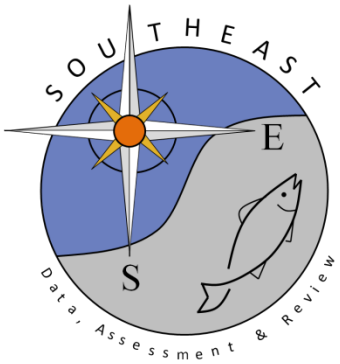
Transition Plan for Gulf State Recreational Fishing Surveys

Gulf of Mexico Subgroup of the MRIP Transition Team

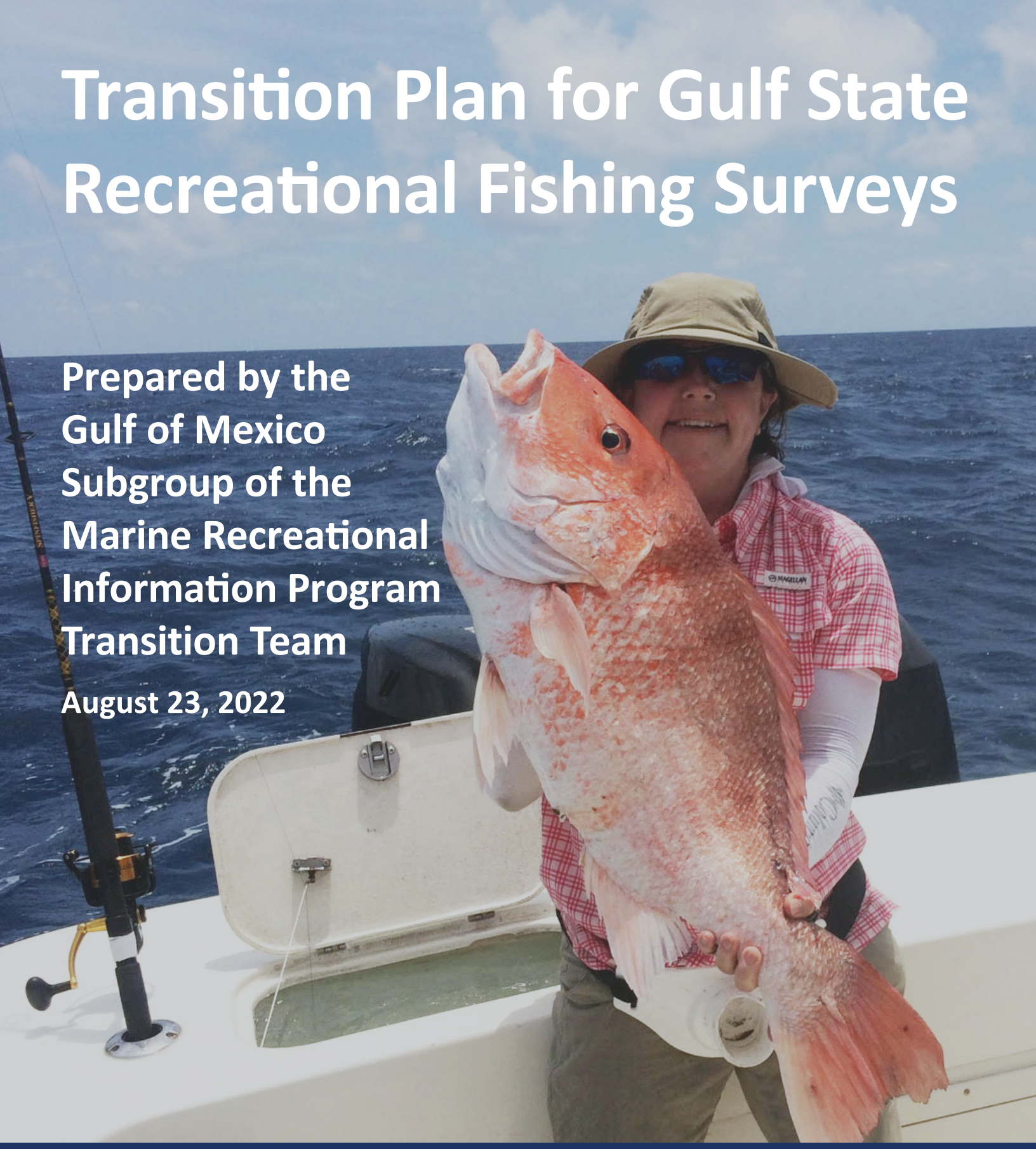
SEDAR79-RD-10

July 2024



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.

Transition Plan for Gulf State Recreational Fishing Surveys



Prepared by the Gulf of Mexico Subgroup of the Marine Recreational Information Program Transition Team

August 23, 2022

Alabama Department of Conservation and Natural Resources
Florida Fish and Wildlife Conservation Commission • Gulf Fisheries Information Network
Gulf of Mexico Fishery Management Council • Gulf States Marine Fisheries Commission
Louisiana Department of Wildlife and Fisheries • Mississippi Department of Marine Resources
NOAA Fisheries • Texas Parks and Wildlife Department

Table of Contents

I. Executive Summary **3**

II. Introduction and Purpose **5**

Background 5

Need for a Gulf State Recreational Catch and Effort Surveys Transition Plan 6

Role of the Gulf of Mexico Subgroup of the MRIP Transition Team 7

III. Description of Approach and Timeline **8**

Approach: Transition Path 8

Short-Term (2022-2023): Interim Calibration and Regional Database Development 8

Interim Calibration 8

Regional Database Development 10

Long-Term (2024-2026): Survey Improvements, Calibration Improvements or Data Integration, and Regional Database Completion and Maintenance 11

Approach: Research Path 11

Short-Term (2022-2024): Developing and Initiating Research Plan 11

Long-Term (2025-2026): Executing Remaining Research Plan Based on Initial Set of Findings 12

Approach: Communications 13

Transition Timeline for Gulf State Recreational Fishing Surveys 13

Transition Path 13

Research Path 14

Communications Support for Both Components 16

IV. Potential Stock Assessment Impacts and Schedule **17**

Data Delays in the SEDAR Stock Assessment Process 17

Datasets Currently Used in SEDAR Stock Assessments That Will be Affected by the Transition 17

Potential Disconnect in Data Streams 18

Assessment Issues due txo the MRIP Transition Process 18

V. Potential Management Impacts and Schedule **20**

VI. Identification of Unknowns **22**

The Degree to Which Non-Sampling Error Can Be Reduced 22

The Corresponding Degree to Which Survey Estimates Can Be More Closely Aligned 23

Is it Possible to Develop a Single, Gulf-Wide Integrated Estimate From the Multiple Surveys? 24

Funding or Other Resource Limitations 26

Process Obstacles 26

VII. Lessons Learned **27**

VIII. References **30**

IX. Appendices **32**

Appendix A: Final Report of the Gulf State Recreational Catch and Effort Surveys Transition Workshop, Including Research Recommendations of the Statistical Consultants A

Appendix B: SRFS Gag Grouper Calibration Terms of Reference and NOAA Fisheries Calibration Review Memorandum B

Appendix C: Statement of Task for Congressionally-Directed Independent Review of Gulf Surveys and Calibration Methods C

Appendix D: SEDAR Stock Assessment Schedule Through 2023 D

Appendix E: Gulf Annual Management Cycle for Federally Managed Stocks With Recreational ACLs E

Appendix F: Transition Plan for Gulf State Recreational Fishing Surveys: Communications Plan F

Appendix G: **R**eview of Texas Private Boat Red Snapper Calibration Ratio G

Cover photo credit: NOAA Fisheries/T. Kennedy

I. Executive Summary

In this plan, the Gulf of Mexico Subgroup of the Marine Recreational Information Program (MRIP) Transition Team (referred to hereafter as the Gulf Transition SG, and which consists of representatives from NOAA Fisheries, the Gulf Fisheries Information Network [GulfFIN], the Gulf of Mexico Fishery Management Council [GMFMC], the Gulf States Marine Fisheries Commission [GSMFC], and Gulf state agencies), details a comprehensive, approximately five-year timeline for transitioning to the use of Gulf state recreational fishing data in SouthEast Data, Assessment, and Review (SEDAR) stock assessments and GMFMC management processes. The plan is designed to meet both short- and long-term stock assessment and management needs, as well as research needs intended to improve all the recreational fishing surveys (hereafter "surveys") in the region.

In the Gulf of Mexico, in-season management needs for red snapper and other species have led MRIP to support the development and certification of state surveys that either overlap with (in Florida, Alabama, and Mississippi), or have replaced (in Louisiana), the MRIP general surveys. (Texas is an exception: the state has operated an independent survey since 1974, and has not participated in MRIP, although NOAA Fisheries conducted the MRIP Fishing Effort Survey in the state in 2016). As a result of differences in sampling designs, state surveys produce estimates that differ substantially from those derived from the MRIP general surveys, and evidence further suggests that the state survey estimates are not comparable to each other.

This outcome is expected, as we know that methodologically different statistical sampling designs are likely to produce different estimates: the errors that are inherent to statistical sampling may affect different surveys in different ways, and may drive estimates apart. However, these differences can generate challenges for stock assessments and management. Sound stock assessments require a full accounting of landings and discards, and consistent historical data, while sound management is best served by consistent, complete, accessible, and statistically comparable data. Bias in one or more survey estimates included in the sum of recreational catch may lead to biases in the resulting stock assessments, sector allocations, and harvest monitoring. This may result in misinformed management decisions such as erring in setting regulations and quotas. It is therefore essential to understand the potential sources of bias in these surveys. Assessment and management cannot halt until this work is complete, however, and there is a strong need to develop rapid solutions that can meet immediate science and management needs while working toward longer term goals.

This Transition Plan includes two parallel paths, a transition path and a research path, with both short- and long-term priorities. In support of both tracks, ongoing communications efforts will keep partners and stakeholders informed of the status of transition plan implementation.

- **In the short-term:** The transition path will make immediate progress on interim calibration of historical catch estimates using currently available data, meeting the most immediate stock assessment and management needs in the region. The research path will initiate a research plan to:
    - Investigate sources of error in all the programs;

    - Develop stable, long-term calibration methods; and
    - Identify and secure funding to complete needed studies and analyses.
- **In the long-term:** The research path will execute a research plan and identify needed improvements to the programs. The transition path will convene an independent review of model-based calibration procedures for use in future stock assessments and management decisions (which can only be completed once sufficient progress has been made on the research path); develop and maintain a centralized, regional database to house raw and processed survey data, metadata, and estimates from all of the Gulf state surveys; and implement any needed changes to the survey programs identified by the research path.

For stock assessments, potential delays in data availability are the greatest risk in the transition process. The current SEDAR assessments have tight deadlines and require catch and effort estimates. They also require raw (i.e. trip-level) intercept and biological data, as well as metadata in review of precision and appropriate application of the data. The federally-managed stocks most affected by this transition (i.e., those that have broad coverage across the state survey programs) are red snapper (assessment ongoing), greater amberjack (assessment planned for 2024), and gray triggerfish (assessment planned for 2025).

For management, the biggest challenge in this transition process is the need for fishery-dependent data used to inform harvest monitoring to be directly comparable to the established catch limits. Key data needs include landings and discards at sufficient temporal scales, catch per angler trip and/or catch per vessel trip, sizes of fish caught, and effort data including target effort in numbers of trips.

There are five key unknowns in this transition process, three of which are technical and two of which are logistical. Technical unknowns include 1) the degree to which bias can be reduced in the surveys, 2) the degree to which survey estimates can be more aligned, and 3) if a single Gulf-wide, integrated estimate can ultimately be produced. Logistical unknowns include 4) funding and resource limitations, and 5) possible federal process obstacles. The logistical unknowns will likely drive the content and timeframe of the research plan, which will elucidate the three technical unknowns. These logistical unknowns may further impact survey implementation or desired survey improvements in the longer-term.

The transition to the use of state survey estimates to fulfill federal stock assessment and fishery management requirements in the Gulf of Mexico has been a protracted and complex process, and has resulted in many lessons learned to apply to future transitions. These include improved state and regional coordination on data use decisions in advance of implementing new surveys; improved communication across NOAA Fisheries, partners, and stakeholders on what the transition process entails (e.g., requirements for peer review and certification of a survey, and calibration needs and endpoints); and more closely aligning the timing of survey transitions with the assessment and management cycles.

II. Introduction and Purpose

This document describes the process and timeline for transitioning to the use of state recreational fishing data—including recreational catch and effort estimates produced by Texas' Coastal Creel Survey (all federally managed species), Louisiana's LA Creel (all federally managed species), Mississippi's Tails n' Scales (red snapper), Alabama's Snapper Check (red snapper), and Florida's State Reef Fish Survey (red snapper and other reef fish caught almost exclusively in Florida, including gag grouper)—in SEDAR stock assessments and GMFMC management processes. When fully executed, this plan will accomplish the following goals:

1. Identify any needed design changes to improve the accuracy of all survey programs in the Gulf of Mexico, thereby minimizing differences in estimates.
2. Incorporate state data into the federal science and management process while maintaining a consistent, regional time series.
3. Develop a single, publicly accessible, standardized database to house all the state recreational survey data in the Gulf of Mexico needed for assessment and management purposes. (Note: In compliance with applicable state and federal laws, this database will exclude all personally identifiable and otherwise confidential information).
4. Ensure that all data under consideration for use are treated and evaluated consistently in the advisory and management processes through adherence to the regional BSIA framework (in accordance with NOAA Fisheries Procedural Directive 01-101-10 [NOAA Fisheries 2019]).
5. Maintain clear and open lines of communication between the Gulf Transition SG and all affected stakeholders about progress toward the above goals.

This plan is designed to meet both short- and long-term stock assessment and management needs, as well as research needs. As such, this plan should be considered dynamic and will be updated over time as progress is made toward the above goals.

Background

MRIP currently serves two core functions:

1. Account for annual landings of all saltwater species at the stock level, providing regionally consistent trend information for stock assessment purposes in accordance with the general requirements of the Magnuson-Stevens Fishery Conservation and Management Act (MSA).
    a. On the Atlantic and Gulf coasts, the general MRIP surveys ,such as the Access Point Angler Intercept Survey (APAIS) and the Fishing Effort Survey (FES), were designed to meet this function.
2. Respond to more specific quota-based management needs for landings information at the state and sector levels over shorter time periods.
    a. In the Gulf of Mexico, in-season, state-level management needs for red snapper and other species have led the program to support the development and certification of

state surveys that either overlap with (in Florida, Alabama, and Mississippi), or have replaced (in Louisiana), the MRIP general surveys. (Texas is an exception: the state has operated an independent survey since 1974, and has not participated in MRIP, although NOAA Fisheries conducted the MRIP Fishing Effort Survey in the state in 2016).

The five state surveys and the MRIP general surveys use different statistical methods to meet the different objectives for which they were designed. Where programs overlap, systematic differences exist among catch and effort estimates, and evidence further suggests that the state survey estimates are not comparable to each other (NOAA Fisheries Office of Science and Technology 2022a). This outcome is likely caused by different types of non-sampling error. Non-sampling error is a challenge for virtually all surveys, and includes, for example:

- **Coverage error**, which occurs when members of a target population are omitted, duplicated, or wrongly included in a sample frame (e.g., a fishing intercept survey that is only conducted at public fishing access sites omits those who fish from private sites).
- **Non-response error**, which occurs when a member of the sample is unable or unwilling to respond to a survey and differs in some key characteristic from those who do respond (e.g., those who don't respond to a fishing activity survey may fish more or less often—or catch more or less fish—than those who do respond.)
- **Measurement error**, which includes data processing errors, data collection errors, errors due to the mode of data collection (e.g., a mail vs. telephone survey), and errors that may arise from questionnaire design (e.g., wording, length, and/or question order).

These differences generate challenges for stock assessments and management: sound stock assessments require a full accounting of landings and discards, and consistent historical data, while sound management requires consistent, complete and accessible, statistically comparable data. Bias in one or more survey estimates included in the sum of recreational catch may lead to biases in the resulting stock assessments, sector allocations, and harvest monitoring, which may result in managers erring in setting regulations and quotas.

Need for a Gulf State Recreational Catch and Effort Surveys Transition Plan

The need to simultaneously meet the national requirements of MSA, as well as the state and regionally specific management needs described above, drove the development of new recreational fishing surveys – for these new surveys to support science and management along with the federal surveys, NOAA Fisheries requires the development and execution of a Transition Plan. As per NOAA Fisheries Policy Directive 04-114 (NOAA Fisheries 2021), a Transition Plan will allow the full use of alternative survey estimates (i.e., state surveys designed to meet objective 2 above) while satisfying the scientific and management requirements under the MSA (i.e., leveraging MRIP's regionally consistent, long-term time series to meet objective 1 above).

The need for a Transition Plan is further motivated by a 2021 request from the House Committee on Appropriations to assess the accuracy and precision of all survey programs in the Gulf, make

recommendations for improvements to all the programs, and suggest how best to calibrate the programs to a common statistical currency. NOAA Fisheries has developed a Statement of Task for this review (Appendix C), but collaborative research is needed before the agency can commission a well-informed, independent review.

Role of the Gulf of Mexico Subgroup of the MRIP Transition Team

The national MRIP Transition Team was formed to develop and recommend standardized processes for transitioning to the use of new surveys for stock assessments and management. The Gulf Transition SG was established to develop and execute this Transition Plan, and consists of representatives from the following member organizations:

- Alabama Department of Conservation and Natural Resources
- Florida Fish and Wildlife Conservation Commission
- Gulf Fisheries Information Network
- Gulf of Mexico Fishery Management Council
- Gulf States Marine Fisheries Commission
- Louisiana Department of Wildlife and Fisheries
- Mississippi Department of Marine Resources
- NOAA Fisheries Office of Science and Technology
- NOAA Fisheries Office of Sustainable Fisheries
- NOAA Fisheries Southeast Fisheries Science Center
- NOAA Fisheries Southeast Regional Office
- Texas Parks and Wildlife Department

The Gulf Transition SG will play an important role in coordinating consistent approaches and methods for the council, interstate commission, and NOAA Fisheries to apply to recreational catch estimates derived from the state surveys for:

- Determining the status of exploited stocks;
- Setting annual catch limits;
- Monitoring catch against catch limits;
- Specifying research, prioritizing implementation of identified improvements to surveys, and monitoring effects of changes (improvements) on those surveys and their outputs;
- Informing potential needs for accountability measures; and
- Conducting analyses leading to the adoption of recreational fishing regulations.

III. Description of Approach and Timeline

The Gulf Transition SG determined an approximately five-year timeline (2022-2026) will be needed for the transition to using Gulf state recreational catch and effort data in federal stock assessments and management. It will consist of two parallel tracks: a transition path and a research path. This five-year timeline is needed to:

1. Ensure that immediate stock assessment and management needs are met;
2. Research and develop a stable, model-based calibration that can better account for variability among the estimates in the Gulf;
3. Allow further research needed to investigate and address non-sampling errors in all the surveys (state and federal) in the region to improve their accuracy and minimize differences among estimates derived from the different programs, and;
4. Implement these design improvements identified by the research.

The transition path will make immediate progress (2022-2023) on interim calibration of historical catch estimates using currently available data, and meet the most immediate stock assessment and management needs in the region. Between 2023-2026, more information will be added to the transition path as parallel progress is made on the research path. Additional longer-term transition path needs include: **a)** convening an independent review of model-based calibration procedures for use in future stock assessments and management decisions (once sufficient progress has been made on the research path), and **b)** developing and maintaining a centralized, regional database for scientists and managers, managed by GulfFIN, that houses the raw and processed survey data, metadata, and estimates from all of the Gulf state surveys.

Similar to the transition path, the research path has short- and long-term priorities. In the short-term, the Gulf Transition SG must fully evaluate the research recommendations made at their February 2022 Workshop (Appendix A) and finalize a research plan that appropriately prioritizes projects that the team can begin immediately, as well longer-term studies that may produce the highest return on investment (e.g., studies that will identify drivers of differences across most survey programs and/or identify shared sources of error that could be used to improve the accuracy of all programs by implementing design improvements). The research plan will be finalized by the end of 2022, and the plan will be executed between 2023 and 2026 depending on the availability of resources.

Approach: Transition Path

Short-Term (2022-2023): Interim Calibration and Regional Database Development

Interim Calibration

For use in the short-term, the Gulf Transition SG has considered several calibration approaches that could be used for ongoing stock assessments and management to make use of available survey data (Table 1). These approaches all rely on ratio-based calibration methods and have various limitations, but are the only immediate options that could make use of all currently available data. Ultimately the team

hopes to pursue a model-based calibration approach (e.g., Erciulescu et al. 2021), which will be investigated as part of this Transition Plan's research path.

