

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

FY 2016 Proposal

David Gloeckner
Created: 10/16/2015

1. Overview

1.1. Sponsor

Vivian Matter

1.2. Focus Group

Information Management

1.3. Background

The Southeast accounts for more than half of the recreational fishing trips in the United States. As a result, recreational fisheries statistics form an important component of stock assessments for many species, and recreational fisheries management is critical for the Southeast Regional Office (SERO) and the three fisheries management councils (SAFMC, GMFMC and CFMC). Additionally the SEFSC provides the scientific support (data and assessment) for Atlantic highly migratory species (tunas, swordfish, billfish and sharks) managed by the Highly Migratory Species division, and for many of those stocks recreational catches represent large fractions of the total removals.

The SEFSC integrates information from multiple recreational surveys for use in stock assessment and management; as a result, the basic information used for those two purposes is the same. Currently the SEFSC integrates information from MRIP, the Texas Parks and Wildlife Department's (TPWD) Creel Survey and the Southeast Regional Headboat Survey (SRHS). The SEFSC expects to incorporate information from the Louisiana Creel Survey (LA Creel) in the coming year (LA Creel was the only survey conducted in LA in 2014 and in the relatively near future is expected to replace MRIP). Additionally in the next few years it is expected that the high interest in implementing charter boat log book reporting in the Gulf of Mexico and in the South Atlantic will result in additional sets of data to be integrated.

The SEFSC uses specific data sets as the definitive data sets for specific strata/data types. For instance MRFSS/MRIP is the definitive data set for VA-ME recreational landings and for non-headboat fisheries for NC-LA. There are numerous other recreational data sets which are not incorporated in the SEFSC system (SC charter boat survey, FL east coast red snapper survey, LA Creel 2013). The MRFSS and MRIP calibrations the SEFSC uses were developed following the procedures defined in the calibration workshops (Salz et al 2012, Carmichael and Van Vorhees 2015). Differences should exist between SEFSC estimates and the source survey estimates where 1) source survey might not provide calibrated estimates and 2) in the weight estimates for which SEFSC uses a standardized procedure for all estimates necessitated by missing estimates for some strata.

The SEFSC system includes a number of components needed to standardize information for assessment and management. For the MRFSS estimates of landings (continental 1981-2003 and Caribbean 2000-2014) there are catches for which landed weight is not estimated due to low numbers of observed weights, there are no landing estimates in weight for the TPWD survey and there will not be any for the LA Creel. The SEFSC system uses a standardized approach to fill these gaps and provide consistent weight estimates over the 35 year MRFSS through MRIP time series. In some cases, special treatment (domain estimates for the MRIP years, alternative treatments for the MRFSS years) is used to assign landings to a management region. In the Caribbean a large number of species are landed and many of them occur rarely; as a result there are multiple instances where the SEFSC system uses weight observations from other members of the same genus or the same family.

Fish released (dead and alive) from recreational fisheries at times represent substantial fractions of the total catch (commercial and recreational) of many of the important stocks managed in the southeast. The TPWD Survey does not provide estimates of released fish. The SEFSC system uses information other Gulf of Mexico recreational fisheries to calculate possible TX recreational releases.

Recreational fishing surveys often change over time while stock assessments and management require long time series of estimates which have been calibrated for survey changes. The SEFSC system incorporates such calibrations and often those calibrations must be made 1) from old surveys / survey methods to new and 2) from new to old so that various needs of customers can be accommodated. For instance managers may need current year MRIP estimates converted to MRFSS estimates if an ACL was calculated several years ago when MRFSS data were used in the assessment. Or conversely historical MRFSS data must be converted to MRIP estimates for use in a current stock assessment.

Currently, computing recreational landings estimates is a lengthy process and involves executing many computing processes manually. Some examples of the conversions that are currently done manually by sequentially running SAS programs are:

- Partitioning the two seasonal estimates from the TPWD recreational fisheries into the standard bimonthly wave estimates used by MRIP
- Partitioning the TPWD aggregated "other species" estimate into individual species' estimates
- MRFSS for hire calibrations
- MRIP re-estimation adjustments
- MRIP APAIS adjustments
- MRFSS post stratification/MRIP domain estimation
- Other Bias adjustments

- SEFSC weight estimation procedures

Additional calibrations are anticipated in the future, including those resulting from the change in the MRIP effort estimation for shore and private modes with the implementation of the MRIP Fishing Effort Survey. There will also be calibrations necessary between the LA Creel survey and MRIP.

Currently some of the data sets are used for:

- Indices by stock assessment scientists, such as estimation of catch rates,
- Size composition analyses for stock assessment
- Bag limit analyses by the regional office, SERO
- Standardized landings history in number and weight from multiple data sets (MRIP, TPWD and SRHS) for use in stock assessment and management.

The current system depends on one federal employee dedicated to both to southeast recreational data integration and dissemination. The fact that only one person is dedicated to this task which is critical for stock assessment and management represents a substantial vulnerability for NMFS.

1.4. Project Description

This project will move the recreational data extraction and integration processes conducted by the Fisheries Statistics Division (FSD) at the Southeast Fisheries Science Center (SEFSC) into an Oracle Data base with reporting capability in APEX.

This project addresses the needs of the Southeast region's stock assessment scientists and managers by streamlining and automating the recreational fisheries data extraction and integration processes. These processes are executed by the SEFSC and not by the survey providers. For example, TPWD provides estimates for two periods per year (Nov-May and May-Nov) . The SEFSC calculates wave and annual estimates from the TX data for compatibility with other surveys, to fulfill calendar year requests, and in some cases to fulfill fishing year requests. Other examples are the MRIP calibrations for the charter mode and other survey methodology changes. Assessment scientists need a consistent time series of estimates and adjustments must be made to the survey estimates. Although the SEFSC works closely with survey providers on solutions, especially MRIP, it has fallen on the SEFSC to execute these important processes for assessment and management needs. If the survey providers desired, access to these modified estimates could be provided.

The resulting database will reflect current Southeast region stock assessment decisions such as,

- the elimination of some modes of fishing from certain species' estimates
- the recommended estimate when competing surveys are available (e.g. MRIP headboat estimates in the Southeast will be excluded when SRHS estimates are available)
- standard substitutions in the early 1980s (e.g. charter and private modes estimated from TPWD data back to 1981)
- post-stratified/domain estimates from MRIP when necessary (e.g. to allocate the Florida Keys to the Atlantic)

1.5. Public Description

1.6. Objectives

The objectives of this proposal are to allow users 1) access to recreational estimates from all programs in one standardized location and 2) access to raw intercept data sets in SEFSC formats. This includes:

- Automation of existing processes
- Development of processes to incorporate new data sets such as LA Creel
- Creation of a data warehouse with storage of data sets and estimates on the ORACLE database
- User access to data sets and estimates via a data warehouse and web portal

Development of a fully integrated system into the ORACLE database would reduce the vulnerability to loss of personnel and increase the feasibility of training backup and/or replacement personnel.

1.7. References

Carmichael, J. and D. Van Voorhees (editors). 2015. MRIP Calibration Workshop II. Report of a workshop on calibrating MRIP catch estimates, September 8-10, 2014. SAFMC/SEDAR, North Charleston, SC.

Matter, V.M. and A. Rios. 2013. SEDAR32-DW02. MRFSS to MRIP Adjustment Ratios and Weight Estimation Procedures for South Atlantic and Gulf of Mexico Managed Species. National Marine Fisheries Service Southeast Fisheries Science Center, Fisheries Statistics Division, Miami, FL and National Marine Fisheries Service Southeast Fisheries Science Center, Sustainable Fisheries Division, Miami, FL.

Matter, V. and S. Turner 2010. Estimated Recreational Catch in Weight: Method for Filling in Missing Weight Estimates from the Recreational Surveys with Application to Yellowedge Grouper, Tilefish (golden), and Blueline Tilefish (SEDAR 22-DW-16),

National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division (SFD-2010-003).

Rios, A, V.M. Matter, J.F. Walter, N. Farmer, and S.J. Turner. 2012. SEDAR31-DW25 Estimated Conversion Factors for Adjusting MRFSS Gulf of Mexico Red Snapper Catch Estimates and Variances in 1981-2003 to MRIP Estimates and Variances. National Marine Fisheries Service Southeast Fisheries Science Center, Sustainable Fisheries Division, Miami, FL, National Marine Fisheries Service Southeast Fisheries Science Center, Fisheries Statistics Division, Miami, FL, and National Marine Fisheries Service Southeast Regional Office, Saint Petersburg, FL.

Salz, R., T. Williams, E. Williams, J. Walter, K. Drew and G. Bray. 2012. MRFSS/MRIP Calibration Workshop Ad-Hoc Working Group Report. Available online at: <http://sedarweb.org/otw-01-calibration-workshop-ad-hoc-group-report-may-2012>.

SEDAR. 2009. SEDAR 16 – South Atlantic and Gulf of Mexico King Mackerel Stock Assessment Report. SEDAR, North Charleston SC. 484 pp. Available online at: http://sedarweb.org/docs/sar/SEDAR16_final_SAR.pdf

SEDAR. 2008. SEDAR 17 – South Atlantic Spanish Mackerel Stock Assessment Report. SEDAR, North Charleston SC. 508 pp. Available online at: <http://sedarweb.org/docs/sar/S17SAR1%20Spanish%20Mackerel%20Final.pdf>

SEDAR. 2011. SEDAR 25 – South Atlantic Black Sea Bass Stock Assessment Report. SEDAR, North Charleston SC. 480 pp. Available online at: http://sedarweb.org/docs/sar/SEDAR25_BlackSeaBass_SAR.pdf

2. Methodology

2.1. Methodology

2.1.1 Project Foundation

A management team will outline the project objectives, critical success factors, and establish project scope boundaries. The project is being proposed as a phased approach. Phase 1 and 2 are proposed in this request.

- Phase 1 involves exploring the current processes and data inputs needed to support those processes. A conceptual design of data and process flow will be developed to guide database design.
- Phase 2 will be to build database structures, load the data and test that processes replicate current processes. Online reporting will be the final product of this phase.
- Future Phase 3 – not part of this proposal will be to build an administrative set of access points (screens or programs) to manage the various tables, parameters, and formulas to run the program.
- Future Phase 4 – not part of this proposal will be to add analytical functions to view the data. Manage versions of the data.

2.1.2 Scalability

The project is scalable (30% reduction in cost) in that some of the basic functionality can be completed in one year, while other features, such as Online reporting, may be put off until additional funding is available. See full text project narrative for cost differences.

2.1.3 Data Exploration

The project team comprised of project stakeholders, SEFSC Staff, and consultants/contractors will participate in data analysis sessions to identify and document a base understanding of the data, (entities and attributes) and the processes involved. It is expected that data providers (MRIP, SRHS, TPWD and GulfFIN) will be fully engaged as well, though given the many years that SEFSC has maintained this SAS system, SEFSC could if necessary accomplish the development independently.

2.1.4 Conceptual Design

The project team will design the data flow process and user interaction workflow for the system lifecycle. Data flow diagrams and entity relationship models will be documented capturing each aspect of the automated process. The data flow will provide the design of the Oracle database storage procedures and packages which will be executed by a scheduler.

2.1.5 Data Loading and User Interface Prototype:

The contractors will document the detailed specifications of the system and develop the system modules to support the data loading and user interface for viewing and making data corrections. The team will build database tables and storage structures using Data Definition Language (DDL) scripts generated from the Entity Relationship Diagram (ERD) tool. The team will identify data-types and storage parameters necessary to generate database structures. The new schema, along with the associated tables, will hold test data copied from existing sources.

2.1.6 System Testing and Online Reporting

This step of the project will test each step of the automated data loading and processing, verify the data accuracy, and record summarized output values for each task. Standardized estimates from Oracle will be compared to estimates from the existing SAS system. Recreational fisheries information management reports will be added to the Southeast Fisheries Reporting system.

2.1.7 Development Methodology

The technical approach for the task order will follow the Information Resource Management (IRM) division of SEFSC standards for developing information systems. The technical solution will be based on technologies currently being utilized or planned for in the near future at SEFSC. Technical Specific Tools, Software, Application Servers and Database Servers associated with the task order:

Software

- RDBMS: Oracle Enterprise Edition 11g
- Oracle PL/SQL, Packages, and Functions
- Oracle Data Modeler
- Oracle Java Developer, Java, and JSP
- Struts or similar Java development framework
- PL/SQL Developer
- Oracle APEX v5

Computer Hardware

- Windows PCs
- Linux-based, J2EE compliant servers

The following diagram depicts at a high level the workflow process for automating the recreational landings estimates.

2.1.8 Data Loading

The data loading effort will move recreational fisheries data from multiple sources, some hosted at GulfFIN and the Office of Science and Technology (S&T), into the SEFSC Data Warehouse staging area.

Data inputs to the system include raw survey datasets and estimates from MRIP, SRHS, TPWD, and other state agencies as required.

- Estimates files (landings and discards, when available)
- Intercept data
 - o MRFSS /MRIP
 - o TPWD target/nontarget files
 - o TPWD fish and party
- Effort estimates
 - o MRFSS/MRIP effort estimates in angler trips
 - o TPWD seasonal effort estimates in angler trips
 - o SRHS angler day estimates, angler trip estimates when available
- Biosample data
 - o GulfFIN size observations from state sampling in FL-TX.
 - o SRHS bioprofiles
 - o TPWD length files

Both GulfFIN and S&T host their data on Oracle database servers. It is anticipated that data can be queried over a database link between the external Oracle database servers and the SEFSC Oracle server.

