GULFFIN RECREATIONAL DATA STANDARDS AND DATA WAREHOUSE WORKSHOP

PREPARED FOR
GULF STATES MARINE FISHERIES COMMISSION

FEBRUARY 4-5, 2025

THE LODGE AT GULF STATE PARK
GULF SHORES, ALABAMA

Report Date: June 17, 2025

GSMFC Publication No. 334
June 2025

GULFFIN RECREATIONAL DATA STANDARDS AND DATA WAREHOUSE WORKSHOP

PREPARED FOR
GULF STATES MARINE FISHERIES COMMISSION

FEBRUARY 4-5, 2025

THE LODGE AT GULF STATE PARK
GULF SHORES, ALABAMA

(this page intentionally left blank)

Table of Contents

APPENDICES

GulfFIN Recreational Data Standards and Data Warehouse Workshop
Gulf States Marine Fisheries Commission
The Lodge at Gulf State Park
Gulf Shores, Alabama
February 4-5, 2025

1.0 INTRODUCTION

The Department of Commerce (DOC) received financial support as part of the Inflation Reduction Act (IRA) to focus on modernization and investment in science and management programs in support of the nation's economy. National Oceanic and Atmospheric Administration (NOAA) National Marine Fisheries Service (NMFS) has allocated $349 million of these funds to support the Climate Ready Fisheries (CRF) initiative to strengthen NOAA Fisheries' science and survey enterprise and to help support the nation's fishing industry and the states, communities, and tribes that depend on it.

As part of the CRF initiative, $20MM has been allocated for red snapper and other reef fish in the Southeastern United States. The intention for the funding is that it be used in cooperation with the Gulf States, management council, and regional partners to implement improvements to state and federal recreational fisheries survey data. To develop workplans and budgets for these projects, the Gulf States Marine Fisheries Commission (GSMFC) held a total of three workshops throughout 2024 and 2025. This report details the background, technologies, and conclusions discussed at the third IRA workshop on February 4-5, 2025. The focus of this workshop was the development of data standards for catch and effort survey data as well as outlining the requirements for a centralized database.

1.1 PURPOSE AND OBJECTIVES

The purpose of the workshop was to develop minimum standards that all partners would strive for in recreational data collection. The focus was on the minimum data elements and formats required for both federal and state science and management. Specifically, discussion centered on the private recreational and state for-hire catch and effort data standards. Standards would apply to shore, private, and state for-hire fishing modes. Specific objectives included the following:

Review and evaluate the applicability of the NOAA Fisheries Office of Science and Technology (OS&T) Recreational Data Standards as a basis for developing GulfFIN Recreational Data Standards.

> Assess whether the seven OS&T Data Standards are sufficient or appropriate to support both federal and state science and management.
>
> If applicable, identify which OS&T Data Standards need to be modified and/or supplemented to address GulfFIN recreational fisheries data needs. Consider both harvest and discard data needs. Provide reasoning for recommended modifications or additions.

Define the criteria needed to create a GulfFIN standardized database that could house all of the Gulf State survey data.

> Evaluate current data collections in each state and weigh the information collected against science and management needs for that information. From that evaluation, develop data standards that each partner must meet to ensure interoperability and standardization of data coming from each partner. Database structure, quality control, and data dissemination methods will rely on these standards.

Discuss the development of standardized variables and formats as well as a process to transform and load state survey data into this unique database format. Develop recommendations based on these discussions.

Discuss whether this standardized database would be used to house estimate datasets or raw field data. Consider efficiencies that could be created for generating summary reports and tools like dashboards for sharing state survey estimates of effort, landings, and discards. Develop recommendations based on these discussions.

Discuss a process for developing supporting documentation for finalized estimates dataset. Identify minimum data elements, data dictionaries, data delivery schedules, and survey design documentation: metadata, data user handbook, custom domain query template programs (R and SAS). Develop recommendations based on these discussions.

1.2 PARTICIPANTS

1.2.1 In-Person Attendees

Anson, Kevin	Alabama Dept of Conservation and Natural Resources
Beckham, Nicole	Alabama Dept of Conservation and Natural Resources
Bray, Gregg	Gulf States Marine Fisheries Commission
Bruger, Catherine	Ocean Conservancy
Cathey, Andrew	NOAA Fisheries - SEFSC
Cermak, Bridget	Florida Fish and Wildlife Conservation Commission
Cheshire, Rob	NOAA Fisheries - SEFSC
Cody, Richard	NOAA Fisheries - OST
Dolinger-Few, Lauren	NOAA Fisheries - OST
Ferrer, Joe	Gulf States Marine Fisheries Commission
Gigli, Eric	Mississippi Dept of Marine Resources
Griffin, Aimiee	Florida Fish and Wildlife Conservation Commission
Harris, Lizzie	Texas Parks and Wildlife Department
Head, Marie	Alabama Dept of Conservation and Natural Resources
Hollensead, Lisa	Gulf of Mexico Fisheries Management Council
Holley, Michael	NOAA Fisheries - SEFSC
Hopper, Tiffany	Texas Parks and Wildlife Department
Hyman, Alexander	University of South Florida
Larkin, Michael	NOAA Fisheries - SERO
Leonard, Nancy	Pacific States Marine Fisheries Commission
Lowther, Alan	NOAA Fisheries - SEFSC
Mareska, John	Alabama Dept of Conservation and Natural Resources
Moncrief, Trevor	Mississippi Dept of Marine Resources
Nuttall, Matt	NOAA Fisheries - SEFSC
Petersen, Andrew	Bluefin Data
Poland, Steve	NOAA Fisheries - SEFSC
Powell, Dalton	Gulf States Marine Fisheries Commission
Ramsay, Chloe	Florida Fish and Wildlife Conservation Commission

