wave is estimated within domains defined by sub-region, state, sub-state region, two-month wave, fishing mode, area fished, species, and catch type. The basic calculation multiplies the catch rates from the APAIS survey by the wave total effort estimates from the FES (the CHTS in prior years) or the FHS+VTR. The standard, total catch per wave is therefore calculated as

$$\hat{Y}_d = \hat{\bar{y}}_d \times \hat{T}_{T\ldots}$$

where $\hat{\bar{y}}_d$ is the APAIS catch rate estimate for the wave (which can be expressed as either the number of individual fish or as a weight in pounds, since both of those data are collected, described in section 2.3.2), and

$\hat{T}_{T\ldots}$ is the wave total effort estimate derived from the FES (the CHTS previously) or the FHS+VTR

($\hat{T}_{T,FES}$, $\hat{T}_{T,CHTS}$, or $\hat{T}_{T,FHS+VTR}$, described in section 6.1).

The variance is estimated using Goodman's formula as

$$V(\hat{Y}_d) = \hat{\bar{y}}_d^{\,2} V(\hat{T}_{T\ldots}) + \hat{T}_{T\ldots}^{\,2} V(\hat{\bar{y}}_d) - V(\hat{T}_{T\ldots}) V(\hat{\bar{y}}_d).$$

The standard annual total catch ($\hat{Y}_D$) and its variance are then estimated as

$$\hat{Y}_D = \sum_{w=1} \hat{Y}_{dl}$$

and

$$V(\hat{Y}_D) = \sum_{w=1} V(\hat{Y}_{dl})$$

where D indicates the annual domain (defined in the same manner as domain d excluding wave), and
l represents the wave index (1 through 6).

6.2.2 Alternative Total Catch Estimation for Small Sample Sizes

Alternative estimation approaches are taken when certain domain estimates (e.g. sub-region, state, sub-state region, mode, and area fished) have low sample sizes. The current threshold for switching from standard to alternative small sample size estimation is when any combination of sub-region, state, sub-state region, and fishing mode has fewer than 100 intercepted angler-trips. These domain estimates are only calculated at the annual level, and no wave estimates are produced. To date, this approach has only been used for charter boat mode for a small number of state-year combinations in the mid-Atlantic and New England regions.

For annual estimates, rather than producing total catch estimates for each two-month wave and then summing the catch estimates across the waves, the alternative approach produces estimates at the annual level directly by calculating catch rates within an annual domain, and multiplying them by the corresponding annual total effort. This approach makes better use of APAIS data when it is limited by estimating fewer catch rates and increasing the number of data points in those catch rates (i.e. producing only one annual domain estimate instead of 5-6 wave domain estimates) which improves the precision of the annual estimate.

The annual level estimation domains are the same as the standard wave estimation domains except wave is of course excluded (defined by year, sub-region, state, sub-state region, fishing mode, area fished, and species). Catch rates are produced from the APAIS data in the same manner as in section 2.3.2 but with two key differences; the two-month wave domains (d) are replaced by annual domains (D'), and the original design strata *H*, are collapsed into pseudo-strata, *H'*, such that there are data for at least two PSUs in every pseudo-stratum. The extent of collapsing across design strata is case-specific, varying with the amount and distribution of available APAIS data. However, a standard process is used that always collapses across strata in the same order: 1) time interval, 2) kind-of-day, 3) month, 4) sub-state region, 5) two waves, 6) three waves, 7) full year. This catch rate estimate and its variance are produced as follows:

$$\hat{\bar{y}}_{D'} = \frac{\sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} I_{D'(h,i,j)} y_{hij}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}\, I_{D'(h,i,j)}}$$

and

$$V(\hat{\bar{y}}_{D'}) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\frac{\left(\sum_{j=1}^{m_{hi}} w_{hij} I_{D'(h,i,j)} (y_{hij} - \hat{\bar{y}}_{D'})\right)}{\sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} I_{D'(h,i,j)}} - \frac{\left(\sum_{i=1}^{n_h} \frac{\left(\sum_{j=1}^{m_{hi}} w_{hij} I_{D'(h,i,j)} (y_{hij} - \hat{\bar{y}}_{D'})\right)}{w_{hij} I_{D'(h,i,j)}} \right)}{n_h} \right)^2$$

where $\hat{\bar{y}}_{D'}$ is the estimated mean catch per angler-trip in domain D', defined by year, sub-region, state, sub-state region, fishing mode, area fished and species;
$h = 1, \ldots H'$ represents the pseudo-stratum index;
$i = 1, \ldots n_h$ represents the PSUs sampled within stratum h;
$j = 1, \ldots m_{hi}$ represents the secondary stage units (angler-trips) sampled in PSU i;
w_{hij} is the final APAIS sample weight described in section 2.3.1;
$I_{D'(h,i,j)}$ is an indicator variable, which equals 1 if (h, i, j) is in domain D' or 0 if otherwise; and
y_{hij} is the number of fish caught on angler-trip j.

Annual total catch is then estimated in a similar manner to the standard methods, where the catch rate is multiplied by the total effort estimate. Here, the annual catch rate ($\hat{\bar{y}}_{D'}$) is multiplied by the annual effort estimate ($\hat{T}_{Annual,FES}$ for private boat and shore mode as described in section 6.1.1; $\hat{T}_{Annual,CHTS}$ for private boat and shore mode pre-2018 as described in section 6.1.2; or $\hat{T}_{Annual,FHS+VTR}$ for for-hire modes as described in section 6.1.3) to obtain the total annual catch estimate ($\hat{Y}_{D'}$):

$$\hat{Y}_{D'} = \hat{\bar{y}}_{D'} \hat{T}_{Annual\ldots}.$$

The variance of is $\hat{Y}_{D'}$ is then estimated using Goodman's formula:

$$\mathrm{V}(\hat{Y}_{D'}) = \hat{T}_{Annual\ldots}{}^2 \mathrm{V}(\hat{\bar{y}}_{D'}) + \hat{\bar{y}}_{D'}{}^2 \mathrm{V}(\hat{T}_{Annual\ldots}) - \mathrm{V}(\hat{\bar{y}}_{D'}) \mathrm{V}(\hat{T}_{Annual\ldots}).$$

6.2.3 Total Catch Estimation for Public Use Datasets

MRIP produces three survey datasets for public use every wave: 1) 'catch,' which includes catch-level data (one record per species for every angler-trip interview) and the variables needed for total catch estimation, 2) 'trip,' which includes trip-level data (one record per angler-trip interview, identified by the variable 'id_code') and the variables required for total effort estimation, and 3) 'size,' which includes fish-length and weight data (one record per fish caught and measured/weighed by an APAIS interviewer) and the variables required for length-frequency (see section 6.3) estimation. External data users can download these files at the MRIP

Data downloads webpage (https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-data-downloads), along with a template R or SAS program to produce their own custom domain estimates.

When using the public use datasets, total catch estimation is consolidated and simplified compared to the standard total catch estimation described in 6.2.1. Rather than calculating weighted catch rates, and then multiplying catch rates by total effort to obtain total catch, total catch is calculated by first rescaling the final APAIS sample weights (described in 2.3.1), using the total effort estimates (described in 6.1), and then calculating total catch as a weighed sum using the new weights. This calculation is essentially an algebraic rearrangement that allows for catch estimates to be produced directly from the public use datasets in a single step, which simplifies the process for external data users since fewer steps and variables are required for estimation.

This algebraic rearrangement can be demonstrated starting with the standard total catch calculation in 6.2.1 ($\hat{Y}_d = \hat{\bar{y}}_d \times \hat{T}_{T\ldots}$) but substituting $\hat{\bar{y}}_d$ for the full catch rate calculation (which is the weighted mean estimator described in 2.3.2):

$$\hat{Y}_d = \frac{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij} I_{d(h,i,j)} y_{hij}}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij} I_{d(h,i,j)}} \times \hat{T}_{T\ldots}$$

which, for readability, can be simplified to

$$\hat{Y}_d = \frac{\sum_d w_d y_d}{\sum_d w_d} \times \hat{T}_{T\ldots}$$

and rearranged as

$$\hat{Y}_d = \frac{\hat{T}_{T\ldots}}{\sum_d w_d} \times \sum_d w_d y_d = \sum_d \left(\frac{\hat{T}_{T\ldots}}{\sum_d w_d}\right) w_d y_d = \sum_d {w_d}^* y_d$$

where ${w_d}^*$ is simply an individual APAIS sample weight within domain *d*, post-stratified (or rescaled) by the total effort divided by the sum of the APAIS sample weights in domain *d*. Conceptually, ${w_d}^*$ can be thought of as the number of angler-trips being represented by an individual angler-trip record. These ${w_d}^*$ analysis weights are provided in the 'wp_catch' field in the catch datasets and the 'wp_int' field in the trip datasets.

Typically, the wp_catch and wp_int values will be equal for the same angler-trip in the catch and trip datasets. However, there can be differences. Most of the differences result from cases where the alternative total catch estimation (detailed in section 6.2.2) is used. These cases are listed on the MRIP Estimates Updates page for 2/21/2017:
https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-estimate-updates
A smaller number of differences are due to applications of the APAIS and FES calibrations (see section 8) to the historical times series of catch and effort estimates.

6.3. Length-Frequencies

Length frequencies are estimated using the 'size' public-use dataset. These frequencies are estimated for each species as domain estimates (e.g. by year, wave, fishing mode, geographic area, and fishing area of interest to the user). They apply to landed fish only (Type A and B1 catch) rather than total catch since length data cannot be collected for discards (Type B2 catch).

Prior to estimating length frequencies, a fourth APAIS sample weighting component is needed to reflect subsampling of catch. The first three weights, described in section 2.3.1, are the stage I weight, the stage II weight and the stage III weight, which are applied to the APAIS catch rate estimates. The stage IV weight (catch weight) is required for estimating length frequencies whenever the number of lengths measured for a particular species on a given angler-trip does not equal the number of landed fish of that particular species. This weight (w_{hijl}) is simply the inverse of the sampling fraction of the landed fish:

$$w_l = \frac{Y_l}{y_l}$$

where Y_{hijl} is the number of fish landed of species l on angler-trip j, in site-cluster-day-time interval i and stratum h; and

y_{hijl} is the number of fish measured for length from angler-trip j in site-cluster-day-time interval i and stratum h.

The final APAIS sample weight is then simply the product of the stage IV weight and the stages I-III weights:

$$w_{hijl} = w_{hijl} * w_{hij}$$

The length-frequencies are then estimated as follows:

$$\hat{Y}_{Dzl} = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} {w_{hijl}}^* I_{D(h,i,j,l,z)} y_{hij}$$

where $\hat{Y}_{Dzl}$ is the number of individuals of species l in length bin z of domain D;

$h = 1, \ldots H$ represents the strata, each of which is defined by year, wave, month, kind-of-day, time interval, sub-region, state, sub-state region, and fishing mode site-group;

$i = 1, \ldots n_h$ represents the PSUs, sampled within stratum h;

$j = 1, \ldots m_{hi}$ represents the secondary stage units (angler-trips) sampled in PSU i;

y_{hij} is the number of fish caught on angler-trip j;

${w_{hijl}}^* = w_{hijl}\hat{Y}_d$, which is the final APAIS sample weight post-stratified by the total landings estimate ($\hat{Y}_d$, estimated in the same manner as total catch in 6.2.1 but only including Type A+B1 catch); and

$I_{D(h,i,j,l,z)}$ is an indicator variable equal to 1 if an individual of species l is in length bin z and domain h, i, j and 0 otherwise.