2.1.9 Data Validation and Linking

The data validation and calibration effort is performed when new data has been imported into the SEFSC data warehouse staging area. Data validation will confirm that raw data was imported properly and completely. Oracle has a well-established database programming language which can perform simple queries as well as detailed analytic functions. SEFSC has successfully completed prior projects using Oracle database analytic functions and PL/SQL to perform complex data calculations, aggregations, and stratification (see Appendix A).

2.1.10 Calibration, Estimate Calculations and Report Data Preparation

The calibration and estimates effort will replace the programming logic which is manually run by multiple SAS scripts today. Logic will include partitioning TWPD into bimonthly wave estimates, partitioning aggregated values into individual species estimates, MRFSS for hire calibrations, MRIP estimation adjustments, MRFSS/MRIP stratification, etc. Report data will be pre-aggregated with drill-down linking.

2.1.11 Data Access Tools for Customers

Users of the recreational estimates data will have online access through the Southeast Fisheries Reporting website. Potential users of the data warehouse are the regional fisheries management councils, NOAA entities, e.g., the Southeast Regional Office, SEFSC stock assessment scientists and social scientists, the Office of Science & Technology, and possibly also the public. Users of the data access tools will be advised of the differences between the SEFSC estimates of recreational catches and those provided by MRIP and links to associated documentation will be provided.

2.1.12 Currently Used Methods and Processes

The SEFSC produces combined recreational estimates to stock assessment scientists and managers in the Southeast region which contain the following modifications to the original survey estimates:

- Partitioning the two seasonal estimates from the Texas Parks and Wildlife Department (TPWD) recreational fisheries into the standard bimonthly wave estimates used by MRIP (SEDAR 16 DW Report, every Gulf stock assessment since)
- Partitioning the TPWD aggregated "other species" estimate into individual species' estimates
- MRFSS for hire calibrations (SEDAR7-AW03, SEDAR16-DW15, SEDAR 17 DW Report, SEDAR 25 DW Report, SEDAR28-DW12)
- MRIP re-estimation adjustments (SEDAR31-DW25 , SEDAR32-DW02)
- MRIP APAIS adjustments (Carmichael and Van Vorhees 2015)
- MRFSS post stratification/MRIP domain estimation (MRIP survey programs)
- SEFSC weight estimation procedures (SEDAR22-DW16, SEDAR32-DW02)

The above described methods have been vetted by the SEDAR review process over several years. We expect the results of the proposed project and system to save the FSD a significant amount of time and effort and reduce reliance on an individual.

2.2. Region

Caribbean, Gulf of Mexico, South Atlantic

2.3. Geographic Coverage

NC-TX and PR

2.4. Temporal Coverage

1981-present

2.5. Frequency

Wave

2.6. Unit of Analysis

species, sub-region, year, wave, state, mode

2.7. Collection Mode

MRIP, MRFSS, state programs, SRHS

3. Communication

3.1. Internal Communication

The SEFSC has a well established system of weekly meetings during program development. This project will follow the same process to communicate status of the project and identify priority areas for the next week.

3.2. External Communication

Project status will be communicated to external (MRIP) interest via a quarterly report.

4. Assumptions/Constraints

4.1. New Data Collection

N

4.2. Is funding needed for this project?

Y

4.3. Funding Vehicle

Contract

4.4. Data Resources

Oracle server and DBA at SEFSC, possible ORACLE resources available at GSMFC for storage of TXPWD recreational datasets.

4.5. Other Resources

On site application development contractor at SEFSC.

4.6. Regulations

This project supports ACL monitoring.

4.7. Other

Process will be built in anticipation of the ability to access TXPWD and LA creel data at GSMFC, but will retain the ability to load these files by hand.

5. Final Deliverables

5.1. Additional Reports

Public access to combined recreational data for the SE region, multiple new reports from APEX

5.2. New Data Set(s)

5.3. New System(s)

Recreational integration system

6. Project Leadership

6.1. Project Leader and Members

First Name	Last Name	Title	Role	Organization	Email	Phone 1	Phone 2
Patrick	Cope	COTR	Team Member	SEFSC	patrick.cope@noaa.gov	305-361-4251	
David	Gloeckner	Chief, Fisheries Monitoring Branch	Team Leader	SEFSC	david.gloeckner@noaa.gov	305-361-4257	
Vivian	Matter	Recreational data analyst	Team Member	SEFSC	vivian.matter@noaa.gov	305-361-4571	
Refik	Orhun	Fisheries data analyst	Team Member	SEFSC	refik.orhun@noaa.gov	305-361-4244	
Steve	Turner	Chief, Fisheries Statistics Division	Team Member	SEFSC	Steve.Turner@noaa.gov	305-361-4482	

7. Project Estimates

7.1. Project Schedule

Task #	Schedule Description	Prerequisite	Schedule Start Date	Schedule Finish Date	Milestone
1	Scope out project needs	none	10/01/2016	10/31/2016	Y
2	Explore data and processes to develop understanding of needs.	Project scoping	10/01/2016	12/31/2016	Y
3	Design data flow and processes to guide database design	Data exploration	01/01/2017	03/31/2017	Y

Task #	Schedule Description	Prerequisite	Schedule Start Date	Schedule Finish Date	Milestone
4	Implement database objects in Oracle	database design	03/01/2017	05/31/2017	Y
5	Test automated loading, processing and develop reporting.	database implemented in Oracle	05/01/2017	09/30/2017	Y

7.2. Cost Estimates

Cost Name	Cost Description	Cost Amount	Date Needed
Consultants/Contracts	Project Foundation and Requirements-Programmer Analyst III	\$6068.40	10/01/2017
Consultants/Contracts	Project Foundation and Requirements-Database Specialist III	\$6120.60	10/01/2015
Consultants/Contracts	Data Analysis-Programmer Analyst III	\$24273.60	10/01/2016
Consultants/Contracts	Data Analysis-Database Specialist III	\$12241.20	10/01/2016
Consultants/Contracts	Conceptual Design & Prototyping -Programmer Analyst III	\$18205.20	01/01/2017
Consultants/Contracts	Conceptual Design & Prototyping -Database Specialist III	\$18361.80	01/01/2017
Consultants/Contracts	Data Loading & Interface Development-Programmer Analyst III	\$30342.00	03/01/2017
Consultants/Contracts	Data Loading & Interface Development-Database Specialist III	\$42844.20	03/01/2017
Consultants/Contracts	System Testing & Reporting-Programmer Analyst III	\$30342.00	05/01/2017
Consultants/Contracts	System Testing & Reporting-Database Specialist III	\$36723.60	05/01/2017
Administrative Support Fees	NOAA Link fee	\$6765.68	10/01/2016
TOTAL COST		\$232288.28	

8. Risk

8.1. Project Risk

Risk Description	Risk Impact	Risk Probability	Risk Mitigation Approach
Delay in GSMFC setting up a process to receive	If the data is not available over a database link to	Medium	Part of the project involves building a process to load

Risk Description	Risk Impact	Risk Probability	Risk Mitigation Approach
recreational data from TXPWD and LA creel. SEFSC expects that we will be able to access this information through a database link to GSMFC, which will speed up processing of data.	GSMFC, SEFSC will have to rely on files sent by FTP or email. This may slow the timeliness of data available for integration.		the files to Oracle by hand, so as to minimize any delay associated with the current process of loading the data by hand.

9. Supporting Documents

"project narrative with figures and appendix", page 1

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

Marine Recreational Information Program

FY 2016 Project Proposal

Develop integration of recreational datasets on SEFSC Oracle server to increase availability of management advice

David Gloeckner

david.gloeckner@noaa.gov

10/15/2015

1. Overview

1.1 Sponsor

Vivian Matter

1.2 Focus Group

Information Technology

1.3 Background

The Southeast accounts for more than half of the recreational fishing trips in the United States. As a result, recreational fisheries statistics form an important component of stock assessments for many species, and recreational fisheries management is critical for the Southeast Regional Office (SERO) and the three fisheries management councils (SAFMC, GMFMC and CFMC). Additionally the SEFSC provides the scientific support (data and assessment) for Atlantic highly migratory species (tunas, swordfish, billfish and sharks) managed by the Highly Migratory Species division, and for many of those stocks recreational catches represent large fractions of the total removals.

The SEFSC integrates information from multiple recreational surveys for use in stock assessment and management; as a result, the basic information used for those two purposes is the same. Currently the SEFSC integrates information from MRIP, the Texas Parks and Wildlife Department's (TPWD) Creel Survey and the Southeast Regional Headboat Survey (SRHS). The SEFSC expects to incorporate information from the Louisiana Creel Survey (LA Creel) in the coming year (LA Creel was the only survey conducted in LA in 2014 and in the relatively near future is expected to replace MRIP). Additionally in the next few years it is expected that the high interest in implementing charter boat log book reporting in the Gulf of Mexico and in the South Atlantic will result in additional sets of data to be integrated.

The SEFSC uses specific data sets as the definitive data sets for specific strata/data types. For instance MRFSS/MRIP is the definitive data set for VA-ME recreational landings and for non-headboat fisheries for NC-LA. There are numerous other recreational data sets which are not incorporated in the SEFSC system (SC charter boat survey, FL east coast red snapper survey, LA Creel 2013). The MRFSS and MRIP calibrations the SEFSC uses were developed following the procedures defined in the calibration workshops (Salz *et al* 2012, Carmichael and Van Vorhees 2015). Differences should exist between SEFSC estimates and the source survey estimates where 1) source survey might not provide calibrated

"project narrative with figures and appendix", page 2

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

estimates and 2) in the weight estimates for which SEFSC uses a standardized procedure for all estimates necessitated by missing estimates for some strata.

The SEFSC system includes a number of components needed to standardize information for assessment and management. For the MRFSS estimates of landings (continental 1981-2003 and Caribbean 2000-2014) there are catches for which landed weight is not estimated due to low numbers of observed weights, there are no landing estimates in weight for the TPWD survey and there will not be any for the LA Creel. The SEFSC system uses a standardized approach to fill these gaps and provide consistent weight estimates over the 35 year MRFSS through MRIP time series. In some cases, special treatment (domain estimates for the MRIP years, alternative treatments for the MRFSS years) is used to assign landings to a management region. In the Caribbean a large number of species are landed and many of them occur rarely; as a result there are multiple instances where the SEFSC system uses weight observations from other members of the same genus or the same family.

Fish released (dead and alive) from recreational fisheries at times represent substantial fractions of the total catch (commercial and recreational) of many of the important stocks managed in the southeast. The TPWD Survey does not provide estimates of released fish. The SEFSC system uses information other Gulf of Mexico recreational fisheries to calculate possible TX recreational releases.

Recreational fishing surveys often change over time while stock assessments and management require long time series of estimates which have been calibrated for survey changes. The SEFSC system incorporates such calibrations and often those calibrations must be made 1) from old surveys / survey methods to new and 2) from new to old so that various needs of customers can be accommodated. For instance managers may need current year MRIP estimates converted to MRFSS estimates if an ACL was calculated several years ago when MRFSS data were used in the assessment. Or conversely historical MRFSS data must be converted to MRIP estimates for use in a current stock assessment.

Currently, computing recreational landings estimates is a lengthy process and involves executing many computing processes manually. Some examples of the conversions that are currently done manually by sequentially running SAS programs are:

- Partitioning the two seasonal estimates from the TPWD recreational fisheries into the standard bimonthly wave estimates used by MRIP
- Partitioning the TPWD aggregated "other species" estimate into individual species' estimates
- MRFSS for hire calibrations
- MRIP re-estimation adjustments
- MRIP APAIS adjustments
- MRFSS post stratification/MRIP domain estimation
- Other Bias adjustments
- SEFSC weight estimation procedures

Additional calibrations are anticipated in the future, including those resulting from the change in the MRIP effort estimation for shore and private modes with the implementation of the MRIP Fishing Effort Survey. There will also be calibrations necessary between the LA Creel survey and MRIP.

Currently some of the data sets are used for:

- Indices by stock assessment scientists, such as estimation of catch rates,
- Size composition analyses for stock assessment
- Bag limit analyses by the regional office, SERO

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

- Standardized landings history in number and weight from multiple data sets (MRIP, TPWD and SRHS) for use in stock assessment and management.

The current system depends on one federal employee dedicated to both to southeast recreational data integration and dissemination. The fact that only one person is dedicated to this task which is critical for stock assessment and management represents a substantial vulnerability for NMFS.

1.4. Project Description

This project will move the recreational data extraction and integration processes conducted by the Fisheries Statistics Division (FSD) at the Southeast Fisheries Science Center (SEFSC) into an Oracle Data base with reporting capability in APEX.

This project addresses the needs of the Southeast region's stock assessment scientists and managers by streamlining and automating the recreational fisheries data extraction and integration processes. These processes are executed by the SEFSC and not by the survey providers. For example, TPWD provides estimates for two periods per year (Nov-May and May-Nov). The SEFSC calculates wave and annual estimates from the TX data for compatibility with other surveys, to fulfill calendar year requests, and in some cases to fulfill fishing year requests. Other examples are the MRIP calibrations for the charter mode and other survey methodology changes. Assessment scientists need a consistent time series of estimates and adjustments must be made to the survey estimates. Although the SEFSC works closely with survey providers on solutions, especially MRIP, it has fallen on the SEFSC to execute these important processes for assessment and management needs. If the survey providers desired, access to these modified estimates could be provided.

