Improving Precision of Data-Limited Recreational Catch Estimates Through Multi-Year Catch Rate Aggregation

NOAA Fisheries Office of Science and Technology

SEDAR101-RD-15

April 2026

This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

NOAA FISHERIES

Office of Science and Technology
Southeast Fisheries Science Center

Improving Precision of Data-Limited Recreational Catch Estimates Through Multi-Year Catch Rate Aggregation

September 8, 2025

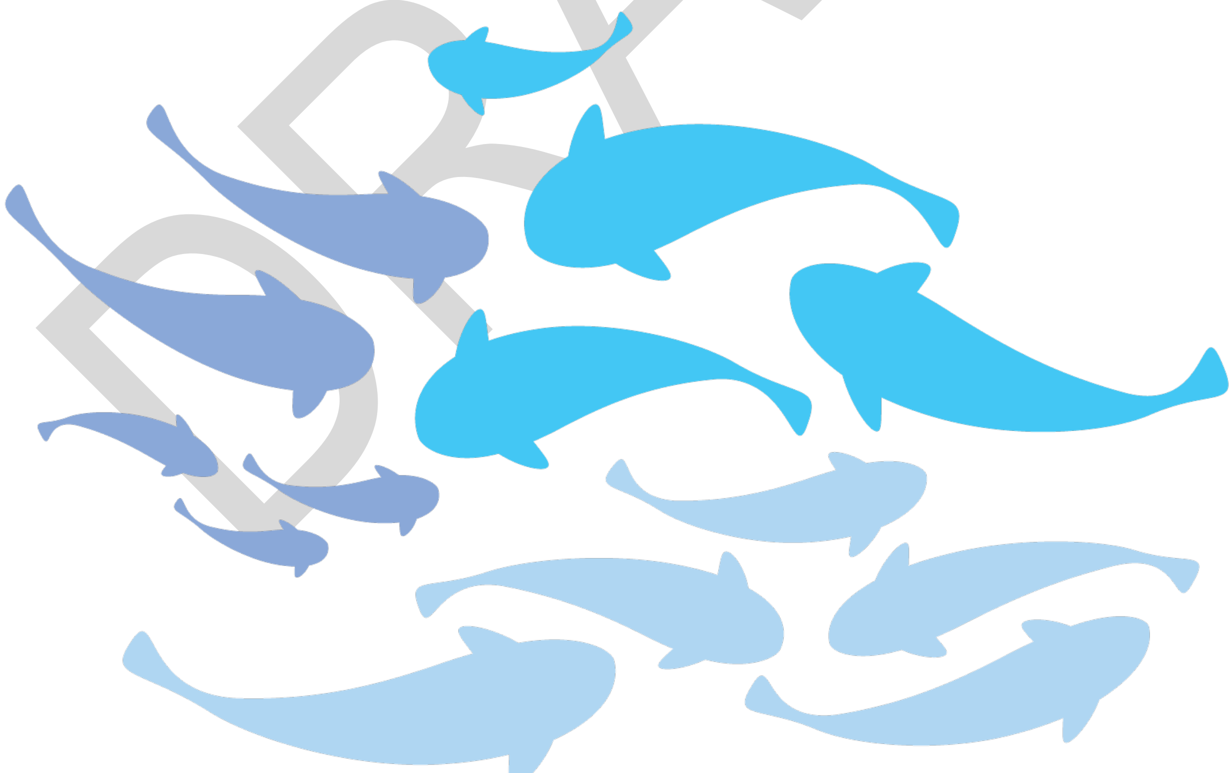


Table of Contents

Executive Summary

The Marine Recreational Information Program (MRIP) leverages general surveys to cover many recreationally caught species to meet a broad range of fisheries stock assessment and management needs in a resource constrained environment. The program's statistical methods have been peer reviewed (NASEM 2017a) and deemed robust; however, challenges for data users exist, especially for priority species that are infrequently encountered in sampling for a variety of reasons (e.g., rarity, temporally or spatially restricted distribution, or restricted seasonal closures). Imprecise recreational catch estimates recur in a number of estimation domains that cause concern among data users on how to best proceed with their work.

To provide some initial guidance, NOAA Fisheries Recreational Fishing Survey and Data Standards (NMFS 2025a), established in 2020, include a precision threshold for statistical reliability (in alignment with broader federal guidelines for statistical surveys): estimates with a percent standard error (PSE) of 50 or greater are not recommended for use in assessment and management without considering aggregation (e.g., across states, geographic regions, fishing modes). However, especially in the case of species infrequently encountered in sampling, data users are often faced with few alternatives in practice due to management needs. Estimates are often used as-is, or users will make ad hoc decisions on how best to handle imprecise estimates for their use-cases. While some flexibility is justifiable (e.g., in assessment models that are robust to recreational catch uncertainty or have built-in mechanisms to account for outliers), there is a desire among the MRIP data user community for guidance on more standardized approaches for treatment of imprecise estimates and across regions that could be shared for both assessment and management applications.

While there is no single, convenient solution to resolving all imprecise recreational catch estimate scenarios, there is a pressing multi-regional, near-term need for interim tools that have the potential to improve precision of at least a subset of estimates. A foundational approach to improving precision is to increase sample sizes, however, steadily rising survey costs make this impractical, underscoring the need to identify cost-neutral options to make more efficient use of the available sample. To that end, a NOAA Fisheries Office of Science and Technology (OST) and Southeast Fisheries Science Center (SEFSC) workgroup has collaborated to propose an interim moving average estimation method to improve the precision of catch estimates that are imprecise when using the standard MRIP design-based estimation method. This alternative method uses data from two-month sampling periods, or waves, over a range of years to calculate catch rates, and then those catch rates are multiplied by standard design-based effort estimates. This method is proposed because it can improve the precision of MRIP catch estimates and:

- Can be rapidly adopted and implemented by a broad cross-section of data users due to its simplicity,

- Provides a consistent framework across regions, species, and fishing modes, and
- Minimizes over-smoothing of time series (i.e., loss of year-to-year trend information).

This report describes the proposed methodology and evaluates its performance through simulations, comparing it to standard design-based estimates as well as several alternative multi-year aggregation approaches. The analysis examines potential precision gains across a range of multi-year averaging periods. Based on the analysis, we offer several general recommendations for applying this method in practice. We recommend implementing this method in cases exceeding the MRIP PSE threshold (PSE > 50) over a 3-5 year aggregation period for compatibility with dynamic fishery conditions and management cycles. We further recommend a centered moving average whenever possible to improve precision while avoiding trend distortion that can arise in lag-based methods. The application should also be automated, standardized while accounting for management history, and objective, with regular re-evaluation of species prevalence to ensure the method remains appropriate as fishery conditions evolve.

Background

With input from NOAA Fisheries partners, including state agencies and regional fisheries information networks, NOAA Fisheries established recreational fishing survey and data standards (NMFS 2025a) in 2020 to promote data quality, consistency, and comparability across the recreational fishing surveys administered and funded through the agency's Marine Recreational Information Program (MRIP). The standards were developed to establish clear criteria for the management of recreational fishing surveys, consistent with federal guidelines and best practices for the publication of statistical information. They define requirements for collecting data and producing and disseminating estimates.

As part of the government's guidelines for statistical programs, the Office of Management and Budget (OMB) requires federal agencies to establish their own criteria for statistical precision (OMB 2006, Guideline 4.1.5). In accordance with these guidelines, MRIP cautions using estimates with a PSE between 30 and 50, does not support the use of estimates with a percent standard error above 50 and in those instances recommends data users consider higher levels of aggregation to the extent possible (e.g., across states, geographic regions, or fishing modes).

This threshold was informed by precision standards applied in other federal statistical programs and by partner efforts, including the Atlantic Coastal Cooperative Statistics Program (ACCSP) Percent Standard Error Workshop (ACCSP 2016), which explored the impact of imprecise estimates on fisheries stock assessments. Workshop participants concluded that estimates with a PSE above 40 should be used with caution due to the increased potential for biases in stock assessment outcomes and subsequent management decisions. MRIP's adoption of a 50 PSE threshold reflects a balance between providing practical guidance to data users and emphasizing the substantial uncertainty associated with estimates that exceed this threshold.

While this standard is intended to improve the interpretation and application of uncertain data in fisheries management, all MRIP estimates are published regardless of PSE. This ensures full transparency, providing stock assessors, fisheries managers, and public users with complete access to the estimates and underlying microdata. Such access allows for flexibility in conducting sensitivity analyses and determining appropriate methods to address data gaps or limitations. Ultimately, it has been the responsibility of stock assessors to decide how best to handle MRIP estimates that do not meet the precision threshold, particularly when geographic or fishing mode aggregation does not satisfy their assessment needs.

To support these decisions, the SEFSC has identified five common scenarios faced by stock assessors in the region, the challenges associated with each, and the typical approaches used to address them. While these do not represent an exhaustive list, they reflect the primary methods applied to date—often on an ad hoc basis and without a formal decision-making framework. In addition to these data-focused approaches, it has also been common in past Southeast Data, Assessment and Review (SEDAR) assessments to use sensitivity analyses to test how alternative treatments of imprecise estimates affect model behavior and outputs.

1. **Highly imprecise estimates at the beginning of time series:** Such estimates can destabilize initial population estimates in assessment models. To address this, the SEFSC has historically applied smoothing techniques, such as averaging neighboring years (excluding years with high uncertainty) or incorporating the true coefficient of variation (CV) during model fitting.