***Table 1**: Suite of calibration options considered for interim assessment and management at the February 2022 Workshop*

Option	Technical Considerations	Source
1. Use the following data streams for both assessment and management: ● In FL, AL, MS: a fully calibrated MRIP APAIS/FES time series (1981-present). ● In LA: a fully calibrated MRIP APAIS/FES time series (1981-2013), and then the LA time series (2014-present), converted to MRIP currency using a peer-reviewed ratio-based calibration developed by LA (GSMFC and NOAA Fisheries 2020). ● In TX: TX time series (1981-present) as-is.	● MRIP APAIS/FES estimates are only available for 2015 and the LA calibration ratio currently relies on one year of overlapping data. ● TX estimates would still be in a different currency, limiting the ability to produce a fully statistically comparable Gulf-wide estimate.	Option 1a in NOAA Fisheries 2020
2. Use the currencies outlined in option 1 for assessments, but convert assessment-based annual catch limits (ACLs) into FL, AL, MS, and LA currencies using the peer-reviewed ratio-based calibration methods developed by each of those states (GSMFC and NOAA Fisheries 2020).	● The ratio calibration methods differ from state-to-state, relying on different years of overlap. ● Texas is still unaddressed, limiting the ability to produce a fully statistically comparable Gulf-wide estimate.	Option 1b in NOAA Fisheries 2020
3. Rescale the MRIP time series to better match estimates based solely on state surveys in FL, AL, MS, and LA and use those data for both assessments and management. In TX: use TX time series as-is.	● Not a statistically valid approach for regional stocks (e.g., red snapper) due to differences among the state surveys, but could be used for stocks that are caught primarily in one state (e.g., gag grouper in FL).	Option 2 in NOAA Fisheries 2020
4. Use a fully calibrated MRIP APAIS/FES time series (1981-present), using ratio calibrations to convert all (FL, AL, MS, LA, and TX) Gulf state estimates to MRIP currency, for use in assessments and management.	● Methods are still in the process of being formally peer reviewed. ● Calibrations for both LA and TX are based on limited information.	Foster methods presented at February 2022 Transition Workshop

The Gulf Transition SG recommended the following approaches for upcoming assessments:

- **Gulf Transition SG Southeast Data, Assessment, and Review (SEDAR) 72 (Gulf of Mexico gag grouper) calibration approach recommendation:** As appropriate, MRIP private boat mode estimates will be calibrated to Florida State Reef Fish Survey (SRFS) estimates (i.e., the 3rd option listed in Table 1 above). Note: this approach included completion of a peer review of the application of the associated calibration factors in this manner (Appendix B). Management advice can then be set in SRFS currency, and FL will continue to monitor catch using their own data.
- **Gulf Transition SG SEDAR 74 (Gulf of Mexico red snapper) calibration approach recommendation:** Use MRIP estimates for Florida, Alabama, Mississippi, and, from 1981–2013, Louisiana. The assessment will also use LA Creel estimates (2014–beyond) and Texas Coastal Creel Survey estimates (1983–beyond) that have been calibrated to the MRIP APAIS/FES time series (i.e., a combination of the 2nd and 4th options in Table 1 above). Note: this approach included completion of a peer review (Appendix G) of the new Texas calibration method (NOAA Fisheries Office of Science and Technology 2022b) and its application. Annual catch limits for all species covered by state surveys (red snapper and others) will be calibrated to the currencies of each state's data collection program, and states will continue to monitor catch using their own data.
- **Gulf Transition SG recommendations for additional upcoming assessments:** Establish a small working group of Gulf Transition SG members to develop generalized Terms of Reference for formal reviews of the application of calibration options for consideration in stock assessments.

Regional Database Development

Efficient access to state survey databases along with associated metadata that explains how surveys are designed, implemented, and how statistical analyses (i.e., estimation, calibration) are accomplished is an essential component of this Transition Plan. GulfFIN is a state/federal cooperative program that works on data management issues that impact state and federal fishery-dependent data collection activities. A working group was formed under the GulfFIN Committee to develop standards and a process for collecting and managing state survey databases through the GulfFIN Data Management System housed at GSMFC. This group is made up of representatives from state fishery management agencies, NOAA Fisheries Office of Science and Technology, Southeast Regional Office, Southeast Fisheries Science Center, GMFMC and GSMFC. The group has identified both short- and long-term objectives for this database development process.

The short-term objectives are to collect survey microdata or raw datasets, estimation datasets, and calibrated estimation datasets as they are currently generated by each state agency. There has been some concern about delivering the raw datasets as they are stored by each state agency — the files can be large with unnecessary variables, including, in some instances, personally identifiable information that must be protected in compliance with applicable state and federal laws. The working group is developing a list of the key variables that would meet the most immediate needs for federal science and management, and states will likely develop an amended version of their survey microdata that will be provided to GulfFIN. The states are also providing their survey standards documentation to GulfFIN that

would be made available through the GulfFIN website. The goal is to have this initial short-term objective completed by December 2022.

The longer-term objective would be to create a standardized database that could house all of the state survey data. This will require coming up with standardized variables and formats and creating a process to transform and load state survey data into this unique database format. This standardization would make it more efficient for federal scientists to obtain and use all state data. We expect to start those discussions in 2023 but also recognize this is a much more complex task and could take up to two years. While the timeline for data availability requires further discussion, we expect that Florida, Alabama, Mississippi, and Louisiana will be able to provide updated datasets every two months, similar to MRIP's estimation timeline. Texas only provides updates twice per year after their low-use and high-use sampling seasons are completed.

Long-Term (2024-2026): Survey Improvements, Calibration Improvements or Data Integration, and Regional Database Completion and Maintenance

Based on the outcomes of the ongoing research, the Congressionally-directed independent review of all the programs, and the outcomes of that review, the Gulf Transition SG will take the following actions (Note: the details of the items listed below are subject to revision, as the execution of the short-term components of this plan may lead to necessary changes to the long-term components of this plan):

1. Make needed survey design and operating modifications to one or more surveys in the region to achieve improved accuracy and precision of estimates, with a goal of minimizing differences as much as possible among estimates derived from the different state and federal monitoring programs. Note: Any decisions to revise methods will be accompanied by side-by-side implementation of the modified and pre-modification methods for a sufficient period of time (as funding permits) to allow comparison of estimates and minimize disruption to management. Changes made to multiple programs will be coordinated to ensure the benchmarking process is standardized and consistent across programs.
2. Either implement a model-based calibration, or, if differential non-sampling errors are reduced sufficiently through the design changes made in action 1, produce a composite estimate that integrates the data from all the state surveys.
3. Finalize and maintain the regional database.

Approach: Research Path

Short-Term (2022-2024): Developing and Initiating Research Plan

In 2022, the Gulf Transition SG will develop a research plan informed by statistical consultant recommendations (Appendix A) that can be accomplished over a five-year time frame, provided sufficient resources are available to fund the research in that time frame and improve the respective

surveys. This plan will address as many of the consultant recommendations as possible in the allotted time frame, prioritizing the following subset:

1. The recommendations related to a) improving individual survey documentation and b) simulation research into drivers of differences in estimates across all survey programs likely have the least funding constraints, and all survey administrators can begin this work immediately and complete it over relatively short timeframes.
2. The recommendations related to the auditing of data and statistical methods may take longer, but are straightforward once complete survey documentation and metadata for all programs in the region are available.
3. The recommendations related to collaborative, experimental research that, after evaluation by the Gulf Transition SG, have the highest perceived return on investment will require the longest amount of time to complete, but have the potential to improve both the comparability and the accuracy of all the surveys operating in the region. Examples are having the potential to identify drivers of differences across the most survey programs (e.g., questionnaire design) and sources of non-sampling errors that are likely to affect the accuracy of all programs (e.g., private access coverage error).
4. The recommendations related to model-based calibration research that use the microdata available from all the different surveys, instead of relying only on the final state and MRIP estimates, would likely take multiple years to develop.

The research plan through 2024 will be finalized in 2022. Its content will include research priorities, cost estimates, and a funding strategy. It is likely that the results of initial studies will inform the remainder of the research plan beyond 2024, which will be developed as we obtain results from that initial research.

Long-Term (2025-2026): Executing Remaining Research Plan Based on Initial Set of Findings

The Gulf Transition SG will continue executing the research plan, monitoring progress regularly and adjusting as needed. During this time, the research needed to develop model-based calibration methods will be completed as well as sufficient non-sampling error research to adequately inform the congressionally-directed independent review. Throughout this period, the Gulf Transition SG will consider a longer-term plan for continuing research, review, and process improvement for all the surveys in the region.

Annually, the Gulf Transition SG will convene a research workshop to discuss progress to-date, prioritize further study, and/or decide that sufficient information is available to initiate the congressionally-directed independent review of all the programs (Appendix C). Once that review has been completed, based on the outcomes, the Gulf Transition SG will convene to make decisions regarding:

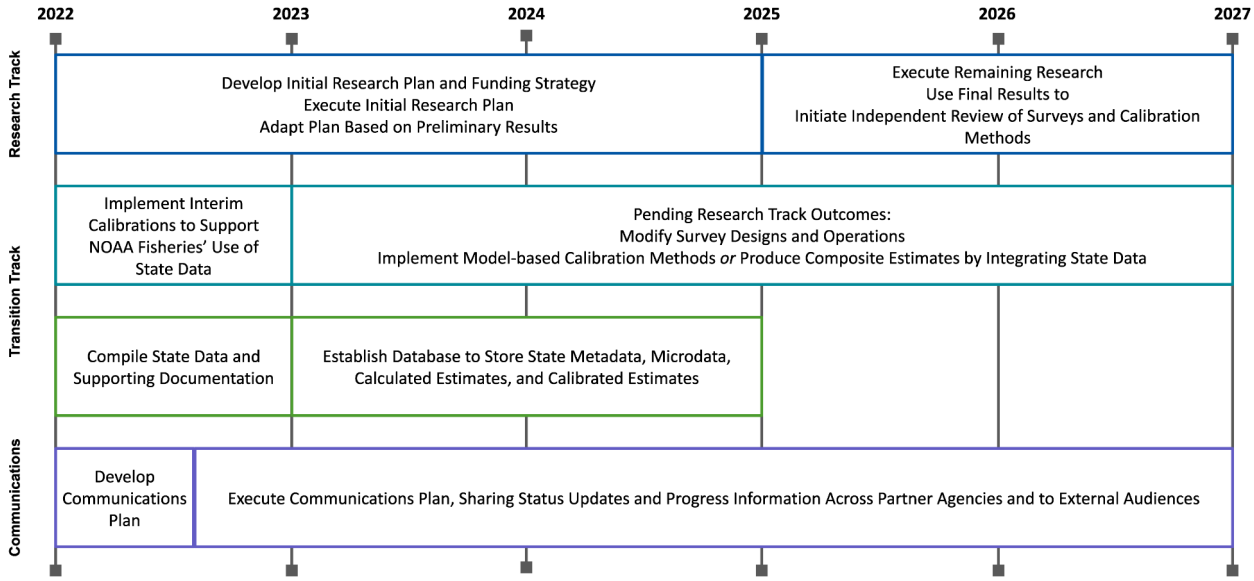
1. Survey design and operating modifications to achieve improved accuracy and precision of estimates;

   a. Implement these revised methods for a sufficient time to enable comparison of estimates with pre-modifed versions of each survey and among the surveys.
2. Calibration methods to be used in the longer-term for regional stock assessments and management, or if possible, methods to calculate a composite estimate that integrates data from the different surveys.
3. Necessary revisions to this Transition Plan to address the decided-upon survey and calibration modifications.

Approach: Communications

Increased and open communications among NOAA Fisheries, state and regional agencies, and stakeholders will be essential to the successful transition to the full use of Gulf of Mexico state recreational fishing data in the federal stock assessment and management processes. With this in mind, the Gulf Transition SG established a cross-agency Gulf Transition Communications Working Group to develop and execute a communications plan to keep partners and stakeholders informed of the status of Transition Plan implementation. This communications plan (Appendix F) reflects the working group's shared goals, activities, and responsibilities.

Transition Timeline for Gulf State Recreational Fishing Surveys



Transition Path

1. **Finalize how to use state data in a) the short-term using calibration, and b) the long-term using either calibration or a composite estimation approach (pending the outcomes of the research path).**

2022:

- Finalize and implement interim calibrations discussed by the Gulf Transition SG at the February 2022 Workshop for use in upcoming assessments and management. For calibrations under consideration in stock assessments, SEDAR will make a case-by-case decision of which calibrated dataset(s) to use, following a peer review of the application of those interim calibration methods.
- Draft generalized Terms of Reference for peer reviews of calibration methods and their application for consideration in stock assessments beyond SEDAR 72 and 74.
- SEDAR 74 data evaluation workshop and assessment process
- SEDAR 81 Gulf of Mexico Spanish mackerel operational assessment

2023-2026: Pending the outcomes of the research path, make needed survey design and operating modifications to one or more surveys in the region; either implement a model-based calibration method, or, if differential non-sampling errors are reduced sufficiently through the design changes resulting from the research path, produce a composite estimate integrating data from all of the state surveys.

2023:

- SEDAR 74 assessment review workshop and assessment report completed (Spring)
- SEDAR 79 Southeastern US mutton snapper

2024:

- SEDAR 86 red grouper operational assessment
- SEDAR 83 vermilion snapper operational assessment

2025:

- Gag grouper operational assessment
- Greater amberjack research track assessment
- Lane snapper operational assessment

2026:

- Tilefish complex assessment
- Gray triggerfish operational assessment

2. Establish a common database for storing state survey data and estimates that meets the needs of stock assessments and management uses.

December 2022: States provide preliminary microdata and supporting documentation to GulfFIN to meet the most immediate federal science and management needs.

2023-2024: GulfFIN completes database development.

Research Path

1. Develop and execute a research plan with a focus on understanding sources of error affecting all survey programs to a) improve their accuracy, b) reduce differences among estimates derived from

the different programs, and c) investigate the development of more sophisticated, model-based calibration options for long-term use in the region.

2022-2026: Evaluate survey error, conduct non-sampling error studies, identify possible survey improvements to reduce error, and research model-based calibration methods.

January 2022: Inventory current survey sponsors' assessment of survey design research completed, in progress, planned and identified as needed.

February 2022:

- Commission MRIP statistical consultants to develop a roadmap for a research program that can identify corresponding design changes to improve accuracy and minimize differences in estimates among the survey programs.
- Prepare a suggested timeline, including time for conducting studies and identify preliminary funding sources.
- Present the information outlined above, discuss among Transition Workshop participants and add/revise as appropriate at Workshop.

March-June 2022: Assemble the Research Path Working Group of the Gulf Transition SG, which will lead and coordinate the research and evaluation of research products.

June-December 2024: The Research Path Working Group identified above will:

- Produce an initial multi-year research plan, including content that discusses the purpose and need for the program and a documented funding strategy.
- Execute the initial research plan.
- Develop the remainder of the research plan based on the findings from initial studies.
- Communicate progress and findings to the Gulf Transition SG on an annual basis.

January 2025-December 2026: The Research Path Working Group will:

- Finish executing the research plan and compile findings.
- Communicate progress and findings to the Gulf Transition SG on an annual basis.

2. Use the outputs of the research plan to inform an independent expert review of the surveys and associated calibration methods to assess relative accuracy of estimates and recommended survey design and implementation improvements.

December 2026 (possibly earlier or later depending on results of research findings and updates to the overall research plan schedule): Contract for an independent expert review of all the surveys and proposed calibration methods (see Appendix C).

Communications Support for Both Tracks

1. **Prepare and execute a communications plan for status updates and progress information sharing across the Gulf Transition SG member agencies and externally.**

February-August 2022: Draft and adopt communications plan.

September 2022-December 2026:

Gulf Transition SG: The Gulf Transition SG will meet quarterly or otherwise as needed to monitor progress, revise and update partner and consultant tasks, and to ensure all partners are fully informed about the status of the transition and research tracks.

Gulf Transition Communications Working Group: Will meet during the last 30 minutes of each Gulf Transition SG meeting, and otherwise as needed to discuss collaborative communications initiatives.

States: Actively work with state partner members of the Gulf Transition SG to monitor Transition Plan progress and to ensure state partner expectations are being met. Actively and regularly communicate with state directors throughout the process, including providing updates during GSMFC meetings.

IV. Potential Stock Assessment Impacts and Schedule

Data Delays in the SEDAR Stock Assessment Process

Given the tight deadlines in SEDAR stock assessments, potential delays in data provision are one of the greatest risks in the transition process. The NOAA Fisheries Southeast Fisheries Science Center constructs a variety of recreational data products used as inputs into regional assessment models, which may incorporate:

- Estimates of both catch and effort, provided at the finest resolution possible (i.e., the strata at which estimates are generated) and in the appropriate currencies (e.g., in native units vs. calibrated estimates).
- Raw intercept data provided at the resolution at which it was collected (i.e., trip-level).
- Raw biological data provided at the resolution at which it was collected (i.e., trip-level).

Delays in the availability of any one of these datasets can affect a wide range of data products needed in SEDAR stock assessments, many of which are outlined below, highlighting the need for these datasets to be widely available and consistently updated (e.g., uploaded to the GulfFIN Data Management System). Survey administrators should implement appropriate quality assurance and quality control (e.g., before uploading data to the GulfFIN Data Management System) and produce survey-specific metadata to provide the context needed to interpret and evaluate these datasets in a timely manner (i.e., descriptions of survey designs, estimation methods, calibrations, and variable lists for submitted datasets). Additional delays may result if these submissions require additional formatting, filtering, and/or analyses before being applied.

Datasets Currently Used in SEDAR Stock Assessments That Will be Affected by the Transition

- Time series of recreational landings, discards, and effort:
    - Estimate files are largely used as provided by the survey administrator, but raw intercept data may be needed to partition estimates into a finer resolution than that provided.
    - Historical catch data – scaling of historical Fishing, Hunting, & Wildlife-Associated Recreation (FHWAR) effort estimates (1955-1980) based on catch and effort estimates from the Southeast Region Headboat Survey (SRHS), MRIP, and Texas Parks and Wildlife Department (e.g., 1981-1985).
    - SRHS headboat discards (e.g., 1981-2004) – scaling of SRHS landings (1981-2004) based on discard:landings rates calculated from MRIP and/or SRHS data.
- Time series of catch-in-weight:
    - Catch-in-weight is calculated as the product of catch-in-number estimates (from above) and average (fish) weight, calculated from raw biological (size) data.
- Uncertainty estimates for catch/effort time series – constructed from raw intercept data.
- Uncertainty estimates for average weight time series – constructed from raw biological data.
- Size/Age Compositions – nominal compositions constructed from raw biological data (size and aging structures), but associated landings estimates are often applied as weighting factors.
- Estimates of assessment model parameters:
    - Growth and morphometrics from biological data that includes length/age pairs.

    - Fecundity estimates from biological data that includes histological samples or other reproductive information.
- Catch per unit effort (abundance) indices – constructed from raw intercept data.

Potential Disconnect in Data Streams

A large component of the MRIP transition process are the calibrations required to convert survey estimates into similar currencies, without which differences in survey design and/or estimation approaches can result in inconsistent data streams that are misconstrued (by the assessment model) as a change in the underlying stock and/or its associated fisheries. These calibrations, therefore, need to be properly reviewed and accepted before their associated surveys are used to inform stock assessments. This is also why tables of calibrated estimates are needed from this process (e.g., uploaded to the GulfFIN database), to ensure the proper estimates are available to be used in generating SEDAR data products.

Similarly, if all required datasets are not available to produce a given SEDAR data product, whether from a delay in data provision or information simply not being collected by a given program, datasets from other surveys may be needed as a substitute. These "substitute" datasets can also result in inconsistent data streams if data collection or estimation procedures differ between surveys, even from random chance alone (e.g., different sampling sites selected from random draws).

Assessment Issues due to the MRIP Transition Process

In the first one to two years of the transition, available data will be as outlined in the Short-Term Transition Path and selected calibration methods will be applied initially for SEDAR 72 (gag) and 74 (red snapper). If the selected calibrations are not accepted through the SEDAR process, the assessments may be subject to more scrutiny going into review by the Science and Statistical Committee (SSC). For gag grouper, the Terms of Reference (which are often finalized more than a year before the assessment takes place) called for the Florida state survey data to be used for a sensitivity run to the model rather than the base run. Trying to incorporate Florida State Reef Fish Survey (SRFS) data to the base run at the council's request circumvents the SEDAR process, which causes problems with both precedent and review. For red snapper, the Texas and Louisiana calibrations were presented to the Data Workshop and underwent scrutiny by both the recreational group and the assessment development team (ADT). The calibration was accepted for use in the SEDAR 74 Gulf of Mexico Red Snapper Research Track Assessment, pending further review at the end of the assessment process by the SSC. The Recreational Working Group recommended that the calibrations, particularly for Texas, be reviewed by the Gulf Transition SG's statistical consultants and that knowledgeable Texas state scientists be included in the transition process. It is unclear if further work recommended in order to accept the Texas and Louisiana calibrations (Appendix A) will be completed in time for the end of the SEDAR 74 Research Track process. If not, the SEDAR participants may reject the Gulf Transition SG's recommendations to calibrate the TX and LA data to MRIP currency in the interim. Red snapper is an example of a species that needs a calibration for each state's data in order to scale everything to a common currency, since red snapper are distributed Gulf-wide. The SEDAR 74 Gulf of Mexico Red Snapper Research Track is set to be completed mid-2023, and the transition process will not be completed for all states at that time.

In the long-term (within the next two to five years), based on the outcomes of the ongoing research, partners may choose to make decisions regarding interim tactics:

1. Survey design and operating modifications to achieve improved accuracy and precision of estimates.
2. Calibration methods, or if possible, calculating a single integrated estimate from the state surveys.
3. Partners implement such revised methods for a sufficient time to allow comparison of estimates with pre-modification versions of each survey and among the surveys.
4. Maintain the GulfFIN regional database.

During this Transition Plan's research path period and beyond, final calibrated estimates should eventually be available for each survey following implementation of recommended survey design and operational improvements, or potentially a single integrated estimate for the stock unit could be available for use. Support will be needed from representatives from individual surveys to present and defend their data/estimates throughout the SEDAR stock assessment process. This will minimize any delays needed to address confusion about the design and/or estimation approaches applied in individual surveys. It also ensures these questions are answered by those most familiar with the survey and that decisions of "best available" are made from accurate characterizations of the state and regional surveys.

Until the transition process is finalized, we anticipate a SEDAR 74-like approach for most stock assessments, where MRIP data were used where available, and where unavailable, state survey data were calibrated to MRIP currency. Each state survey would be calibrated to MRIP currency until all the state data that may be used in the assessment have gone through the transition process.