The resulting database will reflect current Southeast region stock assessment decisions such as,

- the elimination of some modes of fishing from certain species' estimates
- the recommended estimate when competing surveys are available (e.g. MRIP headboat estimates in the Southeast will be excluded when SRHS estimates are available)
- standard substitutions in the early 1980s (e.g. charter and private modes estimated from TPWD data back to 1981)
- post-stratified/domain estimates from MRIP when necessary (e.g. to allocate the Florida Keys to the Atlantic)

1.5. Objectives

The objectives of this proposal are to allow users 1) access to recreational estimates from all programs in one standardized location and 2) access to raw intercept data sets in SEFSC formats. This includes:

- Automation of existing processes
- Development of processes to incorporate new data sets such as LA Creel
- Creation of a data warehouse with storage of data sets and estimates on the ORACLE database
- User access to data sets and estimates via a data warehouse and web portal

Development of a fully integrated system into the ORACLE database would reduce the vulnerability to loss of personnel and increase the feasibility of training backup and/or replacement personnel.

1.6. References

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

Carmichael, J. and D. Van Voorhees (editors). 2015. MRIP Calibration Workshop II. Report of a workshop on calibrating MRIP catch estimates, September 8-10, 2014. SAFMC/SEDAR, North Charleston, SC.

Matter, V.M. and A. Rios. 2013. SEDAR32-DW02. MRFSS to MRIP Adjustment Ratios and Weight Estimation Procedures for South Atlantic and Gulf of Mexico Managed Species. National Marine Fisheries Service Southeast Fisheries Science Center, Fisheries Statistics Division, Miami, FL and National Marine Fisheries Service Southeast Fisheries Science Center, Sustainable Fisheries Division, Miami, FL.

Matter, V. and S. Turner 2010. Estimated Recreational Catch in Weight: Method for Filling in Missing Weight Estimates from the Recreational Surveys with Application to Yellowedge Grouper, Tilefish (golden), and Blueline Tilefish (SEDAR 22-DW-16), National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division (SFD-2010-003).

Rios, A, V.M. Matter, J.F. Walter, N. Farmer, and S.J. Turner. 2012. SEDAR31-DW25 Estimated Conversion Factors for Adjusting MRFSS Gulf of Mexico Red Snapper Catch Estimates and Variances in 1981-2003 to MRIP Estimates and Variances. National Marine Fisheries Service Southeast Fisheries Science Center, Sustainable Fisheries Division, Miami, FL, National Marine Fisheries Service Southeast Fisheries Science Center, Fisheries Statistics Division, Miami, FL, and National Marine Fisheries Service Southeast Regional Office, Saint Petersburg, FL.

Salz, R., T. Williams, E. Williams, J. Walter, K. Drew and G. Bray. 2012. MRFSS/MRIP Calibration Workshop Ad-Hoc Working Group Report. Available online at: <http://sedarweb.org/otw-01-calibration-workshop-ad-hoc-group-report-may-2012>.

SEDAR. 2009. SEDAR 16 – South Atlantic and Gulf of Mexico King Mackerel Stock Assessment Report. SEDAR, North Charleston SC. 484 pp. Available online at: http://sedarweb.org/docs/sar/SEDAR16_final_SAR.pdf

SEDAR. 2008. SEDAR 17 – South Atlantic Spanish Mackerel Stock Assessment Report. SEDAR, North Charleston SC. 508 pp. Available online at: <http://sedarweb.org/docs/sar/S17SAR1%20Spanish%20Mackerel%20Final.pdf>

SEDAR. 2011. SEDAR 25 – South Atlantic Black Sea Bass Stock Assessment Report. SEDAR, North Charleston SC. 480 pp. Available online at: http://sedarweb.org/docs/sar/SEDAR25_BlackSeaBass_SAR.pdf

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

2. Methodology

2.1. Methodology

2.1.1 Project Foundation

A management team will outline the project objectives, critical success factors, and establish project scope boundaries. The project is being proposed as a phased approach. Phase 1 and 2 are proposed in this request.

- Phase 1 involves exploring the current processes and data inputs needed to support those processes. A conceptual design of data and process flow will be developed to guide database design.
- Phase 2 will be to build database structures, load the data and test that processes replicate current processes. Online reporting will be the final product of this phase.
- Future Phase 3 – not part of this proposal will be to build an administrative set of access points (screens or programs) to manage the various tables, parameters, and formulas to run the program.
- Future Phase 4 – not part of this proposal will be to add analytical functions to view the data. Manage versions of the data.

2.1.2 Scalability

The project is scalable in that some of the basic functionality can be completed in one year, while other features, such as Online reporting, administrative tools and analytical functions may be put off until additional funding is available. See section 6.2 and 6.3 for cost differences.

2.1.3 Data Exploration

The project team comprised of project stakeholders, SEFSC Staff, and consultants/contractors will participate in data analysis sessions to identify and document a base understanding of the data, (entities and attributes) and the processes involved. It is expected that data providers (MRIP, SRHS, TPWD and GulfFIN) will be fully engaged as well, though given the many years that SEFSC has maintained this SAS system, SEFSC could if necessary accomplish the development independently.

2.1.4 Conceptual Design

The project team will design the data flow process and user interaction workflow for the system lifecycle. Data flow diagrams and entity relationship models will be documented capturing each aspect of the automated process. The data flow will provide the design of the Oracle database storage procedures and packages which will be executed by a scheduler.

2.1.5 Data Loading and User Interface Prototype:

The contractors will document the detailed specifications of the system and develop the system modules to support the data loading and user interface for viewing and making data corrections. The team will build database tables and storage structures using Data Definition Language (DDL) scripts generated from the Entity Relationship Diagram (ERD) tool. The team will identify data-types and storage parameters necessary to generate database structures. The new schema, along with the

"project narrative with figures and appendix", page 6

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

associated tables, will hold test data copied from existing sources. This will allow developers to test the system with real data.

2.1.6 System Testing and Online Reporting

This phase of the project will test each step of the automated data loading and processing, verify the data accuracy, and record summarized output values for each task. Standardized estimates from Oracle will be compared to estimates from the existing SAS system to ensure consistency with current procedures. Logging results of each task in the process will allow users to review that the unattended job is running accurately. Recreational fisheries information management reports will be added to the Southeast Fisheries Reporting system. Data will also be available from database views for system-to-system data sharing.

2.1.7 Development Methodology

The technical approach for the task order will follow the Information Resource Management (IRM) division of SEFSC standards for developing information systems. The technical solution will be based on technologies currently being utilized or planned for in the near future at SEFSC. Technical Specific Tools, Software, Application Servers and Database Servers associated with the task order:

Software

- RDBMS: Oracle Enterprise Edition 11g
- Oracle PL/SQL, Packages, and Functions
- Oracle Data Modeler
- Oracle Java Developer, Java, and JSP
- Struts or similar Java development framework
- PL/SQL Developer
- Oracle APEX v5

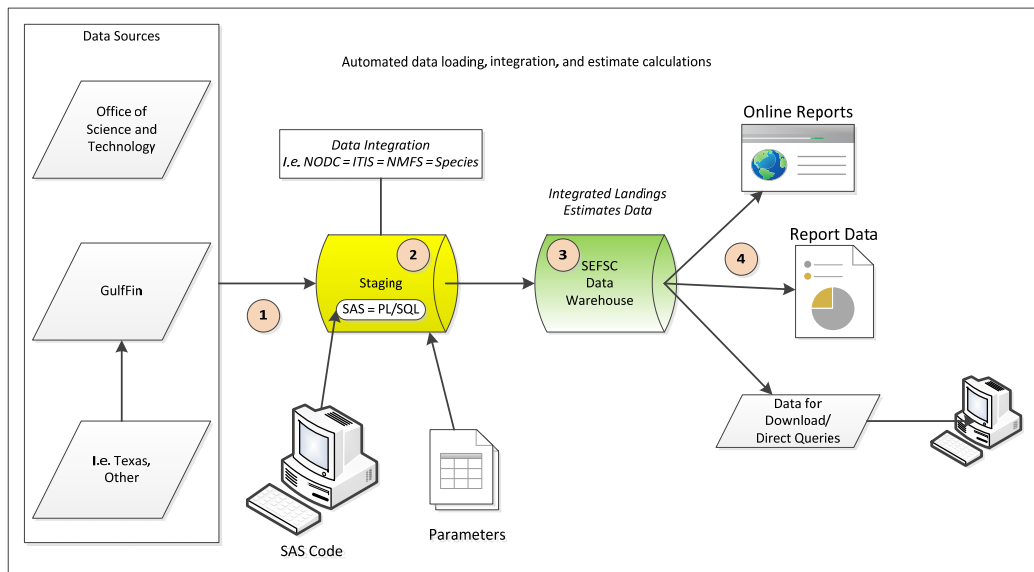
Computer Hardware

- Windows PCs
- Linux-based, J2EE compliant servers

The following diagram depicts at a high level the workflow process for automating the recreational landings estimates.

"project narrative with figures and appendix", page 7

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice



2.1.8 Data Loading

The data loading effort will move recreational fisheries data from multiple sources, some hosted at GulfFIN and the Office of Science and Technology (S&T), into the SEFSC Data Warehouse staging area.

Data inputs to the system include raw survey datasets and estimates from MRIP, SRHS, TPWD, and other state agencies as required.

- Estimates files (landings and discards, when available)
- Intercept data
 - MRFSS /MRIP
 - TPWD target/nontarget files
 - TPWD fish and party
- Effort estimates
 - MRFSS/MRIP effort estimates in angler trips
 - TPWD seasonal effort estimates in angler trips
 - SRHS angler day estimates, angler trip estimates when available
- Biosample data
 - GulfFIN size observations from state sampling in FL-TX.
 - SRHS bioprofiles
 - TPWD length files

Both GulfFIN and S&T host their data on Oracle database servers. It is anticipated that data can be queried over a database link between the external Oracle database servers and the SEFSC Oracle server.

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

Using a database link will minimize the technical complexity of importing data and allow the process to be scheduled for automated running.

SEFSC currently uses data base links to import commercial fisheries landings data on a scheduled process from the ACCSP (1950 to present) and GulfFIN (1999 to present) data warehouses. The processes run unattended with detailed task logging. Each step in the process records summarized record counts, results from validation, and sends stakeholders a summarized email report of success.

2.1.9 Data Validation and Linking

The data validation and calibration effort is performed when new data has been imported into the SEFSC data warehouse staging area.

Data will be imported in batches. Record counts will be recorded and logged. Data validation will confirm that raw data was imported properly and completely. Data which are numbers and dates but stored in string fields will be validated and converted to intended datatypes.

Data linking will perform lookup queries against known lists using conformed dimensions. Linking to common dimensions will allow users of the data to incorporate reference tables for data filtering and stratification. During this process the various sources of data will be transformed where needed to a common standard of codes and filter criteria.

The linking and standardizing will be performed inside the SEFSC Oracle database using Programming Language SQL (PL/SQL). Oracle has a well-established database programming language which can perform simple queries as well as detailed analytic functions. Analytic functions are commonly used to compute cumulative, moving, centered, and reporting aggregates. PL/SQL code and temporary tables will be used to develop any necessary custom functions and procedures. SEFSC has successfully completed prior projects using Oracle database analytic functions and PL/SQL to perform complex data calculations, aggregations, and stratifications, such as Commercial Landings Monitoring, Economic Data Reporting and the Unified Data Processing system which performs advances quality and control validation using historic data and empirical distributions (see Appendix A).

2.1.10 Calibration, Estimate Calculations and Report Data Preparation

After validating and standardizing the source data into staging tables of the SEFSC data warehouse, the automated process will perform data processing replicating logic from current SAS scripts, which were rewritten into PL/SQL database code. Parameter tables, holding user specified values, will be utilized as needed to perform estimate calibrations and adjustments.

The calibration and estimates effort will replace the programming logic which is manually run by multiple SAS scripts today. Logic will include partitioning TWPD into bimonthly wave estimates, partitioning aggregated values into individual species estimates, MRFFS for hire calibrations, MRIP estimation adjustments, MRFSS/MRIP stratification, etc.

Report data will be pre-aggregated with drill-down linking. Users will be able to review report data with the ability to view detailed data.

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

2.1.11 Data Access Tools for Customers

Users of the recreational estimates data will have online access through the Southeast Fisheries Reporting website. Users will be authenticated with logins and defined page level authorization. Users will be able to view and extract data online using interactive reporting with custom fields and data filters. Interactive reporting will enable end users to customize reports. Users can alter the layout of report data by selecting columns, applying filters, highlighting, and sorting. Users can also define breaks, aggregations, charts, grouping, and add their own computations. Users can create multiple variations of a report and save them as named reports, for either public or private viewing.

In addition to online reporting from the Southeast Fisheries Reporting website system users will have the ability to pull prepared data via Oracle SQL views to support system-to-system data sharing.

Potential users of the data warehouse are the regional fisheries management councils, NOAA entities, e.g., the Southeast Regional Office, SEFSC stock assessment scientists and social scientists, the Office of Science & Technology, and possibly also the public.

Users of the data access tools will be advised of the differences between the SEFSC estimates of recreational catches and those provided by MRIP and links to associated documentation will be provided.