2. **Highly imprecise estimates at the end of the time series:** This issue has been relatively rare, with the main exception being imputed estimates from 2020–2021 due to COVID-related sampling disruptions. In such cases, SEFSC has either smoothed the data using the mean of surrounding years or applied the true CV in model fitting.

3. **Highly imprecise estimates in large portions of the time series:** In these cases, SEFSC has either used the data as-is, or substituted these uncertainty measures with a smaller metric for model fitting purposes.

4. **Time series with intermittent catch estimates:** SEFSC assumes a minimal level of removals during years with zero values (e.g., substituting the lowest observed value from the entire time series).

5. **Infrequent or irregular outliers that are highly imprecise:** These are generally left unchanged unless such data spikes cause estimation issues within the model. In these cases, SEFSC smooths these values using immediately adjacent years to reduce model disruptions.

While these approaches have provided practical means to address imprecise estimates in various scenarios, SEFSC desires to shift from ad hoc decision making to a more cohesive, systematic framework to the extent possible. Given the critical importance of improving the precision of recreational fishing data for effective fisheries management, there is a clear need to rigorously evaluate a range of potential methods that could rapidly enhance estimate reliability.

The workgroup therefore explored such methods, assessing their applicability, strengths, and limitations in the context of NOAA Fisheries' immediate operational needs and current data environment in the Southeastern United States.

Evaluation of Potential Methods to Improve Precision

The OST-SEFSC workgroup considered tradeoffs including the practicality of near-term implementation and current availability of methods versus methods that have extended development time. Broadly, three classes of methods were considered to improve the precision of recreational catch estimates: small area estimation (SAE), addressing imprecise outlier catch estimates directly within stock assessment models, and multi-year aggregation. The workgroup ultimately chose to prioritize multi-year aggregation as the most practical and immediately applicable option. Each of these approaches, along with the rationale for the workgroup's focus on multi-year aggregation, is detailed below.

Small Area Estimation

Small area estimation (SAE) encompasses a suite of advanced modeling techniques designed to produce reliable estimates for domains where precise sample-based estimates are not feasible due to insufficient sample sizes. The growing demand across the survey landscape for more granular estimates at finer spatial or demographic scales has driven rapid development in SAE methods in recent years (NASEM 2017b).

Fundamental to SAE is borrowing data from adjacent time periods, geographical regions and relevant external data sources to inform model estimators and improve precision for granular estimation domains. While this approach can introduce bias, particularly if the underlying assumptions about similarity across time and space are violated, it substantially enhances the stability and reliability of resulting estimates (e.g., Quick et al. 2019, Chua & Long 2023).

Within the context of NOAA Fisheries' efforts to improve the precision of recreational catch estimates, SAE represents a promising yet complex methodology that could eventually be incorporated into OST's standard recreational catch estimation processes and yield systematically more robust and precise estimates for infrequently encountered species. However, this is a resource-intensive approach that, with existing resources, will require several years of dedicated development, calibration and validation before operational implementation is feasible. Another critical requirement for small area estimation is comprehensive access to detailed datasets on management actions, as these are necessary covariates to inform models appropriately and ensure regulations (e.g., quotas, fishery closures, size and/or bag limits) are taken into account. SEFSC has an ongoing effort to develop a region-specific management database that could provide a foundation for such future applications, however this will not be available in all regions in the near future.

Given these challenges, and the workgroup's pressing need to address immediate fisheries management priorities, SAE was not pursued further at this time. Its extended development horizon, along with the need for extensive model testing, refinement, and system integration, renders it impractical for near-term operational use. Nevertheless, SAE remains a technique of high potential for addressing sample size limitations and will be prioritized for evaluation by NOAA Fisheries in the longer-term.

Addressing of Imprecise Outlier Catch Estimates within Stock Assessment Models

SEFSC has explored methods of estimating catch within stock assessment models (Williams and Shertzer 2024). The work explored if the models can accurately separate true signals from measurement errors in fisheries catch data, particularly when the data are sparse and prone to anomalous spikes. Several scenarios were tested to evaluate computationally constrained and unconstrained model performance under a suite of data conditions: 1) a false spike, 2) a true spike, 3) a random spike (of particular relevance to practical applications where it is unknown if the spike is true or due to observation error) and 4) data with high accuracy (i.e., a random spike with higher sample sizes/low observation error). Overall, the models performed best under the false spike and high accuracy conditions: in scenarios with a false spike, the unconstrained estimation approach outperformed the constrained approach, while the constrained method yielded better results when a true spike was present. In the random spike scenario, estimates produced by the unconstrained method tended to deviate, although in the high-accuracy data scenario, the unconstrained approach performed reasonably well.

A key insight from these simulations is the importance of understanding the underlying cause of data spikes. If spikes result from observation error, they can often be appropriately estimated; however, if they reflect true changes in the system, estimation becomes considerably more difficult. Additionally, estimates in the terminal year of the assessment period were found to be unreliable. This raises the critical question of whether the method provides a safeguard when the nature of the spike is unknown, which is often the case with imprecise MRIP estimates in sparse time series.

While the method shows promise and offers valuable insights for various applications, the OST-SEFSC working group chose to focus on improving the precision of estimates prior to their use in assessment models for several key reasons. First, the method is applicable primarily to sparse-spike time series and does not address the broader range of imprecise estimate scenarios commonly encountered in the MRIP time series. Second, its performance was limited in handling random spikes—situations that represent a significant portion of imprecise MRIP estimates, particularly when auxiliary data are insufficient to determine whether a spike reflects a true signal or observation error. Finally, there is a critical need for consistency in the estimates used for both stock assessments and management applications, particularly in supporting alignment between assessments and Annual Catch Limit (ACL) monitoring. Using different inputs for these purposes can create challenges in communicating results clearly to stakeholders and the public, potentially undermining confidence in the process.

Multi-year Aggregation

Prior work has been done to investigate methods to refine recreational catch estimates for species that are infrequently encountered by MRIP. Using data from 2004-2018 across 21 species Opsomer et al. 2020 explored both centered moving averages (i.e., using data from both prior and future years to produce an estimate for the target year) and lag moving averages (i.e., using only prior years to produce an estimate for the target year) at both 3 and 5 year aggregations to improve stability and precision of estimates. Results suggested that moving averages of 5 years generally perform best in terms of smoothing and increasing estimate precision, however, the lag moving averages are more useful for improving estimates on an on-going basis. Authors ultimately recommended 5-year multi-year averages for improving historical time series and 5-year lag moving averages for setting catch limits for current years.

In considering the pros and cons of multi-year aggregation at the annual catch level, the workgroup considered what dynamics might be lost by smoothing over data in this manner. One concern was the potential to mask real trends and shifts in recreational fisheries over time, which led the workgroup to explore smoothing catch rates at the wave level instead. The primary rationale is that while effort estimates are generally precise and reflect true trends in fishing activity, catch rates are much more variable and sensitive to small sample sizes, making them the main source of imprecision and instability in the MRIP estimates. From a population dynamics perspective, catch rates should theoretically reflect underlying abundance—being higher when abundance is high and lower when abundance is low—but in recreational fisheries, catch rates are also influenced by a variety of non-biological factors (e.g., fisher behavior, targeting practices), which introduce additional noise. Therefore, smoothing catch rates at a finer temporal scale (i.e., wave level) is reasonable to reduce this variability, as catch per unit effort should be inherently more stable than it can appear in the raw MRIP data. Moreover, the stock assessment scientists in the workgroup highlighted that precise year-to-year catch rate estimates are not critically important for their assessments, further supporting the focus on stabilizing catch rate variability rather than total catch (catch rate X effort).

Methods

We evaluated several multi-year averaging methods against each other and the original MRIP estimation methods (i.e., standard design-based methods). The text below summarizes the original MRIP estimation methods for context and describes the multi-year averaging approaches that were evaluated: lag and centered moving averages of annual catch and centered moving averages of both annual catch rates and wave-level catch rates. For this analysis, we did not include lag approach for the annual and wave-level catch rates because early testing indicated it had a stronger distorting effect on the time series, but retained the lag annual catch moving average to provide a general indication of the performance of the lagged methods.

We further describe the simulation approach and the comparative metrics used to evaluate the performance of the different multi-year averaging approaches. To focus the evaluation, we limited our analyses to private boat estimates.

Original MRIP Estimation Method

Our standard method for estimating private boat recreational catch estimates combines data from two surveys: the Access Point Angler Intercept Survey (APAIS), which is used to produce catch rates (i.e., catch per angler trip), and the Fishing Effort Survey (FES), which is used to produce estimates of total effort (total angler trips). These components are multiplied to produce total catch estimates, landings and releases. For the purposes of our analyses we will focus here on the catch rate X effort and associated variance components—full details on the design-based statistical methods and effort estimation can be found in our survey methodological documentation (NMFS 2025b).

Catch Rate Estimation

Catch rate estimates for two-month waves are produced using a standard weighted mean estimator (SAS Institute Inc., 2020):

$$\underline{\hat{y}}_d = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w I_d y}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w \; I_d}$$

where $\underline{\hat{y}}_d$ is the estimated mean catch per angler trip in domain d;

$h = 1, \ldots H$ represents the strata, defined by temporal and geographic variables, as well as fishing access site group

$i = 1, \ldots n_h$ represents the primary sampling units (site clustered, days and time intervals) sampled within stratum h

$j = 1, \ldots m_{hi}$ represents the angler trips sampled in primary sampling unit i;

w is the APAIS sample weight for angler trip j;

$I_{d(h,i,j)}$ indicator variable (1 if in domain , 0 otherwise) for angler trip j; and

y is the number of fish caught on angler trip j.