The stock assessment schedule is currently set through 2023 (Appendix D), and the schedule for 2024 will be set in the fall of 2022. The only species that have broad coverage across the state survey programs are red snapper, greater amberjack, and gray triggerfish. The next gray triggerfish and greater amberjack assessments in the Gulf of Mexico are slated for 2024 and 2025 respectively. However, additional assessments could be affected by the use of Florida Reef Fish Survey and LA Creel Survey estimates or derivatives thereof.

V. Potential Management Impacts and Schedule

The management levels of Overfishing Limits (OFL), Acceptable Biological Catch (ABC), Annual Catch Limits (ACL) and Annual Catch Targets (ACT) are needed for each stock managed by the regional fishery management councils. The OFL is generated through a stock assessment model with the ABC being set to account for scientific uncertainty, and recommended values for these catch limits are the result of review by the council's SSC. The council cannot implement an ACL that is higher than the SSC's recommended ABC, but can set a lower ACL value to account for management uncertainty. If management uncertainty is considered to be substantial, the council can introduce an additional buffer on the ACL to offset that uncertainty by implementing an ACT. Some stocks are managed with region-wide stock catch limits (e.g., Gulf of Mexico lane snapper), while others have sector-specific catch limits (e.g., Gulf of Mexico commercial and recreational greater amberjack). Fishery-dependent data used to inform harvest monitoring must be directly comparable to the set catch limits to achieve sustainable management. This challenge is illustrated in the management of Gulf of Mexico red snapper, where recreational data are collected using the federal MRIP program and the five Gulf states which each have their own unique recreational red snapper data collection surveys.

A calibration methodology is required if these various recreational surveys are to be input directly in stock assessments, used to set catch limits, and monitor harvest. Currently, for Gulf red snapper, calibrations to the federal MRIP-FES have been generated for state recreational surveys conducted by the states of Florida, Alabama, and Mississippi. Louisiana's state survey was run concurrently with the MRIP-FES survey to generate comparable estimates. To date, recreational catch estimates generated by the Texas state creel survey have been directly used in the stock assessment, uncalibrated to the federal MRIP-FES survey. The primary reason for the Gulf states to initiate their own recreational data collection surveys was their ability to collect and report recreational landings estimates with increased spatial (i.e., state-level management) and temporal resolution (increased reporting frequency), as well as increased precision compared to the two-month increments as performed in MRIP. The latter can result in a lag of over 100 days between the first days of an MRIP reporting wave and when the preliminary data are finally publicly available. Timeliness in reporting must closely align with season duration to result in effective in-season management. Gulf recreational red snapper open seasons have historically been reduced to weeks or days, and monitoring catch in bi-monthly waves increases the likelihood of exceeding catch limits due to the delay in monitoring reporting. Currently, although red snapper is managed and catch limits have been set to reflect a Gulf-wide stock, Amendment 50A to the Gulf of Mexico Fishery Management Council Reef Fish Management Plan delegates some management authority for recreational fishing of red snapper by private anglers in federal waters to the Gulf states. Each state is allocated a portion of the red snapper private angling quota and has authority to set the private angling fishing season, bag limit, and minimum size limit (between 14-18 inches). If the landings of a state exceed that state's quota, the state's quota will be reduced by the amount of the overage in the following year. Calibrations between the available recreational red snapper surveys is required for fishery managers to implement sound catch limits and achieve sustainable fishery goals.

Careful consideration to calibrating recreational data collection surveys in the Gulf is not limited to statistical inference. Fisheries managers setting sector allocations using historical time series of catch rely on its ability to capture anthropogenic trends in fishery use. These comparisons require a relatively standardized data collection survey methodology through time, and a calibration approach has been developed to directly address converting historical recreational landings data. However, the Gulf Council has run into some difficulty in interpreting current catch limit values to those generated by stock assessments using MRIP-FES calibrated historical data. The catch limits resulting from a previous stock assessment are not directly comparable to a contemporary stock assessment analysis inputting MRIP-FES because of this calibration process. The loss of a directly comparable interpretation makes the already arduous process of requirements like determining sector allocations and setting appropriate catch limits all the more difficult. These issues have been observed for the MRIP-FES historical landings time series and it is highly likely these issues will be compounded when accounting for state survey data in future management processes.

Similar to the impacts on stock assessments, delays in data availability are a risk in the transition process. State surveys' recreational landings are needed for amendment analysis. The state surveys need to have sufficient temporal scale (month or two-month wave, or a shorter time period) to evaluate effects of seasonal closures. Landed and discarded catch is needed to evaluate impact of regulation changes. Catch per angler trip and catch per vessel trip for evaluation of bag and trip limit changes. Length of fish caught is needed for evaluation of size limit changes. Effort data are needed, specifically, both target and catch effort in the number of trips. Target effort is specifically needed because that is the primary input into the economic impact models used by the NOAA Fisheries Southeast Regional Office and fishery management councils to evaluate the economic impacts of each sector and from changes in management in each sector.

VI. Identification of Unknowns

The Degree to Which Non-Sampling Error Can Be Reduced

Non-sampling error may be effectively reduced if it can be measured and properly accounted for. For example, fishing effort surveys may employ a stratified sample design to calculate appropriate sample weights that account for known differences in fishing avidities and response rates among groups of participants. After a sample is drawn, post-stratification is another statistically valid method that may be used to account for known sources of potential bias in survey responses. When the known source of bias is due to under-coverage of the angling universe, this may be addressed more effectively with survey design changes or supplemental surveys that improve coverage. When evaluating the best approach, consideration should also be given to preserving long-term time series. For example, the state survey in Texas is only conducted at public access sites, thus it is known that this survey does not measure additional fishing effort from private access fishing sites. This known source of under-coverage could be addressed with a supplemental survey (using a license frame, coastal address frame, etc.) that separately measures fishing effort at private access sites. This approach would not require altering or replacing the historic survey that has consistently served the long-term fishery dependent monitoring needs for that state.

Non-sampling error is more difficult to account for when it is not known. To identify potential unknown sources of bias requires a well thought out research plan. An iterative approach may be necessary to evaluate the accuracy of survey estimates, measure the magnitude of potential bias if present, and home in on the source(s) of non-sampling error that contribute to that bias (Hartill et al. 2012). When there are overlapping surveys, a good first step is to understand how the estimates differ and examine whether this difference is consistent and directional or just random (Hartill and Edwards 2015). For example, are differences spatially or temporally isolated, or more widespread? Are apparent differences significant, or a result of high uncertainty around one or both estimates? This type of evaluation can help identify potential areas to direct additional research.

Landings surveys (dockside catch, discard, biological, etc.) are typically selected from a known universe of sites. However, the individual characteristics of those sites are often poorly characterized beyond location and expected numbers of anglers, and the implicit assumption is made that adequate sampling will ensure that a reasonable representation of the true characteristics of the overall landings from the region will be obtained. With highly variable catch rates for a given site (dependent on seasons, regulations, weather, tides, and a host of other factors), it may not be possible to define those site effects for inclusion in a model. But differences in landings between sites (species composition, biological characteristics, CPUE, etc.) due to site selection differences will contribute to some degree to differences between surveys run in parallel. Methods to minimize those differences in a study designed for calibration purposes may limit the use of one or both of the compared studies to their historic time series, but may better serve the purpose of calibration.

When it is unknown which of the overlapping surveys is contributing to a significant bias, or if no other survey exists for comparison, then it may be necessary to develop new methods to independently verify the accuracy of survey estimates (Holdsworth et al. 2018). This usually requires on-site studies that directly measure effort or catch with a high level of certainty and a very low potential for bias. To keep costs within reason, it may be necessary for such studies to be conducted at reduced scales. For example, an intensive effort may be more readily conducted over a small region within the survey area, and over a short time period. Comparing the number of potential sources of bias and the potential magnitude of each source among overlapping surveys can also help prioritize research efforts by identifying which survey is likely to be the larger source of potential bias. If it is known that one or more surveys produce consistently biased estimates, then the next step is identifying potential sources of error and conducting studies to evaluate whether addressing them reduces bias.

Successful execution of a research plan requires a willingness among all state and federal partners to actively participate in the process, adhere to a scientific approach, and be open and objective to the possibility that the survey they administer could have sources of unknown bias that have yet to be discovered and are causing surveys in the region to not align. Starting the process under an assumption that one survey is the standard that the other should be aligned with is counterproductive and not sound science. While all surveys have some bias, it is plausible for surveys that overlap to differ because there is some inherent unknown bias in either survey that has yet to be realized. The objective of conducting the focused studies is to try to get this issue addressed within a structured scientific and statistical framework. It should be recognized that complete understanding of the differences between surveys is likely not possible without near-infinite resources. Thus, some remaining bias differences are likely inevitable.

The Corresponding Degree to Which Survey Estimates Can Be More Closely Aligned

A key step toward aligning surveys is addressing known gaps in survey coverage. For example, after MRIP was replaced with the LA Creel survey in Louisiana, discards were added to the intercept interview protocols to fill an important data gap for stock assessments and better align with the historic time series. Known data gaps that persist in other surveys include:

- Year-round discards in Texas, and for species covered by specialized state surveys in Alabama and Mississippi.
- Fishing effort from private access sites in Texas.
- Catch and effort from shore mode anglers in Texas and for species covered by Florida's specialized state survey.
- Varied reef fish species not covered in specialized state surveys in Florida, Alabama, and Mississippi.
- Catch and effort from for-hire headboats that do not possess federal permits and thus are not included in the sample frame for any of the surveys, including state surveys, the MRIP For-Hire Survey (FHS), Southeast Region Headboat Survey, or Southeast For-Hire Integrated Electronic Reporting Program (SEFHIER).

The exact set of studies the Gulf Transition SG will undertake to identify the degree to which survey estimates can be more closely aligned will be specified in the research plan. That plan will be informed by the statistical consultant recommendations made at the February 2022 Workshop (Appendix A). To make best use of limited resources, the plan will prioritize researching sources of bias that the Gulf Transition SG determines are most likely to have maximum impact on one or more surveys. In states where there is an absence of benchmarking data with MRIP, a high priority should be placed on quantifying the magnitude of potential bias in order to better direct efforts to improve alignment of estimates across the Gulf of Mexico region:

- In Louisiana, the first step is to verify whether the LA Creel and MRIP surveys produce estimates that are consistently different. With only one year of overlap between them and a relatively low ratio for red snapper landings (LADWF 2022), more years of data (or directed studies to characterize the origins of those differences) could find that estimates from these two surveys are not that far apart. However, it should be noted that the single-year ratio for red snapper discards is much larger, which also warrants further study (LADWF 2022). Also, examination of the factors contributing to the differences in catch rates between the two surveys for offshore species (differences vary in scale and direction between species) should be examined to help understand the reason for such differences (inshore catch rates are similar between the surveys).
- In Texas an attempt was made to account for under-coverage of private access site fishing effort for Red Snapper SEDAR 74. However, this adjustment factor was based on a single year of overlap between the state survey and MRIP's FES (NOAA Fisheries Office of Science and Technology 2022b). It is unknown whether the large ratio adjustment (10.9) is due solely to under-coverage in the state survey or if potential bias in the FES survey is also contributing to this high value. The implication for SEDAR 74 is that data streams from Texas for both landings and discards (which borrows information from LA, since discard estimates are not available) were raised by more than an order of magnitude.
- Long term, statistically defensible methods for integrating surveys in the western Gulf are needed to support regional stock assessments.

Is it Possible to Develop a Single, Gulf-Wide Integrated Estimate From the Multiple Surveys?

It is unclear at this point whether it would be possible to develop a statistically valid, single integrated estimate from the multiple existing Gulf surveys—currently one survey per state plus the regional MRIP survey. However, if a single regional-scale integrated estimate could be developed, current data needs will still require estimates at smaller sub-region, state, and even sub-state scales. Using red snapper as an example, data inputs for the current SEDAR 74 assessment research track require separate estimates of landings and discards across three distinct regions, including western (TX and LA), central (MS, AL and northwest FL), and eastern Gulf (adjacent to west coast of FL peninsula), based on the current understanding of stock boundaries (SEDAR 2021). Further, Amendment 50A to the GMFMC Reef Fish Fishery Management Plan establishes that management of the private recreational sector of the Gulf red

snapper stock occurs at the state level, and in some cases in-season. Thus, a single integrated estimate may not be what is needed to support regional assessment and management.

During workshops to improve recreational catch accounting for red snapper in the Gulf, discussions centered around improved precision of annual catch estimates to better inform regional stock assessments and management, and improved timeliness and precision of landings estimates at the state level to support in-season management. To address these different goals, two methodological approaches were identified: probability surveys with a product estimator (i.e., angler intercepts x effort surveys) and capture-recapture surveys with a ratio estimator. The consultant's report from the first Gulf Red Snapper Workshop (Opsomer et al. 2013) noted that the use of different methods does not mean they are mutually exclusive and hybrid approaches are conceptually possible. However, a consequence of selecting different methods to develop and test in the region is that it increased the level of complexity when it came time to evaluate results and understand the differences between overlapping survey estimates in each state. An even more difficult, if not untenable, technical challenge is that the surveys address different goals. While some state surveys improve estimates of landings and discards at annual and seasonal scales, methods developed in other states were designed more specifically to monitor landings in-season and do not provide annual scale estimates. Given that new permitting and reporting requirements were necessary to test the different methods in each state, it is much more difficult to move toward a single, unified approach now that those surveys are fully implemented. However, if the expectation is that estimates from each state survey be used in regional scale stock assessments and catch advice, it will be necessary for some states to either expand their current survey coverage, supplement it with an additional survey, or replace it with a new methodology that provides year-round estimates of both landings and discards. Addressing this would also facilitate the use of state survey data in Florida, which currently provides year-round estimates at the state level, for species assessed and managed as a Gulf-wide stock.

Nevertheless, there are some shorter-term efforts that could address data gaps and allow the various Gulf surveys to provide broader, more complete, regional-level data. Estimates in the western Gulf are essentially already integrated out of necessity. Recreational discards are not monitored in Texas and the current best practice for SEDAR is to calculate a ratio of harvest:discards using annual estimates from Louisiana that is applied to landings in Texas (SEDAR 2015). Although the method used for Texas has not been peer reviewed or certified through MRIP, similar methods could be explored for use in the eastern Gulf to produce an integrated estimate from state survey data. Similar methods to integrate information collected from different recreational fishing surveys have also been used in New Zealand (Holdsworth et al. 2018). There are likely other examples elsewhere that could be looked at to help develop a path forward in the Gulf.

State surveys in the western Gulf do not overlap with MRIP, and the focus here should not be on replacing surveys that serve the needs of those states. Instead, the goal should be to identify the data needs the state surveys currently do not serve and develop a plan to supplement those surveys for more complete coverage of the entire Gulf fishery. For example, while it may be unreasonable to expect Texas to replace its long-standing survey with a new methodology, there may be support for addressing the

coverage gaps moving forward. A supplemental survey that measures fishing effort at private access sites could be run side-by-side with the existing state survey in Texas to measure that additional source of effort. Consideration of a statistically valid approach to estimate discards is also needed in Texas.

In states where surveys overlap with MRIP but were not designed to be a complete replacement, we need to identify the specific barriers that are preventing estimates from those specialized reef fish surveys from being combined for use in a Gulf-wide assessment. Florida's SRFS data provides year-round estimates of landings and discards for multiple reef fish species, but cannot be combined with state survey data from Mississippi and Alabama for use in stock assessments because they are only conducted in-season, do not measure discards, and are focused on a single species. The questions here should be whether to replace the state surveys that do not have complete coverage or supplement them with additional components that will achieve full coverage of the fishery in each state.

Funding or Other Resource Limitations

Congress has directed NOAA Fisheries allocate $2M of appropriated funds for recreational fisheries data collection for an evaluation of the accuracy, precision, and comparability of the various Gulf recreational fisheries surveys. However, long-term funding to support the survey adjustments or expansions needed to fill data gaps or achieve better alignment in coverage have not yet been identified.

There are some prohibitive limitations attached to federal funds, discussed below, that should also be addressed.

Process Obstacles

One federal process-based unknown that may affect this transition is Paperwork Reduction Act (PRA) clearance. The PRA requires federal agencies to obtain Office of Management and Budget (OMB) approval before requesting most types of information from the public. The intent of this law is to ensure the federal government does not burden the public with unnecessary or duplicative requests for information. It takes approximately six to nine months to obtain such clearance, and there is a chance that clearance may be denied. This is a possible outcome because OMB conducts its own evaluation as to whether the information being collected will be worth the public burden and a good fit for its proposed use.

PRA clearance has possible implications for the surveys themselves and for implementing the research plan. While all ongoing NOAA Fisheries surveys already have PRA clearance, elements of the research plan that NOAA Fisheries participates in or funds (e.g., experimental surveys, or additional questions added to MRIP surveys for research purposes) will also require such clearance. Furthermore, any desired changes to the PRA-cleared surveys based on research findings (e.g., once the Gulf Transition SG has implemented the research plan and identified needed improvements), will have to be processed as new or amended PRA clearances, depending on their magnitude. Also, in certain instances, NOAA Fisheries may be required to obtain OMB-PRA clearance to provide funding for surveys and survey pilot studies conducted by others.

VII. Lessons Learned

The transition to the use of state survey estimates to fulfill federal stock assessment and fishery management requirements in the Gulf of Mexico has been a protracted and complex process. The challenges have largely been due to the number of survey revisions and additions – the transition of MRIP from Coastal Household Telephone Survey (CHTS)-based trip estimates to FES-based estimates, and the addition of four new state surveys, each using a different sampling design with a different scope (along with the legacy Texas survey). During the development period for the Transition Plan, it was necessary to conduct interim assessment and management functions for a highly visible and controversial red snapper fishery. The need for interim calibrations and other analyses to support the ongoing management mandates and other changing conditions, including the COVID pandemic's interruption of data collection functions and a new congressional mandate for independent survey review, all complicated the process. Given that this is the first time NOAA Fisheries and its regional partners have attempted to undertake such a multi-survey transition, these complications have resulted in a number of lessons learned that can help to inform future transitions:

1. The implementation of six different survey designs has significantly complicated the process of inter-calibration of survey estimates into a common currency. In hindsight, seeking consensus among the states and other regional partners for a single supplemental survey design for red snapper (and potentially other predominantly offshore species) estimates could have been a more efficient approach.
2. The process for development, certification, and transition for MRIP supplemental or substitute surveys needs to be better defined and more structured. For example, the process for calibration has been very slow and has lacked a clear set of procedural steps or a defined endpoint.
3. Calibration is a complex subject, and communications regarding calibration are challenging, including:
    a. The purpose and need for calibration, including what it is and what it is not;
    b. The mechanics and differences among calibration methods;
    c. That calibration of estimates from newer surveys to the currency of a legacy survey does not mean that we regard either of them as more accurate;
    d. How calibration relates to catch-history-based allocations;
    e. That calibrations may have to be re-done periodically, even annually, depending on the method used, further changing the quantitative relationship among calibrated estimates, thereby complicating use in management and allocation; and
    f. The communication challenges exist at multiple levels, from Congress to state political leadership to partner agency leadership to the stakeholder community.

    As a consequence, calibration-related communications require significant resources.
4. It is necessary to make partners and stakeholders aware that all surveys are subject to non-sampling error, and to continuously explore those sources and seek design improvements to reduce/minimize them. Significant survey design changes implemented to reduce non-sampling error, which will likely occur following execution of the present Transition Plan's Research Path,

will require updates and revisions to calibrations that partners and stakeholders must be prepared for, including potential multiple adjustments to catch-history-based allocations.

5. More changes to surveys, particularly those that provide for-hire mode estimates (SEFHIER, FHS), are pending and will require additional calibration and transition planning. It will be necessary to prepare transition-weary partners and stakeholders for that certainty as well.
6. If councils continue to base sector allocations exclusively on a prior period's catch history, recreational catch recalibration due to survey changes and associated adjustments to the sector allocation percentages is inevitable. It is critical to improve communication about the cause and effect of such changing percentages, and that they may not reflect a significant change in the current numbers/pounds of a sector's allocation, only its proportion relative to another sector that has had a revised catch history via calibration.
7. Timely completion of this transition plan was impeded by a number of obstacles, including:
    a. Covid-based disruptions to data collection and collaboration opportunities;
    b. The necessary ongoing cycles of stock assessment and management/regulatory rulemaking for which preliminary/interim calibration solutions had to be developed, diverting technical staff and consultant time from developing the final methods for the transition plan;
    c. Challenges associated with certification of the state surveys, including: meeting requirements for complete documentation; the iterative process of review and response/revision; confusion around the meaning of MRIP survey design certification; political pressure to accelerate/evolve the certification process; and
    d. The need to develop plans to address congressional mandates.

   Future transition planning efforts should recognize the foreseeable obstacles and have strategies ready to push through them. For instance, aligning the timing of survey transitions more closely with the assessment and management cycles and communicating early and often with partners and stakeholders about documentation requirements for certification and what to expect from the process are possible strategies to implement for future, smoother transitions.
8. Experience with integrating the Gulf surveys strongly suggests that development and implementation of new recreational fisheries surveys should involve an appropriate planning phase for evaluation of how the data will be used for stock assessment and fisheries management. In particular, the following considerations should be addressed:
    a. While it was understandably necessary to develop surveys specifically for the purpose of facilitating improved management of red snapper in the Gulf (i.e., to allow implementation of Amendment 50 to the Gulf of Mexico Fishery Management Council Reef Fish Fishery Management Plan), if the expectation is to also use those data in regional-scale stock assessments, this should be considered early on, at the survey design stage, and not after a survey is implemented. Such data use decisions should be made in advance because quota monitoring and stock assessments have different data requirements. Quota monitoring requires tracking of landings primarily during the fishing season. Stock assessments require year-round monitoring of landings and discards.

   b. Further, since stock assessment models rely on the magnitude of fishery removals (landings + discards) in time and space (at the unit stock scale) to generate estimates of stock size, it is important that these data be at the same scale (i.e., same units or currency) or can be properly calibrated in time and space.
   c. Calibration is important and a proper benchmarking period should be required before a new survey is used in place of an existing one. One year of overlap is not enough for a statistically valid calibration.
9. The conditional certification of Alabama's Snapper Check survey has resulted in prolonged uncertainty about use of the survey's estimates. As of the time of preparation of this Transition Plan, the conditions of certification have not been met, leaving the Alabama estimates and their calibrations in question. In future survey certification decisions, it may be advisable to resolve all survey performance issues prior to certification to avoid a repeat of the situation that has occurred in this instance.
10. Communication with the public and recreational fishing community could have been better so realistic expectations were established, and compliance and response rates could potentially be higher. Moreover, it would have been preferable to have a communications strategy in place earlier with a goal of having the partners speak in one voice or similar to stakeholders/constituents.
11. The region might benefit from a broader vision of recreational fisheries data collection that is integrative and inclusive of all states. The MRIP Regional Implementation Team could consider incorporating a dialogue on this subject in its ongoing development and updates to its Regional Implementation Plan.