2.1.12 Currently Used Methods and Processes

The SEFSC produces combined recreational estimates to stock assessment scientists and managers in the Southeast region which contain the following modifications to the original survey estimates:

- Partitioning the two seasonal estimates from the Texas Parks and Wildlife Department (TPWD) recreational fisheries into the standard bimonthly wave estimates used by MRIP (SEDAR 16 DW Report, every Gulf stock assessment since)
- Partitioning the TPWD aggregated "other species" estimate into individual species' estimates
- MRFFS for hire calibrations (SEDAR7-AW03, SEDAR16-DW15, SEDAR 17 DW Report, SEDAR 25 DW Report, SEDAR28-DW12)
- MRIP re-estimation adjustments (SEDAR31-DW25 , SEDAR32-DW02)
- MRIP APAIS adjustments (Carmichael and Van Vorhees 2015)
- MRFSS post stratification/MRIP domain estimation (MRIP survey programs)
- SEFSC weight estimation procedures (SEDAR22-DW16, SEDAR32-DW02)

The above described methods have been vetted by the SEDAR review process over several years. The SFD at SEFSC has in-house programmers that are consulted and will be part of team streamlining the recreational fisheries data integration and dissemination process. We expect the results of the proposed project and system, once implemented, to save the FSD a significant amount of time and effort in processing recreational fisheries data.

As already mentioned the project proposes to convert programs/scripts written in SAS code to SQL code. Programmers working for FSD have demonstrated on many occasions (e.g. data analysis for gag grouper and king mackerel Sedars, etc.) that SAS code can be rewritten in SQL code and will produce

"project narrative with figures and appendix", page 10

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

identical results, often with considerable gains in computational efficiency. The natural language of the ORACLE database is SQL and therefore represent the more efficient way of “doing business” in the ORACLE database such as adding, manipulating and extracting data.

Updated or new tools provided by MRIP will be incorporated into the system by writing new code as part of maintenance contract. The proposed system will increase data throughput and FSD’s capability of moving data from collection to end-users such as the stock assessment scientists and fisheries managers (SERO). Once the system is fully adopted into ORACLE, we expect an easier incorporation of changes to MRIP. Automation of data processing within ORACLE will make the transfer of knowledge less complicated than before, and less dependent on the individual. It can become more push-button with less manual and human input and allow more time for quality control. See Appendix A for evidence that methods used to convert SAS to ORACLE have been vetted.

2.2. Region

Caribbean, Gulf of Mexico, South Atlantic

2.3. Geographic Coverage

NC-TX and PR

2.4. Temporal Coverage

1981-present

2.5. Frequency

Wave

2.6. Unit of Analysis

MRIP – species, sub-region, year, wave, state, mode, and area fished

SRHS - species, year, month, and SRHS area

TPWD – species, year, season, mode, and area fished

2.7. Collection Mode

MRIP, MRFSS, state programs, SRHS

3. Assumptions and Constraints

3.1. New Data Collection

N

3.2. Is funding needed for this project

Y

3.3. Funding Vehicle

Contract

"project narrative with figures and appendix", page 11

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

3.4. Data Resources

Oracle server and DBA at SEFSC, possible ORACLE resources available at GSMFC for storage of TXPWD recreational datasets.

3.5. Other Resources

On site application development contractor at SEFSC

3.6. Regulations

This project supports ACL monitoring

3.7. Other

4. Final Deliverables

An Automated process of ACL estimation that is integrated into the Oracle database and a web portal that will allow users access to raw data, annual estimates and other data products.

4.1. Additional Reports

Public access to combined recreational data for the SE region, multiple new reports from APEX

4.2. New Data Sets

4.3. New Systems

Recreational integration system

5. Project Leadership

5.1. Project Leader and Members

Name	Project Role	Organization/ Title	Affiliation	Email	Phone 1	Phone 2
David Gloeckner	Team Leader	Chief, Fisheries Monitoring Branch	SEFSC	david.gloeckner@noaa.gov	305-361-4257	
Steve Turner	Team Member	Chief, Fisheries Statistics Division	SEFSC	Steve.Turner@noaa.gov	305-361-4482	
Refik Orhun	Team Member	Fisheries data analyst	SEFSC	refik.orhun@noaa.gov	305-361-4244	
Patrick Cope	Team Member	COTR	SEFSC	patrick.cope@noaa.gov	305-361-4251	
Vivian Matter	Team Member	Recreational data analyst	SEFSC	vivian.matter@noaa.gov	305-361-4571	

"project narrative with figures and appendix", page 12

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

6. Project Estimates

6.1. Project Schedule

#	Schedule Description	Planned Start	Planned Finish	Prerequisites
1	Scope out project needs	10/03/2016	10/31/2016	none
2	Explore data and processes to develop understanding of needs	10/01/2016	12/31/2016	Project scoping
3	Design data flow and processes to guide database design	01/01/2017	03/31/2017	Data exploration
4	Implement database objects in Oracle	03/01/2017	05/31/2017	Database design
5	Test automated loading, processing and develop reporting.	05/01/2017	09/30/2017	Database implemented in Oracle

"project narrative with figures and appendix", page 13

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

6.2. Cost Estimates-full FY 2016 project

Project Effort and Cost Estimate

Work Schedule

The following outlines an estimated work effort by project phase

Project Phase	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Project Foundation												
Data Loading												
Data Validation and Linking												
Calibration, Calculating Weights and Report Data Preparation												
Testing, Reporting & Data Access Tools for Customers												

Proposed Staffing Labor Categories

The project work will be completed by two staff members at 75% capacity, a Programmer Analyst and a Database Specialist III, with minimal assistance by a project manager.

Role	Description	Rate
Programmer Analyst III	PROGRAMMER ANALYST III - Reviews, analyzes, and modifies programming systems including encoding, testing, debugging and installing to support an organization's application systems. Consults with users to identify current operating procedures and to clarify program objectives. May be expected to write documentation to describe program development, logic, coding, and corrections: Writes manuals for users to describe installation and operating procedures. Must have a working knowledge of relational databases and client-server concepts. A wide degree of creativity and latitude is expected. Requires proficiency in programming languages. May lead and direct others. BS/BA degree in a related area (or equivalent), and 4 to 6 years of related experience.	\$101.14/hr
Database Specialist III	DATABASE SPECIALIST III - Designs and builds relational databases for data storage or processing. Develops strategies for warehouse implementation, data acquisition, and archive recovery. Cleans and maintains the database by removing and deleting old data. May evaluate new data sources for adherence to the organization's quality standards and ease of integration. Familiar with a variety of the field's concepts, practices, and procedures. Relies on experience and judgment to plan and accomplish goals. Performs a variety of complicated tasks. May lead and direct the work of others. Typically reports to a project leader or manager. A wide degree of creativity and latitude is expected. BS/BA degree in a related area (or equivalent), and 4 to 6 years of experience in the field or in a related area.	\$102.01/hr
Project Manager II	PROJECT MANAGER II - Under general supervision, is responsible for all aspects of field and/or task-level project performance (i.e., technical, contractual, administrative, financial). Supervises personnel involved in all aspects of project activity, organizes and assigns responsibilities to subordinates, oversees the successful completion of all assigned tasks, and maintains customer contacts to ensure conformity to all contractual obligations. Exercises independent judgment, as well as a high-level of analytical skill, in solving non-routine technical, administrative, and managerial problems. BS/BA degree in related field (or equivalent) plus 4 to 6 years of related experience, 2 of which must be in a supervisory role.	\$106.91/hr

"project narrative with figures and appendix", page 14

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

The table below contains the estimate of project effort and cost by project phase, consulting labor category, hours, and hourly rate.

Labor Category	Hours	Rate	Sub-total
Project Foundation and Requirements			
Programmer Analyst III	60	\$101.14/hr	\$ 6,068.40
Database Specialist III	60	\$102.01/hr	\$ 6,120.60
Data Analysis			
Programmer Analyst III	240	\$101.14/hr	\$ 24,273.60
Database Specialist III	120	\$102.01/hr	\$ 12,241.20
Conceptual Design & Prototyping			
Programmer Analyst III	180	\$101.14/hr	\$ 18,205.20
Database Specialist III	180	\$102.01/hr	\$ 18,361.80
Data Loading & Interface Development			
Programmer Analyst III	300	\$101.14/hr	\$ 30,342.00
Database Specialist III	420	\$102.01/hr	\$ 42,844.20
System Testing & Reporting			
Programmer Analyst III	300	\$101.14/hr	\$ 30,342.00
Database Specialist III	360	\$102.01/hr	\$ 36,723.60
Sub-Total			\$225,522.60
NOAA Link Fee			\$6,765.68
Cost Estimate			\$232,288.28

Estimate for Travel

Location	Air	Lodge	Auto	Meals	Misc	Subtotal	Count FTE	Total
0	0	0		0	0	0	0	0.00
*Total								\$0.0

*No travel costs expected for Task Order

Estimate for HW/SW

Item	Count	Unit cost	Cost Sub-total
PC Computers/servers for Systems Development	0	\$0.00	\$0
Software	0	\$0.00	\$0
Total			\$0

*No HW/SW Costs expected for Task Order

Estimate Summary Table

Activity	Cost
Effort	\$232,288.28
Travel	\$0
Hardware/Software	\$0
*Cost Estimate Total	\$232,000

*Cost estimate rounded from \$232,288.28

"project narrative with figures and appendix", page 15

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

6.3. Cost Estimates-scaled down project

Project Effort and Cost Estimate

Work Schedule

The following outlines an estimated work effort by project phase

Project Phase	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Project Kick-off												
Data Exploration												
Conceptual Design												
Data Loading & Interface Development												
System Testing & Reporting												

Proposed Staffing Labor Categories

The project work will be completed by two staff members at 75% capacity, a Programmer Analyst and a Database Specialist III, with minimal assistance by a project manager.

Role	Description	Rate
Programmer Analyst III	PROGRAMMER ANALYST III - Reviews, analyzes, and modifies programming systems including encoding, testing, debugging and installing to support an organization's application systems. Consults with users to identify current operating procedures and to clarify program objectives. May be expected to write documentation to describe program development, logic, coding, and corrections. Writes manuals for users to describe installation and operating procedures. Must have a working knowledge of relational databases and client-server concepts. A wide degree of creativity and latitude is expected. Requires proficiency in programming languages. May lead and direct others. BS/BA degree in a related area (or equivalent), and 4 to 6 years of related experience.	\$101.14/hr
Database Specialist III	DATABASE SPECIALIST III - Designs and builds relational databases for data storage or processing. Develops strategies for warehouse implementation, data acquisition, and archive recovery. Cleans and maintains the database by removing and deleting old data. May evaluate new data sources for adherence to the organization's quality standards and ease of integration. Familiar with a variety of the field's concepts, practices, and procedures. Relies on experience and judgment to plan and accomplish goals. Performs a variety of complicated tasks. May lead and direct the work of others. Typically reports to a project leader or manager. A wide degree of creativity and latitude is expected. BS/BA degree in a related area (or equivalent), and 4 to 6 years of experience in the field or in a related area.	\$102.01/hr
Project Manager II	PROJECT MANAGER II - Under general supervision, is responsible for all aspects of field and/or task-level project performance (i.e., technical, contractual, administrative, financial). Supervises personnel involved in all aspects of project activity, organizes and assigns responsibilities to subordinates, oversees the successful completion of all assigned tasks, and maintains customer contacts to ensure conformity to all contractual obligations. Exercises independent judgment, as well as a high-level of analytical skill, in solving non-routine technical, administrative, and managerial problems. BS/BA degree in related field (or equivalent) plus 4 to 6 years of related experience, 2 of which must be in a supervisory role.	\$106.91/hr

"project narrative with figures and appendix", page 16

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

The table below contains the estimate of project effort and cost by project phase, consulting labor category, hours, and hourly rate.

Labor Category	Hours	Rate	Sub-total
Project Foundation and Requirements			
Programmer Analyst III	60	\$101.14/hr	\$6,084.00
Database Specialist III	60	\$102.01/hr	\$6,120.60
Data Analysis			
Programmer Analyst III	240	\$101.14/hr	\$24,273.60
Database Specialist III	120	\$102.01/hr	\$12,241.20
Conceptual Design & Prototyping			
Programmer Analyst III	180	\$101.14/hr	\$18,205.20
Database Specialist III	180	\$102.01/hr	\$18,361.80
Data Loading & Interface Development			
Programmer Analyst III	200	\$101.14/hr	\$20,228.00
Database Specialist III	300	\$102.01/hr	\$30,603.00
System Testing & Reporting			
Programmer Analyst III	100	\$101.14/hr	\$10,114.00
Database Specialist III	125	\$102.01/hr	\$12,751.25
Sub-Total			\$158,982.65
NOAA Link Fee			\$4,769.48
Cost Estimate			\$163,752.13

Estimate for Travel

Location	Air	Lodge	Auto	Meals	Misc	Subtotal	Count FTE	Total
0	0	0		0	0	0	0	0.00
*Total								\$0.0

*No travel costs expected for Task Order

Estimate for HW/SW

Item	Count	Unit cost	Cost Sub-total
PC Computers/servers for Systems Development	0	\$0.00	\$0
Software	0	\$0.00	\$0
Total			\$0

*No HW/SW Costs expected for Task Order

"project narrative with figures and appendix", page 17

Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice

Estimate Summary Table

Activity	Cost
Effort	\$163,752.13
Travel	\$0
Hardware/Software	\$0
*Cost Estimate Total	\$164,000

*Cost estimate rounded from \$163,752.13

"project narrative with figures and appendix", page 18

Appendix A – Supporting Documentation for MRIP FY16 Proposal
Final Report of Completed Project with Similar Technology and Approach

MRIP FY16 Proposals: Develop integration of recreational datasets on SEFSC
Oracle sever to increase availability of management advice

*4. Provide supporting documentation of: the methods used, that the methods have been vetted, and that the appropriate team members are included and/or consulted.
(e.g. rebuilding in Oracle with Apex interface to provide one-stop shop for SERO, Stock Assessment Scientists and the Public..etc.)*

The following documentation is the final report for the economic web-based query tool which is an Oracle with Apex interface providing data for SERO economists. The recently completed project follows a similar development approach as the "MRIP FY16 Proposals: Develop integration of recreational datasets on SEFSC Oracle sever to increase availability of management advice".