The variance is estimated using a Taylor series linearization (Dienes, 1957; SAS Institute Inc., 2020), which calculates the sum of squared deviations of the weighted primary sampling unit totals from the weighted mean within each stratum, adjusted by the number of primary sampling units as follows:

$$\hat{V}(\underline{\hat{y}}_d) = \sum_{h=1}^{H} \left(\frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\frac{\left(\sum_{j=1}^{m_{hi}} w \; I_d (y - \underline{\hat{y}}_d)\right)}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w \; I_d} - \right.\right.$$

$$\left.\left.\frac{\sum_{i=1}^{n_h} \frac{(\sum_{j=1}^{m_{hi}} w \quad I_d(y \quad - \underline{\hat{y}}_d))}{\sum_{h=1}^{H} \quad \sum_{i=1}^{n_h} \quad \sum_{j=1}^{m_{hi}} w \quad I_d}}{n_h}\right) \quad \right)^2.$$

Total Catch (Landings, Releases, Landings and Releases) Estimation

Standard total catch ($\hat{Y}_d$) for a given species, fishing mode and two-month wave is estimated by multiplying the catch rates from the APAIS by the total effort estimates from the FES.

$$\hat{Y}_d = \underline{\hat{y}}_d \times \hat{T}_d$$

where $\underline{\hat{y}}_d$ is the APAIS catch rate for the wave and $\hat{T}_d$ is the total FES effort estimate for the wave.

The variance $\hat{V}(\hat{Y}_d)$ is estimated using Goodman's formula (Goodman 1960).

$$\hat{V}(\hat{Y}_d) = \left(\underline{\hat{y}}_d\right)^2 \hat{V}(\hat{T}_d) + (\hat{T}_d)^2 \hat{V}\left(\underline{\hat{y}}_d\right) - \hat{V}(\hat{T}_d)\hat{V}\left(\underline{\hat{y}}_d\right).$$

Annual total catch, and associated variance, are then estimated by summing the relevant wave estimates.

The PSE is the standard error of the estimate (i.e., the square-root of the variance) expressed as a percentage of the estimate itself.

Multi-year Aggregation Alternatives

Annual Catch - Lag and Centered Moving Average

The estimated simple moving average of annual catch for a given year ($\hat{Y}_y^*$) is calculated as the arithmetic mean of annual catch estimates within a specified n year interval (i):

$$\hat{Y}_y^* = \frac{1}{n}\sum_{y \in i} \hat{Y}_y$$

where $\hat{Y}_y^*$ is the estimated moving average of annual catch for a given year,
$\hat{Y}_y$ is the estimated annual catch for year y included in the moving average calculation,
n is the number of years included in the moving average window (e.g., 3 or 5),
i represents the set of years included in the moving average calculation - this interval is

defined as either a lagging window (e.g., $\{y, y-1, \ldots y-n+1\}$) or a centered window (e.g., for $n = 3, \{y-1, y, y+1\}$).

The variance ($\hat{V}(\hat{Y}_y^*)$) is then estimated as

$$\hat{V}(\hat{Y}_y^*) = \left(\frac{1}{n}\right)^2 \sum_{y \in i} \hat{V}(\hat{Y}_y)$$

where $\hat{V}(\hat{Y}_y)$ is the estimated variance of the annual catch for year y.

Annual Catch Rate - Centered Moving Average

The estimated simple moving average of the annual catch rate for a given year ($\underline{\hat{y}}_y^*$) is calculated as the arithmetic mean of annual catch rate estimates within a specified n year interval (i):

$$\underline{\hat{y}}_y^* = \frac{1}{n} \sum_{y \in i} \underline{\hat{y}}_y$$

where $\underline{\hat{y}}_y^*$ is the estimated moving average of annual catch rate for a given year,

$\underline{\hat{y}}_y$ is the estimated annual catch for year y included in the moving average calculation (derived from APAIS, as described above),

n is the number of years included in the moving average window (e.g., 3 or 5),

i represents the set of years included in the moving average calculation - for the catch rate aggregation, we used solely a centered window (e.g., for $n = 3, \{y-1, y, y+1\}$).

The variance ($\hat{V}(\underline{\hat{y}}_y^*)$) is estimated as

$$\hat{V}(\underline{\hat{y}}_y^*) = \left(\tfrac{1}{n}\right)^2 \textstyle\sum_{y \in i} \hat{V}(\underline{\hat{y}}_y)$$

where $\hat{V}(\underline{\hat{y}}_y)$ is the estimated variance of the annual catch rate for year y.

To produce the total annual catch ($\hat{Y}_y^{**}$), the smoothed annual catch rate is multiplied by the original (un-smoothed) total effort estimate for that year ($\hat{T}_y$):

$$\hat{Y}_y^{**} = \underline{\hat{y}}_y^* \times \hat{T}_y$$

and the variance ($\hat{V}(\hat{Y}_y^{**})$) is estimated using Goodman's formula:

$$\hat{V}(\hat{Y}_y^{**}) = \left(\underline{\hat{y}}_y^*\right)^2 \hat{V}(\hat{T}_y) + (\hat{T}_y)^2 \hat{V}\left(\underline{\hat{y}}_y^*\right) - \hat{V}(\hat{T}_y)\hat{V}\left(\underline{\hat{y}}_y^*\right).$$

Wave-Level Catch Rate - Centered Moving Average

The estimated simple moving average of the catch rate for a given year and wave ($\underline{\hat{y}}_{y,w}^*$) is calculated as the arithmetic mean of wave-level catch rate estimates within a specified n year interval (i):

$$\underline{\hat{y}}^*_{y,w} = \left(\frac{1}{n}\right) \sum_{(y\prime,w\prime)\in i} \underline{\hat{y}}_{y\prime,w\prime}$$

where $\underline{\hat{y}}^*_{y,w}$ is the estimated moving average of the wave-level catch rate for year y and wave w.

$\underline{\hat{y}}_{y\prime,w\prime}$ is the estimated catch rate for year y' and wave w' included in the moving average calculation (i.e., $\underline{\hat{y}}_d$ derived from APAIS described above),

n is the number of years included in the moving average window (e.g., 3 or 5),

i represents the set of years included in the moving average calculation – for the catch rate aggregation, we used solely a centered window ((e.g., for $n = 3, \{(y, w - 1), (y, w), (y, w + 1)\}$).

The variance ($\hat{V}(\underline{\hat{y}}^*_{y,w})$) is estimated as

$$\hat{V}(\underline{\hat{y}}^*_{y,w}) = \left(\frac{1}{n}\right)^2 \sum_{(y\prime,w\prime)\in i} \hat{V}(\underline{\hat{y}}_{y\prime,w\prime})$$

where $\hat{V}(\underline{\hat{y}}_{y\prime,w\prime})$ is the estimated variance of the wave-level catch rate in year y' and wave w'.

To produce the wave-level catch with the smoothed wave-level catch rate, the smoothed wave-level catch rate is multiplied by the original (un-smoothed) total effort estimate for that wave, with the variance estimated using Goodman's formula. The annual catch is then calculated as the sum of those wave estimates, and the corresponding variance is estimated as the sum of the wave-level variance.

Simulation Approach

The simulation approach was parameter-driven and accounted for variation at multiple levels (Appendix B). Within each two-month wave, the simulation incorporated trip-level prevalence of catch, catch-per-trip rates, effort estimates with random jitter, and intercept survey sample sizes, also with random jitter applied to reflect sampling variability. Across waves, an overall effort trend was included to simulate broader temporal changes. The simulations covered a total period of 20 years with 5,000 replicates run to ensure robust estimation. Results were evaluated by aggregating data over different time spans, specifically 3-year and 5-year periods, to assess the impact of temporal aggregation on estimate precision.

To realistically reflect the MRIP catch time series, parameters were tuned such that the simulated "original" estimation method would closely approximate key features of real data. Tuning targets included mean annual catch landings, mean PSE, and the proportion of zero values in 20-year landings series. These targets were based on four real-world, data-limited time series selected to represent a gradient of data sparseness and variability of estimates that do not meet the MRIP precision standard of PSE<50:

- Very low precision scenarios (greater than 80 PSE on average for annual landings estimates from 2003-2022):

- *Florida east coast private boat snowy grouper landings*, encountered in about 3 out of every 10,000 angler intercepts.
- *North Carolina private boat hogfish landings*, encountered in approximately 1 out of every 10,000 angler intercepts.

- Marginal precision scenarios (approximately 50 PSE on average for annual landings estimates from 2003-2022):
  - *North Carolina private boat greater amberjack*, encountered in roughly 2 out of every 1,000 angler intercepts.
  - *Florida west coast private boat scamp*, also at about 2 out of every 1,000 angler intercepts.

The performance of the wave-level catch rate aggregation method was compared to a) a lag moving average of annual landings estimates b) a centered moving average of annual landings estimates, and c) a centered moving average of annual catch rates. Aggregation periods of 3, 5, 7, 9, and 11 years were tested for the very low precision scenarios, while 3- and 5-year periods were used for the marginal precision scenarios. Although 7- to 11-year aggregations would be impractical in real-world applications given that fisheries conditions typically change more frequently, these extended periods were included to explore the maximum potential improvements in precision for the most data-limited cases, particularly how many years of aggregation might be required to meet the MRIP precision standard of PSE<50.