VIII. References

Erciulescu, A., Opsomer, J.D., Breidt, F.J., 2021. A Bridging Model to Reconcile Statistics Based on Data from Multiple Surveys. Annals of Applied Statistics, 15: 1068-1079.

Gulf States Marine Fisheries Commission and National Marine Fisheries Service (2020) GSMFC-NOAA Fisheries Workshop on Red Snapper – Workshop Summary Report. https://gulfcouncil.org/wp-content/uploads/B-8b-Gulf_Calibration_Wrkshp_report_2020_v1.21.pdf

Hartill, B.W., Edwards, C.T.T., 2015. Comparison of recreational harvest estimates provided by onsite and offsite surveys: detecting bias and corroborating estimates. Can. J. Fish. Aquat. Sci. 72: 1379–1389.

Hartill, B.W., Cryer, M., Lyle, J.M., Rees, E.B., Ryan, K.L., Steffe, A.S., Taylor, S.M., West, L., Wise, B.S., 2012. Scale- and context-dependent selection of recreational harvest estimation methods: the Australasian experience. N. Am. J. Fish. Manage. 32 (1): 109–123.

Holdsworth, J.C., B.W. Hartill, A. Heinemann, J. Wynne-Jones, 2018. Integrated survey methods to estimate harvest by recreational fishers in New Zealand. Fisheries Research 204: 424-432.

Louisiana Department of Wildlife and Fisheries. 2022. LA Creel/MRIP Red Snapper Private Mode Landings and Discards Calibration Procedure. SEDAR74-DW-4. SEDAR, North Charleston, SC. 38 pp. Available online at: http://sedarweb.org/docs/wpapers/S74_DW_04_LA_landings_SEDAR74_Updated_v3.pdf

National Marine Fisheries Service. 2019. Procedure 01-110-10: NOAA Fisheries Framework for Determining that Stock Status Determinations and Catch Specifications are Based on the Best Scientific Information Available. Silver Spring, MD. Available online at: https://media.fisheries.noaa.gov/dam-migration/01-101-10.pdf

National Marine Fisheries Service. 2020. Recommended Use of the Current Gulf of Mexico Surveys of Marine Recreational Fishing in Stock Assessments. Silver Spring, MD. Available online at: https://media.fisheries.noaa.gov/dam-migration/94100569.pdf

National Marine Fisheries Service. 2021. Policy 04-114: Implementing Recreational Fisheries Catch and Effort Survey Design Changes. Available online at: https://media.fisheries.noaa.gov/2021-06/04-114_2021.6.9_final%20for%20Doreumus%20Signature_signed.pdf?null

National Marine Fisheries Service Office of Science and Technology. 2022. Evaluating Uncertainty in Gulf Red Snapper Estimates: A Preliminary Sensitivity Analysis of Non-Sampling Errors in the Region's Recreational Fishing Surveys. SEDAR74-DW-11. SEDAR, North Charleston, SC. 39 pp. Available online at: http://sedarweb.org/docs/wpapers/S74_DW_11_OST%20RS%20Sensitivity%20Analysis.pdf

National Marine Fisheries Service Office of Science and Technology. 2022. Methodology Description for a Simple Ratio Calibration of Texas Private Boat Red Snapper Annual Landings Estimates. SEDAR74-DW-10. SEDAR, North Charleston, SC. 5 pp. Available online at: http://sedarweb.org/docs/wpapers/S74_DW_10_TX%20PR%20RS%20Calibration.pdf

Opsomer, J., L. Stokes, J. Breidt, and G. Lesser. 2013. Consultant Report on Red Snapper Recreational Catch Assessment Methods. New Orleans, LA: NOAA Fisheries-Gulf States Marine Fisheries Commission Red Snapper Workshop.

SEDAR. 2015. SEDAR Procedural Workshop 7: Data Best Practices. SEDAR, North Charleston SC. 151 pp. available online at: http://sedarweb.org/pw-07

SEDAR 2021. Gulf of Mexico Red Snapper Stock ID Process Final Report. SEDAR, North Charleston, SC. 87 pp. Available online at: http://sedarweb.org/sedar-74-gulf-mexico-red-snapper-stock-id-process

IX. Appendices

Appendix A: Final Report of the Gulf State Recreational Catch and Effort Surveys Transition Workshop[1], Including Research Recommendations of the Statistical Consultants

[1] Note that the transition workshop proceedings report contains 4 appendices of its own (numbered 1-4). Appendix 3 in that report is the Research Recommendations of the Statistical Consultants which will inform the Gulf Transition SG's research plan.

Marine Recreational Information Program (MRIP)

Gulf State Recreational Catch and Effort Surveys Transition: A Workshop of the Gulf of Mexico Subgroup of the MRIP Transition Team

Summary Report

February 23-25, 2022

Prepared by:

Paul Rago

Mid-Atlantic Fishery Management Council

Scientific and Statistical Committee

June 3, 2022

Table of Contents

Executive Summary 3
Proceedings of Three-Day Meeting 7
Day 1: February 23, 2022 7
Overview 7
Transition Plan 8
Data Requirements for Stock Assessments and Management 10
Survey Research Roadmap (Part 1) 11
Day 2: February 24, 2022 12
Survey Research Roadmap (Part 2) 12
Methods for Calibration (Part 1) 16
Database Storage and Data Management 19
Day 3: February 25, 2022 19
Methods for Calibration (Part 2) 20
Components of the Transition Plan 22
Communications Plan 23
Closing Statements 23
Acknowledgements 24
List of Presentations and Useful Links 24
References 25
Appendices 27
Appendix 1. Agenda 27
Appendix 2: List of Participants 32
Appendix 3: Recommendations of Statistical Consultants 34
Appendix 4: List of Abbreviations 42

Executive Summary

Overview. The Marine Recreational Information Program (MRIP) seeks to continuously evolve in response to the changing needs of recreational harvesters, scientists and managers. Within the Gulf of Mexico region, states where MRIP has historically been conducted have implemented alternative surveys for estimation of effort and catch. These surveys have different degrees of temporal and spatial coverage and provide estimates for various fishing modes, dispositions, and species. These survey designs have been certified to be unbiased in expectation. However, owing to different objectives and sampling designs, these surveys produce estimates that differ from those derived in the MRIP survey. Such differences can create problems for stock assessments and subsequent management because the estimates are of different magnitudes, or have different "currencies." As per NMFS Policy Directive 04-114, a Transition Plan will allow full use of the alternative surveys' estimates while at the same time satisfying the scientific and management requirements under the MSA to use the Best Scientific Information Available (BSIA). Development of the Transition Plan has required the efforts of multiple partners from the federal and state agencies as well as the Gulf Council and Gulf States Marine Fisheries Commission.

This report summarizes the results of a virtual meeting, held Feb. 23-25, 2022, to address critical short- and long-term needs of the Transition Plan. It represents the latest in a series of meetings that have addressed the issue of comparability of alternative estimates. Upcoming assessments for gag grouper and red snapper in the Gulf create additional urgency for this task. This report is the proceedings of that meeting—it is designed to faithfully capture the essence of presentations, and more importantly, the ensuing discussions and recommendations. More than 100 individuals attended the meeting and 50 participated directly in the discussions. Notably, five expert statistical consultants provided recommendations in response to presentations, questions and discussions during the meeting. In addition, the Consultants met after the meeting to craft more synthetic responses to the suite of meeting topics. Their findings are included as an appendix to clearly distinguish topics that were addressed in plenary sessions from those that were addressed outside the meeting.

The meeting was supported by the outstanding efforts of the MRIP Transition Team. Prior to the meeting, webinars were held to update the Consultants on current progress towards responding to recommendations since their initial peer review and certification. The webinars also summarized work that is being done to investigate sources of error and reduce potential bias. The advance meetings gave participants the opportunity to brief MRIP consultants on specific improvements, issues, and ongoing work related to each survey before the Feb 23-25 meeting.

An agenda was distributed before the meeting. Individual presentations set the stage for each topic and were followed by series of clarifying questions and open discussion. To ensure the best use of Consultants' expertise, Consultants' questions and comments were addressed first.

However, all participants were afforded the opportunity to speak. In most instances, the issues could be fully vetted within the allotted time although some topics overflowed to the next day.

The need for a Transition Plan is further motivated by a request from the House Committee on Appropriations (2021) to:

- Assess the accuracy and precision of all programs in the Gulf,
- Make recommendations for improvements to all the programs and
- Suggest how best to calibrate the programs to a common currency.

The Transition Plan requires a pragmatic path to meet short-term assessment and management needs and a research path to address specific issues such as non-sampling error. The Transition Plan must balance scientific innovation with stock assessment needs. The Research Track of the process will require improvements to all surveys. No timeline is specified in the congressional directive, but a period of about five years is likely.

Sound stock assessments require a full accounting of landings and discards, and consistent historical data. Sound management requires consistent, complete and accessible data in a common currency. In-season monitoring of harvests is a primary responsibility of federal managers to avoid overfishing limits. Meeting the demands of BSIA is essential for both sound science and management.

A Research Roadmap was outlined by the Consultants. Specific recommendations for each state and for MRIP are listed in the body of the report and in the Consultants' report in Appendix 3. All parties recognized that further in-depth consultations are required. Information about private dock usage would benefit all programs because it is assumed that patterns of landings and reporting between anglers at public and private access points are similar. There are no guarantees that such information will be sufficient to establish agreement among survey estimates. Consultants cited examples from health care surveys where estimates did not overlap, even though all surveys were peer-reviewed and had validated methods.

Recreational catches in the Gulf states are derived either from 1) capture-recapture surveys with a ratio estimator, or 2) probability surveys with a product estimator (i.e., angler intercepts x effort surveys). Each survey has a different sampling frame such that the target population of anglers differs among states. All the surveys may share many common vulnerabilities to non-sampling error such as coverage error, nonresponse error, and measurement error. Capture-recapture surveys are vulnerable to matching errors and failure of the assumption of independence or reports from anglers on intercepted vs. non-intercepted trips. A detailed list of research recommendations from the Consultants may be found in the body of report and Appendix 3.

Database Storage and Data Management: In the short run, all parties should provide both raw and processed data in a standardized format along with associated metadata[1]. Development of a model-based calibration estimate will require high-resolution data rather than aggregated data typically available. Survey estimates, calibration factors, the raw data tables on angler intercepts and biological data, the metadata for survey design, and in some cases the algorithms used for computation of weights, are required. Not all demands of the data are known in advance, so having the ability to reconstruct estimates as appropriate was judged to be a desirable feature. Provisions for public access to summarized data is also required.

Proper timing of data deliveries is essential for use by scientists and particularly managers. The Southeast Fisheries Science Center (SEFSC) endorsed the concept of expanding the role of the Gulf of Mexico Fisheries Information Network (GulfFIN) for this purpose but noted that timetables must be realistic with respect to available resources. GulfFIN representatives emphasized it could not be responsible for full-scale auditing of each submission. State concerns were similar—they are focused primarily on executing the designs and not sufficiently staffed to handle some of the proposed analyses. All participants agreed that improvements to the data standards process are both desirable and valuable. Sound foundational principles such as the NMFS recreational fishing survey and data standards and the Southeast Data, Assessment and Review (SEDAR) stock assessment guidance documents are excellent starting points.

Calibration: The goals of research for accuracy vs. calibration will not necessarily be the same. Nonetheless, calibration is a requirement for stock assessments to assess the status of populations and the basis for subsequent monitoring of catches by jurisdiction. If side-by-side comparisons reveal large differences in catch estimates among and within states, calibration is required to establish a common currency.

Five different calibration methods were actively considered by the participants. The first two come from the 2020 "White Paper", which identified four primary options for calibration. Two variants of a hybrid approach that built upon the White Paper recommendations were also considered. Ratio methods applied to date have been based on empirical comparisons among surveys but have not attempted to address why the estimates were different. All of the surveys have an angler intercept component and this might be a useful place to look for commonalities. Correlations among surveys at the intercept level could help identify predictor variables. Such patterns could lead to a "new currency" which does not use any single survey (e.g., MRIP) as the frame of reference. This approach was called a model-based approach and will necessarily be part of a longer-range effort.

[1] Metadata includes all data related to the collection of information. This broader definition of metadata is known as paradata.

Simple ratio comparisons between MRIP estimates and state survey estimates for years when both surveys were conducted vary widely by state, ranging from 1.8x for Louisiana to 11x for Texas. The range of ratios have practical implications for calibration:

- Do we use state, MRIP or composite metrics of harvest?
- Are the years used for calibration fixed, or amenable to future changes?
- Do calibration decisions vary by species?
- Will calibration decisions be updated regularly?

Participants acknowledged that such ratio calibrations could, and possibly should, evolve annually. Regardless of the calibration standard used, developing a consistent time series is perhaps the most important criteria. Participants agreed that multiple conversions between assessment results and monitoring estimates could be time consuming, hard to communicate, and possibly error prone. Some states may be able to update their calibration ratios (e.g., FL) but in other states where MRIP is no longer conducted (e.g., TX), the ratios will remain fixed unless auxiliary information can fill gaps.

Participants recognized the need to pursue both short and long-term solutions and affirmed their willingness to work towards imperfect, but necessary, short-term solutions to meet upcoming stock assessment needs. The diversity and complexity of the survey methods heightens the need for detailed specification of data conversion and increases the chances of miscommunication among stakeholders. Participants noted that consistency is a central tenet of BSIA.

Ratio-based calibrations were reviewed during the 2020 Red Snapper V Workshop, but there are nuances of calibrating back in time and handling different fishing modes, particularly shore-based fishing activity. Consultants have previously advised state and MRIP representatives on methods for calibration of means and standard errors. A similar follow-up meeting to discuss Terms of References for stock assessments was proposed. The upcoming gag and red snapper assessments should be illustrative of ratio calibration limitations and considerations. The gag fishery is dominated by harvests from Florida whereas red snapper are harvested by all states in the Gulf.

Calibration Recommendation: After considerable discussion, a proposal to use Option 1b from the White Paper for red snapper was advanced. Participants recognized that the complexities of adjusting five fishing modes within states and varying degrees of overlap with MRIP estimates would make this a difficult task. Additional concerns about the use of Louisiana and Texas data were expressed. Representatives from Texas were not present for this discussion and concerns were expressed about the implications of rescaling the entire time series of red snapper catches to MRIP equivalents. Since the consultants were not provided the opportunity to visit and discuss the details of the Texas survey, the consultants affirmed that they have not reviewed the procedures used in Texas.

Following some additional clarifications, the group consensus agreed to:

- Use 1b for red snapper.

- Use option 2 for gag [see 2020 White Paper for these options].
- Convene a group to set Terms of References and identify individuals to conduct a formal review of the application of the ratio-based calibrations for gag and red snapper at minimum, and others as appropriate. Owing to the complexity of the underlying surveys, it is anticipated that expertise outside the SEDAR process would be necessary for review.

Reaching agreement on a methodology for calibration was viewed as an important accomplishment of the Workshop.

Transition and Communications Plans: The remainder of the meeting was devoted to outlining the components of a Transition Plan and Communication Plan. The Transition Plan outline will be fleshed out and circulated to the Transition Team as soon as possible for review and comment. Participants expressed concerns that the Transition Plan should be finalized as soon as possible.

The overall goals of the Communications Plan are to facilitate development of the Transition Plan and its implementation. The plan will help ensure that workshop outcomes are clear, easy to find and promptly distributed to target audiences. Target audiences include transition team members, leaders of their respective agencies and external stakeholders such as Council members, congressional representatives, anglers and various advocacy groups. Participants should nominate members for the Communications Working Group.

Proceedings of Three-Day Meeting

Day 1: February 23, 2022

Overview

This report is a summary of a virtual workshop to develop a transition plan for integrating the various surveys for recreational catches (See Appendix 1: Agenda) The workshop included representatives from various state, regional and federal programs as well as members of the Gulf States Fishery Management Councils (See Appendix 2: List of Participants).

Recreational fisheries in the Gulf States are monitored by state or federal surveys. In some states, both approaches have been used in the past and continue to the present. In other states, contemporary monitoring is done solely with a state survey (LA) while in Texas, the MRIP survey has never been used comprehensively. This mixture of methodologies creates problems when catch data are used in stock assessments. Every sampling design has varying levels of bias and inherent variability of estimates. Thus, every sampling design imposes constraints on the subsequent use of the data.

Stock assessments attempt to measure population abundance by quantifying known removals and relating these removals to trends in various indices. Estimated removals in stock assessments may be under or over-estimated if one or more of survey estimates included in the sum of recreational catch is biased. The consequences of bias are also important for sector allocations

and monitoring of harvests that are required under the Magnuson-Stevens Act (MSA). Hence, it is essential to understand the potential sources of bias and to develop a sound basis for contemporary science and management.

The purpose of this workshop was to continue the progress toward development of a common currency for short-term stock assessment and management needs, and to emphasize the need for continued efforts toward fully comparable estimates among programs. The workshop builds upon earlier efforts to develop technical guidance. Five statistical consultants with expertise in survey design participated in the workshop to identify candidate approaches for comparing surveys and to develop plans for future research efforts.

This report constitutes the proceedings of the workshop and a summary of key conclusions. The statistical consultants met shortly after the meeting to synthesize their recommendations for future work. Their report (submitted March 28, 2022) is presented verbatim in Appendix 3.

The meeting was convened by Gregg Bray and Richard Cody who gave some opening remarks, expressing appreciation to all participants for their ongoing commitment to the Transition Process. Paul Rago provided an overview of the Workshop purpose, desired outcomes and agenda. The transition process began with the National Academy of Sciences report in 2006, which identified a wide range of needed improvements in the federal survey. MRIP surveys were developed between 2008 and 2015 and fully replaced the earlier MRFSS surveys in 2018. In the Gulf Region, new state programs (LA Creel, AL Snapper Check, FL SRFS, and MS Tails and Scales) were initiated between 2014-2015. These surveys were developed fully or in part to support timely in-season quota monitoring for state-level management of red snapper and are tailored to state-specific fishing patterns and geography.

Comparisons between catch estimates derived from these surveys and MRIP have revealed substantial differences in some cases. As noted above, such differences are problematic for stock assessments. Such differences are mostly attributable to non-sampling errors including differences in coverage. Thus, it is necessary to simultaneously address the underlying survey design differences and need for stock assessment data. The first path (track) requires continuation of long-term research efforts; the second path (track) requires development of interim methods for combining various catch estimates into a reasonably consistent and precise total. Both paths require a transition plan.

Transition Plan

The overarching goal of the meeting was to agree on the elements of a transition plan for both a research path, and a "pragmatic" path. The pragmatic path requires agreement on a flexible calibration approach for all assessed species and allows for different methods among species. All calibration approaches must meet BSIA standards (https://www.ecfr.gov/current/title-50/chapter-VI/part-600/subpart-D/section-600.315). Within this report, the paths are referred to as tracks. Together, they constitute the "hybrid" plan.

Richard Cody presented a summary of milestones to be achieved by the Transition Plan. Notably, NMFS has received a congressional directive from the House Committee on Appropriations (2021), which calls for an independent review that 1) assesses the accuracy and precision of all programs in the Gulf, 2) makes recommendations for improvements to all the programs and 3) suggests how best to calibrate the programs to a common currency. To meet this directive, the following steps (milestones) were identified:

- Develop a research plan to understand differences across all surveys and contract for an independent review.
- Finalize calibration methods, e.g., ratio estimation.
- Recommend interim measures.
- Identify and solve unmet certification issues for state and federal surveys.
- Develop a common database for partners.
- Improve communication plans.
- Make final decisions on calibration methods for future assessments.

Participants noted that the transition plan must balance scientific innovation with stock assessment needs. The research track of the process will require improvements to all surveys. No timeline is specified in the congressional directive, but a period of about 5 years is likely. Following the meeting, the Consultants will be making specific comments on both tracks within the Transition Plan. Conducting longer-term side-by-side survey comparisons may ultimately yield diminishing returns. Participants agreed that such concerns will be important as the Transition Plan is implemented.

Questions were also raised about unmet certification issues. State-specific projects with the Consultants are ongoing and refinement of these issues is expected as more data become available. A number of participants suggested this workshop should focus on the central issue of identifying the primary drivers of differences among surveys, but acknowledged such efforts would likely entail more detailed discussions than might be possible in the plenary meeting. Consultants cautioned there are many instances at other agencies where differences among survey estimates could not be resolved. All agreed that commitment by all parties was essential for moving forward.

The Consultants noted two competing goals under discussion: estimation of overall catch and allocation of shares of that catch to Gulf states. Further discussion on the relative importance of these goals revealed that the overall goal was to obtain the best possible estimates by state rather than specific allocations. Consultants noted that integration of disparate monitoring systems is much more challenging because of different underlying assumptions and potential biases.