Project similarities:

- Data loading of previously SAS maintained data
- Data loading from ACCSP and GulfFIN
- Automated processing
- Data validation and analytic processing
 - Custom PL/SQL code and temporary tables for data aggregations and stratification
 - Use of Oracle analytic functions
- Use of the Southeast Fisheries Science Center (SEFSC) Data Warehouse
- Use of the Southeast Fisheries Reporting system

NOAA FISHERIES - SOUTHEAST FISHERIES SCIENCE CENTER (SEFSC)

Economic Web-Based Query Tool

FY15 Final Report

Larry Perruso - NOAA Federal

9/3/2015

Table of Contents

Table of Contents	2
Project Objective	5
Overview	5
Project goals	5
Development Approach	10
Project Participants and Stakeholders	10
Load and Prepare Data	11
Build Initial Foundation Reports (Fully Functioning Prototypes)	11
Hold Live-Demos of Web Based Query Tool	12
Obtain User and Stakeholder Feedback	12
SEFSC Economic Web-query Tool - List of Desired Functionality	12
Web-Based Query Tool Architecture	14
Hardware	15
Software	15
Data Sets and User Groups (SFR Access Path)	15
User Access Authentication	16
Web-based Query Tool Users	17
Interface Layout	17
Regions	17
Query Parameters/Filters	18
User Customizations	18
Query Results	18
Trip Limit Totals	18
Graphs	18
Detailed Metrics - Descriptive Statistics v.2 Interface	19
Data Sources	19
SAS Dataset (provided by Larry Perruso)	19
SEFSC SAS Dataset Attributes (abbreviated)	20
Southeast Logbooks	20

Economic Web-Based Query Tool

FY15 Final Report [Draft]

Data Warehouses – Accumulated Landings using Trip Ticket data	20
ACCSP – Trip Tickets.....	20
GULFFIN – Trip Tickets	21
Automated Data Import.....	21
Data Preparation.....	22
Inflation Adjustments - Consumer Price Index	23
Define BASE.....	23
Choose CPI Database	23
Choose Seasonally Adjusted or Not	23
Choose Area	23
Choose Alternate or Current Base	24
Create Customized Table	24
Define Periodicity.....	24
Choose Year	24
Consumer Price Index-All Urban Consumers -- Year (Screen 6 of 6)	24
Download CPI Data (Note file names)	24
Web-Based Query Economic Analysis Reports	25
Query Parameters/Filters	25
Filter Species Group	25
Filter Species	25
Filter Region	25
Filter Dates	25
Trip Limits.....	25
Revenue	25
Inflation Adjustment	26
Filter Variable - Gear	26
Filter Variable - State	26
Filter Variable - Counties.....	26
Query Results	26
Query Result Columns and Descriptions.....	27
Trip Limit Results.....	29
Graphs.....	30

"project narrative with figures and appendix", page 22

Economic Web-Based Query Tool	FY15 Final Report [Draft]
<hr/>	
Appendix	31
New Desired Functionality	31

Project Objective

Overview

The Economic Web-Based Query Tool effort is to designing and developing a self-service online reporting application to be used by the Southeast Regional Office (SERO) for filtering and analyzing economic data collected and prepared by the Southeast Fisheries Science Center (SEFSC).

Project goals

The objective of this work is to design and build a web-based data reporting application to view, group, and filter economic (catch revenue, cost and pricing) data that are integrated from data sources collected by the SEFSC. The web-based application is equipped with features that allow the SERO office staff to prepare and analyze self-service reports which are prepared manually today (usually by SEFSC economists). The new web-based online reporting application is shared by the SERO and the SEFSC and accessible by authorized staff, rather than manually prepared spreadsheet reports which are shared via email attachments. The implementation of this interface standardized routine aggregations of economic information regularly requested by SERO staff economists for regulatory assessment needs when performing economic impact analyses.

Through the Web-Based Query Tool SERO users have the ability to setup criteria such as species, year, and location identifiers for filtering the data to the specific dataset they are seeking. The users also have the ability to roll-up the numbers to an aggregated level, group the data by controlled breaks, add, remove, and rearrange report columns, created summation totals, and export data to their local PC.

The browser-based system shall require no additional software to be installed by the users. The Oracle database shall have links to economic data stored in database tables, which shall also need to be developed, with information specific to the type of data reporting required. During the discovery and requirements phases of the effort the contractor shall further outline the specifics of the data and reporting features required.

"project narrative with figures and appendix", page 24

FY15 Final Report [Draft]

Economic Web-Based Query Tool

The following table depicts an example of a manually prepared report.

Year	Number of vessels that caught SATL bluefish (> 0 lbs)	Number of trips that caught SATL bluefish	SATL Bluefish landings (lbs gutted wt)	Dockside revenue from SATL bluefish only (2012 \$)	Other species' landings jointly caught with SATL bluefish (lbs gutted wt)	Dockside revenue from 'other species' jointly caught with SATL bluefish only (2012 \$)	Number of trips that only caught other species	Other species' landings on SATL trips without bluefish (lbs gutted wt)	Dockside revenue from 'other species' caught on SATL trips without bluefish (2012 \$)	Total dockside revenue (2012 \$)
2008	119	714	362,562	\$711,302	564,485	\$1,462,798	3,244	2,931,841	\$7,492,040	\$9,666,140
2009	149	795	435,104	\$817,298	688,642	\$1,680,922	3,806	3,526,472	\$8,079,124	\$10,577,344
2010	131	705	397,165	\$879,655	557,226	\$1,362,821	3,788	3,439,819	\$7,601,958	\$9,844,433
2011	98	320	117,102	\$305,491	355,018	\$946,502	2,974	2,794,739	\$6,161,852	\$7,413,845
2012	125	523	294,254	\$682,699	383,616	\$1,042,293	3,611	2,652,061	\$6,813,035	\$8,538,027
Avg	124	611	321,237	\$679,289	509,797	\$1,299,067	3,485	3,068,986	\$7,229,602	\$9,207,958

Year	Number of vessels that caught SATL bluefish (> 0 lbs)	Number of trips that caught SATL bluefish	SATL Bluefish landings (lbs gutted wt)	Other species' landings jointly caught with SATL bluefish (lbs gutted wt)	Number of trips that only caught other species	Other species' landings on SATL trips without bluefish (lbs gutted wt)
2008	119	714	362,562	564,485	3,244	2,931,841
2009	149	795	435,104	688,642	3,806	3,526,472
2010	131	705	397,165	557,226	3,788	3,439,819
2011	98	320	117,102	355,018	2,974	2,794,739
2012	125	523	294,254	383,616	3,611	2,652,061
Average	124	611	321,237	509,797	3,485	3,068,986

August 2015

Page 6

"project narrative with figures and appendix", page 25

FY15 Final Report [Draft]

Economic Web-Based Query Tool

Year	Number of vessels that caught SATL blueline tilefish (> 0 lbs)	Dockside revenue from SATL blueline tilefish only (2012 \$)	Dockside revenue from 'other species' jointly caught with SATL blueline tilefish only (2012 \$)	Dockside revenue from 'other species' caught on SATL trips without blueline tilefish (2012 \$)	Total dockside revenue (2012 \$)	Average total dockside revenue per vessel (current \$)
2008	119	\$711,302	\$1,462,798	\$7,492,040	\$9,666,140	\$81,228
2009	149	\$817,298	\$1,680,922	\$8,079,124	\$10,577,344	\$70,989
2010	131	\$879,655	\$1,362,821	\$7,601,958	\$9,844,433	\$75,148
2011	98	\$305,491	\$946,502	\$6,161,852	\$7,413,845	\$75,651
2012	125	\$682,699	\$1,042,293	\$6,813,035	\$8,538,027	\$68,304
Average	124	\$679,289	\$ 1,299,067	\$ 7,229,602	\$ 9,207,958	\$ 74,264

Economic Web-Based Query Tool

FY15 Final Report [Draft]

The following Web-based Query Tool images depict the user driven/filtering reporting environment hosted at the Southeast Fisheries Reporting system. Online Reports such as this were developed to provide users direct access to the data for viewing and downloading. The new Web-Based Query Tool streamlines the data sharing process, removes redundancies, includes automation where applicable, and provides a self-services environment with less manual operations.

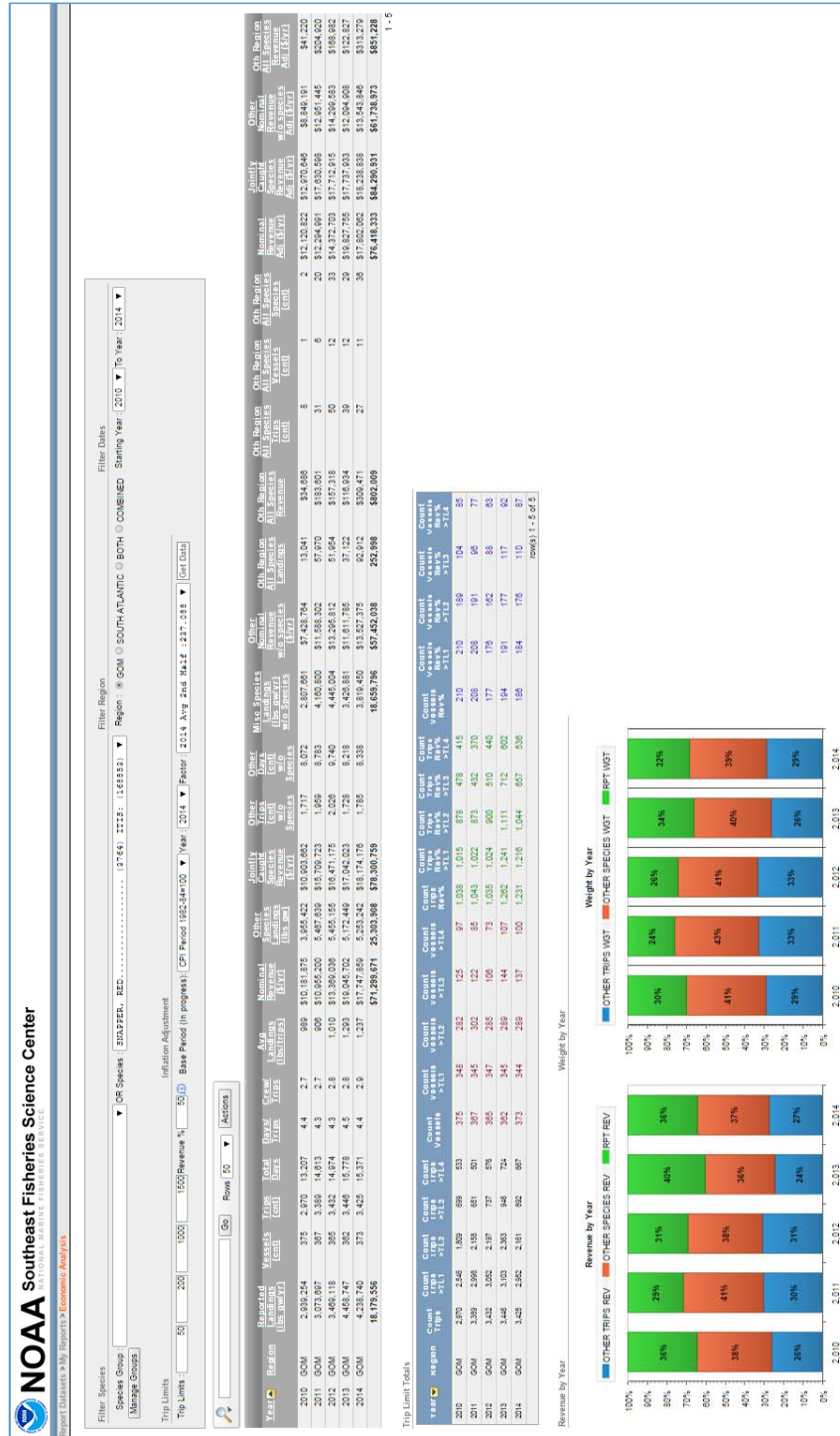


Figure 1: Web-Based Query Tool Online Interface - Economic Analysis

August 2015

Page 8

FY15 Final Report [Draft]

In addition to the parameters and filters described above, running web-based queries using the Descriptive Statistics v.2 interface provide the user with additional data filters such as gear, state, and county. When these additional filters are employed the report interface includes additional page regions (areas) to include detailed metrics information based on the chosen filter/parameters.