All methods were evaluated based on the following set of performance metrics to assess both precision and accuracy of all methods tested:

- Percent standard error
- Bias relative to the true landings series in the simulation
- The proportion of zeros in the time series
- Correlation with the true landings
- Comparisons between the observed and simulated estimates
- Mean PSE and the proportion of zeros compared between the simulated estimates and the true values
- Comparison of annual point estimates and time series trends from each method to both the true catch and the estimates produced using the original estimation method, based on the mean of 5,000 simulation replicates and evaluated across 3-year and 5-year aggregation periods

Results

The simulation replicated real-world conditions across multiple fisheries, as shown by the close alignment between target values and simulation results over a 20-year average across 5,000 replicates (Table 1). For all cases, including landings of NC Hogfish, East FL Snowy Grouper, NC Greater Amberjack, and West FL Scamp, the simulated average annual estimates, PSEs, and proportions of zero-landings years were generally very close to the targets. This consistency across species and regions indicates the simulation was able to capture key dynamics in the observed data.

Table 1: Comparison of Simulation Results to Tuning Targets (Real Private Boat Landings)

Case	Average Annual Estimate (excluding zeros)		Average Annual Estimate PSE (excluding zeros)		Average Annual Estimate (including zeros)		Proportion of Zeros in the Annual Estimate Series	
	Target	Result	Target	Result	Target	Result	Target	Result
East FL Snowy Grouper	5,773	4,946	82.62	89.67	2,886	2,902	0.50	0.41
NC Hogfish	1,369	1,338	93.42	91.14	684	668	0.50	0.50
NC Greater Amberjack	7,313	7,082	53.77	52.73	6,947	7,036	0.05	0.01
West FL Scamp	43,390	43,601	53.91	51.33	43,390	43,395	0	0.005

In terms of precision improvements, all multi-year averaging methods performed similarly across the four cases, with catch rate aggregations providing a slight advantage over annual landings aggregations (Figure 1). For the lowest precision examples, at least nine years of aggregation were required to reduce the PSE below 50 to meet the MRIP precision standard. In contrast, for the marginal precision examples, meaningful improvements were achieved with as little as a three-year aggregation (Figure 1).

With respect to relative bias, all methods performed similarly, introducing minimal bias across scenarios (Figure 2). The centered annual catch rate moving average produced the highest bias, but even this was minor (no greater than 0.10 in any aggregation scenario).

In terms of reducing the proportion of zeros in the time series, the centered aggregation methods, including the centered moving averages of annual landings, annual catch rates, and wave-level catch rates, performed similarly and were effective in filling data gaps (Figure 3). The lagged moving average of annual landings was the least effective in the very low precision scenarios but performed comparably to the other methods in the marginal precision examples (Figure 3).

The wave-level catch rate moving average showed the strongest positive correlation with true landings among all estimators, suggesting the additional smoothing applied in the annual-level aggregations reduced their alignment with expected catch (Figure 4). The lag and centered annual catch moving averages (Methods 2 and 3 in Figure 4) effectively smoothed both catch rate and effort across waves and multiple years. The annual catch rate moving average (Method 4 in Figure 4) smoothed catch rate over waves and years, while effort is only smoothed over waves. The wave-level catch rate estimation method (Method 5 in Figure 4) has the least amount of smoothing—catch rate is smoothed only over years, and effort is not smoothed over waves or years. Overall, there's an inverse relationship between the correlation of an estimation method with true landings and the level or degree of smoothing/aggregation. This pattern holds true even for the catch rate based methods (Methods 4 and 5 in Figure 4), as evidenced by the decline in correlation values observed at 9- and 11-year aggregation periods.

While the simulation results are informative and directionally aligned with observed results in the four example cases, they do appear to be optimistic in terms of expected precision gain (Figure 5). This outcome is likely due in part to incomplete tuning of the simulations to the real-world cases and the limited number of simulation parameters. Despite this, the key findings are consistent: all methods improved precision to a similar extent, but only the highest levels of aggregation (e.g., 9 and 11 years) achieved the MRIP PSE standard in the two very imprecise cases (East Florida snowy grouper and North Carolina hogfish). In contrast, all methods succeeded in meeting the precision standard for the two marginally precise cases (West Florida scamp and North Carolina Greater Amberjack) under both the 3- and 5-year aggregation periods.

Figure 1: Mean Percent Standard Error of Annual Private Boat Landings by Method and Multi-Year Averaging Period: Select Species Examples (20 Year Simulation)	
Very Low Precision Examples (Original PSE > 80)	*Marginal Precision Examples (Original PSE ~ 50)*

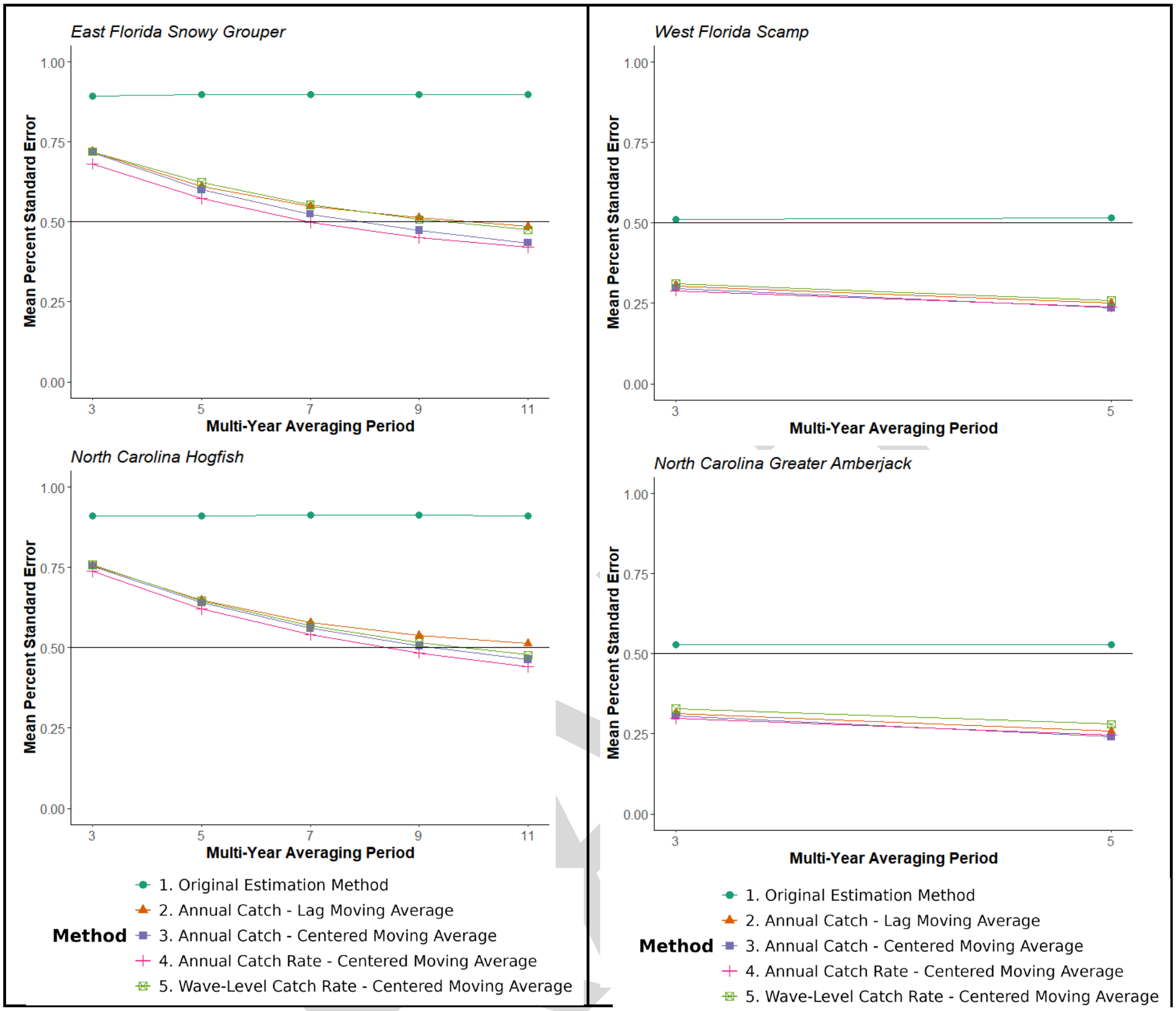


Figure 2: Mean Relative Bias of Annual Private Boat Landings by Method and Multi-Year Averaging Period: Select Species Examples (20 Year Simulation)

Very Low Precision Examples (Original PSE > 80)	*Marginal Precision Examples (Original PSE ~ 50)*