Data Requirements for Stock Assessments and Management

Management of fish stocks under the MSA requires recreational catch data for not only stock assessments but also for a suite of regional management responsibilities. Moreover, the data must satisfy requirements under the BSIA. These topics generated considerable discussion among all participants.

Andy Strelcheck, SERO, outlined the multiple recreational data needs for management. Sound management requires consistent, complete and accessible data in a common currency. In-season monitoring of harvests is a primary responsibility of federal managers to avoid overfishing limits. Additional complexity arises when monitoring must track catches from more than one fishing mode (e.g., private, charter boats). The analyses that support harvest control measures rely on raw intercept data to evaluate effects of changes to bag limits, trip limits and size limits. In turn, these support evaluation of economic tradeoffs. A common database (or warehouse) to support these analyses is viewed as essential.

Consultants noted the complexity of this challenge, especially the difficulties of assigning weights to raw data for a complex survey. Procedures to transfer data processing programs to the warehouse organization/operation would greatly assist post processing of data. Participants asked if data requirements would vary by species. Strelcheck emphasized again the need for consistency and accessibility for all surveys

Recreational data needs for stock assessments were described by Kate Siegfried, SEFSC. Landings and discards, in both numbers and weight, are key ingredients in modern stock assessment models. When combined with additional information such as discard survival rates, and age-length keys, catch data enhance our understanding of population dynamics. Proper estimates of variability in catch facilitate identification of appropriate models for crafting overfishing limits. Siegfried further emphasized the need for consistent historical data. Most models are useless if bias in historical catch data is not accounted for.

Participants noted that some surveys do not estimate out-of-season discards for red snapper, but the importance of this omission is unknown. Ostensibly, this could be a major problem, but other participants noted that closure of the red snapper season also typically signals the end of most reef fishing. These comments helped identify another common thread—surveys are designed to address specific objectives and it is challenging to meet additional objectives after the survey was designed and launched for another purpose. The need for increased coordination among partners was reiterated by participants who also cited a similar conclusion by the 2021 NASEM report on in-season management. Concerns were raised regarding differences between modes within the MRIP as examples of a "mixed currency" system. While such differences in protocols may result in currency differences, these are thought to be small relative to differences in methodologies across states. Participants commented that differences between estimates from private and charter boats were among the motivations for creating alternative state-specific survey methods in the first place.

Under National Standard #2, all recreational data must meet standards under the Best Scientific Information Available (BSIA). Patrick Lynch highlighted these requirements of a six-step process that guides stock assessments, peer review, revisions, SSC deliberations, Council catch specifications and finally NMFS determination of BSIA. No single element is sufficient to define BSIA. The process for evaluating BSIA compliance should not be viewed as competitive among states and MRIP because all programs have strengths and weaknesses. In fact, the utility of various surveys may vary across species (Gag Grouper was cited as an example in which 95% of the stock is sampled by SRFS). BSIA relies heavily on the peer review process within stock assessments. Reviews outside of the stock assessment process might be advantageous particularly in circumstances where time is insufficient for an in-depth review of underlying data.

Survey Research Roadmap (Part 1)

Prior to the Feb 23-25, 2022 meeting, each of the Gulf States (except Texas) and NMFS provided the Consultants with an overview of progress made in response to peer review recommendations and other research efforts to identify non-sampling error. The state-specific, one to two-hour meetings were appreciated by the Consultants. Gregg Bray provided an overview of the major findings of these informational meetings.

Florida: No major concerns were raised during the earlier peer-review and the certified design was expanded statewide in 2020. SRFS was designed to minimize sources of bias and FL has ongoing research to verify accuracy and identify additional sources of bias

Louisiana: LA Creel was designed as a full replacement for MRIP. About 73% of their private angler trips return to publicly accessible docks, and approximately 95% of private angler trips return within sampling timeframe.

Mississippi: Non-sampling errors could be occurring in Tails 'n Scales (e.g., private access sites). Concerns were expressed regarding trends in MRIP general survey harvest estimates that occurred during periods with low survey counts. MS feel this should be addressed and resolved prior to calibrating. MS also welcomes proposed projects to aid in effort estimation, private dock stats and others that will aid in drawing estimates across surveys closer together.

Alabama: A large portion of the annual red snapper fishing effort occurs during first few weekends of the red snapper season. The timing of this period of higher activity relative to timing of MRIP Fishing Effort Survey (FES) questionnaires may contribute to higher fishing effort estimates. Nonresponse rates have decreased since inception of program but they are at or above 50%. In terms of ongoing research, the state is exploring use video cameras and artificial intelligence software to determine offshore fishing activity. In addition, AL is starting a program to subsample on the water – data will be used to compare Snapper Check to the MRIP Access Point Angler Intercept Survey (APAIS).

Texas: No information available.

NMFS: MRIP has responded to certification peer reviews and 2017 NASEM review. These include research into effects of nonresponse and recall error in the FES, and APAIS coverage error. A sensitivity analysis (Foster et al. 2021) suggest non-sampling errors can account for differences in red snapper landing estimates for private boat mode, but understanding of the drivers is incomplete. It was noted the disparity of estimates across states increases with differences in designs and underlying assumptions.

The extended discussion among participants followed Bray's presentation is summarized below.

The Consultants noted that despite the pre-meeting presentations by the States and NMFS, there was insufficient time for the formulation of detailed questions in the one- to two-hour format. Preparation of a list of questions to each state by the Consultants was proposed to address this problem, but that responses would not likely alter their recommendations from this meeting.

Access to private docks could help estimate effort outside the sampling frame. Mississippi for example, attempts to include enforcement data to help estimate license compliance rates. This raised questions about the inclusion of non-probability samples in creel survey analyses as well as other disciplines. It was noted that the IRS has used non-probability sampling designs to audit high-income earners, which could be blended with their regulatory audits.

Collectively, the Consultants provided additional background on the Research Roadmap. A key point was that differences between MRIP and state surveys are not unique in the statistical literature. A study in 2013 compared five national surveys of health characteristics and found little agreement despite each survey being a rigorous, peer-reviewed design. The Consultants highlighted that seemingly unimportant aspects of survey implementation (question ordering for instance) could have major impacts. So called "gate-keeper effects" are thought to account for non-sampling error in telephone surveys, notably the Coastal Household Telephone Survey (CHTS) previously used in MRFSS. The Consultants' experience suggests that any detailed comparative survey analyses will require access to the raw sampling data and extensive modeling. Further discussions suggested that comparisons should also consider the varying uses of catch data for assessment and management, particularly within a season.

Day 2: February 24, 2022

Day 2 began with a review of progress on Day 1 and continuation of Survey Research Roadmap discussion. The Consultants then provided an overview of non-sampling errors in the Gulf. This was followed by plenary discussions and some final recommendations. Details of those recommendations and discussions follow.

Survey Research Roadmap (Part 2)

Overview of Non-Sampling Error in Gulf Surveys

Recreational catches in the Gulf states are derived either from 1) capture-recapture surveys with a ratio estimator or 2) probability surveys with a product estimator (i.e., angler intercepts x effort surveys). Each survey has a different sampling frame such that the target population of angler differs among states. While they share many common vulnerabilities to non-sampling error, they differ in others. Shared non-sampling errors include:

- Coverage error occurs when some sampling units are outside the sampling frame
- Nonresponse error occurs when sampling units (e.g., anglers, mail survey recipients) are unwilling to cooperate or unavailable/unable to respond. The key question is whether the behavior of non-respondents is equal to respondents. Effort surveys with product estimators are particularly vulnerable to this error. Consultants noted that while standard formulas are available to estimate the magnitude of nonresponse bias, it is not easy to do so because the required parameters are not available without special experiments.
- Measurement error arises from sources like questionnaire formats, subsequent data handling and other causes. Both the intercept and effort surveys are thus vulnerable to this source in both estimation systems. Note however, that errors in the angler report data are not a source of measurement error for the capture-recapture estimator, even if angles misreport. This is because these data are not used directly as responses, but rather as auxiliary variables.

Capture-recapture surveys are vulnerable to the following sources of non-sampling error:

- Matching error between the capture and recapture units can be either false positives, false negatives or mismatches. All types will inflate bias, but in different directions.
- Failure of the Independence Assumption occurs when anglers on intercepted trips report differently than non-intercepted anglers.

Key questions from Participants included:

- How is nonresponse characterized in capture-recapture compared to the other probability sampling methods?
    - Intercept surveys typically have a response rate around 80% because of the direct interactive nature of the interview. In contrast, self-administered effort surveys have much lower response rates. Increases in the nonresponse rate increase potential bias. Concerning comparability between compliance rates (capture-recapture) and response rates (probability sample), it was noted that the reports used as the second set of data in capture-recapture survey are not a probability sample. Instead, those data are used to calibrate, rather than measure catch. Inaccurate or incomplete data simply reduce the correlation between reporting and the actual trip. In turn, the variance increases. The nonresponse and measurement error that afflicts the capture-recapture estimator comes from the intercept survey.
- What is the impact of over-coverage, e.g., inclusion of non-fishing households in the FES?

- Over-coverage doesn't matter unless respondents are falsely reporting trips not taken. The focus is the percentage of people taking a trip such that nonresponse bias can come from zeros that should actually be non-zero trips. If none of the FES sampled non-fishing households respond, then the calibrated weights (control totals from the address-based sampling frame) for the responding fishing households would yield totals that are biased high. Estimation of proper sample weights is critical. The FES weighting process uses population-based totals and sample aligned with those calibration totals. There is no "disconnect" in that estimation system.

Discussion about measurement error contrasted the higher level of detail in the FL questionnaire with the FES questionnaire. Such differences enhance the opportunity to understand differences between the surveys (e.g., recall bias) but do not necessarily imply one is better than another. Given sound objectives of the survey program, focus groups can indicate how people respond to multiple data collection methods and facilitate choices consistent with those objectives. The Consultants described the cognitive interviewing process, which usually begins with validated questions from other surveys and progress to more complex topics. An iterative process used for developing the FES required several iterations and revealed the critical importance of the ordering of questions related to fishing activity. People generally want to tell surveyors that they fished in the recent past. A survey that does not address that tendency might be incorrectly filled out if the questionnaire narrowly restricts the period for reporting fishing activity. Incentives for responding (i.e., a small reward) have been shown to be useful, even for long questionnaires. Consultants recommended the use of the American Association of Public Opinion Research (AAPOR) response rate method 1 (AAPOR 2016) for all state and federal programs.

Specific Studies Suggested by Consultants

The Consultants made the following recommendations to the Transition Team:

1. All surveys should review estimator calculation. Consider weighting methodologies used to create final analysis weights. Review them and the program code or software to see if they are correctly applied to account for the sampling design and data being analyzed.
2. All states should review QA/QC procedures
    a. Does each state have a rigorous QA/QC program in place?
    b. Is there consistent training of all interviewers? Consistent within and across all the states?
    c. Is QA done on interviewers assigned to a site, e.g., a supervisor subsampling interviewees to ensure data is collected at the site during the assigned time frame on a regular basis?
    d. Does each state use the same formula for response rates (e.g. AAPOR codes)?
3. All programs that use an offsite effort survey should:
    a. Estimate effort using the intercept data and compare that with effort estimates obtained by the state surveys (FL and LA) and by MRIP through the FES data.
    b. The difference between the intercept-only effort and the offsite survey effort should be an estimate of the public access effort share. Is this ratio plausible? Can it be checked against other sources?

4. MRIP should consider conducting simulations to assess whether the standard formulas for variance estimation work for rare event species with low sample sizes and highly variable observations.
5. FL and LA surveys could be examined through use of split experiments to compare questionnaires with different samples. This could reveal if the difference in questionnaires are contributing to the differences observed in estimates. An example would be to use the MRIP questionnaire on the angler license frame, and the LA Creel/SRFS questionnaire on the FES frame. That could tease out if the differences are due to the sample frame (i.e., coverage error) or the questionnaire (measurement error).
6. Split experiment studies for FL could help address differences between the one-month surveys in SRFS with the two-month surveys in MRIP.
7. The MS program would benefit from an in-depth review of the weighting estimation procedures and use of a split experiments recommended for FL and LA. Further investigation of the utility of law enforcement data could be useful for determining origins of trips launched from public vs. private site. This might also inform differences in reporting rate between on-frame and off-frame anglers.
8. Recommendations for AL include cessation of using catch as a matching variable as this may be biasing the estimates. New matching criteria should consider testing variables to make the trips more distinctive from each other. One experiment that might be useful for future matching would be to ask intercepted anglers if they were aware that they are required to report, if they have already reported their trip or not, and when they became aware that they became aware that samplers were collecting data for this trip.
9. The primary recommendation for TX was to use a probability sample to collect data rather than the non-probability sampling currently in place.
10. All the surveys should assess potential differential reporting rates between public and private access anglers. If the catch per unit of fishing effort (CPUE) estimates are equal, then the potential for bias is reduced. Capture-recapture designs may be especially useful in this regard. Advanced remote sensing analyses (e.g., satellite tracking) might help identify the fractions of trips on and off data collection period.

Participant discussions included questions about weighted estimation, question sensitivity, social desirability bias and methods for estimating coverage error. These discussions later broaden into concerns about next practical steps for the roadmap itself.

The Consultants urged all participants to review the weighting procedures used in their surveys. Weights for site selections and meticulous records of trip counts at each sampling time are essential. Importantly, the Consultants noted that estimating inclusion probabilities (the inverse of weights) is formulaic. Weighting is more difficult for intercept surveys than for effort surveys because of the added complexity of selecting not only the sites but also the time period to visit the sites. Intercept surveys also provide valuable ancillary information. Florida, for example, uses supplemental intercept data to predict fishing effort within season for red snapper.

"Question sensitivity and social desirability bias" can affect the validity of catch estimates. Examples include state of residence, license status, and proper species identification for unobservable catches. In MRIP, the state of residence questions in APAIS affect estimates of

coverage within the effort survey. Hence, bias in one survey can induce additional bias in the effort survey.

The problem of private vs. public access sites affects all survey types. For capture-recapture surveys, biases in estimators can occur due to lack of coverage of private sites unless reporting rates and CPUEs are the same in public and private sites. To examine whether or not this is a problem, special studies must be conducted. Consultants noted that collection of date, time and precise location of interview sites can allow application of more advanced techniques. For example, Alabama's monitoring collects information on whether trips end at private or public docks. Satellite tracking or unique angler ID devices could be valuable for future improvements. These topics led to some discussion of novel effort monitoring studies such as the camera studies in Ocean City, NJ and Alabama ports.

Mississippi reported that marine patrols collect information on private vs. public launch sites at "random" (actually haphazardly), but consultants cautioned that "haphazardly" is not the same as random, and therefore will not produce a probability sample. Haphazard selections cannot be adjusted to perform as a probability sample does without better information about how the selected units were obtained or compare to the unselected ones. A probability sampling design should be adopted prior to the data collection.

Questions were raised about ongoing simulation studies to help refine estimators for rare species. Initial evidence suggests such estimates are more variable than once thought. Sampling theory is based on large samples so explorations of estimator behavior at small sample sizes are valuable. An overall message from this session was that more examination of existing data, coupled with simulation studies, would be valuable for measuring deviations between theory and practice.

Future research efforts are most obviously constrained by funding and time but also by the roles of the partners. A key consideration is whether research should focus on accuracy or comparability questions. It is not clear if improving individual surveys will make surveys more comparable or not. Despite this uncertainty, all partners expressed in interest in executing the research plan.

NMFS has been mandated by Congress to reallocate $2M. Options include funding for consultants, independent reviews, and potentially additional studies. Given the wide disparity in costs of various studies, it was suggested that states identify their highest priorities. States may also be able to leverage funding from other sources.

Methods for Calibration (Part 1)

As the Consultants noted, the goals of research for accuracy vs. calibration will not necessarily be the same. Nonetheless, calibration is a requirement for stock assessments to assess the status of populations and the basis for subsequent monitoring of catches by jurisdiction. If side-by-side comparisons reveal large differences in catch estimates among states, calibration is required to establish a common currency. Richard Cody's presentation noted that calibration is a

longstanding issue with five workshops for red snapper over the past nine years. Owing to the complexity of full model-based comparisons, simple ratio methods have been the preferred/feasible method to date. These were reviewed by the Gulf Council's SSC in 2020 and recommended for implementation in 2023. In October 2021, NMFS and State Directors met to finalize plans for transition. Key decisions included recommendations for a common database, improved understanding of the differences among surveys, and improved communications.

Ratio methods applied to date have been based on empirical comparisons among surveys but do not explain the differences in estimates. Current surveys all have an angler intercept component. This might be a useful place to look for commonalities. Correlations among surveys at the intercept level could help identify predictor variables. Such patterns could lead to a "new currency" which does not use any single survey (e.g., MRIP) as the frame of reference. This approach was used by one of the Consultants to calibrate two versions of the 2016 Fishing, Hunting and Wildlife Associated Recreation (FHWAR) survey collected using different modes (see Erciulescu et al, 2021). A primary difference between the Consultant's referenced calibration work and the current calibration needs for stock assessment purposes is the need for historical comparability. While such methodology might be generalizable to all species, the present focus will be on red snapper. Only MRIP, LA Creel and Texas Coastal Creel Survey estimate catch for all species.

An earlier workshop report (GSMFC & NMFS 2020, hereinafter referred to as the "White Paper" identified four primary options for calibration. Discussants noted that Options 1a and 1b were the most viable given time and resource constraints. There was some additional discussion about the details of currency conversion. In particular, everyone agreed that consistent application of conversion factors to different modes within surveys is essential.

A Hybrid Approach

The needs for stock assessments within the SEDAR process were presented by Kate Siegfried. Catch estimates affect every aspect of status determination, so there are no provisions for competing perceptions of harvest removals—an agreed upon time series of catch is essential. Theory and simulations have confirmed however that an arbitrarily scaled time series of catch is sufficient for status determination and management, as long as the estimation methodology remains the same.

John Foster continued the presentation with a series of examples. Simple ratio comparisons between state surveys and MRIP estimates for years when both surveys were conducted revealed a wide range of ratios (MRIP:State) from 1.8x for Louisiana to 11x for Texas. States with comparable programs, such as Florida and Louisiana had similar ratios, 2.7x and 1.8x respectively. Florida's surveys showed similar trends to MRIP. Alabama had a similar increasing trend, but with more variation. Mississippi also had reasonable agreement but the scaling ratio was 5.6x. For Texas and Louisiana there is only one year of overlap with MRIP so no trend comparisons are possible. Calibrations to alternative state surveys have relatively little effect on the trends, but huge differences in the totals. Moreover, the relative allocations among

states will vary with different currencies. These considerations are far from academic. Practical questions include:

- Do we use state, MRIP or composite metrics of harvest?
- Are the years used for calibration fixed, or amenable to future changes?
- Do calibration decisions vary by species?
- Will calibration decisions be updated regularly?

Participants acknowledged that calibrations could and possibly should evolve annually. Some insights into the magnitude of such changes might be found by examining ratios for individual years in instances where multiple years of overlap occur. Regardless of the calibration standard used, developing a consistent time series is perhaps the most important criteria. Currency conversion can be problematic if changes in calibrations induce changes in stock status.

Other participants discussed the importance of accuracy of the surveys in creating a standard for comparison. Comparisons with the recent independent biomass estimate of red snapper (The Great Red Snapper Count) were a possibility but there was limited discussion. Consultants asked if the Hybrid Approach used weight trimming to reduce outliers. Such trimming is used to exclude weight values above the 95th to 99th percentile. Trimmed estimates are particularly important in management because premature fishery closure based on extreme observations are problematic for all stakeholders.

Further discussions of the Hybrid Method led participants to ask how it differed from Option 1b in the White Paper. MRIP staff explained that while the methods are similar, the Hybrid Method relaxed some of the assumptions about which surveys were more biased. Regardless of the method, participants agreed that multiple conversions between assessment results and monitoring estimates could be time consuming, hard to communicate, and possibly error prone. Moreover, if present patterns continue, some states may be able to update their calibration ratios (e.g., FL) but in other states where MRIP is no longer conducted (e.g., TX), the ratios will remain fixed unless auxiliary information can fill gaps. A variety of scenario permutations were proposed but the primary calibration options discussed boiled down to:

- Conduct a detailed, model-based analysis to generate an aggregate measure of total catch
- Use the Hybrid Method discussed above
- Use Option 1a or 1b from the White Paper.

Consultants noted that the detailed model-based analyses would not be possible in the short term but should be considered in a longer-term research plan.

Time did not permit sufficient discussion of the remaining calibration options on Day 2 of the workshop. Instead, the topic was deferred to Day 3.

Database Storage and Data Management

Day 2 of the workshop concluded with presentations on the development of a common database and warehouse for handling regular updating. Gregg Bray's (GSMFC) presentation highlighted that in the short run, all parties should provide both raw and processed data with associated metadata. GulfFIN already serves this function for LA and TX data sans the metadata. Expansion of their role to the total set of estimates was proposed as a starting point. In the longer term, development of a standardized database and minimal key data requirements are necessary. Development of a model-based calibration estimate would require high-resolution data rather than aggregated data typically available. Gregg Bray repeated their commitment to handling whatever level of resolution partners were willing to provide. Others acknowledged that a regional approach to calibration was best but it would require a "burn-in" period to understand the relationships among the survey data.

Matt Nuttall (SEFSC) reported the Center's need for the finest scale resolution possible including survey estimates, population totals used to calibrate the survey weights, the raw data tables on angler intercepts and biological data, and finally, the metadata for survey design. Not all demands of the data known in advance, so having the ability to reconstruct estimates as appropriate was judged a desirable feature. Therefore, all the information used to create sampling, nonresponse and calibration weights is needed. Finally, it was stressed that timing of data deliveries was essential for use by scientists and particularly managers. SEFSC endorsed the concept of expanding the role of GulfFIN for this purpose.