The weight adjustment component of the above calculation ($w_{hijl}{}^{*}$) can be illustrated through algebraic rearrangement of the standard total catch estimation equation as follows:

$$\hat{Y}_{Dzl} = \hat{\bar{y}}_{Dzl} \times \hat{T}_{T\ldots}$$

$$\hat{Y}_{Dzl} = \left(\frac{\sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hijl} I_{D(h,i,j,l,z)} y_{hij}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hijl}\, I_{D(h,i,j,l,z)}} \right) \hat{T}_{T\ldots}$$

$$\hat{Y}_{Dzl} = \left(\frac{\hat{T}_{T\ldots}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hijl}\, I_{D(h,i,j,l,z)}} \right) \sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hijl} I_{D(h,i,j,l,z)} y_{hij}$$

$$\hat{Y}_{Dzl} = \hat{T}_{T\ldots}{}^{*} \sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hijl} I_{D(h,i,j,l,z)} y_{hij}$$

$$\hat{Y}_{Dzl} = \sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} \left(\hat{T}_{T\ldots}{}^{*} \right) w_{hijl} I_{D(h,i,j,l,z)} y_{hij}$$

$$\hat{Y}_{Dzl} = \sum_{h=1}^{H'} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hijl}{}^{*} I_{D(h,i,j,l,z)} y_{hij}$$

where $\hat{T}_{T\ldots}{}^{*}$ is the post-stratification adjustment (total estimated effort $\hat{T}_{T\ldots}$ divided by the sum of the APAIS sample weights in domain *D*. Again, the total effort used here is $\hat{T}_{T,FES}$ for private boat and shore modes from section 6.1.1; $\hat{T}_{T,CHTS}$ from section 6.1.2 for pre-2018 estimates of private and shore mode effort; or $\hat{T}_{T,FHS+VTR}$ from section 6.1.3 for for-hire modes).

6.4 MRIP Estimation Certification Status

Prior to the full redesign, the MRIP weighted estimation methods were certified in March 2011. The weighted estimation used for the APAIS catch-rates were the specific focus of the 2011 certification. Since the current APAIS was not being used to produce estimates until 2013, interim weighted estimation methods were desired to reflect the complex, probability-sampling design of the MRFSS Intercept Survey. These certified methods were therefore initially used to re-estimate the MRFSS estimates in a manner that was design-unbiased, and then were applied to the APAIS after it began. The traditional MRFSS Intercept Survey estimation methods relied on an unweighted average to produce catch rates, and contained data that were not obtained through a probability sample.

The weighted catch-rate component estimation methods underwent peer-review by three independent sources. After addressing peer reviewer comments (most of which referred to the mix of model-based and design-based methodologies that were necessary for re-estimating historical MRFSS data), the Operations Team and the Executive Steering Committee recommended them for certification in February 2011. The final certification decision was made by NMFS Leadership shortly thereafter.

7. Large Pelagics Survey (LPS): Large Pelagics and HMS Catch and Fishing Effort

The Large Pelagics Survey (LPS) collects fishing effort and catch data for the hand-gear fishery targeting large pelagics and HMS, including tunas, billfishes, swordfish, sharks, dolphins and others in offshore marine waters from Maine through Virginia. Since large pelagic and HMS fishing trips are specialized, targeting specific species using distinct fishing methods, they often fall outside of the sample frames of the APAIS, the FES (the CHTS in prior years) and the FHS. The LPS was designed to fill this gap in the other surveys' coverage of these trips.

The LPS consists of three complementary surveys: the Large Pelagics Telephone Survey (LPTS), the Large Pelagics Intercept Survey (LPIS), and the Large Pelagics Biological Survey (LPBS.) The LPTS is a telephone survey for collecting fishing effort, defined as the total number of vessel-trips where anglers fished with hand-gear (i.e. rod and reel or handline), for large pelagic species. The LPIS, similar to the APAIS, is a dockside intercept survey where detailed trip and catch data are collected from boat captains that have just completed a fishing trip targeting large pelagics and HMS. The LPBS collects additional length and weight data as well as various biological samples to obtain age and life history information about the catch.

The survey is typically conducted from June-October when the majority of the large pelagic and HMS fishing activity occurs. Occasionally the LPS sampling period is extended for an extra two months on either end of the season (May-November) in order to confirm that the current temporal coverage is not missing significant off-season fishing activity. While estimates for all other MRIP surveys are produced every two months, LPS estimates are produced monthly. This difference allows MRIP to provide frequent landings updates to HMS managers for better tracking and monitoring of recreational quotas in-season. In addition, the monthly estimation of HMS catch abides by the International Commission for the Conservation of Atlantic Tunas 2010 measure, which requires monthly reporting of bluefin tuna landings (ICCAT, 2010). At the end of each HMS fishing season, annual catch and effort estimates are produced in a similar manner to the other surveys.

7.1 Large Pelagics Telephone Survey (LPTS): Large Pelagics and HMS Fishing Effort

The Large Pelagics Telephone Survey (LPTS) collects data that are used to estimate the total number of trips in which anglers fished from private or charter boats for large pelagics and HMS using hand-gear. The survey targets vessels with NOAA HMS fishing permits. HMS charter vessels and headboats are covered by the FHS, because the FHS sample frame (described in section 5.1.2) is sorted by HMS and non-HMS permits. During the months that the LPTS operates, the for-hire vessel representatives selected for FHS sampling are asked screening questions (e.g. if they targeted HMS on their recent trips) and a series of LPTS add-on questions in order to collect charter boat information for the LPS. The LPTS-private, conducted as a separate survey, captures effort from private boats targeting HMS.

7.1.1 LPTS-Private Sampling Design

The LPTS-private is stratified geographically by state (although there are three, two-state areas: Maryland and Delaware, Connecticut and Rhode Island, and New Hampshire and Maine) and temporally in two-week reporting periods. The sample frame is created from a comprehensive directory of vessels with the following NOAA issued permits: an Atlantic Tunas General Permit, a Swordfish General Permit, a Tuna/Swordfish General Combination Permit, and an HMS Angling Category Permit. The vessel permit database is updated continuously to maintain accurate records. To be included in the sample frame, vessels in the vessel directory must have the following information: the name, address and telephone number of a vessel representative; the state in which the vessel operates; and a vessel name or identification number.

The sample frame is compiled monthly and sample selection occurs every two weeks. The sample is selected using a stratified random sampling design without replacement. LPTS-private sample sizes vary and are determined by weighing gains in precision against added costs and sampling effort prior to each sample selection.

7.1.2 Data Collection Design

Telephone interviewing for each two-week reporting period is conducted during the seven-day period (Monday-Sunday) immediately following that reporting period. Like the FHS and legacy CHTS, interviewing is done using a CATI system similar to what is described in section 4.2.

One week prior to the two-week reporting period, all vessel representatives receive a letter by mail, notifying them that they have been selected for participation in the survey. The letter includes the dates for which the vessel has been selected to report, as well as the date(s) when calling will be attempted.

A maximum of 10 attempts are made to contact each selected vessel representative. First attempts are made on the Monday immediately following the two-week reporting period, and repeat attempts are distributed among weekend and weekdays, as well as days and evenings. The dialing pattern for each vessel representative includes at least one day attempt (before 5:00PM local time for the area being sampled) and three night attempts (after 5:00PM local time for the area being sampled).

7.1.3 LPTS Charter Boat Sampling General Protocols (Add-on to the FHS)

The HMS charter boats are sampled as part of the FHS. During the LPS sampling months, interviews conducted as part of the FHS ask additional questions related to large pelagic and HMS fishing. Vessel representatives are asked whether or not they have an HMS permit, and whether or not they targeted large pelagics and HMS during the reporting period. If the answer to either question is ‘yes,’ subsequent LPS-specific questions are asked. The LPS-relevant data collected in these interviews are: the number of vessel-trips taken targeting large

pelagics and HMS; the state where each trip took place; the state to which the boat returned; the date and duration of each trip; the fishing gear used; and species targeted.

7.1.4 LPTS Imputation

Missing data is imputed when trips are reported but dates are not given (i.e. either refused or exact date could not be recalled by the vessel representative interviewed). The procedures are primarily logic rule-based because most sampling periods do not cross months. Therefore, when a vessel representative does not report the month of their fishing trips, the month of the sampling period can simply be filled in. When sampling periods cross two months (i.e. the sampling period spans the end of one month and the beginning of another), an algorithm is applied based on how many days of each month are in the sampling period to determine the probability that the trip was taken in either month.

7.1.5 LPTS Estimation Methods

LPTS effort is a calculated as a domain estimate of the number of vessel-trips targeting large pelagic species in a given month, year, state and fishing mode (charter boat or private boat). However, LPTS samples are collected by state, fishing mode and sampling period, which is two weeks for private boat mode, and one week for charter boat modes where FHS sampling methods are used. Therefore, the mean effort is first estimated for the vessels sampled in each sampling period. It is then expanded to the full month, and by the number of vessels in the sample frame, to obtain the domain estimate of the number of in-frame vessel-trips ($\hat{b}_d$)

$$\hat{b}_d = N\alpha\left(\frac{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} b_{hij} I_{d(hij)}}{n_h}\right) = N\alpha\hat{\bar{b}}_{hij}$$

where $\hat{b}_d$ is the total number of vessel-trips taken in domain d,

$h = 1, \ldots H$ represents the strata, defined by year, sample week, state and fishing mode

$i = 1, \ldots n_h$ represents the vessels sampled within stratum h,

$j = 1, \ldots m_{hi}$ represents the vessel-trips taken by vessel i in stratum h,

b_{hij} is the number of vessel-trips taken by vessel i in stratum h.

n_h is the sample size of stratum h,

N is the total number of vessels in the sample frame, and

α is an expansion factor which is the total days in the month divided by the total days in the sample period (for charter boat mode, α=(30 or 31)/7, and for private boat mode, α=(30 or 31)/14.

The variance of $\hat{b}_d$ is estimated as

$$\hat{V}(\hat{b}_d) = \frac{(\alpha N)^2}{(n_h - 1)}\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}}\left(b_{hij} I_{d(hij)} - \hat{\bar{b}}_{hij}\right)^2.$$

7.2 Large Pelagics Intercept Survey (LPIS): Large Pelagics and HMS Catch Rates

The LPIS collects data through dockside interviews with private and charter boat operators who are intercepted after completing fishing trips targeting large pelagics and HMS. Before 2010, incidental catch of large pelagics and HMS were counted as part of LPIS data. However, the LPTS data collection only includes vessel-trips targeting large pelagics and HMS in the sample frame, so to remove the discrepancy in vessel-trip definitions between the two surveys, LPIS stopped sampling incidental vessel-trips in 2010. The survey uses a comprehensive list of fishing access sites, clustered based on estimated fishing pressure, to create the sample frame. The sample is selected using probability proportional to size sampling without replacement. The list of fishing access sites is updated at least monthly based on new information obtained through site visits and interviews.

While the LPIS sampling design is similar to that of the APAIS, it differs in several important ways. First, the LPIS operates at a smaller geographic scale along the northeast Atlantic coast (Maine through Virginia). Additionally, sampling only occurs during peak season (June-October) rather than year-round like the APAIS. Another difference is that only two fishing modes are sampled, private and charter boat modes, and only if large pelagics and HMS were the target of the fishing trip. The LPIS also intercepts vessel-trips, where a single vessel representative is interviewed to collect information about the trip, rather than intercepting individual angler-trips like the APAIS.

The LPIS collects a variety of information about each vessel-trip. The data includes the number of fish caught, landed, released, and sold by species, as well as opportunistic length measurements of landings that are made available to be inspected by an LPIS sampler. For these length measurements, there are no minimum requirements; bluefin tuna are prioritized, but other species are measured as well to be representative of the species diversity of landings. The LPIS further collects the respondent's name, permit status, target species for the vessel-trip, number of anglers that fished, number of lines in the water, hours fished, types of gear, bait and fishing method used, location of fishing, the water depth of their fishing location(s) and temperature of their fishing location(s). Vessel identification information (i.e. HMS Permit Number, State Registration or Coast Guard Documentation Number), time of the vessel's return from its fishing trip, principal port state, and tournament participation information are also collected.

7.2.1 LPIS Sampling Design

7.2.1.1 Sample Frame: LPIS Mater Site Register

The LPIS sample frame is created from the LPIS Master Site Register (MSR), which is a list of fishing access sites to which boats return from fishing trips targeting large pelagics and HMS. The LPIS MSR is very similar to the APAIS public access fishing site register (described in section 2.1.3) including all the same site codes and information (e.g. contact information for the site's supervisor, a street address, the nearest town or city, and latitude/longitude coordinates). The majority of LPIS sites are public access and can overlap with the sites in the APAIS public access fishing site register. However, some LPIS sites are also private access (e.g.

gated marinas or docks) where the LPIS has been granted permission to sample from the site owner.