Ratnam, Ramesh	NOAA Fisheries - SEFSC
Robertson, Charlie	Gulf States Marine Fisheries Commission
Smith, Nicole	Louisiana Dept of Wildlife and Fisheries
Stephen, Jessica	NOAA Fisheries - SERO
Timbs, Jeremy	Mississippi Dept of Marine Resources
Wilms, Sean	Florida Fish and Wildlife Conservation Commission
Wilms, Olivia	Florida Fish and Wildlife Conservation Commission
Yechuri, Amar	NOAA Fisheries - SEFSC

1.2.2 Virtual Attendees

Anderson, Joel
Bellais, Donna
Bianchi, Alan
Bland, Kevin
Cannell, Karen
Conrad, Michele
Crosson, Scott
Dancy, Kiley
Denton, Ryan
Didden, Jason
Drexler, Michael
Edwards, Jason
Fisher, Mark
Gloeckner, Dave
Gordon, Maryellen
Hios, Michael
Hoff, Chris
Huber, Jeanette
Hutt, Cliff
Jimenez, Yanet
Kean, Samantha
Malinowski, Rich
Martin, Randall
McClair, Genine
Menzel, Terri
Papacostas, Katherine
Peterson, Lisa
Stevenson, Rich
Studt, Daniel
Valentine, Deanna
Wharton, Dan
White, Geoff

2.0 DAY 1 – FEBRUARY 4

2.1 VISION FOR CENTRALIZED DATA WAREHOUSE

GulfFIN is a state-federal cooperative program that focuses on collection, management, and dissemination of recreational and commercial fisheries statistics in the Southeast region. The goal of this effort is to make a GulfFIN centralized recreational database that is easily accessible by the federal agencies that need access to the data and to codify the data standards. This would be a new system that is flexible, robust, and allows data to flow between the federal and state levels.

2.2 PACIFIC RECFIN WAREHOUSE APPROACH

The Pacific States Marine Fisheries Commission (PSMFC) is an interstate compact with no regulatory or management authority. They are similar to the GSMFC, but they act in the Pacific region. They work with federal and state partners to develop policy initiatives. The Committees are composed of representatives from program partners. The Technical Committee meets to provide guidance for data management and reporting tools. Their goal was to create a cooperative State/Federal program.

The RecFIN system was established in 1992 and receives data from the Washington Department of Fish and Wildlife (DFW), Oregon DFW, California DFW, and Sportfishing Association of California. Data is collected in similar fashions to the Gulf region. Washington and Oregon have seasonal data collection. California has a year-round data collection. The focus is on the fishery dependent data, and the data they collect has expanded over the years. Confidential data is collected, and summaries are provided. The data has a two-month lag with final refreshes of data in the next year with annual data sets. The database uses an Oracle relational database. This was chosen since the earlier PacFIN system already used the Oracle APEX reporting tools.

There are some considerations to keep in mind for these system designs. The data systems have to be updated as technology progresses. The system has to be available for the different technology literacy levels among the audiences. Flexibility is useful to accommodate the agreements and requirements of the different users. To this end, they provide user guides to act as documentation for the system.

When setting up the RecFIN system, PSMFC learned some key lessons for any similar work. The system relies on building and maintaining trust with data sharing agreements and confidentiality agreements. It should be adjusted to meet the needs of data owners. Proper use and data integrity requires standard vocabulary and documentation of methods and calculations. User guides and video tutorials are useful in this regard. Since the system is meant to support the Commission partners, the data collected should be collaborative and have shared elements. Utilizing electronic data capture and exchange helps to streamline the data management. The data needs to be secure, standardized, and quality checked by continually modernizing systems and conducting validation tests. There is a need for full database backups and stress tests.

2.3 ACCSP WAREHOUSE APPROACH

The Atlantic Coastal Cooperative Statistics Program (ACCSP) was created in 1995. ACCSP began participation in the Marine Recreational Information Program (MRIP) in 2015 with collaboration from GSMFC. The FIN systems were developed to communicate between each other with a consistent approach. FIN collaborates with state, regional, and federal partners to standardize data collection, provide access to data, develop reporting tools, and support new initiatives. All partners must agree to the standards which govern data collection, processing, and dissemination. The requirements, report formats, data collection, and access need to be considered when developing the standards of these systems.

ACCSP uses two database systems to collect and store fishery data. The Standard Atlantic Fisheries Information System (SAFIS) is more focused on the raw data. The ACCSP Warehouse stores the historical data for later analysis. The Data Warehouse hosts the most complete set of fishery data. It includes data from all partners. It also includes a confidential query interface behind a login. The system has automated confidentiality as well.

For commercial data, they have a lag of a year for both the spring load and the fall load. For recreational data, the MRIP data for the previous two months are added to the Warehouse. Biologic and socioeconomic data are loaded annually. The data is checked at the partner level, and participants submit data in ACCSP format and codes. To keep the system functional, maintenance and modernization are critical. For central data, like SAFIS, it includes feedback with the file uploads to ensure the codes and data quality are properly entered. Once data has been added, there is another level of QA checks. Data is checked up to 10 times as it moves from collection to dissemination. Once it has been thoroughly reviewed, data can then be accessed from the warehouse.

When developing these systems, flexibility is important. The requirements for data collection and reporting will require iterations. Additionally, raw records and final reports may require adjustment as focus shifts on different fields of the data. Regarding security, different levels of user access can be incorporated into the system. An easier way they found to track user access was a table-driven approach for electronic approvals, automatic expirations, and the specific access details. Lastly, technology has become standardized for these systems, but it requires a maintained skillset and compatibility with partners.