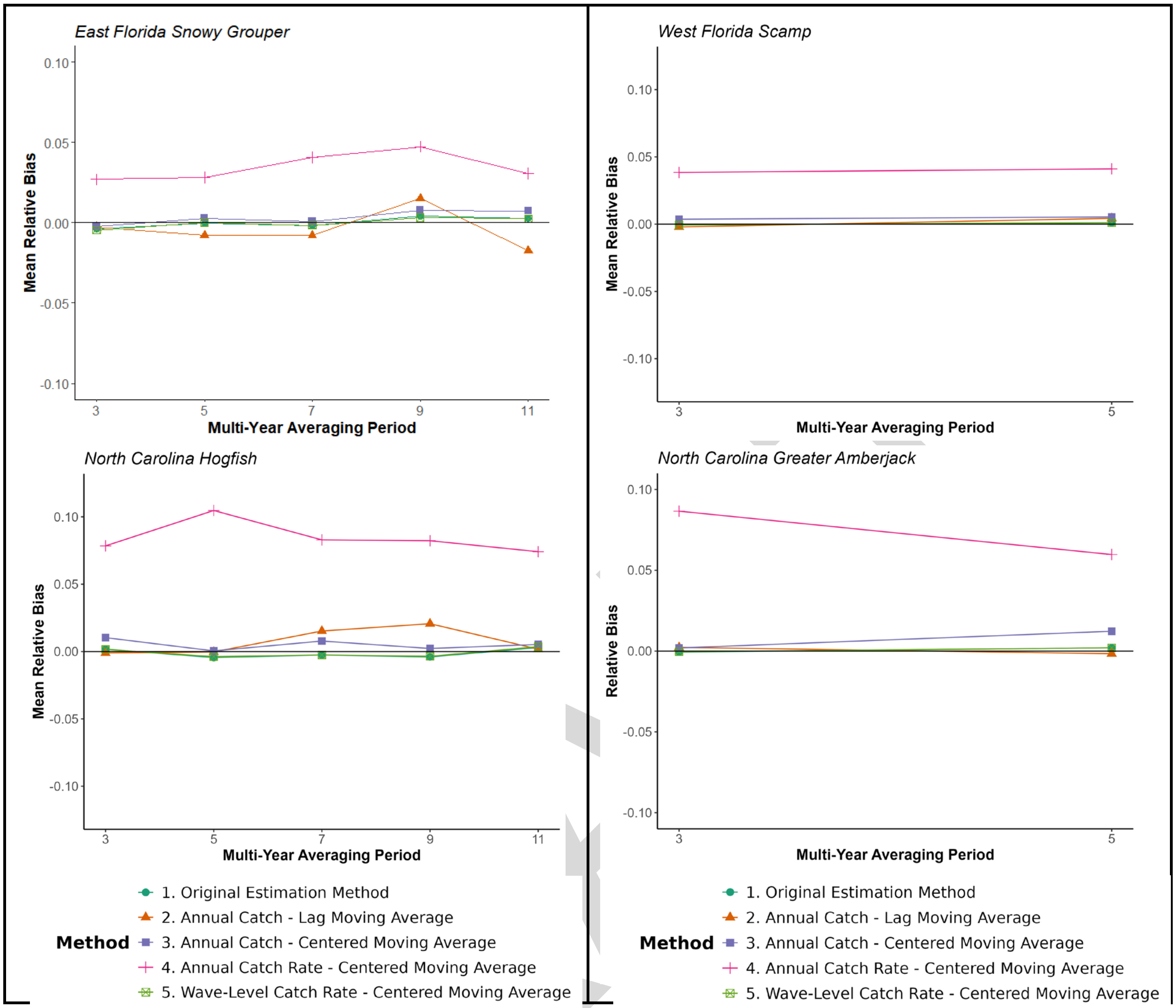


Figure 3: Proportion of Zero Catch Values by Method and Multi-Year Averaging Period: Select Species Examples (20 Year Simulation)	
Very Low Precision Examples (Original PSE > 80)	*Marginal Precision Examples (Original PSE ~ 50)*

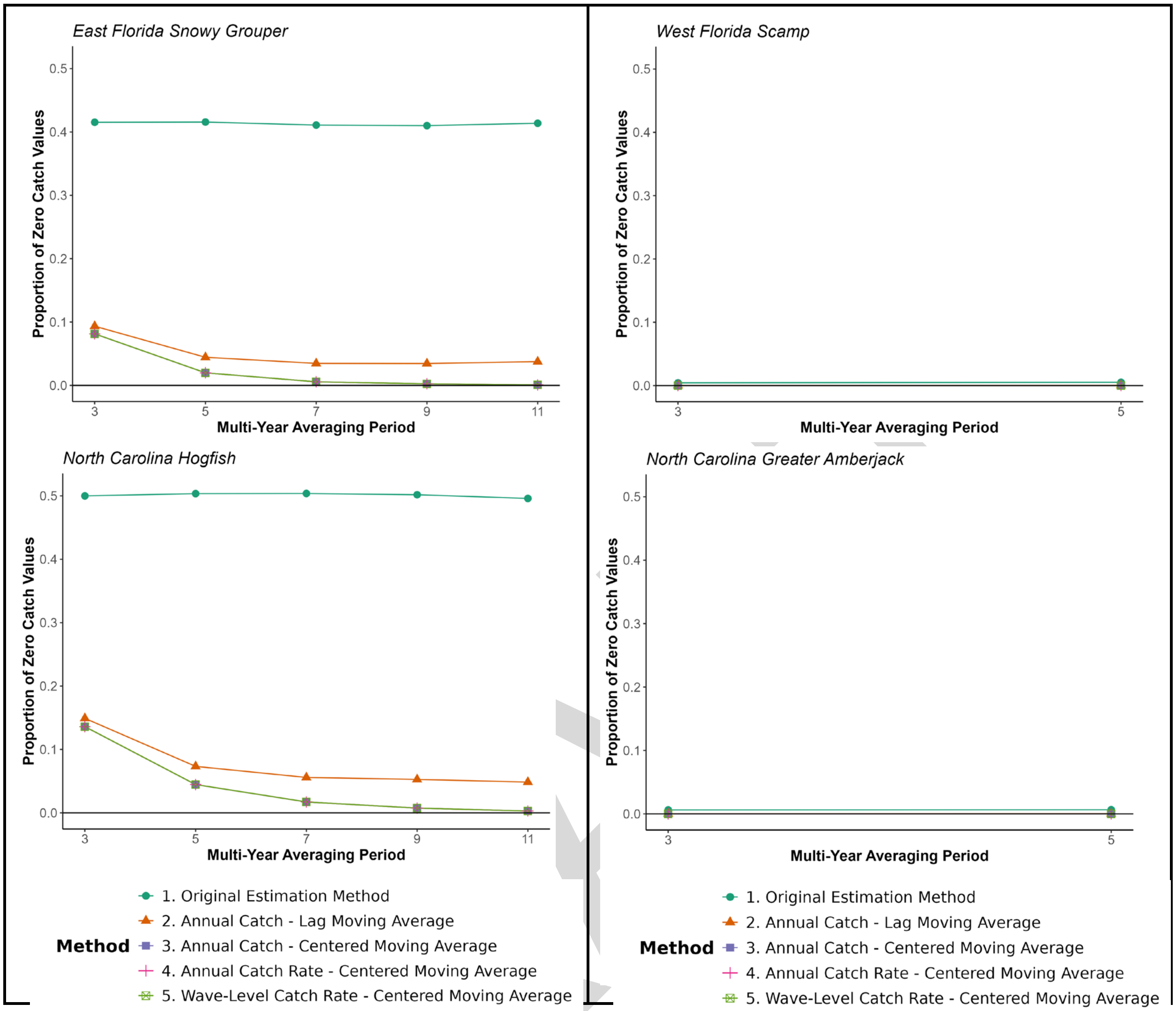


Figure 4: Correlation of Estimated Landings with True Landings by Method and Multi-Year Averaging Period: Select Species Examples (20 Year Simulation)

Very Low Precision Examples (Original PSE > 80)	*Marginal Precision Examples (Original PSE ~ 50)*