Another important distinction between the LPIS MSR and the APAIS public access fishing site register is how fishing pressure is estimated at a site. While in the APAIS, sites are grouped using ordinal fishing pressure categories (based on expected fishing pressure decided by field staff), the fishing pressures used to group sites in the LPIS are continuous values. The LPIS uses average historical sampling productivity (mean interviews obtained per assignment) by mode, month and kind-of-day, as a proxy for fishing pressure, whereas the APAIS uses a categorical scale based on expected numbers of angler-trips per day by mode, month, and kind-of-day. For any new sites added to the LPIS MSR that do not yet have historical data, fishing pressure is estimated based on the productivity of similar sites, or based on site visits made by field supervisors.

Similar to the APAIS site register, the LPIS MSR is regularly updated using historical data as well as inputs from LPIS samplers. All sites in the register are visited twice per year by a field supervisor to update its information in the MSR, regardless of whether the site was selected for sampling. Also like APAIS sites, LPIS sites are never removed from the site register; when a site becomes inactive in all fishing modes, it is coded as 'retired' and is simply excluded from the sample frame. Additionally, hostile sites, or those sites where sampling is unwelcome, are excluded from the sample frame but are always kept in the MSR and periodically assessed for fishing pressure and cooperation. Keeping sites in the register allows for reactivation of the site in the future if fishing activity resumes, or if a hostile site eventually allows sampling to resume.

7.2.1.2 Stratification, Stages of Design, Clustering and Sample Selection

The LPIS is stratified temporally, geographically, and by fishing mode. Temporally, there are three strata: 1) year, 2) month and 3) kind-of-day (which separates weekdays from weekends/holidays). Geographical stratification is by state, although there are three, two-state areas: Maryland and Delaware, Connecticut and Rhode Island, and New Hampshire and Maine. Fishing mode in the LPIS is limited to either private boat or charter boat modes.

Sites are clustered based on fishing pressure to increase sampling efficiency. Individual sites in the LPIS MSR that have low fishing pressure are grouped together to form site clusters, while sites with high fishing pressure remain unclustered. Factors taken into account at each site when forming site clusters are 1) relative fishing pressures, 2) the proximity of the site to others, and 3) historical sampling productivity. Unlike the APAIS where sites within a cluster often vary from month to month, site clusters in the LPIS are generally fixed.

Sampling is conducted in two stages. Site cluster-day is the PSU, which is a selected fishing site or site cluster combined with a selected day (weekday or weekend day/holiday). The SSU is a vessel-trip.

Within strata, a sample of site cluster-days is selected from a sample frame consisting of all possible combinations of site-days by a probability proportional to size without replacement

sampling approach, where the size measure for a given site cluster-day is the historical average number of vessel intercepts obtained per assignment. This approach ensures a relatively high level of sampling productivity because the high fishing pressure site-days will be selected for sampling more frequently than the low fishing pressure site-days.

Tournaments are occasionally hosted at various sites over the course of the LPIS sampling months, and are included in the LPIS sample frame. They often occur at the same site in the same month each year, and the expected increase in activity associated with tournaments is incorporated into the sample selection. Each month when there is a tournament scheduled at a given site, the fishing pressure is estimated to be higher than non-tournament months to reflect the anticipated increase in activity. This estimated increase in fishing pressure is specific to the kind-of-day in which the tournament is occurring. For example, if the tournament is occurring on a weekend/holiday, fishing pressure is increased only for weekend/holidays, and not weekdays, in that month. This change in size measure for tournament sites increases the probability that a tournament site will be included in the selected sample.

Target sample sizes are largely decided by available budget and samplers, but also vary year to year due to the relative unpredictability of the large pelagics and HMS fishery; sample sizes are either decided at the beginning of each season with targets for all sampling months, or are decided one or two months at a time. The target number of interviews can also vary based on changes in the fishery, management priorities, or statistical considerations.

7.2.2 Data Collection Design

Each assignment specifies a site cluster (or a single high fishing pressure site), a date and a fishing mode (private boat or charter boat). The assigned fishing mode is used primarily during the assignment draw process to improve the likelihood of reaching interview targets in each mode. During an individual assignment, however, samplers are allowed to opportunistically interview vessel representatives of either mode. The assigned mode is prioritized only when the sampler cannot intercept all of the vessels arriving at a site during an assignment. Assignments typically last between two and eight hours at one or more sites in the site cluster on an assigned date. The duration of assignments varies, as survey operations staff uses best judgement based on prior experience, knowledge, and local logistics.

7.2.3 LPIS Estimation Methods

The LPIS catch rate estimates are produced similarly to the APAIS catch rates in section 2.3.2, with two important differences. First, LPIS catch rates are estimated as mean catch per vessel-trip, not angler-trip. Second, the mean estimator used for the LPIS is unweighted. This mean catch per vessel-trip ($\hat{\bar{y}}_d$) is produced as a domain estimate, defined by year, month, sub-region, state, mode, species of fish, and catch disposition. Primary catch dispositions for the LPIS include kept (harvested/landed), alive (released alive), and dead (discarded at sea dead). The LPIS catch rate is estimated as:

$$\hat{\bar{y}}_d = \frac{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} I_{d(h,i,j)} y_{hij}}{\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} I_{d(h,i,j)}}$$

where $\hat{\bar{y}}_d$ is the estimated mean catch per vessel-trip in domain d,

$h = 1, \ldots H$ represents the strata, each of which is defined by state, year, month, kind-of-day and fishing mode (private boat or charter boat)

$i = 1, \ldots n_h$ represents the PSUs (site cluster-days), sampled within stratum h,

$j = 1, \ldots m_{hi}$ represents the secondary stage units (vessel-trips) sampled in site cluster-day i,

$I_{d(h,i,j)}$ is an indicator variable, which equals 1 if (h, i, j) is in domain d or 0 if otherwise, and

y_{hij} is the number of fish caught on vessel-trip j.

The variance is then estimated using Taylor series linearization (Dienes, 1957) as

$$V_h(\hat{\bar{y}}_d) = \sum_{h=1}^{H}\left(\frac{n_h}{n_h - 1}\sum_{i=1}^{n_h}\left(\frac{\left(\sum_{j=1}^{m_{hi}} I_{d(h,i,j)}(y_{hij} - \hat{\bar{y}}_d)\right)}{I_{d(h,i,j)}} - \frac{\left(\sum_{i=1}^{n_h}\frac{\left(\sum_{j=1}^{m_{hi}} I_{d(h,i,j)}(y_{hij} - \hat{\bar{y}}_d)\right)}{I_{d(h,i,j)}}\right)}{n_h}\right)^2\right).$$

7.3 LPS Total Effort and Catch Estimation

7.3.1 LPS Total Effort

For total effort targeting large pelagics and HMS, the base effort estimate is adjusted for out-of-frame vessel-trips, calculated from LPIS sampled vessels, much like the APAIS calculation for the FES, CHTS and FHS coverage adjustments (sections 2.3.3, 2.3.4 and 2.3.5). LPIS intercepts are used to identify any vessels falling outside of the LPTS sample frame to adjust the effort estimate accordingly. Intercepted trips are designated as in-frame if the vessel is listed in the LPTS vessel directory for a given state. All vessels not listed in the LPTS vessel directory are designated as out-of-frame.

The coverage adjustment, which is the ratio of total vessel-trips to in-frame vessel-trips ($\hat{R}_i$), is calculated by state, month, and mode (private boat and charter boat) using the following equation:

$$\hat{R}_i = \sum_{i=1}^{I}\frac{X_i}{I_i X_i},$$

where X_i is the total number of vessel-trips of the i-th sampled vessel and

I_i is an indicator variable that equals 1 if vessel i is in-frame or 0 if otherwise.

The monthly total LPS effort ($\hat{b}_T$) is estimated by simply multiplying the base effort ($\hat{b}_d$) by the coverage adjustment ($\hat{R}_i$)

$$\hat{b}_T = \hat{b}_d \hat{R}_i.$$

The variance is then estimated as

$$V(\hat{b}_T) = \hat{b}_d{}^2 V(\hat{R}_m) + \hat{R}_m{}^2 V(\hat{b}_d) - V(\hat{b}_d)V(\hat{R}_m) \qquad \text{(Goodman, 1960).}$$

To obtain an annual total effort estimate at the end of the fishing season, the monthly estimates (for months $w = 1 \ldots W$) are summed

$$\hat{b}_{T,Annual} = \sum_{w=1}^{W} \hat{b}_T$$

and the variance is calculated as

$$V(\hat{b}_{T,Annual}) = \sum_{w=1}^{W} V(\hat{b}_T).$$

7.3.2 LPS Total Catch

Total catch is estimated by multiplying the catch rates from the LPIS by the total LPS effort estimates. Total catch estimates are expressed in numbers of fish. Estimates are produced only for a fixed group of large pelagics and HMS that are of high management importance.

The monthly total HMS catch is estimated as

$$\hat{y}_T = \hat{\bar{y}}_k \hat{b}_T$$

and its variance is estimated as

$$V(\hat{y}_T) = \hat{\bar{y}}_k{}^2 V(\hat{b}_T) + \hat{b}_T{}^2 V(\hat{\bar{y}}_k) - V(\hat{\bar{y}}_k)V(\hat{b}_T) \qquad \text{(Goodman, 1960).}$$

The annual total catch and variance estimates are computed as a sum of the monthly total catch estimates (for months $w = 1 \ldots W$),

$$\hat{y}_{T,Annual} = \sum_{w=1}^{W} \hat{y}_T$$

and

$$V\left(\hat{y}_{T,Annual}\right) = \sum_{w=1}^{W} V(\hat{y}_T).$$

7.4 Large Pelagics Biological Survey (LPBS)

The LPBS collects additional length, weight and other biological information for Atlantic bluefin tuna as well as other priority HMS as specified by NMFS. In addition to collecting length and weight data, the biological sex of each fish is recorded, and otoliths, white muscle tissue, first dorsal spines, and gonads are harvested. Further information including the date, site, vessel name, and vessel type are also recorded as part of the LPBS sampling. While similar to the LPIS, where vessel representatives who have just completed a fishing trip are intercepted, LPBS is a separate data collection effort. The annual target sample size varies based on available funding, but currently LPBS samplers aim to complete 150 assignments in total per fishing season. While collected as part of MRIP, the LPBS data and collected specimens are not incorporated into the recreational catch estimates. The samples are processed, stored and further used by the NMFS Southeast Fisheries Science Center for assessing stock age and growth rates, and garnering additional life history information on those priority species.

Currently, there is no formal probability sampling design for the LPBS; there are only tournament assignments and opportunistic assignments. Tournament assignments provide opportunities to collect large amounts of biological data. Tournaments are randomly selected over the course of the fishing season, and four hours of sampling is completed at those sites during peak hours. Opportunistic assignments are completed at the discretion of a field supervisor or sampler, based on observed trips or tips from site contacts (e.g. vessel representatives or marina staff) that HMS are available for sampling, or based on local knowledge about new, upcoming tournament events. There is no set duration of sampling during opportunistic assignments, and the field samplers collect whatever data is available.

7.5 LPS Certification Status

The LPS design is not yet certified and will be redesigned in the coming years to meet the standards required for certification. Primary issues with the LPS include deviations from formal probability sampling in LPIS field sampling procedures, and the associated lack of weighted estimation that appropriately accounts for the complex design of the LPIS. As such, there is concern over potential bias in the LPS estimates of catch and effort.

The LPS team is in the early stages of the re-design, currently organizing a working group and developing baseline assessments of the current design. The team plans to complete design proposals for the survey by 2020. Subsequent field-testing will then occur over the following 3 years (2020-2022), and certification will be sought by 2023.

8. Time-Series Calibration Methods

With the transition from MRFSS to MRIP, NMFS recognized the clear need to preserve the integrity of the recreational fishing time series dating back to 1981. However, correcting a long time series of survey data after design changes is a challenging statistical issue. The ideal approach involves a benchmarking period where old and new survey methods are conducted side-by-side, and then fitting a model to relate both sets of estimates. This approach was used for the FES/CHTS calibration, but could not be used for the MRFSS Intercept Survey/APAIS calibration. First, there was no overlap period between the MRFSS Intercept Survey and the APAIS, largely due to funding and logistical constraints. (A 2010 North Carolina pilot study did consist of a side-by-side comparison of the APAIS and the MRFSS Intercept Survey, however, and helped to obtain a general sense of how the design changes might affect estimates; see Breidt *et al.*, 2012). Secondly, the CHTS and the FES produce far fewer series of effort estimates, while the MRFSS Intercept Survey and the APAIS produce a much larger number of estimate series (catch by species, disposition, fishing mode, area fished, state, sub-region, year, and wave). A sample weight adjustment approach was therefore selected for the MRFSS Intercept/APAIS calibration, because it provided a means to accomplish all of these numerous calibrations inside a single framework.