2.4 NOAA FISHERIES RECREATIONAL DATA STANDARDS

The standards exist to provide guidance for the quality and improvement of recreational data. Since there are several data collection tools, the standards are aimed at promoting consistency. This goal was pushed by a directive from the Office of Management and Budget (OMB). The survey and data standards are aimed at keeping consistency at a national level.

The first in the listed standards, Survey Concepts and Justification is aimed at planning, compliance with the Paperwork Reduction Act (PRA), and reducing the response burden. Survey Design has to be communicated, especially sample weights and their explanation. The uncertainty also needs to be presented so users understand it. The Data Quality standard is to guide data processing, editing, and quality checks. The standard for Transition Planning for communication of changes, including documentation, needs to be available. Review Procedures need to be codified in the event resource allocation becomes an issue. These procedures include certification, annual reporting, and peer review. The Process Improvement standard includes planned and unplanned modifications. Access and Information Management mostly concerns what the data contains and its presentation. Microdata, detailed data that can be added, allows for future evaluations after it is published.

A plan needs to be in place to address the different precision estimates. File formats should match what is already used or needed. Increasing compatibility between partners will allow different attributes to be incorporated. The standards from NOAA cover different aspects of recreational fisheries survey operations. They are broadly applicable across the different regions. However, it may be necessary for GulfFIN to be more regionally specific with its standards. There is a panel that will evaluate the standards and the key data use. The hope is they will finish the review by the end of the year.

The standards will influence each other as they are implemented, and they will likely evolve as the program develops. It was recommended to keep the program operating and improve standards as needed.

2.5 SOUTHEAST FISHERIES SCIENCE CENTER RECREATIONAL DATA NEEDS

The goal for Southeast Fisheries Science Center (SEFSC) staff is to gain efficiency by obtaining state recreational datasets from one location and to aid in communication by also accessing survey design documentation in the same location. To this end, the SEFSC will need the survey estimates, raw data tables, as well as metadata. Additionally, they will need survey design and calibration information, the list of definitions used in the surveys, the table of associated calibration factors, and SouthEast Data, Assessment, and Review (SEDAR) support. The SEFSC also needs the stratified data provided by the state surveys, and it must be readily accessible. They need estimates on the stratified level for catch and effort, and they need the raw data for intercepts and biological information. The data from the different surveys are added to the Oracle RDI. GulfFIN will ideally be updated every two months. For SEFSC, the LA Creel Survey is a great reference for the data they need.

2.6 SOUTHEAST REGIONAL OFFICE RECREATIONAL DATA NEEDS

The Southeast Regional Office (SERO) will need a time series of landings and effort. This data will need to be separated by state, mode and wave/month. They would like the landings data throughout the year to be available in a timely manner for in-season monitoring of stocks. This data is used for Annual Catch Limit monitoring and to

evaluate the impact of regulation changes, seasonal closures, and trip and bag limits. SERO uses the target effort data for economic analysis, and it is the primary input in the economic models. They would also like the data to be centralized and in a standardized format.

2.7 DISCUSSION OF BASING REGIONAL STANDARDS ON FEDERAL STANDARDS

Matching the federal standards was seen as a goal for the state partners as it would encourage consistency among the data providers. The staff would have to work to implement these standards. However, since the aim for this system is that it will be centralized and accessible in order to assist with federal fisheries management, striving to copy the minimum standards would be a worthwhile endeavor. Additionally, previous efforts to create this documentation for data according to these standards have seen long-term benefits. Once the documentation is in place, it becomes routine to update it with additions and deviations. As the project continues, the committee will have further discussions on which sections of the federal standards they want to adopt as regional standards.

2.8 DISCUSSION OF STANDARDS FOR AGGREGATED ESTIMATE DATA SETS

SEFSC would like data to include time periods, locations, and regional identifying information to know where catch and effort information is coming from. Additionally, having consistent variable names across data sets would make the data processing and assessment process simpler. If a survey is substituted for another or modifications need to be made across multiple data sets, it would be easier with consistent variable names. The hope is the new system will be able to take the state data and translate it into standardized tables without putting the burden on the states. As part of this work, SEFSC wants to confirm if the states have comparable variables or can provide them in a way that can be matched to the system. They will also need data in a wave-level resolution to support their analyses. The ideal system would be a low burden to efficiently enter data. Another goal is to set it up so end users do not need to reach out to data providers to answer questions about the data.

SEFSC also described minimum requirements they would like to see in the data. They explained the minimum time scale needed is monthly estimates. If the information is available on a weekly basis, they will accept it, but it is not a requirement. Regarding spatial strata and regional information, the area fished is important. As an example, Florida is divided into sub-state domains, Louisiana has five Coastal Study Areas (CSA), and Texas is divided into eight bay regions. This identifying information will need to be present in the data. Additionally, species identification information can be provided through either the National Oceanographic Data Center (NODC) codes or the Integrated Taxonomic Information System (ITIS) codes. The system will be built to process either identifier.

Federal assessments and management favor receiving landings in weight and catch in number, both landings and discards. Catch in number is a convention which comes from the MRIP surveys. SEFSC would also like to receive the total landings in number. Separating total landings into observed and unobserved would be great if the resolution exists to separate them, but they will accept it regardless. Total discards in number are also needed. SEFSC has its own average weight estimation approach that calculates the average weight information based on collected biological data, but if the states could also provide it, it would be welcome. One concern with the states collecting biological information is a lack of funding. Mississippi plans to mimic the Louisiana Creel survey where information links back to specific control number surveys and dockside intercept surveys. This information will be uploaded to BioFIN, and the sampling will continue under a different funding source. Alabama plans to continue biological sampling through dockside surveys and providing the microdata from that survey instead of going through BioFIN. It was also noted that LA Creel only collects discards for eleven species (black drum, gray snapper, gray triggerfish, greater amberjack, king mackerel, red drum, red snapper, sheepshead, southern flounder, Spanish mackerel, and spotted seatrout) and does not include gag discards. Since Alabama and Mississippi are adopting surveys similar to LA Creel, the worry is they would also leave out this federally managed species. Mississippi is only cutting out unmanaged species. They would still collect information on all federal species. Alabama plans to collect discard information for all fish caught. In summary, the minimum catch

information the federal partners need are landings in number, discards in number, landings in weight, average weights if available, estimate variances, and uncertainties. It was decided to leave the gutted weight conversions to the SEFSC.