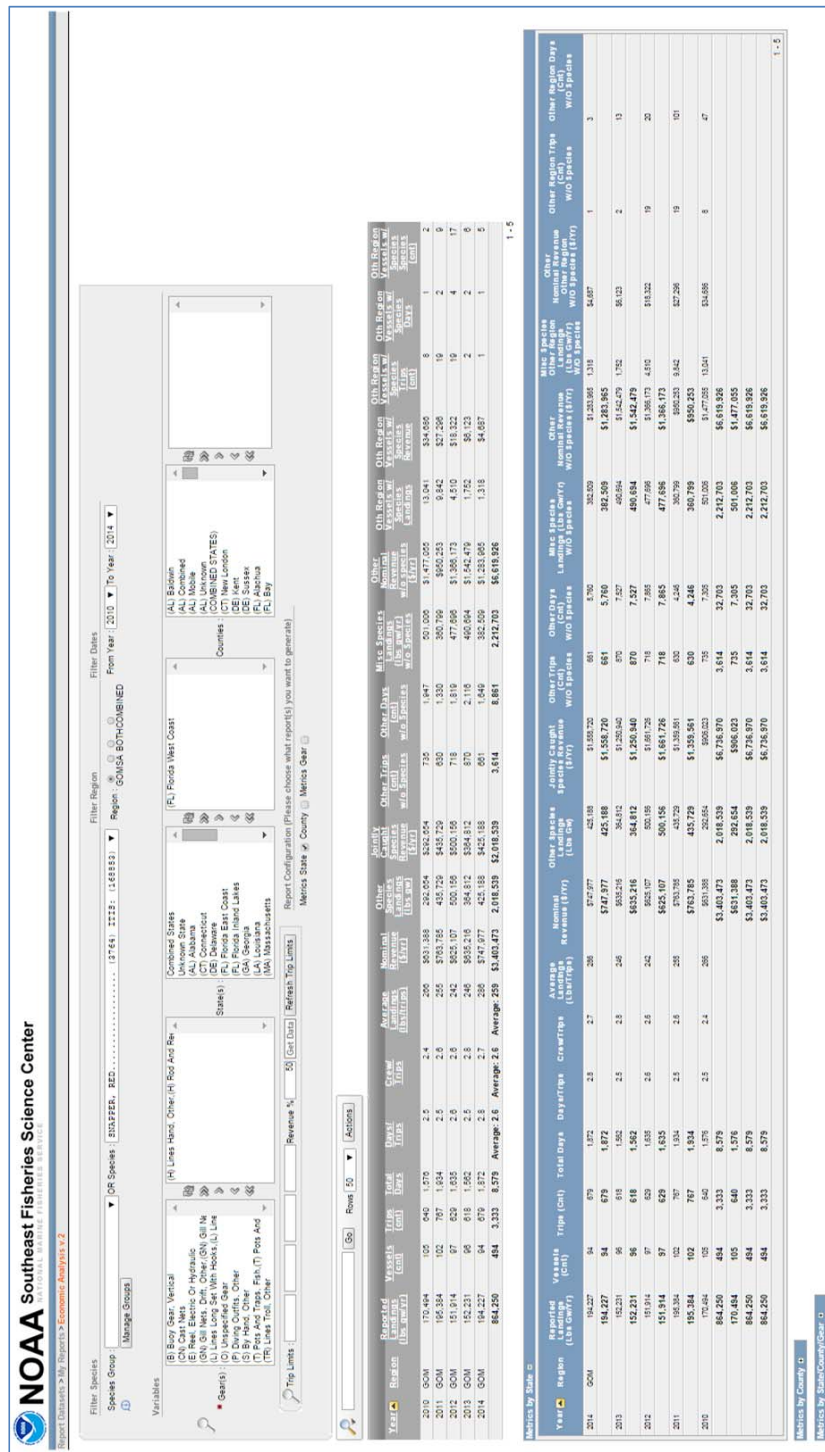


Figure 2: Web-Based Query Tool Online Interface - Descriptive Statistics v.2

Development Approach

The project activities began with a discovery phase which included a definition of the high level project requirements. Initial participants were staff from the Southeast fisheries Science Center (SEFSC) until a working prototype was developed. This approach established a common project foundation from which the detailed reporting functionality would rest.

Project Participants and Stakeholders

The following list contains people who participated in the project as stakeholders, system users, developers, testers, and information sharing through emails and live demonstrations.

Name	Email
Akbar Marvasti	akbar.marvasti@noaa.gov
Assane Diagne	assane.diagne@gulfcouncil.org
Ava Lasseter	ava.lasseter@gulfcouncil.org
Brent Stoffle	brent.stoffle@noaa.gov
Brian Cheuvront	brian.cheuvront@safmc.net
Christina Package-Ward	christina.package-ward@noaa.gov
Christopher Liese	christopher.liese@noaa.gov
David Records	david.records@noaa.gov
David W Carter	david.w.carter@noaa.gov
Denise Johnson	denise.johnson@noaa.gov
Douglas Lipton	douglas.lipton@noaa.gov
Elizabeth Overstreet	elizabeth.overstreet@noaa.gov
George Silva	george.silva@noaa.gov
Guy Dubeck	guy.dubeck@noaa.gov
Jay Oleary	jay.oleary@noaa.gov
Jennifer Cudney	jennifer.cudney@noaa.gov
Joseph Desfosse	joseph.desfosse@noaa.gov
Juan Agar	juan.agar@noaa.gov
Kari Maclauchlin	kari.maclauchlin@safmc.net
Larry Perruso	larry.perruso@noaa.gov
Matthew Mcpherson	matthew.mcpherson@noaa.gov
Michael Jepson	michael.jepson@noaa.gov
Michael Judge	michael.judge@noaa.gov
Mike Travis	mike.travis@noaa.gov
Quigley Kate	quigley.kate@yahoo.com
Randy Blankinship	randy.blankinship@noaa.gov
Rita Curtis	rita.curtis@noaa.gov
Scott Crosson	scott.crosson@noaa.gov
Stephen Holiman	stephen.holiman@noaa.gov
Steve Turner	steve.turner@noaa.gov
Tony Lamberte	tony.lamberte@noaa.gov
Yanet Jimenez	yanet.jimenez@noaa.gov

Load and Prepare Data

The current process for preparing data for the Web –Based Query Tool includes manual steps. The economic data currently resides in SAS on a computer desktop with network access.

A future version of the system proposes to migrate SAS code calculations from the PC to the Oracle database and use ACCSP and GULFFIN accumulated landings (Trip Tickets), and Southeast Logbooks.

Build Initial Foundation Reports (Fully Functioning Prototypes)

The design and prototype (report development) were completed in an iterative process with presentations to users independently and through review meetings. Once the technical foundation and the data were prepared, the development of reports using the online Web-Based Query Tool began.

1. Develop online reports
2. Demo reports and make available for online access
3. Users tested report queries and provided feedback
4. Development team refined, updated, and built new report queries

Building on the iterative process allowed the users to quickly provide feedback throughout the construction process. The use of the iterative approach allowed project team members to see a visual representation of the product at an early stage in a fairly complete form.

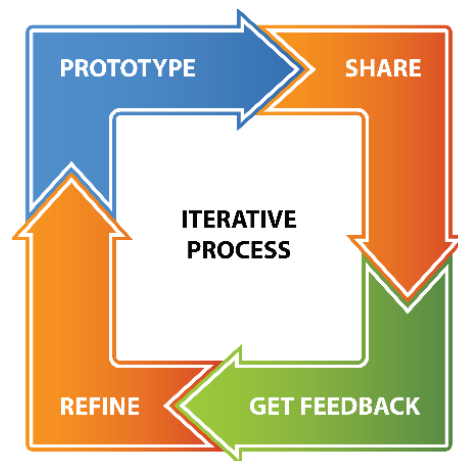


Figure 3: Iterative Development Process

Hold Live-Demos of Web Based Query Tool

A website we prepared and moved to the Internet for users to access the Web-Based Query Tool hosted at SEFSC. Authentication with username and password are required to access the site.

Figure 4: Southeast Fisheries Reporting Logon

Live demonstrations were held for the project team along with one-on-one demos and electronic communications via the website, email, and chat.

Twenty-six users were created under the Southeast Fisheries Reporting web-based query application with economic data access.

Obtain User and Stakeholder Feedback

Throughout the project the development team communicated with the stakeholders and users to obtain feedback, develop the new requested functionality, and present the results on the website for further review.

The following table lists the latest design functionality requested by the users. Some items have been completed and some are in development.

SEFSC Economic Web-query Tool - List of Desired Functionality

Function	Description
Cross region functionality	Ability to capture a vessel's activity in another region. For example, if the query is for vessels that harvest/land king mackerel in the Gulf, their other activity output (trips w/o Gulf king mackerel) will capture fishing/landing that may have exclusively occurred in the South Atlantic. This is particularly relevant for vessels in some fisheries.
Custom species grouping	Allow users to select more than one species and produce estimates for vessels that landed any of them.

Economic Web-Based Query Tool

FY15 Final Report [Draft]

Dealer queries	If we link it to the 'raw' ALS data (vs. what Larry currently works up), then multiple dealer queries should be supported. like find all dealers that purchase a species/group, then show where they are, how much, how much of other species, etc. This would be an expansion of the "number of dealers" request.
Number of dealers	Provide a count of dealers located within different units of geography
Aggregate data at community (city) level	Provide option to select data presented at the community (city) level including a drop down list similar to the county list.
Trip limit correction	It may be more useful to show the total number of vessels that landed over each trip limit, rather than showing the number of vessels in each trip limit range (i.e. those that landed over one limit but not the next). The idea is we want to show what percentage of the fleet will be affected at each trip limit level, so the way it's currently set up we would have to add them together. Granted vessels that are over the limit by a lot would be expected to be more affected by a trip limit change than those that are only over by a little, but I don't think we will typically try and interpret the data at that level, rather we just want to identify those likely affected, and drill down deeper in the RFA as needed, probably with custom SAS code. This would also apply to the Trip limit revenue statistics (i.e. total number of vessels that landed over each trip limit by any amount and that also derived 50% of their revenue from the selected species). Also, as a side note, the description of these variables in the website documentation for the Vessels TL1 through TL5 says these vessels earned a plurality of their revenues from this species, but I think that only applies to the Revenue TL1 through TL5 variables.
Cross gear functionality	Ability to capture a vessel's activity with other gears. For example, if the query is for vessels that harvest/land black sea bass in the South Atlantic with pot gear, their other activity output (trips w/o South Atlantic black sea bass harvested with pots) will capture fishing/landing of any/all other species with any other gears.
Breakdown by gear	Provide option to pull statistics by species and gear type. Gear types could be broad categories (i.e. trawl nets, gill nets, longlines, etc.)
Breakdown by permit type	Retrieve species-specific and/or general revenue statistics for permit types. Note that this should include the identification/inclusion of vessels without a/the federal permit. This will/could be significant with the inclusion of state trip ticket data in the accessible domain of data.
Search by vessel or permit number	Useful to identify total landings of an owner's fleet
Inclusion of US Caribbean query capability	Similar to capabilities for other regions, as modified by data content differences.

Breakdown by revenue category	Pull statistics by total annual vessel-level revenue ranges for selected species (i.e. Total \$ = 0 to 50K, 50K to 100K, 100K to 200K, 200K to 300K, 300K to 500K, 500K to 700K, 700K to 1mill, and 1mill+). This could help get at industry structure. Open to suggestions on ranges.
Breakdown by vessel length category	Pull statistics by vessel length using vessel length ranges (i.e. 0 to 30 ft, 30 to 50 ft, 50 to 75 ft, 75+ feet). Ranges could vary by species groups based on vessel size distribution (i.e. reef fish vessels vs. shrimpers). Open to suggestions on ranges.
Inflation adjustment	Provide option to select and apply a price deflator to revenue estimates for desired dollar year, including Consumer Price Index (CPI), Producer Price Index (PPI), and Gross Domestic Product (GDP) deflators
Reverse query capability	Instead of identifying all vessels that harvest a species/group, query for all vessels that do not.

Web-Based Query Tool Architecture

The Economic Web-Based Query Tool was developed using PL SQL Developer (All Around Automation), Oracle Application Express framework, Programming Language SQL (PL/SQL), and SQL.

Oracle Application Express (Oracle APEX) is Oracle's primary tool for developing Web applications with SQL and PL/SQL. Using primarily web browser technology, the team was able to develop and deploy Web-based applications for desktops and mobile devices. Oracle Application Express is a fully supported, no cost option of the Oracle Database, and is installed by default in all editions of the Oracle Database. Oracle Application Express was the chosen database-centric development tool used to build and publish the Economic Web-based query Tool.

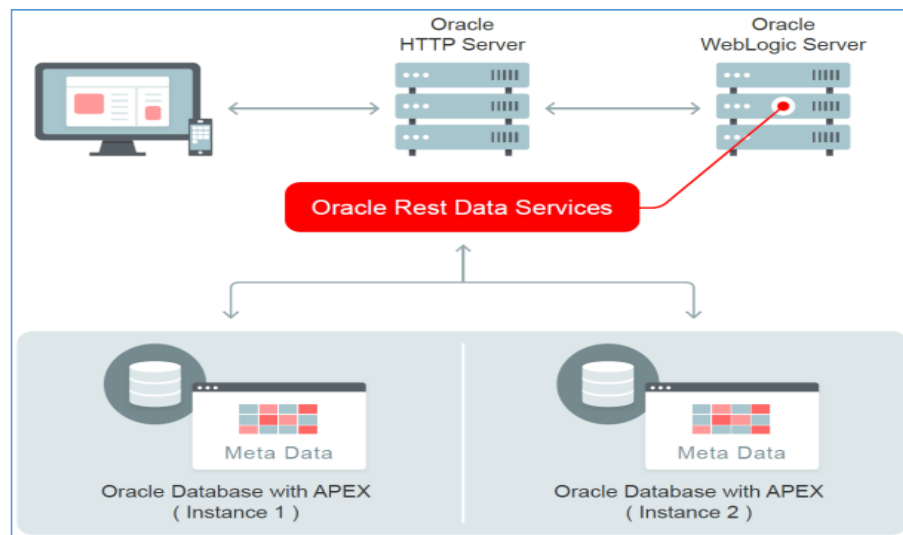


Figure 5: Oracle APEX Technical Architecture

Hardware

The Southeast Fisheries Science Center hosts the Economic Web-Based Query Tool website. The application and database servers use LINUX on Dell Xeon servers. Developers use windows based desktop machines.