In the ensuing discussion, the Consultants emphasized the ancillary data associated with the design including the weights, stratum information, PSU data and so forth. At the finest scale, it is important to have the actual processing code used by the partners. Proper weighting of observations is perhaps the most important aspect of survey design; having sufficient information for auditing of this process was acknowledged difficult but essential. Proper use of multiple databases requires a deep understanding of the algorithms used for distillation of estimates. It was acknowledged that this aspect of data storage is beyond current mandates and expertise. No single group would likely have the resources to implement such a comprehensive system.

Lisa Hollensead (GMFMC) emphasized that Gulf Council's needs for management are linked to timely decision making. Moreover, such data should be readily available to all parties including the general public. The focus here is the end products rather than the atomic components of the estimates.

Day 3: February 25, 2022

Discussions about data management needs continued on Day 3. GulfFIN representatives emphasized again that it could conduct some auditing responsibilities but could not be responsible for full-scale auditing of each submission. Consultants noted that the intensive auditing was a precursor to having a useful integrated database. Without such auditing, the warehouse would become a collection of disparate, hard to compare databases. Some states

echoed concerns similar to GulfFIN—they are focused primarily on executing the designs and not sufficiently staffed to handle some of the proposed analyses.

MRIP is currently responsible for detailed documentation of the design, weighting methodology, estimation methodology, data processing designs, and quality control. The purpose of this documentation is to ensure that data are being collected according to the design. Current standards (NMFS 2020) also require annual reports which summarize data collection procedures and within year changes for each year.

All participants agreed that improvements to the data standards process are both desirable and valuable. Sound foundational principles such as the Recreational Fishing Survey and Data Standards and SEDAR stock assessment guidance documents are excellent starting points.

SERO highlighted its responsibilities for in-season monitoring and support for development of regulations for landings and discards. These data include daily catch rates, size limits, bag limits and economics of individual trips. SERO noted that many of its responsibilities spanned the entire Gulf of Mexico rather that state specific estimates.

Methods for Calibration (Part 2)

Participants and Consultants returned to the unfinished business from Day 2 to select a calibration method. The primary options under consideration are summarized in the table below. Viable options included two from the 2020 review (Options 1a and 1b, White Paper), and three proposed at this meeting. Foster's two proposals address the alternative currency issue by either converting all state-based catch estimates upward by the ratio of recent MRIP to state estimates. Conversely, one could convert historical MRIP estimates by state to their new values by multiplying them by the ratio of recent state to MRIP estimates (i.e., the inverse transformation). Model outputs would be expressed in either MRIP currency or summed state currencies. The option based on Opsomer's proposal (Erciulescu et al. 2021) to conduct an in-depth model-based analysis of each survey would be an important advance. However, because it would likely take several years to complete, this option was not considered further for short-term assessment needs. Instead, the Opsomer proposal should be considered as part of a research track for longer-term needs.

Option	Model Inputs by State		Model Outputs	Monitoring	Comments
	MRIP Data	State Data			
1a	Fully calibrated(FES, APAIS) series	La Creel converted to MRIP units. TX not addressed	MRIP units	MRIP units	LA only adjusted by ratio of overlap between years
1b	Fully calibrated(FES, APAIS) series	La Creel converted to MRIP units. TX not addressed	Converted to State units	State units	Adjusted by ratio of overlap between years
Foster: Calibrate to MRIP	Fully calibrated(FES, APAIS) series	**Converted to MRIP currency.** Used for to improve estimates for 2017+	MRIP units	MRIP units in FL, AL, MS. Ratio adjusted catches in LA, TX	State data raised by ratio in overlap period. Limited info for LA & TX. Population estimate greater.
Foster: Calibrate to State	Fully calibrated (FES, APAIS) series **converted to state currency.** Used to improve estimates for 2017+.	State only in 2017+. MRIP converted to State units before 2017. Total Catch= sum over all states	Sum of scaled State Units	State units	MRIP data reduced by ratio in overlap period. Model output is linear combination of MRIP adjusted catches. Population estimate lower.
Opsomer, AFWA	Model adjusted	Model Adjusted	New Units	New Units	New currency. Time>2 yrs

Participants recognized the need to pursue both short- and long-term solutions and affirmed their willingness to work towards imperfect but necessary short-term solutions to meet upcoming stock assessment needs. The diversity and complexity of the survey methods heightens the need for detailed specification of data conversion and increases the chances of miscommunication among stakeholders. Discussion among participants focused on identifying specific recommendations in light of these overarching concerns.

Participants noted that consistency is a central tenet of BSIA. It was argued that if state currencies were to be used over the entire assessment period, a peer review would be needed. The inability to meet this requirement in the short run pushes the group toward option 1a or 1b. Even option 2 in the White Paper document would require a peer review. Similar sentiments were expressed about the options proposed by Foster. The current meeting was not designed to be a review of the methodology. If either of the Foster approaches were to be used, an additional peer review would be necessary.

Further discussions helped clarify the scope of needed reviews. Ratio based calibrations have been reviewed in the White Paper, but there are nuances of calibrating back in time and handling different fishing modes, particularly shore-based fishing activity. Consultants have previously advised state and MRIP representatives on methods for calibration of means and standard errors. A similar follow up meeting to discuss Terms of References for stock assessments was proposed. The upcoming gag and red snapper assessments should be illustrative of the problems of calibration. The gag fishery is dominated by harvests from Florida whereas red snapper are harvested by all states in the Gulf.

After considerable discussion, a proposal to use Option 1b from the White Paper for red snapper was advanced. Participants recognized that the complexities of adjusting five state-specific

survey estimates that have varying degrees of overlap with MRIP estimates would make this a difficult task. Some surveys provide year-round estimates of total landings and discards, whereas others only provide in-season harvest estimates, making it difficult to produce a Gulf-wide calibrated time-series for released catch in state survey currencies. Additional concerns about use of Louisiana and Texas calibration ratios based on a single year of overlap with MRIP data were expressed. Presently, there is no agreed upon method for converting Texas catch data into MRIP units. Foster's approaches did include a proposed conversion factor. Representatives from Texas were not present for this discussion and concerns were expressed about the implications of rescaling the entire time series of red snapper catches to MRIP equivalents. Since the consultants were not provided the opportunity to visit and discuss the details of the Texas survey, consultants affirmed that they have not reviewed the survey procedures used in Texas.

Additional discussions by the participants led to further consideration of Option 2 for gag grouper. The above text table (with five options) reflects what was presented to the group at the meeting. Details of Option 2 may be found in the White Paper (2020). Option 2 would rescale the MRIP time series to the different state survey currencies. The White Paper (2020) noted that this premise "would be very difficult to defend on statistical grounds." Gag grouper, however, are caught primarily (~95%) in Florida and monitored by SRFS, so the need for comparable estimates across multiple states is less relevant for this stock. Hence, the group recommended that Option 2 be used for gag grouper.

Following some additional clarifications, the group consensus is as follows:

- Use Option 1b for red snapper.
- Use Option 2 for gag [see 2020 White Paper for these options].
- Convene a group to set Terms of Reference and identify individuals to conduct a formal review of the ratio-based calibration for gag and red snapper at minimum, and others as appropriate.

Components of the Transition Plan

Richard Cody gave a presentation of the Transition planning required under NMFS Policy Directive 04-114. An important aspect is the distinction between certifying a design and implementing it. Plans that are implemented as designed are capable of producing unbiased estimates. An initial draft outline of the Transition Plan was presented to the group. The basic outline is as follows:

I. Executive Summary
II. Introduction and Purpose
III. Description of Approach and Timeline
IV. Potential stock assessment impacts and schedule
V. Identification of Unknowns
VI. Lessons learned
VII. Appendices

Section III outlines the needs for a Transition Track and a Research Track as well as ongoing Communications needs. The Transition Track has both short and long-term needs. Short-term needs are dominated by the responsibilities of getting data and interim calibrations to stock assessment scientists as soon as possible. Longer term needs include convening an independent review of final calibration procedures for use in stock assessment and management, and maintenance of the FIN database. The Research Track involves progressive improvements of survey methodologies and ancillary studies.

Section V on Identification of Unknowns addresses the potential reduction in non-sampling errors, the potential alignment among surveys, and whether full integration of the survey data (e.g., composite estimation) is possible. Funding and resource limitations are a major unknown. Process obstacles such as the Paperwork Reduction Act are non-trivial.

The Plan outline will be fleshed out and circulated to the Transition Team in a timely manner for review and comment. Participants echoed the urgency of finalizing the Transition Plan. Avoiding species- and assessment-specific approaches should also be a goal of the Transition Plan. Owing to the complexity of the underlying surveys, it is anticipated that expertise outside the SEDAR process would be necessary.

Communications Plan

An outline of a communications plan was presented to the group by Catherine Krikstan (OST). Overall goals of the plan are to facilitate development of the transition plan and its implementation. The plan will help ensure that workshop outcomes are clear, easy to find and promptly distributed to target audiences. Progress on milestones will be regularly published. Target audiences include not only transition team members and leaders of their respective agencies but also external stakeholders: Council members, congressional representatives, anglers and various advocacy groups.

Immediate next steps will be development of short-term tactics and development of a formal communications plan for inclusion in the Transition Plan. Participants should nominate members for the Communications Working Group.

Closing Statements

Participants were reminded of a number of remaining tasks. Notably, the Consultants planned to convene following the meeting to develop a list of final recommendations. Their report has been included its entirety in Appendix 3. Consultants noted that not all of future decisions and recommendations will be statistical. Participant consensus will be essential for those non-statistical elements driven by assessment or management needs.

The initial goals of the workshop were largely met. The primary exception was development of a statement of task for the final independent review of methods. This will be addressed later by the transition team after the major tasks and studies have been identified.

Participants and consultants appreciated the planning that occurred before the workshop and the presentations by the states and MRIP staff to the Consultants immediately prior to the workshop. This ensured a better understanding among all participants of the distinctive methods and concerns of each group. Presentations during the meeting were uniformly excellent and stimulated highly engaged discussions among the Participants and Consultants. There was a general sense that the workshop had been successful, especially with respect to agreeing on a framework for satisfying data needs for the upcoming gag and red snapper assessments.

Acknowledgements

The workshop chair, Paul Rago, expresses his appreciation to the members of the Transition Team for their helpful consultations before, during and after the Workshop. Richard Cody and Gordon Colvin are thanked for overall guidance on the project and abundant background. John Foster and Katie Siegfried provided survey and stock assessment expertise, respectively. Matthew Titlow coordinated the many moving parts. Catherine Krikstan and Katherine Papacostas provided communications and editorial assistance. In particular, Katherine Papacostas prepared extraordinary notes from the meeting and in-depth reviews of earlier drafts of this report. She somehow managed to capture all the salient points from the meeting. It is hoped that this distillation of the proceedings adequately captures the flow of the discussions, the major concerns, and the primary conclusions of the meeting. A special thanks is appropriate to the organizers, the participants, and the Consultants for the informative questions and debates over the course of the three days.

List of Presentations and Useful Links

All of the presentations and a recording of the meeting may be found at: https://www.fisheries.noaa.gov/event/gulf-state-recreational-catch-and-effort-surveys-transition-workshop

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

National Academies of Sciences, Engineering, and Medicine (2006) Review of Recreational Fisheries Survey Methods. Washington DC: The National Academies Press. https://www.nationalacademies.org/our-work/review-of-recreational-fisheries-survey-methods

National Standard 2 https://www.ecfr.gov/current/title-50/chapter-VI/part-600/subpart-D/section-600.315

References

American Association for Public Opinion Research. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR. https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Erciulescu, A., J.D. Opsomer, and F.J. Breidt (2021). "A Bridging Model to Reconcile Statistics Based on Data from Multiple Surveys." *Annals of Applied Statistics,* 15: 1068-1079 (DOI: 10.1214/20-AOAS1437).

Foster, J., Papacostas, K., Krikstan, C., Andrews, R., Kitts-Jensen, R., Cody, R., Colvin, G., Dolinger Few, L. (2021) Evaluating Uncertainty in Gulf Red Snapper Estimates: A Preliminary Sensitivity Analysis of Non-Sampling Errors in the Region's Recreational Fishing Surveys. https://media.fisheries.noaa.gov/2022-02/NOAA-Fisheries-OST-Preliminary-Sensitivity-Analysis-Non-Sampling-Errors-Gulf-Recreational-Fishing-Surveys.pdf

Gulf States Marine Fisheries Commission and National Marine Fisheries Service (2020) GSMFC-NOAA Fisheries Workshop on Red Snapper – Workshop Summary Report. https://gulfcouncil.org/wp-content/uploads/B-8b-Gulf_Calibration_Wrkshp_report_2020_v1.21.pdf.

House Committee on Appropriations (2021) Joint Explanatory Statement of the Committee of Conference on H.R. 133: Title I, Division B. House Rules Committee on Senate Amendment to the Consolidated Appropriations Act of 2021. *Congressional Record* 166:218 (Book III). https://www.govinfo.gov/content/pkg/CREC-2020-12-21/pdf/CREC-2020-12-21-house-bk3.pdf.

Magnuson-Stevens Fishery Conservation and Management Act, P.L. 94-265, as amended through P.L. 109-479. (2006) https://www.congress.gov/bill/109th-congress/house-bill/5946.

National Academies of Sciences, Engineering, and Medicine (2021) Data and Management Strategies for Recreational Fisheries with Annual Catch Limits. Washington, DC: The National Academies Press. https://doi.org/10.17226/26185.

National Marine Fisheries Service (2020) Recommended Use of the Current Gulf of Mexico Surveys of Marine Recreational Fishing in Stock Assessments. Silver Spring, MD. https://media.fisheries.noaa.gov/dam-migration/94100569.pdf.

National Marine Fisheries Service (2020) Recreational Fishing Survey and Data Standards. Silver Spring, MD. https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-survey-and-data-standards

National Marine Fisheries Service (2021) Policy 04-114: Implementing Recreational Fisheries Catch and Effort Survey Design Changes. https://media.fisheries.noaa.gov/2021-06/04-114_2021.6.9_final%20for%20Doreumus%20Signature_signed.pdf?null

Paperwork Reduction Act, P.L 104-13. (1995) https://www.congress.gov/104/plaws/publ13/PLAW-104publ13.pdf

Appendices

Appendix 1. Agenda

WORKSHOP ON TRANSITION TO USE OF STATE SURVEY CATCH ESTIMATES IN GULF OF MEXICO FISHERIES

Workshop Dates: February 23-25, 2022; all times approximate

AGENDA

Feb. 23: 9:00 am - 5:00 pm EST

I. 9:00 am - 9:15 am: Meeting logistics, introductions: **Gregg Bray**, **Richard Cody**
   A. Virtual connection instructions and where to get help.
   B. Self-introduction of attendees

II. 9:15 am - 9:45 am: Review Workshop Purpose, Desired Outcomes, Agenda: Workshop Chair, **Paul Rago**
   A. Overarching goal: Agree on the elements of a transition plan that will achieve goals 1-5. (This plan will be written and executed following the workshop).
   B. Long-Term Goals of the Transition Plan
      1. Identify any needed design changes to improve the accuracy of all survey programs, thereby minimizing differences in estimates.
      2. Incorporate state data into the federal science and management process while maintaining the needed consistent, regional time series.
      3. Develop a single, publicly accessible, standardized database to house all the recreational fishing data streams in the Gulf of Mexico.
      4. Develop guidelines and best practices to inform future decision-making regarding BSIA when overlapping statistically valid data streams exist (note: while the transition team will be developing guidance, it does NOT have the authority to make BSIA determinations).
      5. Maintain clear and open lines of communication between the Transition Team and all affected stakeholders about progress toward the above goals.
   C. Specific Desired Outcomes of the Workshop
      1. Toward Goal 1:
         a) Adopt a research and analysis plan for understanding the drivers of differing catch estimates among the Gulf red snapper surveys.
         b) Adopt Terms of Reference for a Congressionally directed, independent peer review (that will be commissioned after the research and analysis plan has been implemented).

2. Toward Goal 2:
    a) Decide on a calibration approach to be used to 1) inform new stock assessments happening this year (e.g. ratio-based methods), and 2) to inform stock assessments and management in the longer term (e.g., model-based methods).
    b) Describe and evaluate a hybrid approach that puts the transition on two tracks: (1) initiate transition using available data for stock assessments in the near term; (2) complete research plan and revise data collection to inform future assessments.
3. Toward Goal 3:
    a) Determine technical and funding requirements to create the unified, regional database and assign specific roles and responsibilities to NMFS, states, and regional partners to initiate its development.
    b) Identify needed updates to data flow processes for stock assessments.
4. Toward Goal 4:
    a) Review data requirements of National Standard 2.
5. Toward Goal 5:
    a) Agree on the goals and components of a communications plan (to be adopted following the workshop).

III. 9:45 am - 10:15 am: Break

IV. 10:15 am - 11:30 am: Transition Plan Content: **Richard Cody**
- A. Present Work Group's list of Milestones that are to be addressed in a Transition Plan;
    1. Include description of hybrid approach (see also VI C below) that puts the transition on two tracks: (1) initiate transition using available data; (2) complete research plan and revise data collection designs/calibration/integration as indicated by the results;
- B. Group discussion facilitated by Workshop chair. Intent is to achieve a general consensus of the Transition Team Subgroup members on the content of the transition plan, and timing for completion of the various milestones.

V. 11:30 am - 2:00 pm: Lunch Break

VI. 2:00 pm - 3:00 pm: Review of Data Requirements to meet NMFS requirements for management and stock assessment:
- A. Management requirements: **Andy Strelcheck**
- B. Science requirements, including SEDAR 74 specific needs: **Kate Siegfried**
- C. BSIA: **Patrick Lynch**

VII. 3:00 pm - 3:30 pm: Break

(Start on item VIII in the afternoon of Day 1, then continue the discussion in the AM on Day 2)

VIII. 3:30 m - 5:00 pm: Survey Research Roadmap: **MRIP Consultants** and Workshop Chair **Paul Rago**

A. Independent review of surveys called for by Congress
    1. Alignment of the mandate with requirements for transition
B. Summary of state and NMFS pre-briefings to consultants: **Gregg Bray**

Feb. 24: 9:00 am - 5:00pm EST

(VIII cont.) 9:00 am - 11:00 am

C. Overview of non-sampling errors and their effects on survey estimates: **MRIP Consultants**
    1. General overview of non-sampling errors and effects with examples from Gulf surveys
D. **MRIP Consultants** will present a proposed Roadmap that identifies research, pilot studies, and analyses--related to sources of non-sampling error and bases for differences in estimates among the Gulf Surveys (see detailed statement of task for this product appended to this agenda).
E. Group discussion of Roadmap facilitated by Workshop Chair:
    1. Achieve consensus of attendees regarding scope and needs for completion of identified research.
    2. Agree on roles for partners in executing research plan.
    3. Discuss funding options (limiting the discussion to identifying totals that might be needed to fund research projects), sequencing, and timing.
    4. Discuss Terms of Reference Draft prepared by NMFS/OST

IX. 11:00 am - 11:30 am: Break

X. 11:30 am - 12:30 pm: Calibration I

A. Review of current status of calibration method development and use: **Richard Cody**
B. Overview of calibration methods, limitations and future use considerations: **MRIP Consultant(s)**, including requirements for enabling integration of calibrated state estimates into a single Gulf-wide estimate for the affected species: **Jean Opsomer; Lynne Stokes**, **other member(s) of the consultant team**?
C. Potential Hybrid Approach whereby all available rec catch series are used in assessments while research is ongoing, including different ways the data might be incorporated into the assessment models or how outputs from separate model runs might be integrated into final results (**Foster, Siegfried, Nuttall**).

XI. 12:30 pm - 2:00 pm: Lunch Break

XII. 2:00 pm - 3:30 pm: Calibration II

A. Facilitated discussion of preferred options for pursuing integration of catch estimates: **Paul Rago**, facilitator
    1. Seek group's consensus on option(s) to implement.
    2. Identify requirements and partner needs for implementation. Note that Reconciliation of completing needs will need to be a primary consideration. There may be options to produce a Gulfwide estimate that considers all of the data and allows us to make progress on transitioning while attending to the research needs that will help identify improvements to be made. Potential options (maybe at the assessment level) beyond standard statistical integration methods and associated assumptions and requirements will need to be evaluated.
    3. Identify calibration requirements for species with state specific or restricted distributions (e.g., gag).
B. Facilitated discussion of preferred options for calibration, including phased approaches (e.g., using a ratio based approach in the interim while a model based approach is being developed): **Paul Rago**, facilitator
    1. Seek group's consensus on option(s) to implement.
    2. Identify requirements and partner needs for implementation.

XIII. 3:30 pm- 4:00 pm: Break

XIV. 4:00 pm - 4:30 pm: Additional Requirements for Completing Transition

A. Database Storage and Data Management via FIN
    1. Input and Data Management Requirements: **Gregg Bray**
    2. Output Requirements: SEFSC (**Matthew Nuttall**); SERO: **SERO** requirements for data outputs will be covered in Andy Strelcheck's presentation under VI.A on 2/23--additional questions can be addressed here; Gulf Council (**Lisa Hollensead; Ryan Rindone**?)
    3. Roles and responsibilities of Data Providers/State Survey Managers
B. Identify and resolve how and when to address any unmet requirements for use of estimates from State Survey Certification Decision Memos **Richard Cody**:
    1. Generally, these include the following for each survey:
        a) integration (see VII.A.2: not completed for any of the Gulf surveys at present,
        b) calibration with the legacy time series,
        c) peer review: Note that, for the Gulf state surveys, this requirement occurs at two levels:
            (1) Initial independent expert review conducted for MRIP Certification, completed for FL, AL, MS, LA surveys;
            (2) independent review of all the current Gulf surveys required by Congress per agenda item V above.

d) apply to time series

2. Also address any individual, specific next steps from Certification Memos, including SnapperCheck Conditions of Certification requirements.