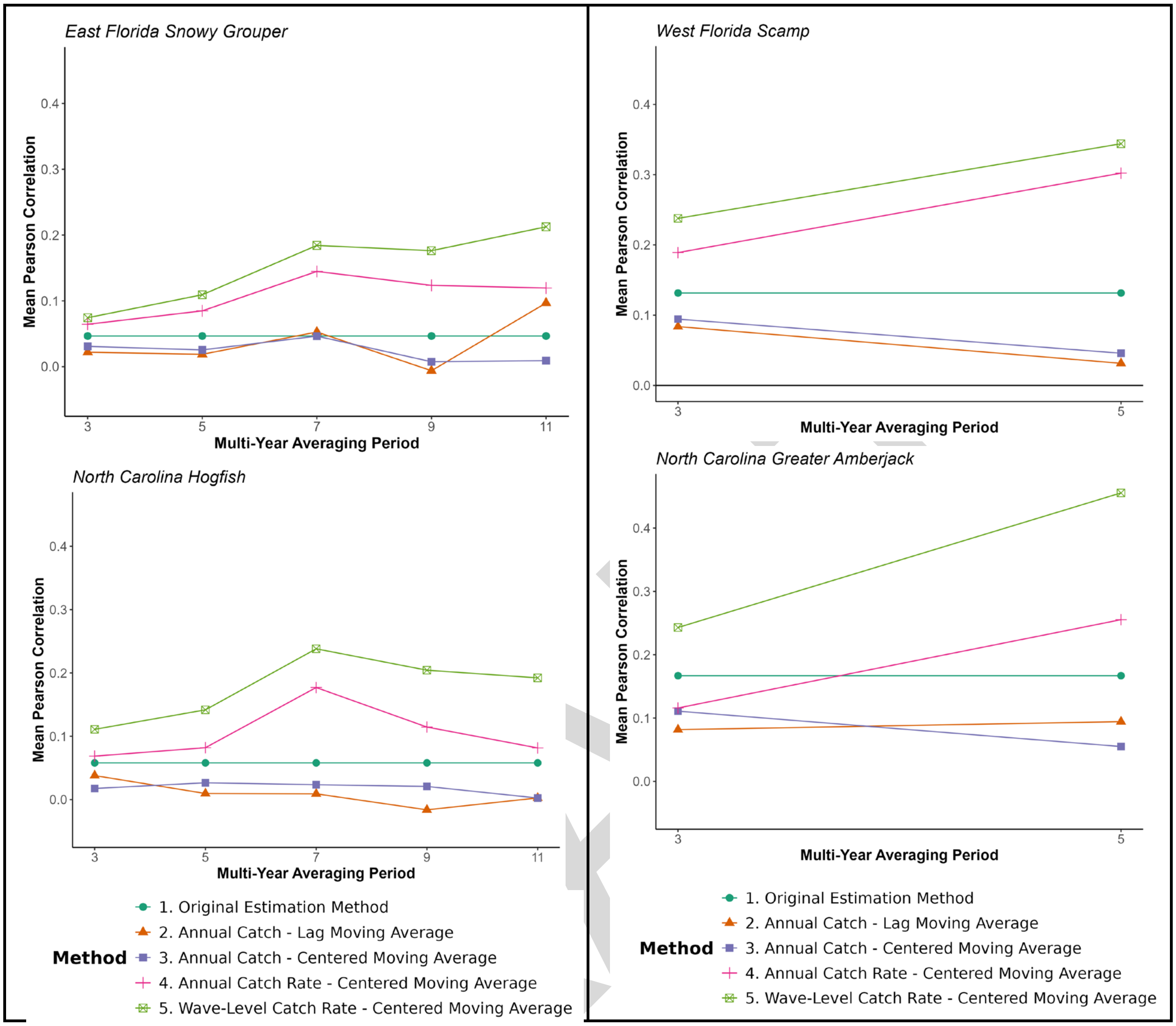
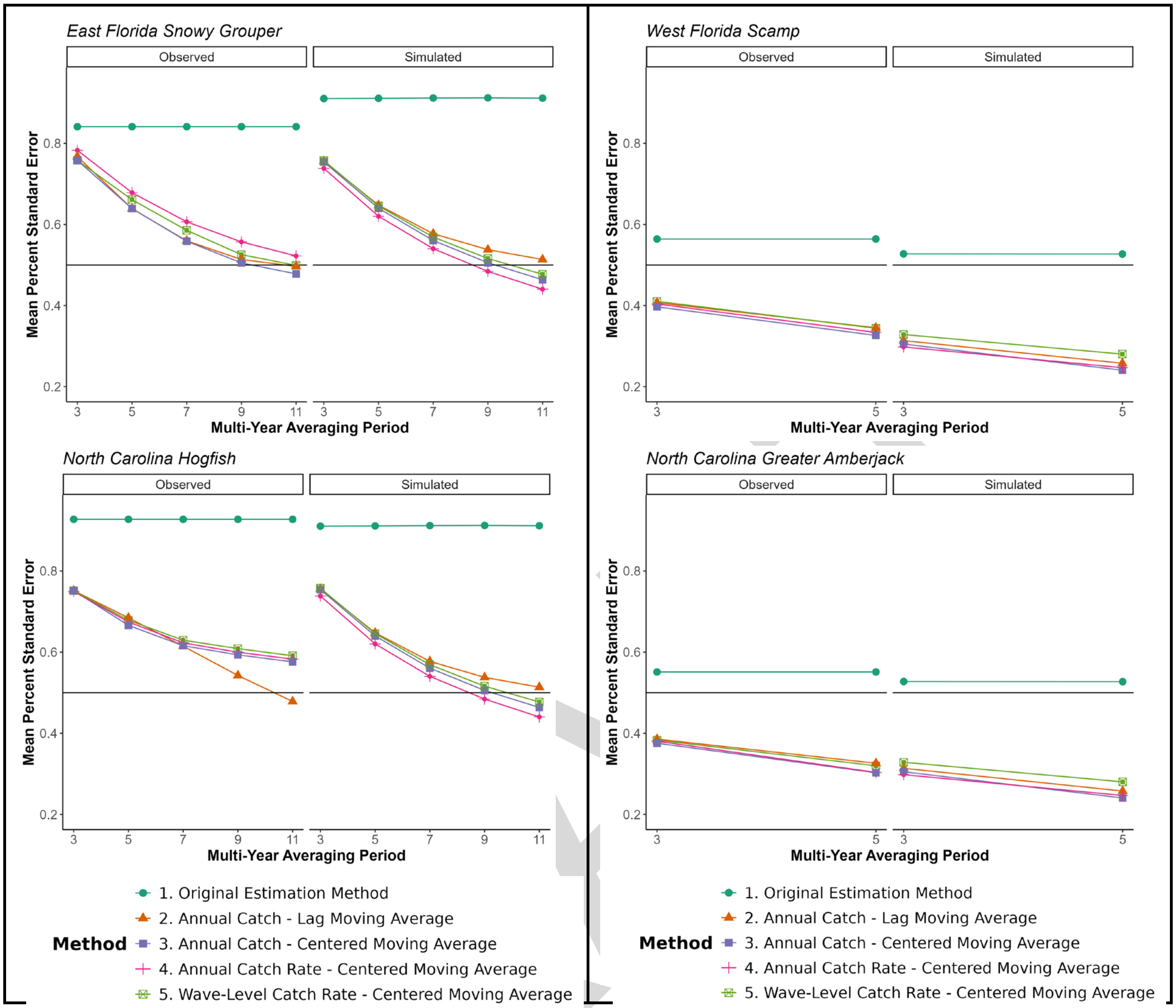


Figure 5: Mean Percent Standard Error (Observed vs Simulated) of Annual Private Boat Landings by Method and Multi-Year Averaging Period: Select Species Examples (20 Year Simulation)	
Very Low Precision Examples (Original PSE > 80)	*Marginal Precision Examples (Original PSE ~ 50)*



Effect of Multi-year Aggregation Methods on Annual Point Estimates and Time Series Trends

In evaluating the performance of the four multi-year aggregation methods relative to both the original catch estimation method and the true catch series, the original estimation method and the wave-level catch rate centered moving average consistently show the closest alignment with the true catch time series over the 20-year simulation period (Appendix A). Among the smoothing methods, the annual catch centered moving average (Method 3 in Appendix A) also performs well in capturing overall trends but introduces a higher degree of smoothing, which can reduce sensitivity to interannual variability. The annual catch lag moving average (Method 2 in Appendix A), as expected, displays a delayed response to annual fluctuations and often underrepresents turning points in the catch trajectory.

Extending the aggregation window from three to five years increases smoothing and reduces noise but also further dampens the ability to detect finer-scale interannual changes. This reinforces the trade-off between variability reduction and temporal responsiveness observed in prior analyses.

A key insight from this annual look at bias—rather than the relative bias examined over the entire time period (Figure 2)—is that the wave-level aggregation method's annual means are consistently closer to the true values each year. This is a notable advantage over the other aggregation methods, which, despite achieving a similar level of low bias across the full time series, could be much further from the truth in any given year.

Discussion

The simulation results suggest that multi-year averaging methods can improve the precision and stability of recreational catch estimates in data-limited situations. While all methods demonstrated gains in precision and effectively reduced the proportion of zeros in the time series, the wave-level catch rate moving average maintained the strongest correlation with true landings relative to the other estimators evaluated. This suggests that the wave-level catch rate moving average may be the most promising approach, given it was well-suited to smoothing variability in sparse datasets while preserving key signals in the data.

While the differences in performance among methods were relatively modest, the wave-level catch rate moving average has conceptual advantages that make it a compelling choice. Specifically, it avoids over-smoothing real year-to-year fluctuations, which is an important consideration for interpreting meaningful trends in a fishery and/or fish stock. The more heavily smoothed annual-level approaches, while delivering the largest raw gains in precision, often obscure these effects. This distinction, while subtle in results, plays out meaningfully in the logic and integrity of catch estimation.

The benefits of any method are influenced by species prevalence and aggregation period. Species that appear more frequently in the data benefit more substantially from multi-year averaging, with precision improvements evident even at the 3- to 5-year aggregation levels. In contrast, rare-event species saw more limited gains, particularly under short aggregation windows. While the prevalence of these rare-event species is typically fixed based on ecological, fisheries- and sampling-related realities, shifts in prevalence can occur over time. As such, it is appropriate to routinely evaluate these trends for changes, recognizing that such evaluations may not be feasible for all rare-event species due to ongoing data limitations.

Of note, the wave-level catch rate moving average is not a universal solution to all issues associated with estimating catch in rare-event or low-precision datasets. The method cannot fully overcome the limitations posed by extremely sparse data or very small sample sizes, particularly for stocks with inherently rare landings. Moreover, precision gains for the sparsest data scenarios required aggregation periods of at least nine years to meet the MRIP precision standard. This level of aggregation is likely impractical in most real-world management settings where fishery conditions, stock status, and regulatory measures change more frequently.

To address these realities while offering a practical path forward, we propose a focused and standardized application of the wave-level catch rate moving average.

Recommended Use of the Wave-Level Catch Rate Moving Average

We recommend the following for the application of this method:

Use in cases exceeding the MRIP PSE threshold (PSE > 50): This method can readily bring estimates that are already near the precision threshold below it, increasing stability and confidence in those values. While this method will not enable all imprecise MRIP estimates to meet the precision threshold, applying this method will still offer valuable benefits by smoothing time series, reducing interannual variability, and improving the interpretability of imprecise estimates.

Limit aggregation to 3-5 years: While longer periods may yield greater gains, they are typically incompatible with dynamic fishery conditions and management cycles.

- The decision of using 3- or 5-year aggregation should be informed by species life history, fishery dynamics, and data availability. Consideration of management cycles for a particular fishery may also play a factor, as well as management history—a major regulatory change, for instance, could warrant adjusting the specific years used. Generally:
    - **3-year aggregation** offers more responsiveness to recent changes, making it appropriate for short-lived species, rapidly changing populations, or estimates already close to the precision threshold of 50 PSE.
    - **5-year aggregation** provides greater smoothing and stability, better suited for long-lived, slower-growing species or situations with very low precision and particularly sparse data.