For the data uploads, the federal partners will also need the calibration factors or some way to identify them. Further conversations are needed to codify how to provide them as the work team develops the project. Whether the work falls on the states or federal partners, there will need to be rules or documentation in place to keep data as transparent as possible.

For effort estimates, SEFSC needs the point estimate of effort, the variance of the effort, and any associated calibration factors. Most states provide effort estimates based on angler trips, but Texas provides hours fished. SEFSC uses angler trips for assessments, but they have trip level microdata from Texas which they can use to calculate a similar proxy to angler trips. The data and the value-added fields from Texas and Louisiana would be translated into the system for the federal partners to get the information they use for stock assessments. If effort variance is there, they want to see it. Otherwise, they can get at it through raw microdata.

2.9 DISCUSSION OF STANDARDS FOR DISAGGREGATED ESTIMATES

When displaying imprecise estimates in the past, NOAA Fisheries decided not to censor them. They added language explaining they supported them at a PSE level above 30%, according to best practices for probability surveys. The language expressed that caution should be exercised when using these estimates and that it was better to aggregate if possible. There could be a need for the quality of the available microdata to look at different levels of aggregation rather than going to a situation where it is rolled up to the next level of stratification to get an estimate that may or may not be usable.

Assessments have shown they can use this data beyond that 30% PSE level. Evidence shows that a PSE level of 40% is adequate for some assessment purposes, but mainly this data is to give analysts the information they need to evaluate its uncertainty. Additionally, analysts have employed unfamiliar models in the past, and the received feedback was that masking lower quality data was unhelpful. Masking should be avoided unless they can provide significant justification. Furthermore, it is informative to see what the original data was. Seeing the original estimate provides the opportunity to see what is driving the imprecision. In terms of adjustments, there is some ongoing work for guidelines and standards. There are no firm recommendations for those adjustments at this time. Finally, there is a benefit to providing the data in order to avoid the perception of a lack of transparency when masking data.

Estimates will be loaded to the system and presented based on the statistics generated with cautionary guidance based on high levels of precision. All states would proceed with uploading the estimates and allowing the data to be displayed.

2.10 DISCUSSION OF MINIMUM STANDARDS FOR RAW INTERCEPT DATA

When this database is developed, the raw survey microdata could be housed within a confidential access system where only those who have authorization can access it. The system could also include public summary tables based on the raw data with confidential information removed. To accept all raw data, GulfFIN would work with the state partners to find common variables and fields that can be used to create a standardized table for dockside data and effort data. In places where fields do not correspond between the different survey methods, these could be included with explanations for their source and applicability.

The states agreed to provide raw microdata for dockside surveys and sample weights. However, Alabama will be unable to provide vessel registration numbers for charter vessel intercepts unless they can resolve it with their legal department. At this time, there is not a short term need to receive and warehouse raw effort survey data. The

work team will instead prioritize asking the states to provide survey elements from their dockside surveys and build a standardized table around them and include additional fields that do not overlap.

It was cautioned that once the federal effort survey is no longer used in the Gulf, scrutiny might focus instead on the state effort surveys. Collecting and presenting the raw microdata for the effort surveys along with the dockside survey data would provide full transparency. Additionally, housing all the data together would facilitate inquiries for analysis.

Once the work team has outlined a potential structure for the database using the dockside surveys, the committee will have further conversations on whether to include the raw effort survey data. Due to the lack of immediate need for analysis with the effort survey microdata, the work of harmonizing the three effort survey methods can be a future development.

2.11 DISCUSSION OF MINIMUM STANDARDS FOR PROVIDING SAMPLE WEIGHTS

Sample weights will be included with raw dockside survey data. Guidance on how they should be applied will need to be provided. The necessary sample weights are the final weight for the final adjustment and the initial weight.

Nonconfidential sample frames are also needed by the federal partners. One way to cover this need is to provide the site registers. The SEFSC uses the site register to confirm any outliers as a useful verification tool. The site register would also need to include site pressure since that affects the sample weights.

Regarding the effort data, the states could provide annual summaries of what the sample frame was. This would be more of a reporting metric in addition to a site registry.

2.12 DISCUSSION OF MINIMUM STANDARDS FOR TIMELY DATA AVAILABILITY

In previous regional meetings, increased timeliness was a high priority for recreational survey data. The same priority was adopted for the state recreational surveys.

One potential hurdle is the creel surveys do not have electronic data collection in the field. Implementing electronic data collection could speed up the data uploads. An additional challenge is Texas experiences high-use and low-use seasonality. The high-use season sees all personnel going for dockside surveys. Entering the data from this season into their database takes several months of review. They do biannual data drops, and their weighting changes depending on the data reflecting a weekday or weekend, the bay system, the high-use or low-use season, and the year.

The standard was set with the understanding it could be adjusted once the current capabilities are documented. Having a capability to provide raw survey data on a monthly basis was seen as beneficial. GulfFIN would offer assistance if there were something they could do to help make data cleaner, faster, and more readily available.