8.1 CHTS/FES Calibration

For a three-year period from January 2015-December 2017, the CHTS and FES were conducted simultaneously prior to discontinuing the CHTS in 2018. This benchmarking period was used to compare the data from both surveys, and revealed that the effort estimates produced from the FES were, on average, several times larger than the CHTS estimates. A number of factors contributed to this large discrepancy. First, research has shown that the difference in survey mode can result in very different responses to survey questions. Mail surveys are self-administered, and FES respondents have the benefit of being able to read the questions and think prior to answering, which can result in improved recall (De Leeuw, 2005; Dillman *et al.*, 2009). Since the CHTS was conducted by phone, respondents had to remember fishing activity on the spot without the added benefit of memory cues. Additionally, the greater coverage and higher response rates (approximately 35% for the FES vs. <10% for the CHTS) likely also contributed to the difference between FES and CHTS estimates. Further, the increase in cell-phone use and prevalence of wireless-only households likely had an effect on the CHTS survey error, whereas the FES is unaffected by those trends.

Using the three-year benchmarking data, a statistical model was developed to calibrate CHTS estimates to FES estimates and vice versa (Breidt *et al.*, 2017). In summary, the model assumes that while both the mail and telephone estimates target the actual, 'true' fishing effort, each survey estimate is distorted by both sampling error and non-sampling error. Sampling error is the error that arises when statistical characteristics of a population are estimated from a sample, or subset, of that population; because a sample can never perfectly represent a full

population, sampling error is always a factor in surveys. However, sampling error is well understood and can be estimated from probability samples (e.g. Assael and Keon, 1982; Groves, Robert, 1990; Cui, 2003). Non-sampling error is due to factors like coverage error (where the target population is not captured fully by the sample frame), non-response error (where some people do not respond to the survey), or measurement error (where the measured value is not the true value, due to either random or systematic error). The FES calibration therefore includes variables that contribute to the 'true' effort common to both surveys, as well as those variables that influence both the sampling and non-sampling error that are unique to each survey.

The methodology follows Fay-Herriot small area estimation, a well-established procedure originally developed to produce model-based estimates of per capita income for small areas in the United States with small populations (Fay and Herriot 1979). The approach uses linear mixed modeling, a class of statistical models that include both fixed and random effects. Fixed effects models are very common in basic statistical modeling and analyses, where the interest lies in estimating one true effect of one or more independent variables on a dependent variable. Random effects models assume a range of potential effects on the dependent variable, and the goal is to estimate not just one true effect, but the mean of a distribution of possible effects. The Fay-Herriot model is therefore essentially a multiple linear regression that includes fixed effects connecting direct estimates to both auxiliary variables (i.e. variables that aid in making estimates on incomplete data) and random effects (that are meant to capture any remaining variation not explained by the auxiliary variables). Sections 8.1.1 and 8.1.2 describe the model variables and the procedures used to calibrate the effort time-series.

The CHTS/FES calibration was completed using both SAS (SAS Institute Inc., 2016). and R (R Development Core Team, 2013). The following series of R packages were called into SAS using the SAS IML Procedure:

- the 'maps' package (Becker *et al.*, 2018), which is used to pull in Federal Information Processing Standard (FIPS) codes and state abbreviations;
- the 'sas7bdat' package (Shotwell, 2015), which allows SAS datasets to be imported into R;
- the 'nlme' package (Pinheiro *et al.*, 2018), which allows the user to fit/compare linear and nonlinear mixed effects models, and is used to model the FES/CHTS design variances; and
- the 'sae' package (Molina and Marhuenda, 2015), which is used for Fay-Herriot estimation.

8.1.1 Auxiliary Variables

8.1.1.1 Covariates – variables that contribute to 'true' effort captured by both surveys

The effort calibration model firstly attempts to capture the spatial and temporal variability of the true effort shared by both surveys. True fishing effort varies from year-to-year, season-to-season, and state-to-state. In addition to these fixed covariates, there are also random covariates that can contribute to changes in true effort in any given state, season, or year (i.e. any distinctive occurrence such as a large storm that disrupts ordinary fishing activity, or a significant change in

fishing regulations). As such, the three main covariates that are shared by both the FES and the CHTS are 1) annual variation 2) seasonal variation, and 3) 'irregular' variation. Irregular variation is a random variable that encompasses any other variation, unrelated to annual and seasonal variation that contributes to changes in effort over time (Table 5).

Annual, seasonal, and irregular variation differs across states. For instance, one state's fishing activity may change more from year-to-year than another state; a southern state may have less seasonal variation in fishing activity than a northern state; or a fishing regulation on one species affects the states within that species' distribution, but obviously has no effect on other states. These potential interactions among variables are taken into account in the model as well.

Proxy variables are used for annual and seasonal variation, and irregular variation is modeled using standard statistical methods for random variables. Annual variation in fishing effort occurs largely due to state-level population changes (e.g. a higher concentration of people living in a coastal state causes fishing effort to increase) so state-level population is used as a proxy variable for annual variation. This state-level population is obtained from the U.S. Census Bureau (U.S. Census Bureau, 2016). Indicator variables for each state and wave are used as a proxy for seasonal variation. Irregular variation is modeled as independent, normally distributed random variables with a mean of zero and a variance that needs to be estimated (since, again, random variables are assumed to represent a broad population of potential values).

Table 5: Summary of shared variation in the FES and the CHTS that contributes to similarity of effort estimates

Variable	Type of Variable	What it is	How it is incorporated into the model
Annual	Fixed	Changes in fishing effort from year to year	Using state-specific population sizes (a primary factor that contributes to annual changes in effort) from the U.S. Census Bureau as an auxiliary variable
Seasonal	Fixed	Changes in fishing effort from season to season (e.g. summer vs winter fishing activity)	Using indicators for the six two-month waves in each state as auxiliary variables
Irregular	Random	Any other variation distinct from seasonal and annual (e.g. changes in fishing regulations or unusual storm patterns that disrupt normal fishing activity in a given season)	As a random variable with a normal distribution, a mean of zero, and unknown variance to be estimated as part of the Fay-Herriot model building procedures.

8.1.1.2 Unshared variables - sampling and non-sampling error unique to each survey

In addition to incorporating shared variation, the effort calibration model incorporates estimates of sampling error for each survey, along with the major differences between the surveys due to non-sampling error. In particular, the model attempts to capture method effects

(i.e. non-sampling error related to coverage, response and measurement error) including the effects of survey mode (mail vs. telephone) and the changing coverage properties of the CHTS due to increases in wireless-only households over time.

The sampling error for each survey is estimated using the variances of the fishing effort estimates. They are assumed independent of one another, since the samples were selected and the surveys were conducted independently. Sampling error is also assumed to be normally distributed; both the CHTS and the FES have moderate to large sample sizes, which, according to classical probability theory, has central-limiting effects (e.g. Le Cam, 1986; Prokhorov and Statulevičius, 2000).

Method effects unfortunately cannot be completely disentangled from the 'true' effort since they may also change both spatially (state-by-state) and temporally (annually and seasonally). However, given the three-year overlap in the CHTS and the FES from 2015-2017, the difference in method effects can be estimated and extrapolated over time by using any covariates that are specific to each survey. The primary, measurable covariate that has changed over the course of the CHTS is the prevalence of wireless-only households; the National Health Interview Survey has estimated proportions of wireless-only households from 2007-2014 (Blumberg and Luke, 2016). A simple model was fit to this 2007-2014 time series to expand the covariate from 1981-2017 so it could be used across the full effort time series. The National Health Interview Survey data was empirical logit transformed as follows:

$$elogit(\hat{p}_s) = \log\left(\frac{\hat{p}_s}{1-\hat{p}_s}\right)$$

where $\hat{p}_s$ is the proportion of wireless-only households in state s. The logits were then fit state-by-state to a simple linear model with a slope change in 2010 that corresponded with a significant increase in wireless-only households. The fitted model (R^2=0.9948) indicated that the proportion of wireless-only households were approximately zero in all states before the year 2000. This finding indicated that the wireless-only households covariate should only be included in the modeled time series after 2000, since its effects were negligible in years prior.

As with the other covariates described in section 8.1.1.1, the wireless-only covariate's interactions with state, season and annual changes in population also need to be taken into account in the model. This assumption is supported by the fact that all FES estimates from 2015-2017 are consistently higher than the CHTS estimates.

8.1.2 Model

The traditional Fay-Herriot model is written as

$$\hat{Y}_d = \theta_d + e_d$$
$$\theta_d = x_d'\beta_d + v_d$$

where $\hat{Y}_d$ is the design-unbiased survey estimator in domain d;

θ_d is the true value of the variable of interest for domain d, contorted by method effects;

e_d is the sampling error, an independent random variable with an assumed known design variance $\sigma^2_{e,d}$;
x'_d represents a vector of auxiliary variables $(x'_1, x'_2 \ldots x'_n)$ related to the variable of interest;
β_d represents a vector of unknown regression coefficients $(\beta_1, \beta_2 \ldots \beta_n)$ for x'_d to be estimated, and;
v_d represents random effects variables, assumed to be distributed with a mean of zero and an unknown variance of ψ to be estimated.

The models for the CHTS ($\hat{T}_d$) and FES ($\hat{M}_d$) survey estimates are therefore

$$\hat{T}_d = \theta_d^T + e_d^T$$
$$\theta_d^T = x'_{Td}\beta_d + v_d$$

and

$$\hat{M}_d = \theta_d^M + e_d^M$$
$$\theta_d^M = x'_{Md}\beta_d + v_d$$

where $\hat{T}_d$ is the CHTS fishing effort estimate in domain d (i.e. in a specific state and year-wave combination), natural log transformed to normally distribute the data.
$\hat{M}_d$ is the FES fishing effort estimate in domain d (i.e. in a specific state and year-wave combination), natural log transformed to normally distribute the data.
θ_d^T is the true fishing effort for domain d contorted by CHTS method effects.
θ_d^M is the true fishing effort for domain d contorted by FES method effects.
e_d^T and e_d^M and are the sampling errors for the CHTS and FES respectively (again, these are assumed to be independent of one another since the surveys are conducted independently, and assumed to be normally distributed based upon the central limit theorem [e.g. Le Cam 1986]), with log design variances $\sigma^2_{eT,d}$ and $\sigma^2_{eM,d}$.
x'_{Td} represents a vector of auxiliary variables $(x'_1, x'_2 \ldots x'_n)$, which include:

- the logarithm of the state-specific population sizes as a proxy for annual variation;
- wave indicator variables as a proxy for seasonal variation;
- state indicator variables to capture geographic variation;
- two interaction terms accounting for annual and seasonal changes in geographic variation: log(state-specific population sizes)×state indicator variables, and wave indicator variables×state indicator variables;
- the estimated difference in CHTS-FES method effects (modeled by the increase in wireless-only households from 2000 on);
- A set of interaction terms between the method effects and all of the above auxiliary variables;
- An indicator variable for the presence of an FES estimate and its interactions with log(state-specific population sizes), wave indicators, and state indicators (this variable is only relevant to the short 2015-2017 benchmarking period so no interaction term with the CHTS method effects was taken into account)

- x'_{Md} represents a vector of the same auxiliary variables of x'_{Td}, including estimated the difference in FES-CHTS method effects (modeled by the decrease in wireless-only households back to 2000);

β_d is a vector of unknown regression coefficients of the fixed auxiliary variables shared between the CHTS and the FES.

v_d represents the random effects irregular term, assumed to be distributed with a mean of zero and an unknown variance of variance ψ to be estimated.

These two models are used to convert between FES and CHTS estimates. To produce FES-equivalent estimates prior to 2015, $\widehat{M}_d$ is used to produce shore and private boat fishing effort time series. To calculate CHTS equivalent estimates beyond 2017 (after the CHTS was discontinued), $\widehat{T}_d$ is used. Note that in estimating $\widehat{T}_d$, the CHTS method effects and all related interaction terms are dropped in the model for the 1981-2000 time series since there were very likely no wireless-only households prior to 2000 (as described in section 8.1.1.2).