Apply a centered moving average whenever possible: Centering reduces bias by incorporating data both before and after the year being estimated, avoiding the trend distortion that can arise in lag-based (one-sided) methods.

- **For terminal year estimates:** A centered moving average cannot be computed for the most recent years of a time series, as the required future data points are not yet available. To produce a timely, smoothed estimate for these terminal years—which are often critical for ACL monitoring and other management applications—the calculation should **transition to a lag moving average**. This hybrid approach would allow for the use of a more stable and balanced centered average for the historical part of a given time series while still providing a timely, smoothed estimate for the most recent years needed for management.

Automate and standardize its application while accounting for management history: The goal is to deploy this method consistently across species to enhance understanding and transparency in how catch estimates are derived, facilitate routine updates across assessment cycles, and support more efficient review and communication of results with stakeholders. However, standardization should not override critical evaluation. A key step within this framework should be to consider whether management history (which is typically not consistent across species) may affect the chosen averaging period. For instance, analysts should check if significant management changes have caused drastic shifts in catch rates within a potential year range before applying the moving average. While this adds a layer of expert review, it would ensure the automated application is appropriate and robust.

Re-evaluate species prevalence over time: Monitoring for changes in data sparsity ensures the method remains appropriate as fishery conditions evolve. Significant shifts (e.g., increased landings or declining encounter rates) may warrant reassessing the chosen aggregation period or whether the moving average approach remains appropriate under evolving fishery conditions.

Consider the full time series context when deciding where to apply the method: The choice to apply this method to part or all of a time series should reflect the overall pattern of data quality rather than isolated precision issues. For instance, if only one estimate in an otherwise precise time series exceeds the PSE threshold, applying the moving average may not be warranted for the entire series. Conversely, if the majority of estimates in a time series are imprecise but one year appears precise, averaging across the full series may be more appropriate. In intermediate cases, expert judgment should guide the decision. This judgment should be based on:

- **Distribution of PSE values across the time series**: If most years fall just below the PSE threshold and only a few exceed it, it may be appropriate to apply the method only to those outlier years. However, if a majority of years exceed the threshold, the method should be applied across the full series to enhance interpretability and reduce variability.
- **Magnitude and pattern of variability**: When high PSE values are isolated and surrounded by more precise estimates, selective application may be justified. If, however, imprecision is part of a recurring pattern (e.g., every other year or entire clusters of years), applying the moving average to the full time series is likely warranted to provide a clearer view of underlying trends.
- **Underlying causes of variability**: If variability is due to temporary or well-understood disruptions (e.g., a one-time sampling gap or a brief fishery closure), it may be reasonable to limit averaging to affected years. Conversely, if the causes are persistent or structural (e.g., low encounter rates, shifting effort, or geographic redistribution of effort), smoothing the full series can better support stable interpretation over time.
- **Importance of preserving temporal resolution**: For fisheries where year-to-year changes are highly relevant (e.g., short-lived species or fast-changing stocks), it may

be preferable to apply the method only where necessary. In contrast, for long-lived species or relatively stable fisheries, applying the method to the full time series may be more appropriate, especially where imprecision is widespread.

- **Intended use of the data**: For applications where clear and stable communication is key, broader application of the method may be useful. For assessment inputs, where methodological transparency and data resolution are critical, selective application, paired with clear documentation of decisions, may be more appropriate.

Maintain objectivity and transparency in decision-making: To preserve the credibility of the estimates and support sound science, the decision to apply (or not apply) the moving average must be based solely on data quality, methodological consistency, and the needs of the analysis. The perceived favorability of the resulting estimates should not be a factor in determining whether or how the method is applied.

Practical Implications for Assessment and Management

A centered, wave-level catch rate moving average may be particularly valuable for species or regions experiencing high interannual variability in recruitment or effort. The ability to dampen estimate variability without obscuring real trends supports more accurate stock assessments and enhances management advice. However, realizing the full benefit of this method will likely require management bodies to accommodate some degree of delayed data integration, especially to leverage centered methods that rely on post-season inputs.

Ultimately, the wave-level catch rate moving average should be regarded as a practical, interim tool that enhances the interpretability of catch estimates. It helps to address precision-related impacts to catch estimates, which often stem from challenges like increasing survey costs and limited sample sizes. It is crucial to recognize that this approach is not a one-size-fits-all solution. Its value lies in its logic, transparency, and ease of implementation, especially when deployed within a standardized framework that enables assessments and management timelines to remain on track despite ongoing data limitations.

Importantly, this method does not consider other information managers may have at their disposal, which can provide additional, important context for the interpretation of catch information. Consequently, its interpretation, and the decisions derived from it, necessitate integration with auxiliary data sources (e.g., fishery-independent surveys, biological indicators) to ensure a holistic understanding of the fishery system grounded in the best available science.

References

Atlantic Coastal Cooperative Statistics Program. (2016, February 26). *Proceedings of the workshop on percent standard error (PSE) of recreational fishing data (PSE Workshop Proceedings)*. https://www.accsp.org/wp-content/uploads/PSE_Workshop_Proceedings_FINAL_WEB-1.pdf

Chua, N. J. Y., & Long, B. Y. B. (2023). It's time to build: A small area estimation methodology for time-to-event data. *Statistical Journal of the IAOS, 39*(4), 793–802. https://doi.org/10.3233/SJI-230075

Dienes, P. (1957). *The Taylor Series: An Introduction to the Theory of Functions of a Complex Variable*. Dover Publications.

Goodman, L. A. 1960. *On the Exact Variance of Products*. Journal of the American Statistical Association 55(292):708-713. https://doi.org/10.1080/01621459.1960.10483369

National Academies of Sciences, Engineering, and Medicine. (2017b). Appendix C: Small-area modeling in space and time with multiple data sources. In *Improving crop estimates by integrating multiple data sources* (p. 125). Washington, DC: The National Academies Press. https://doi.org/10.17226/24892

National Academies of Sciences, Engineering, and Medicine. (2017a). *Review of the Marine Recreational Information Program*. The National Academies Press. https://doi.org/10.17226/24640

National Marine Fisheries Service. (2025a). *Recreational fishing survey and data standards*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. Retrieved July 22, 2025, from https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-survey-and-data-standards

National Marine Fisheries Service. (2025b). *Marine recreational information program survey design and statistical methods for estimation of recreational fisheries catch and effort*. https://www.fisheries.noaa.gov/resource/document/mrip-survey-design-and-statistical-methods

Office of Management and Budget. (2006). *Standards and guidelines for statistical surveys (Statistical Policy Directive No. 2 - Guideline 4.1.5)*. https://nces.ed.gov/fcsm/pdf/OMB_Standards_Guidelines_Statistical_Surveys.pdf

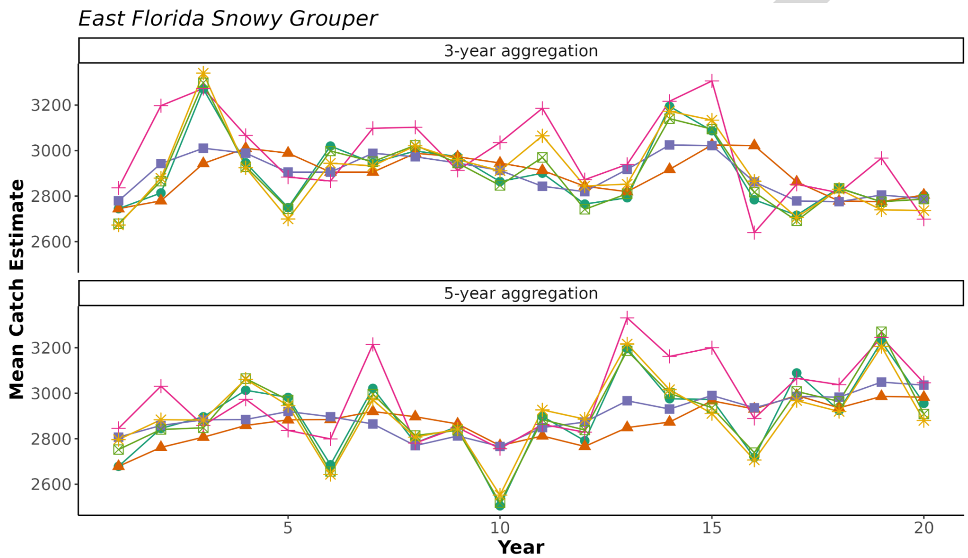
Opsomer, J., Chen, A., & Guillo, C. (2020, March 10). Improving estimation of rare event species in MRIP (SAFMC Rec Workgroup A3a). South Atlantic Fishery Management Council. https://safmc.net/wp-content/uploads/2022/06/A3a_Improving_Rare_Event_Estimates.pdf