Explaining further difficulties with this standard, Texas could provide disaggregated data, but it would be at the end of their two seasons. In addition, Florida collects raw dockside data uses the SRFS supplemental surveys and APAIS survey. They are deciding whether to move away from APAIS for the SRFS species. Florida is also on the MRIP estimate timeline with a few additional weeks for review. In this way, Florida would likely provide both data sets, and the system would need to flag repeated entries to drop.

The expectation is that every month, raw intercept data and the prior month's uncalibrated estimates will be uploaded. To track these updates, the database could include time stamps when the data was last updated. Refining this idea, data within the last year could show just a time stamp, but anything updated further back

would need to include explanatory documentation. The group will have further discussions on what calibrated estimates will look like in the system and how they could be developed in time to meet needs.

3.0 DAY 2 – FEBRUARY 5

3.1 MINIMUM REQUIRED VARIABLES FOR CALCULATING DIRECTED TRIPS

The federal partners need the target species included with the raw intercept data. If a species were targeted but was unable to be caught, it provides them valuable information such as an indicator of problems with the stock or an economic indicator of the for-hire and commercial industries. Additionally, economists often use the targeted trips data field. At the least, the federal partners need the primary targeted species. They would like primary and secondary targets, but they will compromise for just primary target species if it would create too many data fields.

The states plan to provide a field in the raw dockside intercept data with this information. NOAA will generate their own estimates for targeted trips because they will know which to select of the multiple approaches that calculate these estimates.

3.2 PRESENTATION ON AT-SEA OBSERVER WAREHOUSE

Three states are collecting at-sea data to improve the understanding of released catch on for-hire boats. Florida has been a longtime participant in the program. With the monitoring funding from the Return 'Em Right program, Alabama and Mississippi have been included. At the start of 2025, the program is being expanded to Louisiana and Texas. The data collected by onboard observers include rates of release of offshore managed species and reef fish species in addition to the release methods.

In order to house data from the three states, GulfFIN contracted development of a system. Development took between six and eight months to construct the system, but a firmer understanding of what they needed from the beginning would have likely shortened the development time. The system was built around the file formats as the states provide them, and it is locked to the formats in their current state. The data is transformed by the system to fit a standardized structure. When data is uploaded, the system runs validation checks for warnings and errors. The major achievement is the system is an easier method for states to provide data with some additional quality checks. In addition, the system includes a reporting dashboard with filterable data summaries.

A similar system for the scope of this workshop could be developed to house data for all five states with easy access for end users like the federal partners. It could provide summarized reporting tools, program documentation, and contact information of state data providers. The hope is data like the at-sea observer data, recreational estimates, and biological data would be available in one place within the GulfFIN Recreational Data Management System for ease of access.

The goal is to have this data available for stock assessment. Making the data publicly available will require a summary tool and further discussion on adhering to confidentiality practices like the Rule of Three.

3.3 IMPROVEMENTS FOR STATE & COMMISSION DATA MANAGEMENT SYSTEMS

One of the opportunities presented by this one-time funding is to make improvements within the commission and at the state level for the data management systems to meet the needs for the future data warehouse system. Improvements would be aimed at things that lead to cleaner data being turned in sooner (e.g. hardware, software, programming time, people writing better quality control processes, telecommunication lines, servers, electronic

data collection, etc.). The funding is not a long-term source, but it provides benefits now. In addition, similar improvements could be applied to the system that will be developed so it has a better start for success.

Furthermore, the commercial and recreational data sets are independent enough that these improvements to recreational data will only benefit the recreational side. If they augment the system in general to help recreational data, it is an added benefit that could apply to both data sets. However, the funding is intended to be focused on recreational data.

They are using the funding to improve the timeliness, availability, and quality of recreational survey data. If a transition to mobile electronic data collection in the field would accomplish that task, similar to what was seen when APAIS made the same change, and allowed for more data to be collected with additional quality checks, it would be possible to use this funding for that purpose. The money could be used for the startup cost, hardware acquisition and software development, but ongoing maintenance costs for the programs would need to come from another funding source. Additionally, upgrades for system security and technology will also be available within this funding.

GSMFC will have further discussions with the individual states over the extent of the IRA funding. They will work over the next two years to identify the highest priority needs, to explore their proposed costs, and to implement improvements through agreements between the commission and the states in an effort to benefit the overall process of getting data into the system.

3.4 QUALITY CONTROL PROCESSES

To help the development and programming for the system to import state data and run quality control checks, all states will document what they already do with their own data through program design documentation. For the centralized system to run quality checks, it could incorporate flags on potential errors. One error that will be eliminated through the centralized system is the issue of duplicate entries. The system will be developed to reject duplicate records from uploading. One potential problem was raised by Florida. Since Florida uses three labs to provide aging data, they will need to include in their documentation the scenarios where data should be pulled for the GulfFIN Bio database.

There are common errors and mistakes the system can look out for as another level of checks. In the development process, GulfFIN will work closely with each state partner to import data and develop quality control processes to be sure it meets their needs. Knowing what the states currently do will be beneficial, and that will be part of the later development discussions. The development process will also help them pinpoint specific features they need for the system.

3.5 SHARING PROGRAM DESIGN DOCUMENTATION

In order to communicate program design documentation to the end user, a downloadable hyperlink will be provided. Additional contact details for the appropriate staff will be added to the system. They can develop a more complex system if it is needed, but this approach should meet the current needs.

3.6 SUMMARY REPORTING DASHBOARD

When development begins and a contractor is selected, having clearly identified features for a reporting dashboard will help to speed up the development process. One piece of advice was to avoid overwhelming the user with data. Knowing the intended audience will help to avoid throwing too much into the dashboard. Information could also be broken into smaller chunks and spread across a few dashboards depending on how the group wants to present the data. GulfFIN could also include links to more detailed documents. The feeling is the scientific

community will be one audience, but there is an eye towards information the public would be interested in seeing. The summary dashboard will be more public facing while the more detailed data could be behind confidentiality logins or download requests. GulfFIN could also include a metric dashboard to show the success of the project. These metrics could be generated annually or semi-annually. Eventually, calibrated estimates will need to be incorporated in the dashboard once they have a firm plan for them and their presentation.