Computer Hardware

- Windows 7
- Linux application server
- Linux database server

Software

An Oracle Relational Database Management System (RDBMS) is used to support the Economic Web-Based Query Tool. Along with the database, and typical desktop publishing tools, the development team uses data modeler tools, PL/SQL developer tools, and web browsers.

Software

- RDBMS: Oracle Enterprise Edition 11g
- Oracle PL/SQL, Packages, and Functions
- Toad Data Modeler
- Application Express (Apex) Workspace
- Allround Automation – PL/SQL Developer

Data Sets and User Groups (SFR Access Path)

The Southeast Fisheries Reporting (SFR) website was developed to host the Economic Web-based Query Tool as well as other Southeast Fisheries data sets and reporting interfaces.

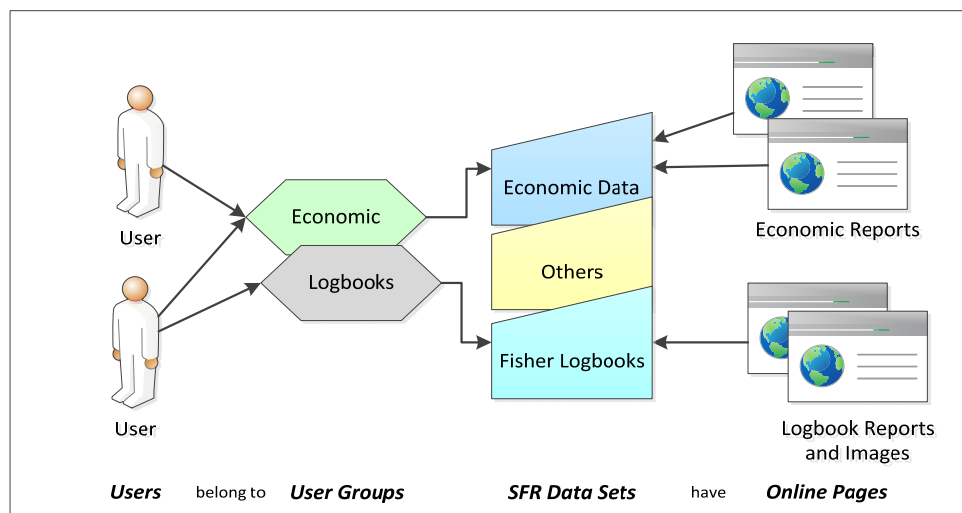


Figure 6: SRF Authorization Access Mapping

Authorization to access economic reports (or other data sets) is accomplished through associating users to users groups, and online interfaces to SFR datasets. SFR supports an unlimited number of users groups and an unlimited number of data sets.

User Groups

Administrators of the SRF have the ability to create new users in the system. However this does not grant access to data in itself. A user must be added to a user group through the online maintenance module to have access to the online interface reports.

SFR Data Sets

User interface pages (online reports or maintenance modules) are associated to SFR Data Sets. A data set can contain an unlimited number of online interface pages.

User Groups to SFR Data Sets

A user is granted access to online interface pages when one or more of their users groups are associated to a data set. When a user is associated to a data set they have the ability to interact with the entire population of interface pages associated with that data set.

User Access Authentication

To access the Economic Web-based query tool, a username and password are required. The online system users page level security/authorization. A user may log into the system and only see the online reports they have been granted access.

If a user attempts to modify the browser URL (Uniform Resource Locator) in an attempt to access an unauthorized page or a user bookmarks a page deep within the website which they no longer have access, they are presented with a access denied message.

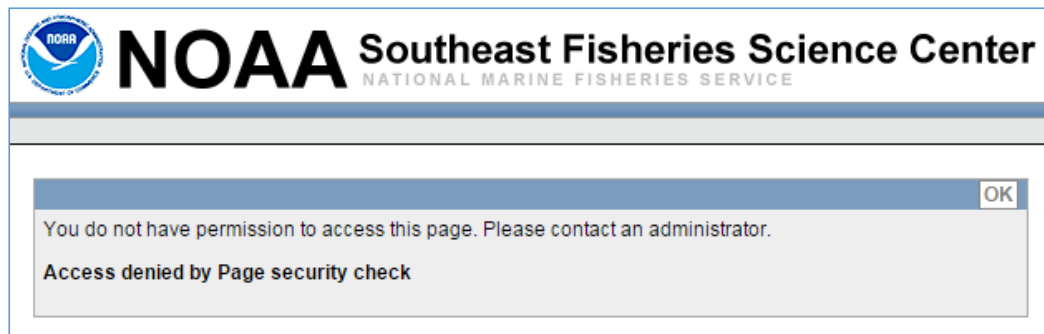


Figure 7: SRF Page Level Security

Web-based Query Tool Users

At the time of this writing, the following people are associated to the user groups and data sets for the Economic Web-Based Query Tool

Akbar Marvasti	Douglas Lipton
Assane Diagne	Jay Oleary
Ava Lasseter	Juan Agar
Brent Stoffle	Kari Maclauchlin
Brian Cheuvront	Larry Perruso
Christina Package-ward	Matthew Mcpherson
Christopher Liese	Michael Jepson
David Carter	Michael Judge
David Records	Mike Travis
Denise Johnson	Quigley Kate

Interface Layout

The Web-based Query Tool and the SRF web-application engine renders user interface pages by combining templates with application components defined by the developer and data in the database.

The Web-based Query Tool overall framework (or structure of a page) is determined by the page template chosen. For example, the page template controls if a page uses tabs and a navigation bar. It can also define if a page includes a bar on the left side that serves as a placeholder for navigation or secondary content. The Web-based Query Tool uses database generated navigation and breadcrumbs. A "breadcrumb trail" is a type of secondary navigation scheme that reveals the user's location in the Web application. (The term comes from the Hansel and Gretel fairy - drop breadcrumbs to form a trail back to home).

The Web-based Query Tool template also includes definitions of region positions, which enable precise control over placement of regions using HTML tables or style sheet definitions. The page template itself is composed of HTML combined with substitution strings, which are substituted with the appropriate components at run time. The team added data and content to the user interface pages by creating regions

Regions

A region is an area on a page that serves as a container for content, such as report data or data entry parameters. The Web-based Query Tool is made of user interface pages with many regions.

The development team created page regions to group page controls (such as items or buttons) as well as elaborate regions that frame content within HTML tables or images. Some regions display standard report data and some regions are interactive.

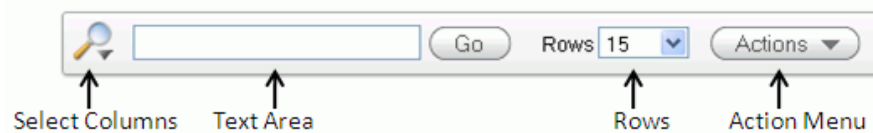


Figure 8: report Customization Tools

Interactive report regions were used by the development team to define queries that can be highly customized by end-users. The economic users can customize the layout of the data by choosing the report

"project narrative with figures and appendix", page 36

Economic Web-Based Query Tool

FY15 Final Report [Draft]

columns they wish to see, apply filters, highlight data, and sorting. They can also define breaks, aggregations, computed columns and include a chart of the query results. Users can create multiple variations of the report and save them as named reports, output to comma-delimited files, and print to PDF documents.

Query Parameters/Filters

The Economic Web-based Query Tool requires user defined filters and parameters. Users are initially presented with form fields which are used to filter the report data. There are three primary filters used by the web-query tool: Species, Region, and Date Range.

User Customizations

The interactive report region controls allow users to highly customize query result reports.

Query Results

The results of the query are presented in this page region based on user defined filters and parameters.

Trip Limit Totals

The Trip Limit Totals region of the page show what percentage of the fleet will be affected at each trip limit level defined in the query parameter.

Graphs

Two graphs are presented in this page region; one for Revenue by Year and one for Weight by Year.

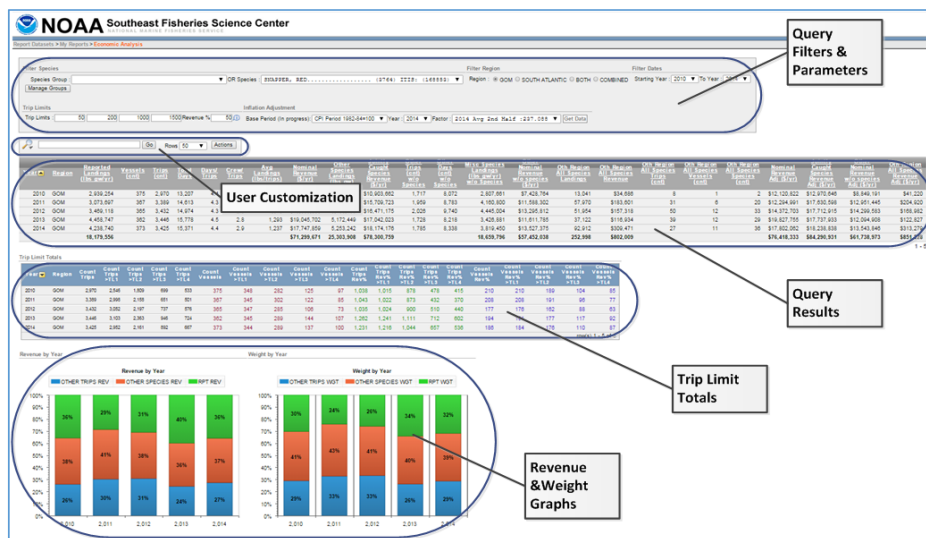


Figure 9: Web-Based Query Tool Interface Region Layout

Through the web-base queries of the Descriptive Statistics v.2 interface users have the option to include additional parameter filters for gears, states, and counties. When a user chooses one or more of these additional filters the associated Detailed [State] Metrics page region is populated with detailed data base on the users choice.

Report Configuration (Please choose what report(s) you want to generate)

Metrics State ☒ County ☐ Metrics Gear ☐

[illegible]

Data Sources

The Southeast Fisheries Science Center Data Warehouse hosts several data sources from different Southeast data collection programs. The southeast database contain: logbook data, trip ticket data, observer data, Port agent interview data, and others.

SAS Dataset (provided by Larry Perruso)

Southeast Fisheries Logbook data and Accumulated Landings data are imported into SAS programs running on a desktop computer at the Southeast Fisheries Science Center (SEFSC). SAS computer programs merge data from these two data sources to prepare a dataset of vessel information, species information, landed weights, and revenue, and more.

SEFSC SAS Dataset Attributes (abbreviated)

Fields(s)	Description
VESID	Fishing vessel identifier
SCHEDULE	Fisher's Logbook identifier
LANDED, YEAR, MONTH, DAY	Date of landing
STATE, COUNTY	Location of species landing
DEALER	Dealer identifier
TOPGEAR, MULTGEAR, TOPNUMGEAR, TOPEFFORT	Gear information
LENGTH, YEAR_BUILT, HORSEPOWER, GROSS_TONS, NET_TONS, HOLD_CAPACITY, FUEL_CAPACITY, FUEL_TYPE, HULL_TYPE, REFRIGERATION, HOME_PORT_CITY, HOME_PORT_STATE	Vessel attributes
DEPTH_NMFS	Depth information by species (NMFS code)
LBS_NMFS	Weight information by species (NMFS code)
REV_NMFS	Revenue information by species (NMFS code)

The SAS dataset is based on data from the Southeast Fisheries Logbook program which is merged with Accumulated Landings data to calculate price values by year, month, state, and county.

Southeast Logbooks

The Southeast Logbook data is the foundation data sources for the Web-based Query Tool. The Southeast Logbook program collects commercial fishing trip information from captains on coastal and highly migratory species (HMS) fishing trips.

Data Warehouses – Accumulated Landings using Trip Ticket data

The Southeast Fisheries Science Center imports dealer reports data (trip ticket) for the ACCSP (Atlantic) and GULFFIN (Gulf of Mexico) data warehouse.

ACCSP – Trip Tickets

The Atlantic Coastal Cooperative Statistics Program (ACCSP) is a cooperative state-federal program that designs, implements, and conducts marine fisheries statistics data collection programs and integrates those data into a single data management system.

The ACCSP maintains accumulated landings fishing trip data for the Atlantic region. The Southeast Fisheries Science Center maintains a copy of the state approved landings data from the ACCSP Oracle database.

GULFFIN – Trip Tickets

The Gulf States Marine Fisheries Commission (GSMFC) is an organization of the five states (Texas, Louisiana, Mississippi, Alabama, and Florida [west coast]), whose coastal waters are the Gulf of Mexico. The Fisheries Information Network (FIN) is a state-federal cooperative program among agencies to collect, manage and disseminate statistical data and information on the commercial fisheries of the Southeast Region.

The GULFFIN is a source of accumulated landings fishing trip data for the Gulf of Mexico. The Southeast Fisheries Science Center maintains a copy of the state approved landings data from the GULFFIN Oracle database.

A Future version of the Web-based Query Tool proposes to use ACCSP and GULFFIN data, which is house at SEFSC.

Automated Data Import

An unattended scheduled nightly process performs an automated data import from the ACCSP and GULFFIN data warehouses.