Prior to running either of these two models, however, a suite of additional parameters need to be estimated beyond those described in section 8.1.1: $\sigma^2_{eT,d}$ and $\sigma^2_{eM,d}$ (CHTS and FES design variances on a log scale), ψ (the variance of the irregular term v_d), and β_d (vector of regression coefficients of the auxiliary variables shared between CHTS and the FES). The FES and CHTS variance estimates are known (described in sections 3 and 4). These variances, which are on the original effort scale, are essentially 'converted' to the log scale to estimate $\sigma^2_{eT,d}$ and $\sigma^2_{eM,d}$ using generalized variance function estimation (see Wolter 2007). The variable ψ is estimated using Restricted Maximum Likelihood (REML), and β_d is estimated using Empirical Best Linear Unbiased Predictor (EBLUP). These two methods are standard statistical approaches to estimating unknown model parameters; REML is a general method in linear mixed modeling for estimating variance components. EBLUP relates regression coefficients to the known auxiliary variables for each of the domains that constitute the target population.

For further, more technical details on the effort calibration, see Breidt *et al.*, 2017.

8.2 MRFSS Intercept/APAIS Calibration

The MRFSS Intercept/APAIS calibration involves adjustments to sample weights within several time-periods (Figure 6) using a raking methodology (see Foster *et al.*, 2018), which is completed using SAS (SAS Institute Inc., 2016). From 1981 to 2013, the MRFSS Intercept Survey was used to sample angler-trips for detailed trip and catch information. Starting in Wave 2 of 2013, the new MRIP APAIS was implemented with the weighted estimation methodology described in section 2. While the MRFSS Intercept Survey design and sampling procedures were better documented from 2004 and beyond, detailed information dating back to 1981 was incomplete, and even more limited in earlier years; design changes in early years were not thoroughly documented, sample sizes varied considerably over time, and exact sample sizes per wave are unavailable for those years.

Prior to the transition to the APAIS in 2013, a weighted estimation procedure was developed in an initial attempt to preserve the time series, at least back to 2004. This adjustment created 'pseudo-weights' for MRFSS Intercept Survey data collected between January 2004-

March 2013. However, the pseudo-weights modeled components that were not entirely comparable to the fully design-based sample weights used in the MRIP APAIS design. In order to preserve the integrity of the entire time series, the MRIP team had to develop a method that addressed two challenges in calibrating the estimates:

1) Supplement the 2004-2013 estimate adjustment methods to increase the validity of the initial pseudo-weighting approach; and
2) Produce reliable estimates from 1981-2003 that are robust to the lack of survey design and sampling information from those earlier years.

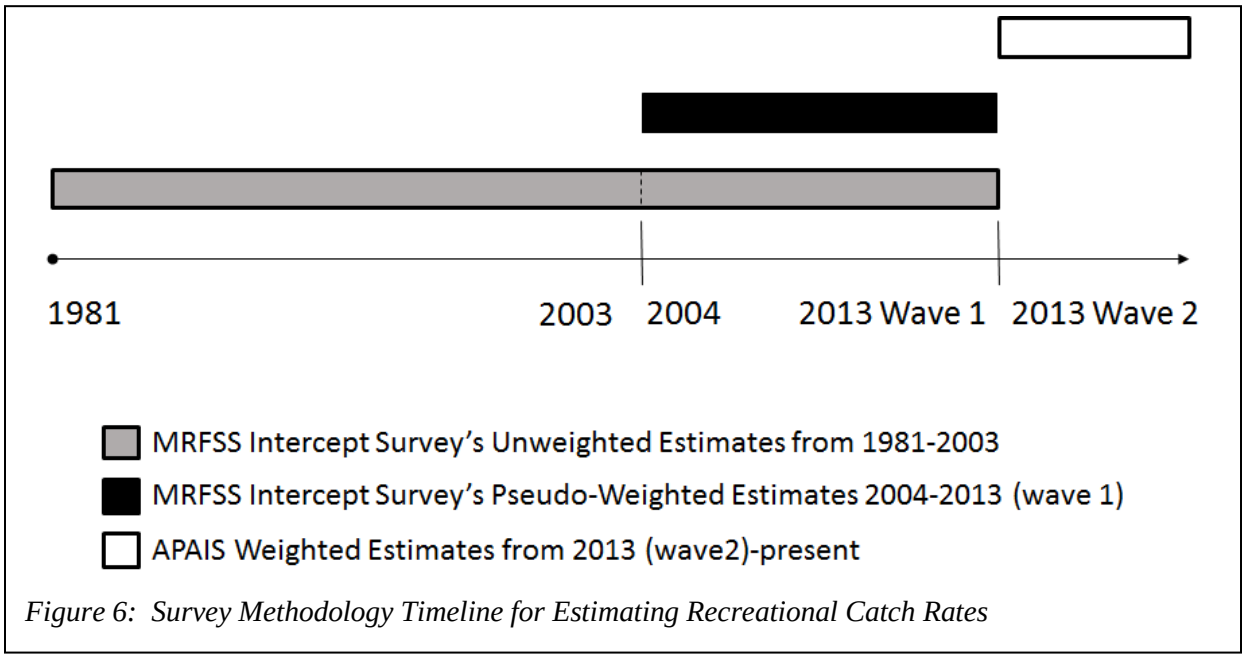

Figure 6: Survey Methodology Timeline for Estimating Recreational Catch Rates

8.2.1 Initial Pseudo-Weighting Approach used from 2004 - wave 1, 2013

An initial adjustment procedure was developed prior to the transition from the MRFSS Intercept Survey to the MRIP APAIS, which involved applying pseudo-weights to the 2004-2013 (wave 1) time series. One of the major issues that needed to be dealt with in weighting the 2004-2013 estimates was that the MRFSS Intercept Survey allowed samplers visit up to two alternate sites in addition to their assigned site in an effort to maximize the number of interviews obtained per dollar spent. As such, while inclusion probabilities of the assigned sites were known from a formal sample selection process, inclusion probabilities for these alternate sites were unknown. Because these weights were unknown, a modeling approach was used to develop the pseudo-weights using historical frequencies of alternate and assigned site visits. This approach involved a logistic model to estimate alternate site inclusion probabilities, so that the final site-day sampling weights could be calculated as a combination of assigned and alternate site inclusion probabilities.

In addition to modeling alternate site inclusion probabilities, an adjustment was also needed to account for the fact that 1) in some cases, only a subset of all the anglers during a

sampled day were being observed at a selected site, and 2) the time spent on-site was only a part of the 24-hour day. The first issue of only interviewing a subset of total anglers at a site was relatively simple to address, in that the MRFSS Intercept Survey collected counts of missed trips of individual anglers. Further, if a trip was from a boat, interviewed anglers were asked how many other anglers were on the boat that day, which could be used to estimate the number of anglers per boat per site-day. The larger challenge was expanding trips to the 24-hour period. MRIP consultants developed a circular normal model to estimate the proportion of daily trip end times by hourly time-intervals using the data from the CHTS, since one of the questions asked during the CHTS interviews involved recalling the time that fishing trips ended. This information provided an idea of the distribution of fishing trips throughout the day by fishing mode, and this input was used to feed into the model approach (see Hernandez-Stumpfhauser *et al.*, 2016 for more details).

The major assumption underlying the validity of this approach is that the trips occurring during the time-period the sampler was on-site were representative of those that took place during the full day. However, this assumption is only satisfied if either 1) the time the sampler spent on-site was randomly selected within a 24-hour period, or 2) trip characteristics are not related to time-of-day. Neither of these conditions were met. Firstly, the sampling was not random throughout the day, since the MRFSS Intercept Survey emphasized maximizing the number of interviews obtained per dollar spent rather than strictly adhering to formal probability sampling. Traditional MRFSS Intercept Survey sampling procedures instructed samplers to visit an assigned site during the assigned day's peak fishing activity period, leading most site visits to be made between the hours of 10am and 4pm. Secondly, trip characteristics, including landings per angler-trip, were in fact related to the time of day for many species (see MRIP Staff, 2014). For these reasons, the procedure described below in 8.2.2 was developed to supplement the initial weighting of the 2004-2013 data by accounting for differences in trip characteristics between those observed during site visit intervals and those outside of those intervals.

8.2.2 Calibration Approach for the 2004 - 2013 Data

A raking ratio adjustment method was used to calibrate the 2004-2013 (wave 1) period to the 2013 (wave 2) – 2016 period (Foster *et al.*, 2018). Raking ratio adjustment, also called iterative proportional fitting or sample balancing, is a widely used approach in survey calibration (originally proposed by Deming and Stephan [1940]). Raking is a sequential process that adjusts sample design weights repeatedly until weights converge (i.e. stop changing), based upon known population characteristics. The starting calculation for the raking algorithm used in the MRFSS Intercept Survey/APAIS calibration is as follows:

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

where w_j^* is the adjusted sample weight of angler-trip j;
w_j is the initial sample weight of angler-trip j;
$\widehat{N}_{D,new}$ is the average of the domain estimates (e.g. by state, sub-

state region, year, wave, fishing mode, area fished, coastal/non-coastal household, and for-hire boat frame membership) of intercepted angler-trips from 2013 (wave 2)-2016; and

$\widehat{N}_{D,old}$ is the average of the domain estimates of intercepted angler-trips from 2004-2013 (wave 1).

This ratio adjustment can correct for the overall under or over-representation of trips in a domain under the MRFSS Intercept Survey design relative to the new APAIS methods, which improves estimates for key variables in the domain that is being adjusted.

The component estimates of $\widehat{N}_{D,new}$ (i.e. the estimated angler-trip totals from wave 2 of 2013-2016, or $\widehat{N}_{D,2013w2} - \widehat{N}_{D,2016}$) and $\widehat{N}_{D,old}$ (i.e. $\widehat{N}_{D,2004} - \widehat{N}_{D,2013w1}$) are calculated by summing the weights of the trips in that domain as follows:

$$\widehat{N}_{D,YYYY} = \sum w_j \, I_{D(j)}$$

where $\widehat{N}_{D,YYYY}$ is the estimated number of angler-trips taken in domain D in a given year *YYYY*;

w_j is the weight of angler-trip j (note that for $\widehat{N}_{D,new}$, this value is actually the final APAIS sample weight w_{hij} described in section 2.3.1, and for $\widehat{N}_{D,old}$ this value is the pseudo-weight described in section 8.2.1); and

$I_{D(j)}$ is an indicator variable, which equals 1 if (j) is in domain D or 0 if otherwise.

The raking is done by making sequential ratio adjustments on a set of coarse domain estimates (Table 6) until the weights converge (Figure 7). The coarse domains were decided upon because it was impractical to implement all possible small domain estimates (i.e. estimates from every combination of year, wave, state, sub-state region, fishing mode, fishing area, coastal vs. non-coastal household). These domains contain important trip characteristics that were found to contribute to differences in the MRFSS Intercept Survey and APAIS data (MRIP Staff, 2014).