The public-facing information would include reports on the number of landings and the number of trips sampled. The overall effort in the Gulf is helpful because it characterizes the fisheries. The public will also be more interested in landings estimates for their species of choice in that state or across the Gulf. To provide this information, GulfFIN will include a simple report that allows users to quickly filter landings for species, state, and year. This kind of summary is anticipated to be highly popular. Along with the summaries for landings, a caveat should be included to explain that the data could be different from what was used for management.

Anything the work group decides to incorporate will be discussed with partners to be sure nothing is missed. If they reach the point where all survey data is housed under GulfFIN, it could be possible to generate summary dashboards for all the information thus contained.

3.7 DATA DISSEMINATION RULES AND PROCESSES

All raw survey microdata and estimate tables will be scientifically and publicly accessible. Federal partners will continue to have access through a database link. While states work to upload finalized survey microdata, the estimates could include language that they are based on preliminary numbers and are subject to change. Having the preliminary summaries available will also help federal partners when they need to make quick decisions. If raw effort data is available, the system might limit some fields or put it all behind some confidential accessibility system.

For data downloads, people sometimes bog down similar systems with large data requests. To avoid this, the technology team will work to either explain on the site how to pull smaller chunks of data or to limit the amount of information that can be queried at once. The team will also need to discuss what type of output they want to make available for these downloads. Current projects just allow someone to download an ASCII file that they can adapt to meet their needs. Presenting the data in the same format as it is stored would be a short-term solution for people to run their own analysis. The query reports could show criteria that led to this data and the time stamps of when the query was run. As this project proceeds, the work team could develop the system so it provides a public access file in a specific format with guidance on how to use it along with the sample weights to allow for more custom analyses. The work team could ask for some information from the people who first start requesting data downloads. In this way, they would have a better idea of the audience to inform development.

Another aspect of the system that will need to be addressed is whether it will allow queries to pull data from multiple states at once. The summaries can be grouped together, but restricting the microdata will require limiting the capabilities of the query. Placing language and guidance warning against treating raw microdata as equivalent across all states is unlikely to completely curtail misuse, but it will help inform end users on what they are accessing and precautions they should take when using the data sets. The system could also include a data use agreement before allowing a download in order to guide the end user. Furthermore, the user guidance could benefit from explaining how blank entries should be handled in data sets.

The work group will be comprised of people from multiple committees in GulfFIN with input from states to feed development back to the GulfFIN system. To change any data formatting or data collection once the system is in place, it will need to be discussed with GulfFIN to decide how it will impact everything in the system and whether they should implement it. Everything falls under the GulfFIN program which means both state and federal collaborative discussions will be held to find the best path forward. Federal partners would like to be part of these

discussions in case some changes impact the region. In this way, they could potentially implement improvements in the federal standards.

Andrew Peterson of BlueFIN Data brought up a concern with posting database links for flat files. Although they are a convenient method, these links pose a security risk. He suggested using a database programming interface (API) for these files. In addition, APIs are automatic and would reduce burdens on staff.

Catherine Bruger of Ocean Conservancy wanted to build on her comment letter (see Appendix A) that using the reporting dashboard to cover annual reports would be very efficient. She also suggested a change of terminology in calling data final. Public perception is that final means static, and the changes made to historical data could cause those estimates to change.

Karen Cannell said data downloads of microdata or all data should be handled differently from report or query requests. They are different in intent and volume, regardless of whether they are looking at one state or all states. She also agreed with Andrew Peterson that they should consider using APIs.

3.8 CONCLUSIONS

- The warehouse system for state survey data must use Oracle APEX.
- The committee will continue discussions on which sections of federal standards will be adopted as regional standards.
- Further discussions are needed to specify what the aggregate estimate calibration factors will look like in the system.
- The states will upload disaggregated estimates to the system with cautionary guidance.
- All states will work with the work team to develop standardized tables for their raw dockside survey microdata.
- The committee will have further discussions on whether to include raw effort survey microdata.
- The plan is that every month, raw intercept data and the prior month's uncalibrated estimates will be uploaded to the new system.
- The states will provide a field in the raw intercept survey data that includes the targeted species.
- GSMFC will discuss with each state the budgets and availability of funds to improve data management systems.
- GulfFIN will work with each state to develop quality control processes based on their existing practices and needs.
- Program design documentation will be available in the system along with relevant staff contact information.
- The features of the summary dashboards will be verified with partners to be sure they include all needed information.
- Raw data and estimate tables will be scientifically and publicly available from the system.
- Language will be included on the download and query tool pages in the new system to educate and guide users on how to understand the data and its uses.
- The work group will be a new committee formed from the existing technical work groups, every state, and pertinent offices from NOAA Fisheries and the council.

APPENDIX A

WRITTEN COMMENTS

600 1st Avenue North
Suite 301
St. Petersburg FL 33701

Ocean Conservancy

727.895.2188 Telephone
727.895.8652 Facsimile
www.oceanconservancy.org

January 29, 2025

Mr. Dave Donaldson, Executive Director
Gulf States Marine Fisheries Commission
2404 Government Street
Ocean Springs, MS 39564

RE: Standards for Recreational Fishing Surveys and Data

Dear Mr. Donaldson:

Ocean Conservancy[1] offers the following comments to the Gulf States Marine Fisheries Commission (GSMFC) in support of the development and implementation of survey and data collection standards and a regional centralized data warehouse.