C. Others?

XV. 4:30 pm - 5:00 pm: Components of the Transition Plan

A. Identify a small group to prepare the Transition Plan Outline, based on the work group's discussion, in an overtime session the evening of Day 2.

B. Review and discuss the proposed outline

Feb. 25: 9:00am - 1:00pm EST

XV. 9:00 am - 10:30 am: Components of the Transition Plan (continued)

C. Presentation of the Transition Plan Outline by the assigned group.(XV.B)

D. Discussion of the proposed components to seek general concurrence with that list and of a process and timing to scale it up to a full Transition Plan for submission to NMFS by the Gulf Subgroup.

E. Revisit long-term goals and desired outcomes of workshop

XVI. 10:30 am - 11:00 am: Break

XVII. 11:00 am - 12:00 pm: Communications Plan: Discuss, agree on plan components: **Catherine Krikstan**

A. Among Transition Team Members

1. Goals, purpose of plan
2. Key tactics
    a) Regular (monthly?) ad hoc State Directors' Meetings
    b) Bimonthly meetings of MRIP CET Regional Work Group
    c) Assemble Gulf Transition Team Communications Work Group consisting of representatives from NMFS, Gulf states, Gulf council, and GulfFIN.
    d) Briefings at Gulf Council and GSMFC meetings
3. Other?

B. Stakeholders: same items as above?

XVIII. 12:00 am - 1:00 pm Wrap Up

A. Final advice and recommendations from MRIP Consultant Team

B. Closing comments by Transition Team Gulf Subgroup members

Appendix 2: List of Participants

First Name	Last Name	Affiliation
Jason	Adriance	Louisiana Department of Wildlife and Fisheries
Rob	Andrews	NOAA Fisheries OST
Kevin	Anson	Alabama Department of Conservation and Natural Resources
Patrick	Banks	Louisiana Department of Wildlife and Fisheries
Scott	Bannon	Alabama Department of Conservation and Natural Resources
Luiz	Barbieri	Florida Fish and Wildlife Conservation Commission
Dave	Bard	NOAA Fisheries OST
Harry	Blanchet	Louisiana Department of Wildlife and Fisheries
Kevin	Bland	Louisiana Department of Wildlife and Fisheries
Gregg	Bray	Gulf States Marine Fisheries Commission
Mike	Brick	Consultant, Westat
Richard	Cody	NOAA Fisheries OST
Gordon	Colvin	NOAA Fisheries OST
David	Detlor	NOAA Fisheries OST
Jill	Dever	Consultant, RTI International
John	Foster	NOAA Fisheries OST
John	Froeschke	Gulf of Mexico Regional Fishery Management Council
David	Gloeckner	NOAA Fisheries SEFSC
Marie	Head	Alabama Department of Conservation and Natural Resources
Matt	Hill	Mississippi Department of Marine Resources
Lisa	Hollensead	Gulf of Mexico Regional Fishery Management Council
Evan	Howell	NOAA Fisheries OST
Cliff	Hutt	NOAA Fisheries OSF
Catherine	Krikstan	NOAA Fisheries OST
Michael	Larkin	NOAA Fisheries SERO
Virginia	Lesser	Consultant, Oregon State University
Ty	Lindsey	Louisiana Department of Wildlife and Fisheries
Alan	Lowther	NOAA Fisheries SEFSC
Patrick	Lynch	NOAA Fisheries OST
Vivian	Matter	NOAA Fisheries SEFSC
Jessica	McCawley	Florida Fish and Wildlife Conservation Commission
Trevor	Moncrief	Mississippi Department of Marine Resources
Matt	Nuttall	NOAA Fisheries SEFSC
Jean	Opsomer	Consultant, Westat
Katherine	Papacostas	NOAA Fisheries OST

Clay	Porch	NOAA Fisheries SEFSC
Paul	Rago	Workshop Chair, Scientific and Statistical Comm., MAFMC
Ryan	Rindone	Gulf of Mexico Regional Fishery Management Council
Michael	Ruccio	NOAA Fisheries OSF
Bev	Sauls	Florida Fish and Wildlife Conservation Commission
Kate	Siegfried	NOAA Fisheries SEFSC
Carrie	Simmons	Gulf of Mexico Regional Fishery Management Council
Joe	Spraggins	Mississippi Department of Marine Resources
Lynne	Stokes	Consultant, Southern Methodist University
Andy	Strelcheck	NOAA Fisheries SERO
Matt	Titlow	NOAA Fisheries OST
Mike	Travis	NOAA Fisheries SERO
Joe	West	Louisiana Department of Wildlife and Fisheries
Chris	Wright	NOAA Fisheries OSF
Xinan (Adrian)	Zhang	Louisiana Department of Wildlife and Fisheries

Appendix 3: Recommendations of Statistical Consultants

Summary of Recommendations Based on Transition Team Gulf Workshop Meeting
February 23-25, 2022

Mike Brick, Jill Dever, Virginia Lesser, Jean Opsomer, Lynne Stokes
March 28, 2022

Overall Comments

The goal of the state surveys is to increase the precision of estimates of catch for one or more species of high interest to the state. The state surveys have mostly achieved this goal, in the sense of increasing the sample sizes and greatly reducing the variance of the estimates. However, sizeable non-sampling errors, which vary between the states, may be present in the estimates. Virtually all surveys conducted today are subject to non-sampling errors, so this does not imply that these surveys are not scientifically valid. Nevertheless, the fact that the types and magnitude of the non-sampling errors differ between states and are different from those in MRIP greatly complicates directly comparing the estimates obtained from the surveys as well as obtaining comparable data that when combined across the states allow production of Gulf State-wide estimates.

There are two broad approaches to improving comparability between the surveys: (1) identifying and removing sources of non-sampling error, and (2) harmonizing the methods (questionnaires and other data collection methods, sampling design, sampling frame, estimator construction, etc.) used in the surveys so that the *differences* in non-sampling error are reduced. Approaches (1) and (2) can be pursued in parallel. We believe, however, that implementing (1) alone is unlikely to sufficiently remove differential impacts of non-sampling errors.

One approach to eliminate differences among the states is to adopt a single sampling design for all states. Using an identical design has the best chance of all options available to eliminate the differential errors across states that create challenging comparability problems. In order to also achieve the precision goals of the states, the sample sizes could be increased under this design relative to those currently used in MRIP, while maintaining comparability. While making the designs identical is likely not realistic, we still want to state this option, to make clear that obtaining directly comparable estimates is achievable, at least in principle.

In the remainder of this report, we have assumed this option is not available, and formulated our recommendations accordingly.

Recommendations

1. *Differences in estimates*: As discussed at the meeting, the differences between estimates obtained under various methods used by the states and MRIP is not unusual. Other agencies have seen this issue and have addressed it. Some agencies report their estimates and recognize another agency, with different methods, reports another estimate. Examples of discrepant measures by federal agencies are unemployment estimates from states and federal (BLS) and from the population (Census).

   Another example is one where different federal agencies measure a parameter differently such as poverty measures; sometimes American Community Survey (ACS) estimates have gone in the opposite direction from the Current Population Survey Annual Social and Economic Supplements (CPS ASEC) official estimates. Census has chosen to handle this by advising users to work with the CPS ASEC estimates for national numbers, the ACS for state numbers, and yet a third source, the (modeled) Small Area Income and Poverty Estimates (SAIPE), for substate estimates. SAIPE estimates are used for Title 1 formula funding; ACS estimates are used in some programs; and CPS estimates are not used in funding (although they are the "official" estimates) because the sample size is too small. As indicated, this is an administrative decision.

   Drawing a parallel to the recreational fishery situation, one could use the MRIP estimates to allocate fractions of annual catch to the states (with LA and TX still handled separately), because it is the survey that is most comparable across states, while a different approach is used to determine total annual catch, for instance based on the state surveys. We are not necessarily advocating this particular option, but want to point out that it is possible to “pick and choose” among the estimates for different purposes, as long as it is consistently done and well documented.

2. *Review weighting methodology for creating final weights used in estimation*: This includes reviewing the process for MRIP (both APAIS and FES) and the details used to generate weights for each state, to make sure that they correspond to the sampling design and correctly adjust for nonresponse. This would include both the effort and intercept weights for those states that use both. Simple checks, such as weights should sum to universe size, should be made. Weight trimming for unusually large weights should be considered and documented if not already implemented.

   The population of saltwater anglers is not defined in official statistics. However, MRIP survey weights are adjusted to align with official population-based household estimates, consequently producing estimated population counts by state. In contrast, most states use sampling frames based on licenses and/or permits. Nevertheless, it might be useful to compare estimated population totals from the states and those from MRIP to determine if they align. Substantively large differences that cannot be explained fully by license frame under coverage (e.g., state regulations may not require a license for certain age groups) would suggest an evaluation of the weighting adjustments for the

states.

3. *Variance estimation procedures*: In addition to a review of the weighting methodology, the associated variance estimation procedures should be reviewed, to ensure that they appropriately reflect the stratification and multi-stage nature of the sampling design. The survey estimates and associated measures of precision should be computed using software that accounts for the design. The SAS "survey" procedures or the R survey package are examples of such software.

4. *Study documentation*: Each state should review the NOAA Fisheries reporting standards and provide the corresponding information for the programs used in their state. There were a number of questions we had from the state presentations. Nearly all of our questions would be answered if we saw such details. The standards appear on this website.

   https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishingsurvey-and-data-standards

5. *Review of measurement error*: These are questions that would be addressed in the documentation mentioned in #4 above.
   a. Does each state have a QA/QC program in place and is a document available for review? We suggest this documentation, if not already, be available for review.
   b. Are the same methods used by all states to train the interviewers?
   c. Is there a quality assurance check on interviewers assigned to a site? For example, does a supervisor subsample a set of interviews for each interviewer regularly to assure the data are collected at the selected site at the assigned time.
   d. Does each state have a protocol for evaluating and addressing survey responses for extreme values?

   Since we do not know how the procedures differ from state to state, we do not have evidence to suggest whether or not these procedures could be a substantive source of differences. However, if discrepancies do exist, they are examples of the types of procedures that could be harmonized across states to reduce differential non-sampling errors.

6. *Response rates*: All states and MRIP should use the same formula for calculating response rates and be documented. See the American Association of Public Opinion Research website for standard methods. We propose to use RR1. Refer to the following website for this information on response rates.

   https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx

7. *Programs with both effort and intercept survey (Florida, Louisiana, and MRIP)*: To further evaluate effort estimates, compare the questionnaires used by each state with the FES questionnaire. Consider splitting a sample into two groups as part of an experiment to examine the effect of questionnaire differences. For example, send the FES questionnaire to half of the sample and the state questionnaire to the other half of the sample. Compare effort estimates for both approaches. Another approach to examine effort is to send the state questionnaires to a sample of households (the current FES frame) and compare the effort estimates (run at the same time) with those households receiving the FES questionnaire. This type of surveys will likely require additional size to ensure sufficient precision of both the original and the experimental results.

8. *Finding comparable components within the state surveys and MRIP*: The estimates produced by the state surveys differ substantially from the MRIP estimates, because of differences in sampling designs, estimation methods and/or survey modes, making it difficult to identify and quantify all the possible sources of non-sampling error. It might be possible to identify *components* of the state surveys that are more comparable, and which could be used to create harmonized estimates. The most promising such component is the intercept surveys, which all the states and MRIP are conducting in person, and which have mostly similar sampling designs for site-days. It is possible to obtain estimates of catch directly from these intercept surveys; these estimates are not as precise as either the capture-recapture or the two-survey estimates, but they are likely to be less impacted by the differential non-sampling errors. We recommend investigating whether this would provide a way to create a consistent benchmark for the relative catch of the states.

9. *Large Pelagic Survey model*: As already mentioned in the general comments above, another option to eliminate the comparability issues is to have a single survey that covers all states. A possible model for this is the Large Pelagic Survey (LPS) on the Atlantic, which is a separate survey targeted at a small number of rare species of high interest that are not adequately addressed by the general MRIP surveys.

10. *Florida*: Because of the similarity of the Florida and MRIP data collections systems on the intercept survey, the state's data provide an excellent testbed to examine the causes of the varying effort estimates. Experiments such as those described in #7 above to compare GRFS and FES should be especially revealing here. Examine each questionnaire and conduct side by side studies changing the questionnaire slightly to examine the impact of a question wording, of the order of questions, etc. Such experiments would provide valuable information on how sensitive the results are to these questionnaire design aspects. Subsequently, it would be possible to select a "standard" questionnaire that can be used across states for eliciting fishing effort.

11. Mississippi survey

a. Review the estimation process to ensure that weights (both base weights and any nonresponse adjustments) are calculated and used correctly for the estimators themselves and the standard error estimates.
b. One source of non-sampling error that is of concern for areas with large fraction of trips from non-public sites, as Mississippi has, is coverage error. The capture recapture estimator relies for approximate unbiasedness on the assumption that both reporting rate and CPUE are identical for trips from public and private sites. To identify whether or not a violation of these assumptions is causing substantial non-sampling error, it is imperative that they be examined. Consider methods for using the law enforcement data for this purpose. We realize these data are based on a non-probability sample. However, we encourage the state to set up a protocol for law enforcement to collect data from a sample of boats that are as representative as possible for a short period. For example, perhaps they could be instructed to collect data from a systematic sample (every *n*th boat) or a rotating location for a specified data collection period. From these stops, data allowing estimation of reporting rate and average catch should be collected, as well as their return site (public/private). From these data, estimates of differences can be produced to assess the impact on non-sampling error due to coverage error. (See Stokes *et al* (2021) for method).

12. Alabama

a) Do not use catch as a matching variable. When catch is used as a matching variable, it eliminates the possibility of adjustment for misreporting of catch by anglers. Since it is unlikely that anglers never misreport, this provides an inaccurate measure of uncertainty for the final estimates. Evaluate the impact on estimates for removing that criterion as a matching variable.
b) A source of non-sampling error for capture recapture estimators is matching error. Misclassification of trips as reported or not can have a large effect on the bias, so minimizing the misclassification errors is critical for this methodology. This is not a problem for Mississippi's system because it requires pre-trip registration, so is a unique problem among the Gulf States for Alabama's program. When catch is eliminated, there will be fewer matching variables and therefore less distinguishable trip profiles. If this is a problem, consider adding additional distinguishing variables to the reporting app. Two examples are: (1) which trip of the day (which would also have to be added to the intercept survey); (2) return site (from a dropdown list in the app).
c) Another source of non-sampling error for capture recapture estimators is lack of independence between the two reports. In the capture recapture implementation, this can only occur if reporters are aware that they are to be intercepted before they report (since the "recapture" sampling units are actually randomly selected). Collecting data about how often this occurs by observation alone is impossible since electronic communication among anglers may be possible. One possible approach is to conduct an experiment in which these data are collected by survey. That is, add

questions to the intercept survey about whether anglers are aware that they are required to report; if they reported their trip prior to interception; and when they became aware that samplers were collecting data for a particular trip. A comparison of the reporting rate and reported catch profile of those with and without prior knowledge of intercept could provide insight into whether this source of non-sampling error is a problem. Since the questions are sensitive, some experimentation to discern what is the most effective way to elicit accurate responses to these questions will be necessary.

d) As noted in #11, the differential between reporting rate and CPUE between public and private sites is a potential source of bias for the capture recapture estimators. Alabama does not have the ability to use law enforcement data to investigate this, so some additional approaches are suggested here. These would also be available approaches for Mississippi.

(1) Ask reporters to provide location of their return in a way that allow analysts to know if the site is or is not on the intercept frame. This would allow comparison of reported CPUE for public and private sites, at least among reported trips. Some experimentation is likely needed to elicit this information in a format that is understandable to the angler and usable by the analyst for categorizing sites. Two possibilities are to request latitude and longitude, or to select Other on a drop down list of sites.

(2) Assessing differential trip reporting rates between public and private sites is more difficult, since data from reporters as suggested in (i) clearly will not allow reporting rates to be compared. To make such a comparison, some alternative source of data is required. One possibility is a passive data collection system for a sample of boats, such as that of the pilot survey conducted by CLS America in the Gulf of Mexico in 2016-17. (See Stokes et al (2021) for discussion of how these data were used for this purpose.) A second possibility is an address-based (or license frame based) survey asking anglers to report return location and reporting status for recent trips.

13. Texas

a. Since a review of the sampling design and procedures was not conducted for Texas, we have no details to assess whether the methodology used by the State of Texas is accurately describing a probability sampling design (Nuttall and Matter, 2020). The lack of detailed documentation on the procedures used by Texas and the inability to obtain a statistical review of the Texas methodology is limiting the ability to obtain gulf-wide estimates and the partitioning, since this number is not available. Additionally, we are unclear if data are collected for private areas or shore fishing.
b. Given the inability to assess the methodology, the estimates from Texas are only an index and should not be considered a scientifically based estimate.
c. Since we have no sampling error to evaluate, we also have no non-sampling error to evaluate.

14. *Simulations*:
    - *a)* Estimate effort using the intercept data (associated with public-access only) and its site-day sampling design, and compare it with the effort estimates obtained by the effort surveys. This estimate should link to public access effort. Discuss how different these two approaches are and whether the differences are plausible.
    - *b)* *Rare species*. Define rare species: this should include rare in terms of time, space and sample size.

        i. For all surveys using the product estimator to obtain catch (e.g., FL, LA and MRIP), determine a rare species. This should be done for one group, so it is all done the same way. Assuming the data and the weights currently being used, produce estimates and the standard error using the formulae adopted in MRIP. Are the estimates subject to large variability? Are the standard errors appropriate for these rare species using the sample sizes typically observed? With the weighting structure currently in place, when does it give "good" and "bad" answers, from the perspective of useability and comparability of the estimates.

        ii. For states and MRIP that use a product estimator, it is possible to have an outlier for the product that is not an outlier for either one of the surveys. A weight given on one of these estimates is then multiplied by the other estimate which could create an outlier. Investigate the circumstances for this to occur in the simulations.

        iii. Using the simulations, how well does the product estimation work? Are there conditions which give rise to unstable estimates?

15. *Calibration*. We understand some of the concerns regarding the use of a simple ratio for calibrating state-level estimates. It does not account for the differences in precision and non-sampling errors between the different state surveys. Because of differences in data availability on which to compute adjustment ratios, even the adjusted estimates are not directly comparable: unlike the other states, Louisiana has a single historical year of overlap between MRIP and its state survey and Texas has no overlap at all (nor a valid state survey to begin with). Despite these concerns, a simple ratio adjustment provides a way forward in the short term, and is also relatively transparent to data users. We suggest further examination of this approach, including the development of a variance estimator and, at least for the states continuing with MRIP, an investigation of what time and location scales to use as the basis for the ratio (annual vs. by wave, multi-year or not, state vs. substate).

    It is possible to use a more rigorous approach to develop a set of estimates that fully take advantage of all the available data sources. This involves creating a statistical model that uses the *data* available in the different surveys, instead of relying only on the

final state estimates. This is similar to what was done for the FES-CHTS calibration as well as for a more recent project for the 2016 FHWAR (Erciulescu *et al*, 2021), in which two versions of the survey were combined to create a new set of model-based estimates that were compatible with earlier versions of the survey. Creating such a model is likely a challenging and time-consuming statistical endeavor, so that estimates would not be available for 1-2 years while the method is being developed (this is a guess on our part at this point). However, subsequent estimation would not take as long, since it involves re-fitting the same (or very similar) model and ensuring that its model diagnostics remain acceptable. An advantage of this modeling approach is that it would provide a statistically rigorous way to combine the data from the different surveys, so that the estimates take advantage of the larger sample sizes and precision of the state surveys, while also being available in a single "currency."

References

Erciulescu, A., J.D. Opsomer, and F.J. Breidt (2021). "A Bridging Model to Reconcile Statistics Based on Data from Multiple Surveys." *Annals of Applied Statistics,* 15: 1068-1079 (DOI: 10.1214/20-AOAS1437).

Nuttall, M.A., Matter, V.M. (2020) Texas Parks and Wildlife Department's Marine Sport-Harvest Monitoring Program Metadata. SEDAR70-WP-03. SEDAR, North Charleston, SC. https://sedarweb.org/docs/wpapers/S70_WP_03_TXPWD_metadata.pdf

Stokes, L., B. Williams, R. McShane, and S. Zalsha (2021). "The Impact of non-sampling errors on estimators of catch from electronic reporting systems." *Journal of Survey Statistics and Methodology* 9, 159-184 (2021). (doi.org/10.1093/jssam/smz042)

Appendix 4: List of Abbreviations

AAPOR American Association of Public Opinion Research

APAIS Access Point Angler Intercept Survey

BSIA Best Scientific Information Available

CPUE Catch Per Unit (of Fishing) Effort

FES Fishing Effort Survey

FIN Fisheries Information Network

CHTS Coastal Household Telephone Survey

GMFMC Gulf of Mexico Fishery Management Council

GSMFC Gulf States Marine Fisheries Commission

GulfFIN Line item designation for Fisheries Information Network in Gulf of Mexico

MAFMC Mid-Atlantic Fishery Management Council

MRIP Marine Recreational Information Program

MSA Magnuson Stevens Act

NMFS National Marine Fisheries Service

NOAA National Oceanographic and Atmospheric Administration

OSF NOAA Fisheries Office of Sustainable Fisheries

OST NOAA Fisheries Office of Science and Technology

SEDAR Southeast Data, Assessment, and Review

SERO Southeast Regional Office

SEFSC Southeast Fishery Science Center

SRFS (Florida) State Reef Fish Survey

Appendix B: SRFS Gag Grouper Calibration Terms of Reference and NOAA Fisheries Calibration Review Memorandum

Terms of reference (TOR) for the use of calibrated estimates for stock assessment and management

April 21, 2022

The following provides guidance on species-specific simple ratio based survey estimated calibrations for use in stock assessment and management. The TORs distinguish between review requirements for model based approaches and other data treatments that may impact microdata as well as resulting estimates and the application of a simple ratio based scalar to survey catch estimates. The TORs described herein pertain to the latter only.