Table 6: Domain estimates used for raking approach

Domain	**Estimated Intercepted Angler-Trips (for 2013 wave 1 - 2016)**	**Estimated Intercepted Angler-Trips (for 2004 - 2013 wave 1)**
Area Fished, State, Wave, and Fishing Mode	$\widehat{N}_{D,new,AF}$	$\widehat{N}_{D,old,AF}$
Household Status (i.e. Coastal or Non-Coastal), State, Wave and Fishing Mode	$\widehat{N}_{D,new,HS}$	$\widehat{N}_{D,old,HS}$
For-hire frame status (i.e. vessels on the for-hire sample frame or not), State, Wave, and Fishing Mode	$\widehat{N}_{D,new,FH}$	$\widehat{N}_{D,old,FH}$
Sub-State Region, State, Wave, and Fishing Mode	$\widehat{N}_{D,new,RE}$	$\widehat{N}_{D,old,RE}$

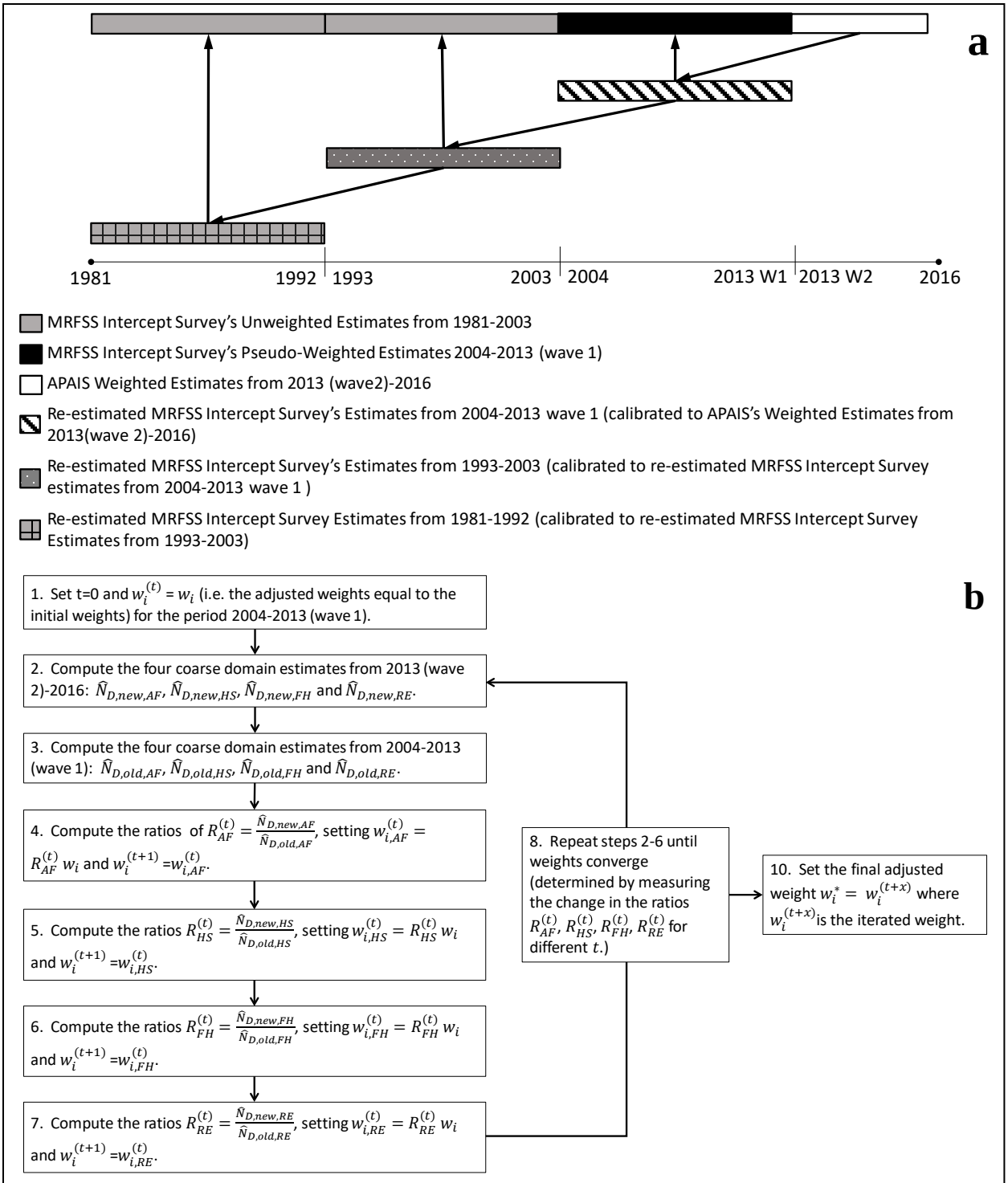

Figure 7: Sequential approach to MRFSS Intercept Survey/APAIS calibration of the 1981-2013 time series (a) and the step-by-step raking procedure (b).

The only issue with this method is that it assumes that the differences between the 2004-2013 (wave 1) estimates and 2013 (wave 2)-2016 estimates can be attributed solely to changes in

the survey design. Such an assumption is valid if trip characteristics remain constant over the time periods considered. Given the lack of benchmarking period, however, it cannot be fully verified that the differences are entirely due to the design change, or actual changes in the fishery over time. In these scenarios, raking would result in a weight adjustment that is too large, because it removes both the design-induced changes and the changes in the fishery.

To reduce the risk of over-adjustment when real changes in the fishery may have occurred over time, a test is applied to identify any significant temporal 'drift' in important trip characteristics before adjusting weights. Prior to raking, estimates are calculated for each year and wave between 2004 and 2013 (wave 1) using the four domains described in Table 6 (area fished, household status, for-hire frame status, and sub-state region), resulting in four time series, each consisting of 145 data points. A linear regression of the totals over time for each of these time series is then performed, and their slopes are tested for significant differences from zero at the 97.5% confidence level. For those time series with slopes that are not significantly different from zero, the raking procedure is applied as described above. However, if a significant temporal trend is found for a given domain, the raking algorithm is modified for that domain so that only the years 2010-2013 (wave 1) are used in the computation of $\widehat{N}_{D,old}$. In other words, the ratio adjustment for that domain is only based upon the most recent years instead of the full time period.

8.2.3 Calibration Approach for the 1981 - 2003 Data

The weight adjustments for data collected prior to 2004 are performed following similar computations to those described in section 8.2.2. The 20-year time period, however, is divided into two intervals prior to applying the weight adjustments: 1981-1992 and 1993-2003. This division is made to minimize the chances of any over-adjusted weights masking both actual changes in the fishery, and unaccounted-for changes in the MRFSS Intercept Survey design over time.

The major difference, in the 1981-2003 calibration compared to the 2004-2013 (wave 1) calibration is in the methods for setting initial sample weights, since the exact sampling design in these prior years is unknown. The number of angler-trips cannot be estimated for this time period, so the initial sample weights used for this older time series capture the relative changes in site-day sampling intensity over time. These angler-trip weights (w_j) are set as a ratio of site-day tallies that had angler intercepts in domains defined by year, wave, state and mode to the maximum site-day tallies across years:

$$w_j = \frac{C_{D,max}}{C_{D,YYYY}}$$

where $C_{D,YYYY}$ is the total count of site-days in domain D in year YYYY; and
$C_{D,max}$ is the maximum count of site-days in domain D across all years in the time series interval (either 1981-1992 or 1993-2003).

These weights are greater than 1 for all domains except for that which corresponds to the year with the maximum site-day count; this domain has an angler-trip weight equal to 1.

After constructing initial weights, the raking algorithm is implemented to create adjusted weights as described in 8.2.2. The starting ratio used for these older time period calibrations is the same ($w_j^* = \hat{N}_{D,new}/\hat{N}_{D,old}\, w_j$), except that $\hat{N}_{D,new}$ for the 1993-2003 time series is the average of the adjusted angler-trip estimates by domain from 2004-2013 (wave 1), and the $\hat{N}_{D,new}$ for the 1981-1992 time series is the average of the adjusted angler-trip estimates by domain for 1993-2003 (Figure 7).

The linear regression testing for a significant time series trend described in section 8.2.2 is also performed for both the 1981-1992 estimates and the 1993-2003 estimates, but for both $\hat{N}_{D,new}$ and $\hat{N}_{D,old}$ rather than just $\hat{N}_{D,old}$. If a significant trend is detected in $\hat{N}_{D,old}$, the adjustment ratios are computed using only the final three years of the time period instead of the full time period. Similarly, if a significant trend is detected in $\hat{N}_{D,new}$, the adjustment ratios are computed using only the first three years of the time period instead of the full time period.

The two following coarse domains were also added to this raking procedure to control for unobserved design effects:

- KOD (state, wave, mode, and kind of day): $\hat{N}_{D,new,KOD}$ and $\hat{N}_{D,old,KOD}$
    - This domain corresponds to the weekend/weekday strata that are part of both the MRFSS Intercept Survey and the APAIS designs.
- AC (state, wave, mode and site activity class): $\hat{N}_{D,new,AC}$ and $\hat{N}_{D,old,AC}$
    - The site activity classes consist of two groups (high activity and low activity), based on annual counts of intercepts by fishing mode. Sites with counts above the annual mean (within domains defined by state, mode, year, sub-state region) were classified as high; sites at or below the mean were classified as low.

In the first version of these methods to calibrate the 1981-2003 series, a third domain MG (state, wave, mode, and month-group) was included, which consisted of three classes of months based upon historical fishing activity. However, this third domain exhibited too much variability to be useful, and was dropped in the final calibration methods that were implemented in July 2018.

For more detailed information about the MRFSS Intercept Survey/APAIS calibration methods, see (Foster *et al.*, 2018).

9. Historical and Anticipated MRIP Design Changes

Since its initiation in 2008, MRIP has been constantly improving its data collection and estimation methods in an effort to disseminate the highest quality information possible to managers, scientists and the public. To this end, MRIP has developed a process that allows for evaluation of current methods, development and testing of new methods, peer-review and certification of scientifically sound surveys. To-date both the APAIS and the FES have been certified and plans are in place to move the FHS and the LPS designs towards certification in the future. Toward the goal of achieving nationwide consistency in data quality, MRIP also supports certification of survey methods administered by state and regional data collection partners.

In 2017 the MRIP underwent a second program review, which concluded that 'significant improvements' had been made to the program, particularly in the statistical soundness of the new survey designs, and made further recommendations for improvement (National Academies of Sciences, Engineering and Medicine 2017). As these recommendations are addressed in the upcoming years, they will be reflected in subsequent versions of this document. Below summarizes the nine-year evolution of these MRIP surveys, transitions that are in progress, and improvements that will be made in the near future.

9.1 APAIS Changes

9.1.1 Historical Changes to the APAIS

The APAIS design has been implemented as detailed in section 2 since March of 2013. The APAIS was a drastic change from the MRFSS Intercept Survey design that was used in prior years. Described below, the major alterations from the MRFSS Intercept Survey as it transitioned into the APAIS were in the sample frame, stratification, sample selection, several sampling protocols and the weighting methods for catch rate estimation. Since 2013, only three additional, minor changes have been made to the survey, including the design of the temporal stratification, the design of the fishing mode stratification, and the maximum number of sites allowed per cluster.

First, the site register used for the MRFSS Intercept Survey was historically a private list only accessible by APAIS staff and affiliates, but is now publicly accessible on the MRIP website (https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp). This change allowed for public input and review, which helped increase the accuracy of the APAIS sample frame and confidence in its comprehensiveness.

Changes have also been made in both temporal and fishing mode stratification. Two key alterations to the APAIS temporal stratification from the MRFSS Intercept Survey was the addition of the night-time interval and interval P (the 11AM-5PM interval corresponding with peak hours of fishing activity). Sampling for the MRFSS Intercept Survey was not done at night,

and the survey inaccurately assumed that night and day fishing activity were similar. The full 24-hour coverage of the current APAIS design reduced that bias. Sampling at peak hours also increased the productivity of the survey, resulting in an increase in the number of completed interviews without compromising the existing design. Inclusion probabilities were adjusted to account for this overlap in the two other daytime sampling intervals. One final temporal stratification change was made in July 2013, where Fridays were moved from the weekday to weekend stratum, since the observed high fishing activity on Fridays was more comparable to weekend activity. The fishing mode stratification went from individual fishing modes (private, shore mode, charter boat) to site group strata to increase the interviewing productivity of the survey, as detailed in section 2.1.1. This change was completed in May 2014 for charter and private boat modes, and in early May 2016 for shore mode.

While the current APAIS design optimizes the sampling of the PSUs rather than maximizes target number of interviews, the MRFSS Intercept Survey had specified target samples of angler-trip interviews for any given stratum. The MRFSS Intercept Survey therefore had three types of assignments: 1) fixed assignments, which had to be completed regardless of interview targets; 2) flexible assignments, which were not completed if interview targets for the wave had been reached; and 3) reserved assignments, which were decided upon by samplers in the field, and only used if interview goals had not been met by the fixed and flexible assignments. Further, the MRFSS Intercept Survey sample selection used an entirely random PPSWOR approach, which was difficult to adhere to in practice due to sampling logistics. As such, assignments could be rescheduled due to bad weather or unexpected sampler problems, and alternate sites could be visited at a sampler's discretion to increase the angler interviewing productivity. The multiple types of assignments, the rescheduling, and the use of alternate sites to those obtained through the probability sample, caused difficulties in calculating accurate inclusion probabilities. The current APAIS design eliminated these issues by making all assignments fixed, prohibiting rescheduling, and implementing the replication-based selection that controls the PPSWOR based on logistical constraints. These alterations made it much more straightforward to calculate the inclusion probabilities for use in the weighted catch rate estimation, resulting in higher confidence in the estimates.