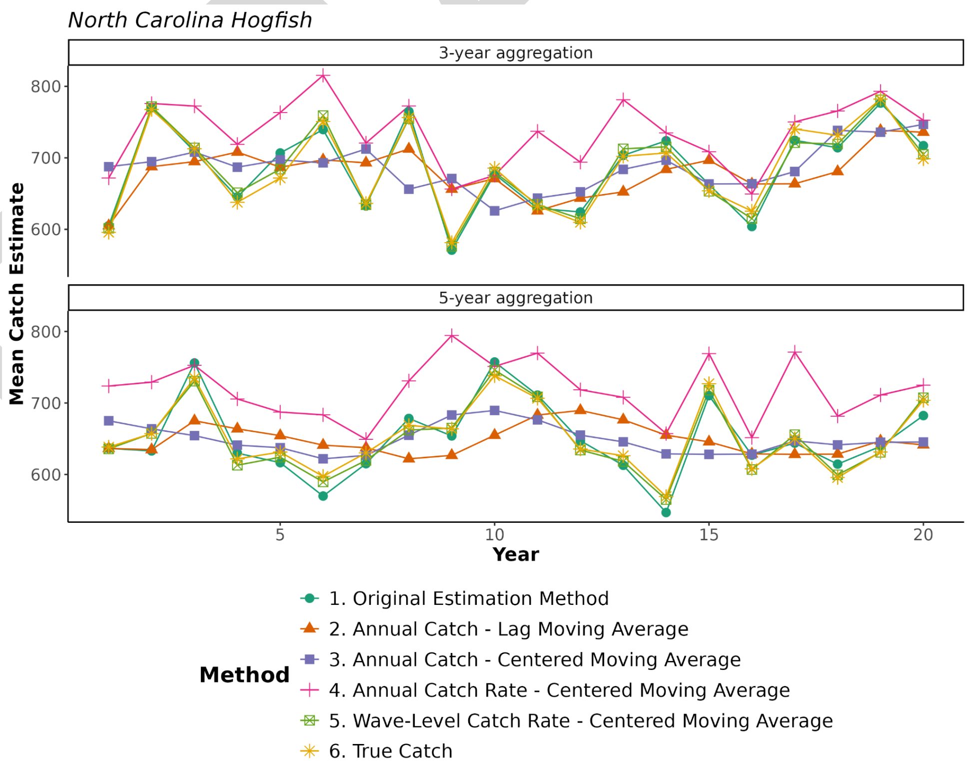
Quick, H., Hirsch, J. A., Moore, K. A., Terloyeva, D., & Diez Roux, A. V. (2019). *Borrowing information from neighbors: Estimating health in small areas*. Drexel University Urban Health Collaborative. https://drexel.edu/uhc/resources/briefs/Small%20Area%20Estimates%20Brief/

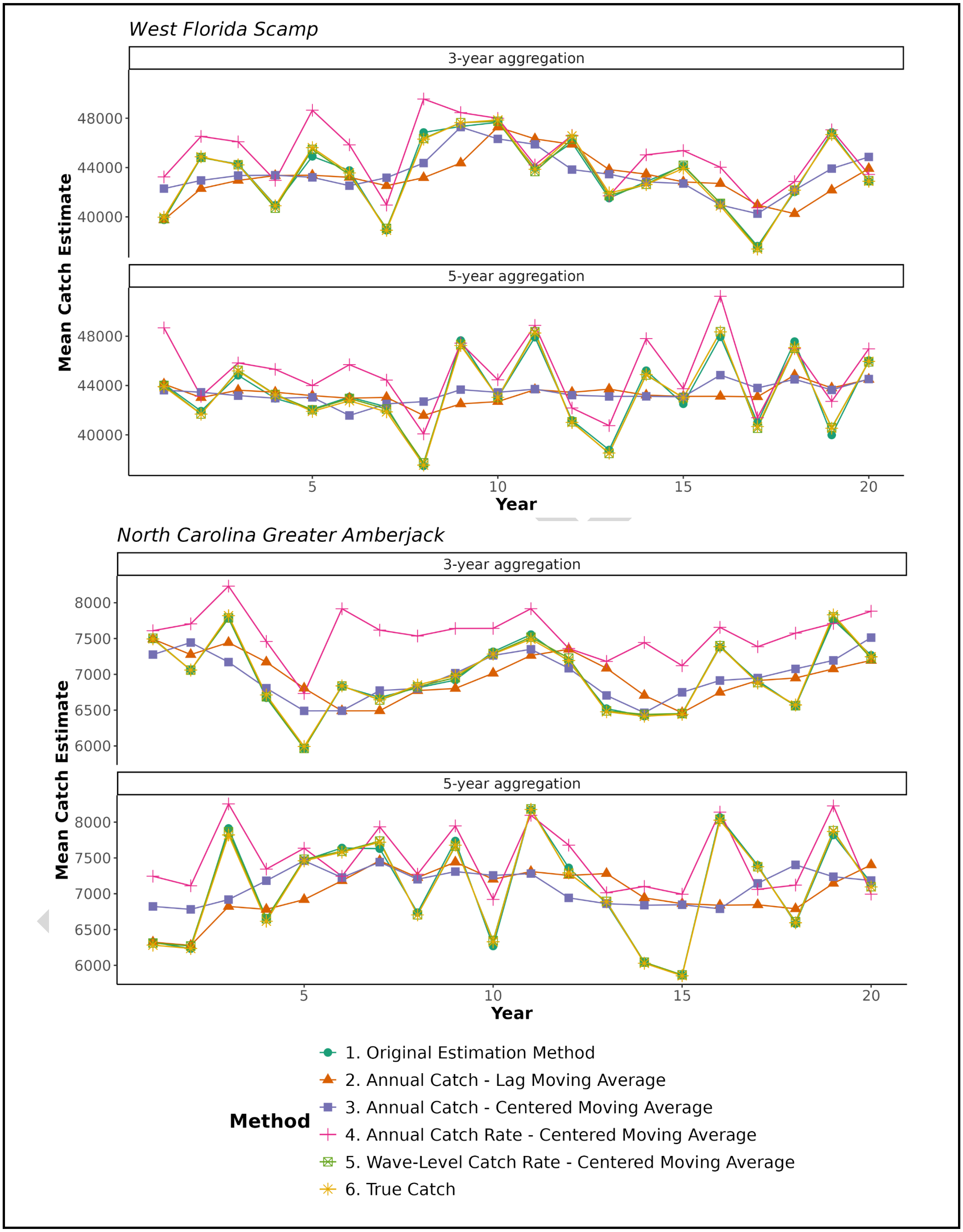
SAS Institute Inc. (2020). *SAS/STAT® 15.2 user's guide*. Cary, NC.

Appendix A

Catch Estimates by Estimation Method and Aggregation Period: Time series of mean catch estimates (across the 5000 replicates) generated using 3-year (top panel) and 5-year (bottom panel) aggregation periods, compared against the true catch and the original estimation method. Results are shown for four example species, selected to represent different levels of original estimate precision (East FL Snowy Grouper, NC Hogfish, West FL Scamp, and NC Greater Amberjack).






West Florida Scamp
3-year aggregation
5-year aggregation
Mean Catch Estimate
48000
44000
40000
5
10
15
20
Year
North Carolina Greater Amberjack
3-year aggregation
5-year aggregation
Mean Catch Estimate
8000
7500
7000
6500
6000
Method
1. Original Estimation Method
2. Annual Catch - Lag Moving Average
3. Annual Catch - Centered Moving Average
4. Annual Catch Rate - Centered Moving Average
5. Wave-Level Catch Rate - Centered Moving Average
6. True Catch

Appendix B: Simulation Description

The simulation was completed using SAS 9.4 and was designed to assess the performance of several multi-year averaging methods against the standard "original" annual estimate. The simulation generated a known "true" catch for a fish population over 20 years and then mimicked the process of sampling that fishery to produce catch estimates. The simulation was repeated 5,000 times for each scenario to ensure the results were statistically robust and to adequately characterize the distribution of potential outcomes.

Parameter Setup

The simulation is built around several key parameters that define the fishery being modeled. For each species case, these parameters define the underlying population and sampling characteristics for each of the 6 waves (bimonthly periods, denoted by subscript w) within a year (denoted by subscript y). Core parameters include:

- **Prevalence** (p_w): The probability of encountering the species on a given fishing trip, defined by the *&prev* macrovariable in the SAS code.
- **Catch per positive trip** (CPT_w): The expected number of fish caught on a trip that successfully encounters the species, which is a random variable drawn from a zero-truncated Poisson distribution (to ensure that a "positive trip" always reflects a catch of at least one fish). This is defined by the *&c_p_t* macrovariable in the code.
- **Fishing Effort** $(E_{y,w})$: The total number of fishing trips occurring in a given period (defined by *&effort* macrovariable in the code).
- **Sample size** $(n_{y,w})$: The number of trips sampled for the data (defined by *&n* in the code).

Core Logic

The simulation first calculates a **"true" total catch** (C_{true}) for each wave, which represents the real, unobserved total catch in the simulated fishery as:

$$C_{true.w} = p_w \times CPT_w \times E_{y,w}$$

Next, it simulates the real-world sampling process by generating observed catch $(C_{obs,w})$ from a sample of $n_{y,w}$ trips. Catch for trips with catch is drawn from a shifted Poisson distribution with a mean of CPT_w (i.e., the specified true catch per trip parameter). The total observed catch from the sample is then used to calculate the estimated catch for the wave $(\hat{C}_w)$ and year $(\hat{C}_y)$ which

$$\hat{C}_{y,w} = C_{obs,y,w} \times \frac{E_{y,w}}{n_{y,w}} \text{ and } \hat{C}_y = \sum_{w=1}^{6} \hat{C}_{y,w}$$

The simulation also calculates the variance for this estimate using established statistical formulas (described in the main text), to properly account for variability arising from both the catch rate and the fishing effort.

The core of the analysis is to compare this standard annual estimate with four alternative methods that use moving averages to smooth the data over a multi-year window, and evaluate all estimates against the known true catch to evaluate performance, which are all described in the main text of the report.