As a first step, **Ocean Conservancy encourages alignment on terminology surrounding development of "standards"** as there appears to be ongoing miscommunication on the distinction between "performance standards" and "technical specifications." Looking to recent literature, we suggest defining performance standards as guidance that promotes data quality and statistical rigor and which ensures alignment of the overarching purpose of performance objectives, while remaining sufficiently flexible to accommodate state-specific data collection programs.[2] Technical specifications, on the other hand, relate to specific data elements and processes that would inform data warehouse development. Both are needed to ensure the execution of high quality data collection and management programs and development of a functional database that supports those programs.

Second, **Ocean Conservancy supports implementation of the National Recreational Fishing Survey and Data Standards ("Standards")[3], which were developed by NOAA Fisheries Office of Science and Technology (OST), as a minimum level of performance standards to be achieved**, noting that a peer-review is underway by the National Academies of Sciences[4] and will be conducted by the Fisheries Data

[1] Ocean Conservancy envisions a healthier ocean, protected by a more just world. We are working to protect the ocean from today's greatest global challenges and, together with our partners, we create evidence-based solutions for a healthy ocean and the wildlife and communities that depend on it.
[2] Melissa Garren, Forrest Lewis, Laura Sanchez, Daniella Spina, Annie Brett. How performance standards could support innovation and technology-compatible fisheries management frameworks in the U.S. Marine Policy. Volume 131. 2021. *Available at:* https://doi.org/10.1016/j.marpol.2021.104631
[3] Recreational Fishing Survey and Data Standards. *Available at: https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-survey-and-data-standards*
[4] Peer Review of the Marine Recreational Information Program Data Standards. *Available at:* https://www.nationalacademies.org/our-work/peer-review-of-the-marine-recreational-information-program-data-standards

Governance Committee.[5] Proposed revisions or additions may be recommended as a result of these reviews.

Based on our initial assessment of the current supplemental data collection programs ("State Surveys") in the Gulf, most of them either adhere to or are actively working toward achieving these standards (*see enclosure*) as a condition of the Marine Recreational Information Program (MRIP).[6] We welcome any corrections or additions to our initial assessment of achievement of the standards by current programs. Adhering to the standards promotes consistency, ensures transparency, and enhances data quality, which provides state fishery managers and stakeholders with a reasonable level of assurance that their reported data will be used in the fishery management process. Further, **Ocean Conservancy encourages the GSMFC and workshop participants to identify what hinders current data collection programs from meeting the baseline MRIP standards and identify pathways to achieve those standards to ensure expedited certification, transition planning, and long-term survey funding**.

Specifically, Standards 1-6 focus on survey methodology *documentation*, and are crucial to ensuring long-term viability, robustness, and adaptability for data collection programs over time. Standard 7 relates to data access and information management, how each state will functionally share their data, and is the only standard that prescribes "technical specifications." **It is Standard 7: Access and Information Management, that seems to be the issue of greatest interest and direct relevance to development of the GSMFC FIN data warehouse** Specifically, discussions around Standard 7 should focus on data elements that relate to data warehouse development (e.g., survey data, key statistics, publication schedule, aggregation) and the limiting factors for inclusion currently. As mentioned, many of the state programs already meet or are in the process of achieving these standards. Therefore, **Ocean Conservancy supports implementation of Standards 1-6 by all recreational fishery data collection surveys. Ocean Conservancy encourages a robust discussion and consideration of "technical specifications" which would be *additive* to Standard 7**, to enable and support the development of a centralized database[7] and calibration metrics,[8][9] to promote transparency and understanding of how the data relate to management, and to streamline integration of state data into the stock assessment.

Finally, the development of a database supported by sufficient performance standards and technical specifications can allow for a semi-automated MRIP certification and reporting process and eventually allow for in-season optimization of quotas across the region. If additional standards are developed or

---

[5] Fisheries Modernization Strategy. From Data to Decisions. 2023-2026 Implementation Plan. *Available at:* https://repository.library.noaa.gov/view/noaa/66639

[6] NATIONAL MARINE FISHERIES SERVICE PROCEDURE 04-114-02. Guidance and Procedures for the MRIP Certification Process. *Available at: https://media.fisheries.noaa.gov/dam-migration/04-114-02.pdf* Page 10.

[7] Red Snapper Data Improvements Under the Inflation Reduction Act. Updated December 4, 2024. *Available at:* https://www.fisheries.noaa.gov/climate/red-snapper-data-improvements-under-inflation-reduction-act

[8] Project 3.1: Finding comparable components within state surveys and MRIP. Gulf States Survey Transition Research Plan. September 24, 2024. *Available at:* https://www.fisheries.noaa.gov/s3//2024-09/Gulf-Transition-Research-Plan-2024-508-compliant-9.6.24.pdf

[9] Project 3.2: Collect side-by-side data for additional benchmarking in states where multiple programs do not exist. Gulf States Survey Transition Research Plan. September 24, 2024. *Available at:* https://www.fisheries.noaa.gov/s3//2024-09/Gulf-Transition-Research-Plan-2024-508-compliant-9.6.24.pdf

proposed, Ocean Conservancy recommends GSMFC include a robust public review process before they are incorporated.

Again, we welcome your feedback and input on our initial assessment of adherence to the MRIP standards by current state data collection programs and we thank you for your continued work on these issues. We value the efforts of the Commission, agencies, and staff throughout the region and look forward to supporting these important endeavors. Please contact us with any questions or concerns.