Another change, made in July of 2013 that improved sampling logistics was allowing a maximum of two-site clusters instead of up to three. Three site clusters were initially allowed from March 2013-July 2013, but were quickly found to be difficult to sample in practice.

Finally, the APAIS sample weighting process detailed in section 2.3.1 is a new feature to the statistical methodology, as the MRFSS Intercept Survey used unweighted averages that did not fully reflect the multi-stage sampling design.

9.1.2 Future Improvements

The APAIS is a certified survey, and no drastic changes to the design are planned in the immediate future. However, the National Academy of Sciences (NAS) made five primary technical recommendations (National Academies of Sciences, Engineering and Medicine 2017) for the APAIS in their 2017 peer review that will be addressed in the coming years:

- Integrate electronic data-collection into the APAIS design to increase efficiency and accuracy and decrease costs.
- Evaluate angling activity at private access points; currently, an unverified assumption is made that public and private fishing activities are similar. The NAS review suggested validating that assumption.
- Develop validation programs for discard mortality to be able to estimate fates of discarded catch.
- Explore alternative methods (e.g. small area estimation procedures) to produce domain estimates at a finer scale.
- Explore the accuracy of the PPS sampling design and explore alternatives if needed to reduce variance estimates.

Research will be initiated and appropriate changes to the APAIS methods will be made as NAS recommendations are addressed.

9.2 FES Changes

There have been no significant changes in the FES design since its certification in 2015.

9.2.2 Future Improvements

Some of the key NAS recommendations (National Academies of Sciences, Engineering, 2017) for the FES were:

- Evaluate the two-month reference period to ascertain the quality of the data.
- Continue to monitor nonresponse and its effect on the data quality of the survey.
- Explore electronic data collection, such as web questionnaires or smartphone apps, to reduced costs and improve data quality.
- Explore alternative methods for estimating variance.

In the effort to address these recommendations thus far, the MRIP team has begun evaluating the two-month reference period of the FES in comparison to alternative designs, and is beginning to explore electronic reporting options. A pilot study was conducted to examine the impact of reference period on recall, and found no indication of omission error in the FES relative to one-month reference periods (Andrews *et al.*, 2018). In the final two waves of 2018, a 'push to web' pilot study for the FES was conducted in order to test the feasibility of implementing a web-based FES questionnaire. The team plans to continue exploring other potential sources of bias in the design, as well as additional methods to improve the survey, in the coming years.

9.3 FHS Changes

9.3.1 Historical Changes to the FHS

The FHS was initially developed to address undercoverage of charter and headboats in the CHTS. The survey was initiated as a pilot project in 1997, then referred to as the 'Charter Boat Survey' and was only conducted in Gulf of Mexico states (Florida, Alabama, Mississippi and Louisiana). This Gulf FHS became official in 2000 for estimating charter boat fishing effort, and in that year, research was undertaken on expanding it as a method for both charter and headboat effort. In 2003, the current version of the FHS started being implemented along the Atlantic Coast.

The only significant change to the FHS since its initiation was in the estimation methods developed to integrate the FHS and VTR data into the Atlantic for-hire estimates; the dual-frame estimation methodology was first implemented in 2004. Initially, VTR data was used only for the annual estimates due to concern over some late submissions from captains that could not be incorporated into the wave estimates. At the end of each year, the VTR data was integrated with the FHS data to produce an annual for-hire effort estimate, but then also was used retroactively to re-estimate that year's wave estimates. However, this re-estimation led to large changes to the for-hire effort estimates at the end of each year, so the methods were changed in May of 2017 to use the VTR for all initial wave estimates as well as for the annual estimates. The change in methods better maintains consistency of the wave and annual estimates, although the effects of late submissions on the wave estimates are still being examined.

9.3.2 Future Improvements

While the FHS was not discussed in the level of detail of the APAIS and FES in the NAS 2017 review, the review did echo recommendations from a consultant's report on the FHS (Chromy *et al.*, 2009). The primary statistical recommendation was to examine the coverage properties of the FHS sample frame in order to better understand and account for coverage error in the survey. Another recommendation from the NAS was to explore the potential for collecting both for-hire catch and effort data through electronic logbook programs and other electronic data reporting technologies. The for-hire angling community has shown a keen interest in electronic data collection procedures, and MRIP pilot studies are currently underway to explore methods of self-reported catch data, and evaluate the statistical validity of those data for the for-hire estimates. These are high priorities to improve the FHS in the coming years.

Another development for the FHS is a new online FHS vessel directory that was partially launched in late 2017 and made fully operational in early 2018. The new vessel directory was developed to improve the management of information quality, functional processes and reconciliation of data from different sources. It closely resembles the Public Access Fishing Site Register that is used for the APAIS, but is privately accessible to MRIP staff and partners via login. A few features of the new directory include the following:

- Version control and tracking of changes that are made by different users;
- Increased sample selection and data processing efficiency;

- Improved functions for identifying vessels operating in multiple states; and
- Improved functions for validating/adding sites to the register.

9.4 LPS Changes

9.4.1 Historical Changes to the LPS

The LPS had been ongoing since 1986, but only became associated with the MRFSS in 2005; for the first 19 years of the survey, the LPS was managed by the NMFS Sustainable Fisheries HMS Management Division, but in 2005 was transferred to the Office of Science and Technology Statistics Division that directed the MRFSS and now manages the MRIP. Prior to 2005, the LPS consisted of a telephone survey (that became the LPTS), an intercept survey (that became the LPIS), and a dockside mark-recapture survey that was used to ascertain information about population sizes of high priority HMS. The dockside mark-recapture survey was discontinued in 1995 and no federal biological survey was conducted on HMS until the LPBS was trialed in 2000. There was then another gap in biological data collection on HMS until the MRIP team restarted the LPBS in 2006.

After the change in LPS management in 2005 the program design was revised, but has remained largely unchanged since MRIP's initiation in 2008. The only edits that have been made since then are the exclusion of incidental catch of large pelagics and HMS from the LPIS and a change in the number of LPBS assignment types. As mentioned in section 7.2, the LPIS excluded incidental catch of large pelagics and HMS in 2010 from sampling to remove discrepancies in vessel-trip definitions between the LPIS and LPTS. Currently, LPBS assignments are either tournament assignments or opportunistic assignments as described in section 7.4. However, the 2013 version of the LPBS included two other assignment types: fixed date/fixed site assignments and fixed date/roving assignments. Fixed date/fixed assignment types were those in which NMFS specified the sites and dates for an assignment during the four-hour periods corresponding with peak hours where vessels returned from fishing trips targeting large pelagics and HMS. Fixed date/roving assignments were those in which NMFS specified the date and state for the assignment, and samplers chose sites that maximized the probability of sampling large pelagics and HMS. Now however, the current goal of the LPBS is to maximize sampling of large pelagics and HMS (which tournament and opportunistic assignments alone achieve). Since there was no formal probability sampling design for the fixed assignments anyway, they were removed from the design.

9.4.2 Future Improvements

The LPS is currently being redesigned, and plans are in place to have a new, certified version of the LPS ready for implementation by 2024. During this seven-year timeframe, design development, field-testing of potential designs, survey certification, and calibration of the LPS time series will be completed.

10. References

AAPOR. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. The American Association for Public Opinion Research.

Andrews, R., Brick, J. M., and Mathiowetz, N. A. 2013. Continued Development and Testing of Dual-Frame Surveys of Fishing Effort: Testing a Dual-Frame, Mixed-Mode Survey Design. Silver Spring, MD. https://www.st.nmfs.noaa.gov/mdms/public/finalReport.jsp?ReportID=831 (Accessed 17 February 2017).

Andrews, R. 2015. Continued Development and Testing of Dual-frame Surveys of Fishing Effort. Silver Spring, MD. https://www.st.nmfs.noaa.gov/pims/main/public?method=DOWNLOAD_PROPOSAL&record_id=1309 (Accessed 21 April 2017).

Andrews, W. R., Brick, J. M., Mathiowetz, N. M., and Stokes, L. 2010. Pilot Test of a Dual Frame Two-Phase Mail Survey of Anglers in North Carolina. https://www.st.nmfs.noaa.gov/mdms/public/finalReport.jsp?ReportID=355 (Accessed 9 March 2017).

Andrews, W. R., Papacostas, K. J., and Foster, J. 2018. A Comparison of Recreational Fisheries Surveys with One- and Two-Month Reference Periods. North American Journal of Fisheries Management, *accepted*.

Andridge, R. R., and Little, R. J. A. 2010. A Review of Hot Deck Imputation for Survey Non-Response. Int Stat Rev., 78: 40–64. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3130338/pdf/nihms299474.pdf (Accessed 5 April 2017).

Assael, H., and Keon, J. 1982. Nonsampling vs. Sampling Errors in Survey Research. Journal of Marketing, 46: 114–123. https://www.jstor.org/stable/pdf/3203346.pdf?refreqid=search%3A7597554707e367f7375af9affaa1842a (Accessed 22 May 2018).

Author, C., Bieler, G. S., and Williams, R. L. 1993. Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Carcinogenicity. Source: Biometrics, 49206112: 793–801. http://www.jstor.org/stable/2532200 (Accessed 26 April 2017).

Barnett, V., and Lewis, T. 1994. Outliers in Statistical Data. Wiley. 584 pp. https://ideas.repec.org/a/eee/intfor/v12y1996i1p175-176.html (Accessed 21 April 2017).

Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P., and Deckmyn, A. 2018. Package 'maps'. The Comprehensive R Archive Network (CRAN). https://cran.r-project.org/web/packages/maps/maps.pdf (Accessed 2 October 2018).

Blumberg, S. J., and Luke, J. V. 2015. Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July–December 2015. (Accessed 14 February 2017).

Blumberg, S. J., and Luke, J. V. 2016. Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July–December 2016. https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201705.pdf (Accessed 24 July 2017).

Breidt, F. J., Lai, H. L., Opsomer, J. D., and Van Voorhees, D. A. 2012. A Report of the MRIP Sampling and Estimation Project: Improved Estimation Methods for the Access Point

Angler Intercept Survey Component of the Marine Recreational Fishery Statistics Survey.
Breidt, F. J., Liu, T., and Opsomer, J. D. 2017. A Small Area Estimation Approach for Reconciling Mode Differences in Two Surveys of Recreational Fishing Effort. Silver Spring, MD.
Brennan, K. 2010. Southeast Region Headboat Survey Program Description. SEDAR 24-RD56, South Atlantic Red Snapper Data Workshop, May 2010.
Brick, J., and Kalton, G. 1996. Handling missing data in survey research. Statistical Methods in Medical Research, 5: 215–238. Sage PublicationsSage CA: Thousand Oaks, CA. http://smm.sagepub.com/cgi/doi/10.1177/096228029600500302 (Accessed 23 February 2017).
Brick, J. M., Andrews, W. R., Brick, P. D., King, H., Mathiowetz, N. A., and Stokes, L. 2012. Methods for Improving Response Rates in Two-Phase Mail Surveys. Survey Practice, 5.
Brick, J. M., Andrews, W. R., and Mathiowetz, N. A. 2016. Single-phase Mail Survey Design for Rare Population Subgroups. Field Methods, 28: 381–395.
Brodziak, J., Richards, B., DiNardo, G., Richards, B., and DiNardo, G. 2012. Stock Assessment Update of the Status of the Bottomfish Resources of American Samoa, the Commonwealth of the Northern Mariana Islands and Guam, 2012. Honolulu, HI. 96822-2396 pp.
Bureau, U. S. C. 2016. American Community Survey. http://www2.census.gov/programs-surveys/acs/ (Accessed 1 January 2017).
Chromy, J. R., Holland, S. M., and Webster, R. 2009. Consultant's Report: For-Hire Recreational Fisheries Surveys.
Cui, W. W. 2003. Reducing Error in Mail Surveys. Practical Assessment, Research and Evaluation, 8. http://pareonline.net/getvn.asp?v=8&n=18 (Accessed 22 May 2018).
Curtin, R., Presser, S., and Singer, E. 2005. Changes in Telephone Survey Nonresponse over the Past Quarter Century. Public Opinion Quarterly, 69: 87–98. Oxford University Press. https://academic.oup.com/poq/article-lookup/doi/10.1093/poq/nfi002 (Accessed 14 February 2017).
Czaja, R. 2005. Sampling With Probability Proportional to Size. *In* Encyclopedia of Biostatistics. John Wiley & Sons, Ltd, Chichester, UK. http://doi.wiley.com/10.1002/0470011815.b2a16066 (Accessed 5 December 2016).
De Leeuw, E. D. 2005. To Mix or Not to Mix Data Collection Modes in Surveys. Journal of Official Statistics, 2: 235–255. http://hbanaszak.mjr.uw.edu.pl/TempTxt/deLeeuw_2005_To Mix or Not to Mix Data Collection Modes in Survey.pdf (Accessed 24 July 2017).
Dienes, P. 1957. The Taylor Series: an introduction to the theory of functions of a complex variable. Dover Publications. 552 pp. https://www.lib.noaa.gov/uhtbin/cgisirsi/?ps=ryv1XYIWQh/SILVERSPRG/156290028/8/287246/The+Taylor+series+:+an+introduction+to+the+theory+of+functions+of+a+complex+variable+%5E2F (Accessed 23 January 2017).
Dillman, D. A., Smyth, J. D., Christian, L. M., and Dillman, D. A. 2009. Internet, mail, and mixed-mode surveys : the tailored design method. Wiley & Sons. 499 pp.
Fay, R. E., and Herriot, R. A. 1979. Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. Journal of the American Statistical Association, 74: 269. Taylor & Francis, Ltd.American Statistical Association. http://www.jstor.org/stable/2286322?origin=crossref (Accessed 22 August 2017).
Foster, J., Breidt, F. J., and Opsomer, J. D. 2018. APAIS data calibration methodology report. Silver Spring.