Guidance and Procedures for the Transition Process for Modification of Recreational Fishing Catch and Effort Methods can be found in Procedural Directive 04-114-01 "Implementing Recreational Fishery Catch and Effort Survey Design Changes" which is available at: https://www.fisheries.noaa.gov/national/lawsand-policies/policy-directive-system

The following terms of reference pertain to development and application of simple ratio based scalars to adjust the scale of annual catch estimates. The terms of reference provide guidance to the data provider and reviewer on documentation deemed necessary for a review of the development and application of calibrations to rescale estimates from one survey standard to the other.

1. Provide "fit for purpose" documentation for the development of calibrations (ratio scalars), where "fit for purpose" documentation is defined as inclusive of all elements required to reproduce the calibrated time series.
    a. Generally, documentation will include a complete description of calibration procedures, terms and time series application, datasets related to the development of calibration, source datasets (annual catch estimates) used to calculate ratios, metadata and other data sets, program code for the generation and application of calibrations.
        i. Calibrated estimates should be reproducible by a third party, using the information provided.
    b. Describe how the method is intended to be used in future years when new data become available, or how it is expected to be modified.
    c. For variance estimates, please describe the methods used, for example, Taylor's series approximation (linearization), jackknife or other replication method, other alternatives (e.g., Second or Multiple Derivative Methods, Goodman's).
    d. Evaluate whether the time series is continuous and whether the estimated variances reflect temporal variation in precision. Are there any particular biases in the time series?
2. Identify underlying assumptions for developing, and applying calibrations to the recreational catch time series of landings and discards

    a. Assumptions should pertain to the choice of years selected, the relationship of survey estimates (for example but not limited to temporal, geographic and other coverage considerations such as fishing mode and catch type)
    b. List justification of why the specific years were selected for adjustment and others were not selected.
    c. For the purposes of development and application of calibrations, are estimation domains aligned spatially and temporally to provide equivalent ratio terms?
    d. Describe specific assumptions related to the application of scalars to unaligned domains (e.g., assumptions related to but not limited to the application of ratio scalars to uncovered modes, catch types or effort).
3. Identify underlying assumptions for development of variance approximations
    a. Assumptions should pertain to the choice and application of methods, relationship of survey estimates (dependence), the treatment of covariance terms (where applicable) in the generation of estimators
    b. Justify why one variance method was chosen over another
4. Is the methodology consistent with the simple ratio based approach that was presented and deemed reasonable for use in the Fifth Red Snapper Workshop (2020)?
    a. If not, please describe modifications or deviations
        i. The description should indicate where changes have been applied to the time series and include justification for said changes.
5. Provide a review report summarizing the Review Panel's evaluation of the calibration methodology and documenting whether each TOR was met.

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL MARINE FISHERIES SERVICE
Silver Spring, MD 20910

July 7, 2022

To: Carrie Simmons, Ph.D.
Executive Director, Gulf of Mexico Fishery Management Council

From: Evan Howell, Ph.D.
Director, NOAA Fisheries Office of Science and Technology

Clay Porch, Ph.D. PORCH.CLARENCE.EDWARD.1365852720 Digitally signed by PORCH.CLARENCE.EDWARD.1365852720 Date: 2022.07.07 10:21:37 -04'00'
Director, NOAA Fisheries Southeast Fisheries Science Center

Andy Strelcheck STRELCHECK.ANDREW.JAMES.1365863152 Digitally signed by STRELCHECK.ANDREW.JAMES.1365863152 Date: 2022.07.07 10:25:37 -04'00'
Director, NOAA Fisheries Southeast Regional Office

cc: Katie Siegfried, Jack McGovern, Richard Cody, John Foster, Rob Ahrens

Subject: Florida State Reef Fish Survey gag (*Mycteroperca microlepis*) catch estimates calibration review

As requested in the Gulf Survey Transition Workshop in February 2022, terms of reference were developed for a review of the Florida Fish and Wildlife Conservation Commission's (FWC) ratio based calibrations to convert federal survey estimates of catch to State Reef Fish Survey (SRFS) equivalents throughout the historical time series prior to the implementation of the SRFS. The review coordinated by NOAA Fisheries Office of Science and Technology, was completed in May 2022 and findings subsequently evaluated by NOAA Fisheries Office of Science and Technology (OST) and the Southeast Fisheries Science Center (SEFSC) staff. There were no major concerns identified in the review that would preclude the use of the calibrations for their intended purpose. NOAA Fisheries OST will follow up with FWC staff and reviewers to address minor questions related to documentation formatting and presentation and clarification of justification for variance approximation methodology. Florida is developing a response to reviewer feedback which will be made available once both parties are satisfied that questions raised have been adequately addressed. In email consultation between FWC, NOAA Fisheries OST and SEFSC, FWC confirmed to NOAA Fisheries that to the best of their knowledge, the calibrations provided by their agency were correctly calculated as presented. The Southeast Fisheries Science Center confirmed that the stock assessment could proceed using the SRFS calibrated estimates.

Attached supporting documentation for consideration by the Council SSC includes:

- collated terms of reference developed following the February 2022 workshop, for the review of calibrations and completed reviewer reports;
- Documentation provided by FWC (Notation, text versions of R-related files and data input and outputs;
- February 2022 Gulf Transition Workshop Proceedings.

Appendix C: Statement of Task for Congressionally-Directed Independent Review of Gulf Surveys and Calibration Methods

Senate language:
"*The agreement includes $2,000,000, from within the funding provided to support State management of red snapper, for NMFS to contract with a non-governmental entity with expertise in statistics and fisheries-dependent data collection to provide the following: (1) an independent assessment of the accuracy and precision of both the Federal and State recreational catch data programs in the Gulf of Mexico; (2) recommended improvements to be made to the Federal and State recreational catch data programs in the Gulf of Mexico to improve accuracy and precision; and (3) an independent assessment, based on the results of the two prior items, of how best to calibrate the Federal and State recreational catch data programs in the Gulf of Mexico to a common currency*."

Part I: Survey Evaluation

Request the CIE (or the Survey Research Methods Section of ASA, or another independent provider of expert reviewers) to empanel a group of 3 to 5 survey methodology experts (not necessarily fisheries survey experts) who have no history of involvement with NOAA's MRIP program, the Gulf States' marine fisheries programs, or with Gulf of Mexico fisheries research or management. Panelists' qualifications should include expertise in survey methodology, sampling design, estimation, non-sampling errors, modeling, and calibration.

Statement of Task:

- Primary scope of task: Determine the relative strength and weakness of each current survey in providing the components of catch rates, effort and catch of red snapper in Gulf of Mexico private boat mode fisheries.
- Review available documentation of survey design and estimation methodology for the following surveys:
    - NOAA MRIP Fishing Effort Survey (FES) and Access Point Angler Intercept Survey (APAIS);
    - Florida Gulf Reef Fish Survey;
    - Alabama Snapper Check Survey;
    - Mississippi Tails 'n Scales Survey;
    - Louisiana LA Creel Survey;
    - Texas Parks and Wildlife Coastal Creel Survey.
- Review available independent review reports for the surveys, including NOAA MRIP Certification files.
    - Review the alignment of estimation methods and survey designs as implemented
- Review completed and ongoing studies of the performance of the surveys, including studies related to sources of non-sampling error.

    - Include an evaluation of the status of implementation of the surveys, and whether survey design requirements and required assumptions are being met.
- Identify, assess, and characterize known and potential sources of non-sampling error for each survey, including characterization of the probability and relative magnitude of such errors.
- Recommend actions to address non-sampling error for each survey, including:
    - Studies to be carried out to better understand the sources and magnitude of non-sampling error;
    - Survey design, sampling, and estimation modifications that could reduce the magnitude of error.
- Examine, compare, and characterize the precision of private boat mode red snapper catch estimates for each survey. Assess the accuracy of the surveys' estimates of precision. Identify and describe sources of error and uncertainty of the precision estimates.
- Recommend improvements in survey design and estimation procedures that would reduce variance of estimates and improve the accuracy of the estimates of precision for each survey.
- Determine whether each survey can provide the components of catch and effort estimates that are necessary to achieve agency mandates for fishery stock assessments and fishery management functions for the federally-managed Gulf of Mexico red snapper fisheries, and note relative strengths and weaknesses of each survey in meeting such requirements. (Note: Study Sponsor will provide guidance on the required components.)

Part II: Calibration Methodologies

With the same or a new panel, review available information related to the calibration of Gulf of Mexico red snapper catch estimates for the surveys reviewed in Part I.

Statement of Task:

- For reference, consider information provided by the Sponsor including NOAA workshop reports, associated methodology descriptions, the July 2019 "White Paper" on Recommended Use of the Current Gulf of Mexico Surveys of Marine Recreational Fishing in Stock Assessments, and NOAA calibration reports for calibration of the:
    - NOAA MRIP APAIS pre-2012 to post-2012 survey designs;
    - NOAA MRIP Coastal Household Telephone Survey (CHTS) to FES;
    - NOAA MRIP CHTS-based calibrated time series to each state survey;
    - NOAA MRIP FES-based calibrated time series to each state survey.
- Recommend preferred short and long term calibration methods to be applied to enable use of properly calibrated state catch estimates time series.

- Recommend alternate or modified calibration methodologies from those previously identified and reviewed.

Format of Report: Preferred format is a consensus report including the above-listed components and an executive summary that identifies the principal conclusions and recommendations and summary statement of rationale. If consensus is not possible, a majority and minority report each with the same components is acceptable.

Appendix D: SEDAR Stock Assessment Schedule Through 2023

Region	Calendar Year	Task/Meeting
Gulf	2021/2022	SEDAR 75 Gulf of Mexico Gray Snapper Operational Assessment
Gulf	2021/2022	SEDAR 68 Gulf of Mexico Scamp Operational Assessment
Gulf	2021/2022	SEDAR 74 Gulf of Mexico Red Snapper Research Track
Gulf	2022	SEDAR 81 Spanish Mackerel Operational Assessment
Gulf	2023	SEDAR 85 Yellowedge Grouper Operational Assessment
Gulf	TBD	SEDAR 74 Red Snapper Operational Assessment
Gulf	TBD	SEDAR 87 Shrimp (Pink, White, Brown) Research Track
South Atlantic	2021/2022	SEDAR 76 Black Sea Bass Operational Assessment
South Atlantic	2021/2022	SEDAR 68 South Atlantic Scamp Operational Assessment
South Atlantic	2021/2022	SEDAR 78 South Atlantic Spanish Mackerel Operational Assessment
South Atlantic	2022/2023	SEDAR 82 Gray Triggerfish Research Track
South Atlantic	2023	SEDAR 83 Vermilion Snapper Operational Assessment
South Atlantic	2023	SEDAR 86 Red Grouper Operational Assessment
Caribbean	2021/2022	SEDAR 80 Caribbean Queen Triggerfish Operational Assessment
Caribbean	TBD	SEDAR 84 Yellowtail Snapper, Stoplight Parrotfish
Caribbean	TBD	SEDAR 84 Yellowtail Snapper
Highly Migratory Species	2021/2022	SEDAR 77 Hammerhead Shark Research Track
Highly Migratory Species	TBD	SEDAR 77 Hammerhead Shark Operational Assessment
Florida	2021/2022/2023	SEDAR 79 Southeastern US Mutton Snapper

Appendix E: Gulf Annual Management Cycle for Federally Managed Stocks With Recreational ACLs

After a stock assessment is completed, which usually takes about a year, it is reviewed by the Gulf of Mexico Fishery Management Council's (Council) Scientific and Statistical Committee (SSC). If the stock assessment is accepted, the SSC provides a recommendation for the overfishing Limit (OFL) and acceptable biological catch (ABC) of the stock. Next, the Council specifies the annual catch limit (ACL) and annual catch target (ACT), which are implemented through rulemaking by NOAA Fisheries. The Council will also specify any additional regulations (i.e. size limits, bag limits, trip limits) if needed to adjust harvest of the stock, which would also be implemented through rulemaking by NOAA Fisheries. It takes about a year from when the stock assessment is completed and accepted by the SSC to NOAA Fisheries implementing new ACL, ACT, and regulations. Landings for a stock are monitored and compared against the OFL, ABC, ACL, and ACT every year until a new stock assessment results in new management benchmarks.

Appendix F: Transition Plan for Gulf State Recreational Fishing Surveys: Communications Plan

Transition Plan for Gulf State Recreational Fishing Surveys: Communications Plan

Background

Increased and open communications among NOAA Fisheries, state and regional agencies, and stakeholders will be essential to the successful transition to the full use of Gulf of Mexico state recreational fishing data in the federal stock assessment and management processes. With this in mind, the Gulf of Mexico Subgroup of the Marine Recreational Information Program Transition Team established a cross-agency Gulf Transition Communications Working Group to develop and execute a communications plan to keep partners and stakeholders informed of the status of Transition Plan implementation. This communications plan reflects the working group's shared goals, activities, and responsibilities.

Gulf Transition Communications Working Group Members

Member Agency	Representatives
Alabama Department of Conservation and Natural Resources	Kevin Anson
Florida Fish and Wildlife Conservation Commission	Bev Sauls
Gulf of Mexico Fishery Management Council	Carly Somerset Emily Muehlstein
Gulf States Marine Fisheries Commission	Gregg Bray
Louisiana Department of Wildlife and Fisheries	Jason Adriance
Mississippi Department of Marine Resources	Trevor Moncrief
NOAA Fisheries: Office of Communications	Tim Sartwell
NOAA Fisheries: Office of Science and Technology	Dave Bard
NOAA Fisheries: Office of Sustainable Fisheries	Chris Wright
NOAA Fisheries: Southeast Regional Office	Mike Larkin
NOAA Fisheries: Southeast Fisheries Science Center	Keeley Belva
Texas Parks and Wildlife Department	Julie Hagen

Audiences

1. **MRIP Transition Team Gulf of Mexico Subgroup Members**

   These include staff from:
   - NOAA Fisheries Office of Science and Technology, Office of Sustainable Fisheries, Southeast Regional Office, and Southeast Fisheries Science Center
   - Gulf States Marine Fisheries Commission
   - Texas Parks and Wildlife Department
   - Louisiana Department of Wildlife and Fisheries
   - Mississippi Department of Marine Resources
   - Alabama Department of Conservation and Natural Resources
   - Florida Fish and Wildlife Conservation Commission
   - Gulf of Mexico Regional Fishery Management Council

2. **MRIP Transition Team Gulf of Mexico Subgroup Member Agency Leadership**

   These include the administrators and executive directors of the agencies and organizations listed above.

3. **Stakeholders**

   These include:
   - Gulf of Mexico Regional Fishery Management Council Members
   - Gulf States Congressional Delegation
   - Recreational and Commercial Fishermen in the Gulf Region
   - Recreational and Commercial Fishing Advocacy Groups in the Gulf Region

Goals and Tactics

1. **Ensure the outcomes of the Gulf State Recreational Catch and Effort Surveys Transition Workshop are clear, easy to find, and distributed promptly to our target audiences.**

   In support of this goal, all Communications Working Group members will:
   - Deliver workshop products to their respective agency leadership through their preferred communications channels (e.g., direct email, virtual or in-person briefings, etc.). (Audience 2).

   The Gulf of Mexico Fishery Management Council will:
   - Upon council member request, facilitate the delivery of status updates at council meetings. Work with the MRIP Transition Team Gulf of Mexico Subgroup and/or Gulf Transition Communications Working Group to select an appropriate presenter, identify the topics to be discussed, and avoid redundancies in updates.

   NOAA Fisheries' Office of Science and Technology will:

- Post workshop products (e.g., presentations, proceedings, and research recommendations) to the workshop webpage and distribute via email to all workshop participants, including MRIP Transition Team Gulf of Mexico Subgroup members (Audience 1).

2. **Ensure stakeholders understand what this Transition Plan will accomplish and what is beyond its scope.**

   In support of this goal, all Communications Working Group members will:
   - Identify common questions about the transition process to inform the potential improvement of transition-related communications products.
   - Deliver new and/or updated transition-related communications products to their respective agency leadership through their preferred communications channels (e.g., direct email, virtual or in-person briefings, etc.). (Audience 2).

   NOAA Fisheries' Office of Science and Technology will:
   - Review its transition-related web content to ensure relevant policies, processes, and procedures are clear. If and when transition-related communications products are updated and/or developed, distribute via email to MRIP Transition Team Gulf of Mexico Subgroup members (Audience 1).

3. **Support the publication and dissemination of the Transition Plan.**

   In support of this goal, all Communications Working Group members will:
   - Contribute to the development of the key messages, talking points, and Q&As that will inform the "roll out" of the Transition Plan.
   - Amplify key messages, talking points, and Q&As to stakeholders as appropriate (Audience 3).

   NOAA Fisheries' Office of Science and Technology will:
   - In accordance with internal policies and procedures, develop a Roll Out Plan for the Transition Plan, to include key messages, talking points, Q&As, and the identification of mechanism(s) (e.g., stakeholder webinars) through which the final Transition Plan will be formally delivered to MRIP Transition Team Gulf of Mexico Subgroup members (Audience 1), MRIP Transition Team Gulf of Mexico Subgroup member agency leadership (Audience 2), and stakeholders (Audience 3).

4. **Publish status updates to ensure all of our target audiences are aware of progress toward key milestones.**

   In support of this goal, all Communications Working Group members will:
   - Monitor the updates that are shared at MRIP Transition Team Gulf of Mexico Subgroup meetings, State Directors meetings, and other venues. Work together

to determine how these updates should be amplified to agency leadership (Audience 2) and stakeholders (Audience 3) through their preferred communications channels (e.g., direct email, virtual or in-person briefings, a transition timeline, etc.).

The Gulf of Mexico Fishery Management Council will:

- Upon council member request, facilitate the delivery of status updates at council meetings. Work with the MRIP Transition Team Gulf of Mexico Subgroup and/or Gulf Transition Communications Working Group to select an appropriate presenter, identify the topics to be discussed, and avoid redundancies in the updates that are delivered.

NOAA Fisheries' Southeast Regional Office and the Gulf States Marine Fisheries Commission will:

- Convene State Directors meetings at which status updates can be shared (Audience 2).

Processes and Procedures

The Gulf Transition Communications Working Group will hold joint meetings with the MRIP Gulf of Mexico Transition Team Subgroup, using the last 30 to 60 minutes of each meeting to discuss communications needs and activities. The communications working group will also hold meetings on an as-needed basis (e.g., to develop shared messaging or to coordinate efforts to amplify the status updates that will indicate progress toward key milestones).

When possible, and as time and internal policies allow, transition-related communications products published by any MRIP Transition Team Gulf of Mexico Subgroup Member Agency—whether on its own or in collaboration with another member—should be shared with all Gulf Transition Communications Working Group members for their review twice: First when a draft is available, and second between the time working group input has been incorporated and the time the final product is published. (The second round of review should focus on ensuring factual accuracy rather than influencing editorial choices.) Additional subject matter experts should be included in this review process where appropriate. (For example, products that mention SEDAR should be shared with SEDAR Program Manager Julie Neer.)

If an agency or organization's policies prohibit the external review of an anticipated product, notice of the product should be given; if possible, in advance. Gulf Transition Communications Working Group members agree that, as much as feasible, transition-related communications products should reflect the collaborative nature of this transition process and the effort from all involved to improve data collection.

Appendix G: Review of Texas Private Boat Red Snapper Ratio

Review of Texas Private Boat Red Snapper Ratio

Lynne Stokes (Southern Methodist University), Mike Brick (Westat)

July 2022

After reviewing the materials provided to us by NOAA staff (S74DW10_TX PR RS Calibration Notes.docx and the project report on the FES sampling in TX), we have the following comments on the calibration (ratio) approach. We also provide our opinion about how and when the calibrated series should be used.

1) *The ratio seems reasonable given the materials available. We have a few observations on the specifics below.*
    - The biggest concern is that the Texas estimate and sampling error come without any documentation of the methods used to generate them. In our other efforts, we have never been given documentation to determine if the estimates are from a valid sample design and estimation strategy. Since the calibration approach in the note depends on this assumption, we can only conditionally agree that the ratio is appropriate provided the Texas estimates are from such a design.
    - The table at the end does suggest that the ratio is within the range that might be expected, but those ratios come from very different designs and vary substantially in both point and variance estimates as a result. Thus, the information is valuable in terms of context but given the uncertainty about the Texas sample estimates they are not strong evidence. The fact remains that the ratio is based on one data point and is likely unstable.
    - Minor points. The covariance of $\hat{C}_{y,TPWD}$ and $\hat{E}_{y,TPWD}$ mentioned between equations (4) and (5) is not needed and this line can be deleted. The ratio is for the baseline year not the year of projection. The Goodman variance is exact, and the estimate is unbiased so no discussion of whether it is conservative is needed. The number of digits in the computations gives a false impression of the precision. The last two columns of Table 2 should have MRIP* to distinguish these from the actual estimates.
2) If the recreational private boat catch estimates for each state are to be calibrated or scaled to the FES "currency" for use in the next stock assessment for Gulf of Mexico red snapper (SEDAR 74), would use of the original landings series provided by the Texas Parks and Wildlife Department Coastal Creel Surveys program or the calibrated series calculated using the described methodology better support this objective?
    - The original landings series provided by Texas appears to have the same difference from the FES estimate as the other states shown in Table 2. In other words, unadjusted landings estimates are substantially and consistently lower than FES adjusted estimates. If the Texas private boat catch estimates are to be used in stock assessments for the Gulf of Mexico red snapper along with other states and those other states are using the FES “currency” then the ratio adjusted Texas estimates would very likely have considerably less bias than the unadjusted landings data.