Sincerely,

Catherine Bruger	**Michael Drexler, PhD**	**Michele Conrad**
Manager, Fish Conservation	Scientist, Fish Conservation	Oceanbeat Consulting

ENCLOSURE - Recreational Fishing Survey and Data Standards Documentation Checklist

CC – Gregg Bray (Gulf States Commission); Mr. Andy Strelcheck (SERO); Russ Dunn (SERO); Dr. Clay Porch (SEFSC); Dave Gloeckner (SEFSC); Dr. Evan Howell (OST); Dr. Richard Cody (OST); Dr. Katherine Papacostas (OST); Lauren Dolinger Few (OST); Dr. Carrie Simmons (GMFMC)

Recreational Fishing Survey and Data Standards Documentation Checklist

State recreational fishing surveys and data collection programs,[1] which are included in the national Marine Recreational Information Program (MRIP), should adhere to the Recreational Fishing Survey and Data Standards ("Standards")[2] as required by the Certification Procedural Directive.[3] While many of the Gulf States have data collection programs that largely achieve the MRIP Standards, it is unclear whether and where there are challenges or barriers to meeting these standards. Ocean Conservancy has developed an initial checklist for each of the state recreational fishing surveys that describes the extent to which each program, to our knowledge, is consistent with the Standards and where more work needs to be done, and we welcome any corrections or additions to the checklist.

[1] Recreational Fishing Surveys. *Available at:* https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-surveys

[2] Recreational Fishing Survey and Data Standards. *Available at: https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-survey-and-data-standards*

[3] NATIONAL MARINE FISHERIES SERVICE PROCEDURE 04-114-02. Guidance and Procedures for the MRIP Certification Process. *Available at: https://media.fisheries.noaa.gov/dam-migration/04-114-02.pdf* Page 10. *"Pursuant to the MRIP Data Standard, once a certified survey is implemented, the survey administrator will be required to submit an annual report within three months of the conclusion of each survey year. Annual reports must provide an overview of data collection procedures, including questionnaires, data collection schedules, and sample sizes, as well as response rates and key survey estimates within the survey year. These annual reports will be primarily used to ensure that certified surveys are adhering to MRIP standards, but will also be reviewed to ensure that the survey design is being implemented as it was certified. If an annual report contains evidence that a certified survey is being implemented with edits to the design that were not part of its certification, or that survey design assumptions are being violated, a review will be triggered to determine if the survey should maintain its certified status and continue to receive funding, as well as if the statistics derived from the survey data should continue to be eligible for use in stock assessments and management actions. Further, if it is apparent that the survey may no longer fully meet the Terms of Reference from its pre-certification review for any reason, such a review will likewise be triggered."*

		Supplemental Survey					Role of GulfFIN Centralized Warehouse
Standard	**Detail**	**AL – Snapper Check**	**FL – SRFS**	**LA – LA Creel**	**MS – Tails n' Scales**	**TX – TX Creel**	
Standard 1: Survey Concepts and Justification	1.1: Planning	Met - 11/2/2021	Met – 10/5/2018	Met – 11/2/21	*In progress* - 9/28/23	Does not exist	A centralized warehouse could support a clearinghouse for survey documentation and Gulf States could partner in ensuring adherence to standards.
	1.2: Paperwork Reduction Act Compliance	*Each state varies with respect to this standard according to state regulations*					
Standard 2: Survey Design	2.1: Sampling	Met - 11/2/2021	Met – 11/2/21	Met – 11/2/21	*In progress* - 9/28/23		
	2.2: Data Collection						
	2.3: Estimation						
	2.4: Evaluation						
Standard 3: Data Quality	3.1: Processing, Editing, and Quality Control	Included in Certification Packet	Included in Certification Packet	Included in Certification Packet	Included in Certification Packet		
	3.2: Quality Assurance						
Standard 4: Transition Planning Per Policy Directive 04-114-01		Gulf State Transition Plan initially released: October 12, 2022 Gulf Transition Research Plan (September 24, 2024)					Clear identification of transition needs can be used to prioritize funding in both Transition and Implementation Plans.
Standard 5: Review Procedures	5.1: Certification Per Policy Directive 04-114	*Conditionally* Certified – June 25, 2018	Certified – October 5, 2018	Certified – December 21, 2017	Certified - June 15, 2018	Does not exist	Including meta and micro data and methodologies and formulas will eliminate back and forth of data and methodology exchanges.
	5.2: Annual Reporting	*Unknown*	*Unknown*	2022: Met – 3/16/2023 2023: Met – 11/27/2024	*Unknown*	*Unknown*	Inclusion of meta and micro data can result in automated performance reporting and possibly eliminate the need for manual reporting.

	5.3: Peer Review. *Peer-Review Procedures have been implemented for FES/APAIS.*	*Unknown.*	*Unknown*	*Unknown*	*Unknown*	*Unknown*	Clear identification of areas that do not adhere to standards and can be used to prioritize funding.
Standard 6: Process Improvement	6.1: Process Improvement Plan	*Should be included in annual reports.*					Clear identification of proposed changes can be used to prioritize funding in both Transition and Implementation Plans.
	6.2: Unplanned Modifications	*Should be in annual report (e.g., hurricane reduces sampling). Scale of modifications could trigger re-certification, program is supposed to submit a letter.*					
Standard 7: Access and Information Management	7.1: Microdata	*Unknown*	*Unknown*	Available online: LA Creel Data Query | Louisiana Department of Wildlife and Fisheries Includes PSEs and warnings (Standard 7.2.2)	*Unknown*	*Unknown*	Can publish data as they become available to support meeting the standard. Convert to consistent data currencies for ACL comparisons
	7.2: Estimates	*Unknown*	*Unknown*		*Unknown*		
	7.3: File Formats	*Unknown*	*Unknown*	*Unknown*	*Unknown*		
	7.4: Attribute Values and Formats	*Unknown*	*Unknown*	*Unknown*	*Unknown*		
	7.5: Information Management	*Unknown*	*Unknown*	*Unknown*	*Unknown*		

this page intentionally left blank
end of report