Fuller, W. A. 1975. Regression Analysis for Sample Survey. Sankhy, Series C, 37: 117–132.
Fuller, W. A. 2009. Sampling Statistics. Wiley, Hoboken, New Jersey. 475 pp.
Glasser, G. J., and Metzger, G. D. 1972. Random-Digit Dialing as a Method of Telephone Sampling. Journal of Marketing Research, 9: 59. http://www.jstor.org/stable/3149607?origin=crossref (Accessed 29 December 2016).
Goodman, L. A. 1960. On the Exact Variance of Products. Journal of the American Statistical Association, 55: 708. http://www.jstor.org/stable/2281592?origin=crossref (Accessed 30 November 2016).
Google. 2016. Google Maps API Distance Matrix Service. https://developers.google.com/maps/documentation/javascript/distancematrix (Accessed 12 July 2016).
Gray, G. W., Kline, L. L., Osborn, M. F., Salz, R. J., Van Voorhees, D. A., and Witzig, J. F. 1994. MRFSS User's Manual: A Guide to Use of the National Marine Fisheries Service Marine Recreational Fisheries Statistics Survey Database. http://www.asmfc.org/uploads/file/sr37MRFSS__UsersManualDec94.pdf (Accessed 30 November 2016).
Green, L. M., and Campbell, R. P. 2010. Trends in Finfish Landings of Sport Boat Anglers in Texas Marine Waters, May 1974-May 2008. Austin, TX. 1-645 pp.
Groves, Robert, M. 1990. Theories and methods of Telephone Surveys. Annual Review of Sociology, 16: 221–240. https://www.jstor.org/stable/pdf/2083269.pdf?refreqid=search%3A3eb7f35d37e28a7606eda4b85f5abff2 (Accessed 22 May 2018).
Henry, K., and Valliant, R. (n.d.). Methods for Adjusting Survey Weights When Estimating a Total. http://sitesusa.s3.amazonaws.com/wp-content/uploads/sites/242/2014/05/Henry_2012FCSM_V-A.pdf (Accessed 16 March 2017).
Hernandez-Stumpfhauser, D., Breidt, F. J., and Opsomer, J. D. 2016. Hierarchical Bayesian small area estimation for circular data. Canadian Journal of Statistics, 44: 416–430. Wiley-Blackwell. http://doi.wiley.com/10.1002/cjs.11303 (Accessed 28 June 2018).
Horvitz, D. G., and Thompson, D. J. 1952. A Generalization of Sampling Without Replacement From a Finite Universe. Source Journal of the American Statistical Association, 47: 663–685. American Statistical Association. http://www.jstor.org/stable/2280784 (Accessed 21 April 2017).
ICCAT. 2010. International Commission for the Conservation of Atlantic Tunas: Report for Biennial Period 2010-2011. https://www.iccat.int/Documents/BienRep/REP_EN_10-11_I_1.pdf (Accessed 9 May 2017).
Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. 1983. Optimization by Simulated Annealing. Science, 220: 671–680. https://www.jstor.org/stable/1690046?seq=1#page_scan_tab_contents (Accessed 13 December 2016).
Lai, H. L., and Foster, J. 2008. Surveys and Statistical Methods for Estimation of Catch and Effort in U.S. Marine Recreational Fisheries Draft 1.0. 1-106 pp.
Le Cam, L. 1986. The Central Limit Theorem Around 1935. Statistical Science, 1: 78–91. Institute of Mathematical Statistics. http://projecteuclid.org/euclid.ss/1177013818 (Accessed 4 January 2018).
Lessler, J. T., and Kalsbeek, W. D. 1992. Nonsampling error in surveys. Wiley. 412 pp.
Lumley, T. 2004. Analysis of Complex Survey Samples. Journal of Statistical Software, 9: 1–19.

http://www.jstatsoft.org/v09/i08/ (Accessed 23 January 2017).
Lumley, T. 2010. Complex surveys : a guide to analysis using R. John Wiley. 276 pp.
MacKellar, D. A., Gallagher, K. M., Finlayson, T., Sanchez, T., Lansky, A., and Sullivan, P. S. 2007. Surveillance of HIV risk and prevention behaviors of men who have sex with men--a national application of venue-based, time-space sampling. Public health reports (Washington, D.C. : 1974): 39–47. http://www.ncbi.nlm.nih.gov/pubmed/17354526 (Accessed 5 December 2016).
Magnuson-Stevens Fishery Conservation and Management Act, as amended through January 12, 2007. (n.d.). . http://www.nmfs.noaa.gov/ia/slider_stories/2016/03/msa_amended_2007_.pdf (Accessed 14 February 2017).
Molina, I., and Marhuenda, Y. 2015. sae: An R Package for Small Area Estimation. The R Journal, 7: 81–98. http://cran.r-project.org/package=sae. (Accessed 28 September 2018).
MRIP Staff. 2014. A Descriptive Analysis of the Access Point Angler Intercept Survey 2013 Design Change. Silver Spring, MD.
Muhib, F. B., Lin, L. S., Stueve, A., Miller, R. L., Ford, W. L., MPH Wayne Johnson, M. D., and Smith, P. J. 2001. A Venue-Based Method for Sampling Hard-to-Reach Populations. Public Health Reports, 116.
National Academies of Sciences, Engineering, and M. 2017. Review of the Marine Recreational Information Program. National Academies Press, Washington, D.C. https://www.nap.edu/catalog/24640 (Accessed 29 June 2017).
National Marine Fisheries Service. 2016. Fisheries of the United States 2015. https://www.st.nmfs.noaa.gov/commercial-fisheries/ fus/fus15/index (Accessed 20 December 2016).
National Research Council. 2006. Review of Recreational Fisheries Survey Methods. National Academies Press, Washington, D.C. http://www.nap.edu/catalog/11616 (Accessed 30 November 2016).
Ogunyinka, E. O., and Lavergne, D. 2009. 2008 Louisiana Recreational Fisherman and Health Advisory Survey Report. Baton Rouge, Louisiana.
Papoulis, A. 1984. Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York. 576 pp. http://www.worldcat.org/title/probability-random-variables-and-stochastic-processes/oclc/9757552 (Accessed 21 April 2017).
Parsons, J. T., Grov, C., and Kelly, B. C. 2008. Comparing the Effectiveness of Two Forms of Time-Space Sampling to Identify Club Drug-Using Young Adults. Journal of Drug Issues, 38: 1061–1081. SAGE Publications. http://jod.sagepub.com/lookup/doi/10.1177/002204260803800407 (Accessed 5 December 2016).
Pinheiro, J., Bates, D., DebRoy, S., EISPACK authors, Heisterkamp, S., Van Willigen, B., and R Core Development Team. 2018. Package 'nlme'. The Comprehensive R Archive Network (CRAN). https://cran.r-project.org/web/packages/nlme/nlme.pdf (Accessed 2 October 2018).
Potter, F. J. 1988. Survey of Procedures to Control Extreme Sampling Weights. Proceedings of the Section on Survey Research Methods, American Statistical Association: 453–458.
Potter, F. J. 1990. A Study of Procedures to Identify and Trim Extreme Sampling Weights. Research Triangle Park, NC.
Prokhorov, Y. V., and Statulevičius, V. 2000. Limit Theorems of Probability Theory. Springer

Berlin Heidelberg, Berlin. http://link.springer.com/10.1007/978-3-662-04172-7 (Accessed 24 July 2018).
R Development Core Team. 2013. R: A language and environment for statistical computing. Vienna, Austria.
RTI International. 2017. SUDAAN Statistical Software for Weighing, Imputing and Analyzing Data. https://www.rti.org/impact/sudaan-statistical-software-analyzing-correlated-data (Accessed 15 May 2017).
SAS Institute Inc. 2016. 14.2 User's Guide. SAS/STAT, Cary, NC.
Shotwell, M. 2015. Package 'sas7bdat'. The Comprehensive R Archive Network (CRAN). https://cran.r-project.org/web/packages/sas7bdat/sas7bdat.pdf (Accessed 2 October 2018).
U.S. Census Bureau. 2016. State Population Totals Datasets: 2010-2016. https://www.census.gov/data/datasets/2016/demo/popest/state-total.html (Accessed 21 September 2017).
Uspensky, J. V. 1937. Introduction to Mathematical Probability. McGraw-Hill Book Co., New York; London. http://www.worldcat.org/title/introduction-to-mathematical-probability/oclc/565356955 (Accessed 20 April 2017).
Vijayan, K. 1968. An Exact π PS Sampling Scheme: Generalization of a Method of Hanurav. Journal of the Royal Statistical Society B , 30: 556–566.
Wang, J. 2003. Data Mining: Opportunities and Challenges. Idea Group Inc, Hershey, PA. https://books.google.com/books?id=Nyvaaq9pI40C&pg=PA183&lpg=PA183&dq=cold+deck+imputation&source=bl&ots=9iK_7A0-Bu&sig=3jFwF-T5BUq1DJdgrzqK9_YcmOA&hl=en&sa=X&sqi=2&ved=0ahUKEwiTmdWBvcXTAhVB7IMKHdVWBagQ6AEIhgEwEg#v=onepage&q=cold deck imputation&f=fals (Accessed 27 April 2017).
Woodruff, R. S. 1971. A Simple Method for Approximating the Variance of a Complicated Estimate, 66. http://www.jstor.org/stable/2283947 (Accessed 30 November 2016).
Wright, T. 2014. A Simple Method of Exact Optimal Sample Allocation under Stratification with Any Mixed Constraint Patterns. https://www.census.gov/srd/papers/pdf/rrs2014-07.pdf (Accessed 17 March 2017).

11. Acronym Glossary

Acronym	Description
ABS	Address-based samples
APAIS	Access Point Angler Intercept Survey
CATI	Computer Automated Telephone Interviewing system
CDS	Computerized delivery sequence program
CHTS	Coastal Household Telephone Survey
FES	Fishing Effort (Mail) Survey
FHS	For-Hire (Telephone) Survey
GARFO	Greater Atlantic Regional Fisheries Office
GIS	Geographic Information System
HMS	Highly Migratory Species
LPBS	Large Pelagics Biological Survey
LPIS	Large Pelagics Intercept Survey
LPS	Large Pelagics Survey
LPTS	Large Pelagics Telephone Survey
MRFSS	Marine Recreational Fisheries Statistics Survey
MRIP	Marine Recreational Information Program
MSE	Mean Square Error
MSR	Master Site Register
NMFS	National Marine Fisheries Service
NSAR	National Saltwater Angler Registry
PPSWOR	Probability proportional to size sampling without replacement (statistical method for sample selection)
PSU	Primary Stage Unit
QSU	Quaternary Stage Unit
RDD	Random digit dialing (standard method for conducting telephone surveys)
SSU	Secondary Stage Unit
TSU	Tertiary Stage Unit
USPS	United States Postal Service