Proposed Future Phase

- Migrate SAS Calculations to Southeast Data Warehouse
- Replace SAS Data Source with Southeast Data Warehouse (ACCSP&GULFFIN)

Data Preparation

The current data preparation process contains the following steps:

1. Receive SAS Results
 - a. Accumulated Landings data is opened for SAS processing on the desktop
 - b. SAS code is used to calculate economic pricing and costs
 - c. SAS prepared data are exported to a delimited file
2. Upload CSV file to temporary tables in Data Warehouse Staging area
 - a. Delimited landings data file is uploaded to an Oracle database table
 - i. ECON_TRIP_COST
 - b. Database process transpose and normalizes economic data into trip and catch attributes
 - i. ECON_TRIP (report table)
 - ii. ECON_TRIP_CATCH (report table)
 - c. Validation is performed on the file (numbers are numbers, dates are dates, etc.)
 - d. Data is linked to existing lookup values in the SEFSC Data Warehouse
 - i. Link economic data with species dimensions for ITIS code mapping
3. Upload delimited data for inflation adjustments
 - a. Consumer Price Index (CPI)*
4. User defined species groups
 - a. Users maintain specific species groups for the Web-based Query Tool
5. Data is made available for the Web-Based Query Tool

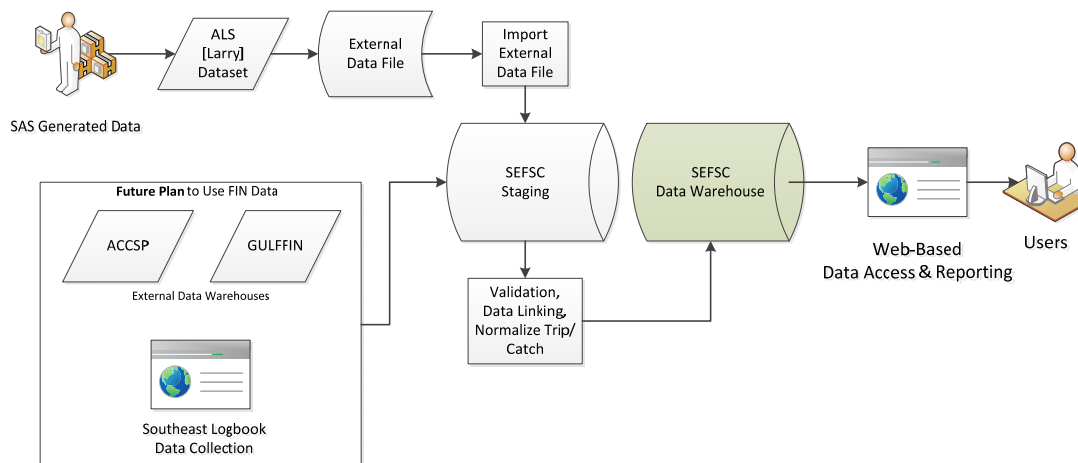


Figure 12: Manual and Future Data Sources Processing

* The inflation adjustments of Consumer Price Index (CPI) are obtained from the Bureau of Labor Statistics website. <http://www.bls.gov/cpi/data.htm>

"project narrative with figures and appendix", page 41

Inflation Adjustments - Consumer Price Index

Data for the CPI or other inflationary adjustments can be obtained via the Bureau of Labor Statistics using the following steps:

Define BASE

Indicates the designated reference date from which price change is measured, where the "current" base year is 1982-84=100 or more recent (S) and the "alternate" base year (A) is prior to the current base year (kept for contract purposes).

Choose CPI Database

CPI Databases

Databases

Database Name	Special Notice	Top Picks	One Screen	Multi-Screen	Tables	Text Files
All Urban Consumers (Current Series) (Consumer Price Index - CPI)		 TOP PICKS	 ONE-SCREEN DATA SEARCH	 MULTI-SCREEN DATA SEARCH	 TABLES	 TEXT FILES
Urban Wage Earners and Clerical Workers (Current Series) (Consumer Price Index - CPI)		 TOP PICKS	 ONE-SCREEN DATA SEARCH	 MULTI-SCREEN DATA SEARCH	 TABLES	 TEXT FILES
All Urban Consumers (Chained CPI) (Consumer Price Index - CPI)	 SPECIAL NOTICE	 TOP PICKS	 ONE-SCREEN DATA SEARCH	 MULTI-SCREEN DATA SEARCH	 TABLES	 TEXT FILES
Average Price Data (Consumer Price Index - CPI)	 SPECIAL NOTICE	 TOP PICKS	 ONE-SCREEN DATA SEARCH	 MULTI-SCREEN DATA SEARCH	 TABLES	 TEXT FILES

Choose Seasonally Adjusted or Not

Consumer Price Index-All Urban Consumers -- Seasonal (Screen 1 of 6)

You have a total of **5372** series to query for.

Choose: ☐ Seasonally Adjusted ☒ Not Seasonally Adjusted

Choose Area

Consumer Price Index-All Urban Consumers -- Area (Screen 2 of 6)

Your query has been narrowed to **5060** series.

Choose Area:

0000 U.S. city average
0100 Northeast urban
0200 Midwest urban
0300 South urban
0400 West urban
A000 Size Class A (more than 1,500,000)

Choose Alternate or Current Base

Consumer Price Index-All Urban Consumers -- Base (Screen 3 of 6)

Your query has been narrowed to 494 series.

Choose Base: ☐ Alternate ☒ Current

Create Customized Table

Consumer Price Index-All Urban Consumers -- Item (Screen 4 of 6)

Your query has been narrowed to 491 series.

Choose Item:

SA0 All items

SA0E Energy

SA0L1 All items less food

SA0L12 All items less food and shelter

SA0L12E All items less food, shelter, and energy

SA0L12E4 All items less food, shelter, energy, and used cars and trucks

Define Periodicity

Consumer Price Index-All Urban Consumers -- Periodicity (Screen 5 of 6)

Your query has been narrowed to 2 series.

Choose Periodicity: ☒ Monthly ☒ Semi-Annual

Choose Year

Consumer Price Index-All Urban Consumers -- Year (Screen 6 of 6)

Download CPI Data (Note file names)

Series Id: CUUR0000SA0
Not Seasonally Adjusted
Area: U.S. city average
Item: All items
Base Period: 1982-84=100

Web-Based Query Economic Analysis Reports

The Web-Based Query Tool is hosted at the Southeast Fisheries Science Center on their Southeast Fisheries Reports website: <http://www.sefsc.noaa.gov/apex/f?p=110>

Query Parameters/Filters

Uses of the Economic Web-based Query Tool are first presented with form fields used to filter the report data. There are three primary filters used by the web-query tool.

Filter Species Group

The species group filter lists user defined species groups. If a group does not exist the user may create a new one. The report data is filtered by the chosen species group.

Filter Species

The species filter lists the available species including the species name, the National Marine fisheries Service - NMFS code, and the Integrated Taxonomic Information System - ITIS code (linked during data loading). The report data is filtered by the chosen species. Either a species or a species group may be chosen.

Filter Region

The region filter allows the users to choose either the Gulf of Mexico (GOM), the South Atlantic, Both, or Combined. The chosen region will include species which were landed in that region. When Both is chosen, the report data will include a separate report line for each of the two regions. When Combined is chosen, the report will combine the data into a single total.

Filter Dates

The date filter allows the user to provide a date range where landings data are included if the landing date is between the start date and an end date.

Trip Limits

Trip Limit parameters are a useful way to show the total number of vessels that landed over each trip limit starting from zero. Initially developed to show the number of vessels in each trip limit range (i.e. those that landed over one limit but not the next) during demonstrations and user feedback the Trip Limit was changed to be encompassing. The idea is to show what percentage of the fleet will be affected at each trip limit level. Trip Limits help to identify those vessels likely affected, with an eventual drill down to allow deeper review in the RFA as needed. The Trip Limit also applies to the Trip limit revenue statistics (i.e. total number of vessels that landed over each trip limit by any amount and that also derived 50% of their revenue from the selected species).

Revenue

The Revenue Percentage value will filter the report results to only include those vessels which received at least the percentage of revenue specified from their total revenue while catching the species of interest.

Inflation Adjustment

The Inflation Adjustments provide an option for users to select and apply a price inflator or deflator to revenue estimates for the desired dollar year such as Consumer Price Index (CPI).

Future versions will allow for additional adjustments such as Producer Price Index (PPI), and Gross Domestic Product (GDP) deflators, etc.

Filter Variable - Gear

The Filter Variable of Gear(s) allows the users to select one or more gears to filter the report data. When the data is filtered by gear, in addition to the yearly query results, a new page region is populated showing the detailed metrics for the chosen gear(s).

Filter Variable - State

The Filter Variable of State(s) allows the users to select one or more states to filter the report data. If a state is chosen it is assumed that all counties in the state are included unless specifically selected by the user. When the data is filtered by state, in addition to the yearly query results, a new page region is populated showing the detailed metrics for the chosen state(s).

Filter Variable - Counties

The Filter Variable of counties allows the users to select one or more counties to filter the report data. If a state is chosen it is assumed that all counties in the state are included unless specifically selected by the user. The user may choose a specific county in a state and/or choose an additional county to be added to the counties of a chosen state. When the data is filtered by states and counties, in addition to the yearly query results, a new report page region is populated showing the detailed metrics for the chosen states and or counties.

The date filter allows the user to provide a date range where landings data are included if the landing date is between the start date and an end date.

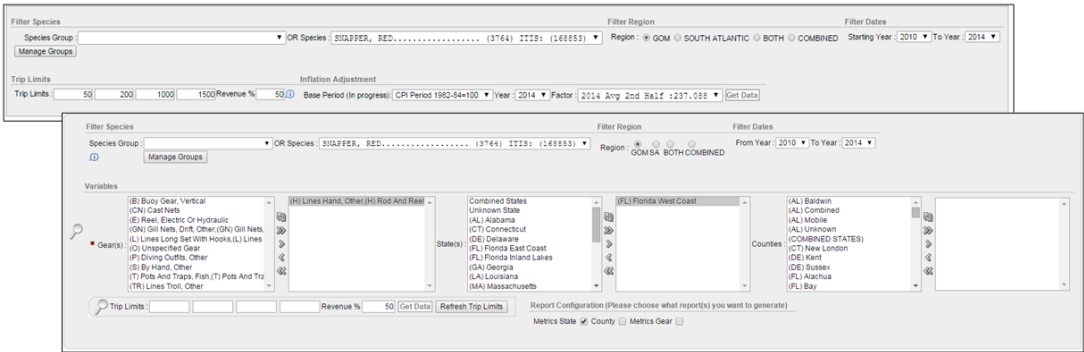


Figure 13: Web-Based Query Tool Filters and Parameters

Query Results

The Web-Base Query Tool presents aggregated results by year and region. The comprehensive report includes a customization toolbar allowing users to choose their columns of interest in the order they wish.

"project narrative with figures and appendix", page 45

Economic Web-Based Query Tool

FY15 Final Report [Draft]

Year	Region	Reported Landings (lbs gw/yr)	Vessels (cnt)	Trips (cnt)	Total Days	Crew/Trips	Avg Landings (lbs/trip)	Nominal Revenue (\$/yr)	Other Species Landings (lbs gw)	Other Jointly Caught Species Revenue (\$/yr)	Other Trips (cnt) w/o Species	Other Species Revenue (\$/yr)	Other Jointly Caught Species Revenue (\$/yr)	Other Trips (cnt) w/o Species										
2010	GOV	2,839,254	375	2,870	13,207	4.4	2.7	\$89	\$10,181,875	\$395,422	\$10,363,862	1,717	\$,072	2,807,669	\$7,429,704	13,041	\$24,088	8	1	2	\$12,120,822	\$12,970,848	\$8,948,151	\$41,220
2011	GOV	3,271,887	387	3,380	14,813	4.3	2.7	\$58	\$10,365,200	\$,487,038	\$10,756,723	1,859	\$,783	4,160,885	\$11,586,382	\$7,870	\$18,851	31	8	20	\$12,284,881	\$17,850,588	\$12,281,445	\$254,835
2012	GOV	3,488,118	385	3,432	14,974	4.3	2.8	1,010	\$13,369,038	\$,485,193	\$14,474,175	2,026	\$,740	4,445,004	\$13,298,812	\$1,854	\$187,318	50	12	33	\$14,372,703	\$17,712,815	\$14,298,583	\$168,862
2013	GOV	4,455,747	382	3,448	15,778	4.3	2.8	1,200	\$19,045,702	\$,172,449	\$17,342,023	1,728	\$,216	3,428,881	\$11,811,788	\$7,122	\$16,834	38	12	26	\$18,827,785	\$17,737,853	\$12,094,808	\$122,827
2014	GOV	4,236,740	373	3,425	16,375	4.4	2.8	1,237	\$17,747,888	\$,283,242	\$16,174,176	1,785	\$,338	3,818,460	\$15,527,376	\$2,882	\$308,471	27	11	16	\$17,802,862	\$18,238,838	\$15,543,848	\$191,278
		18,179,656						\$71,279,671	\$5,363,888	\$76,396,739				18,859,796	\$57,452,838	\$52,888	\$902,099				\$76,418,333	\$84,296,931	\$65,738,973	\$881,288

Figure 14: Web-Based Query Tool Report Results

Query Result Columns and Descriptions

Column Name	Description
Year	From the year(s) chosen in the filters
Region	The region(s) chosen from the filter
Reported Landings (lbs gw/yr)	Trips having landed 1 lbs or more of the chosen species, year, region combination
Vessels (cnt)	The Vessels (cnt) is the count of vessels which landed 1 lbs or more of the chosen species, year, region combination
Trips (cnt)	Count of fishing trips having landed 1 lbs or more of the chosen species, year, region combination
Total Days	Sum of all Days Away for trips having landed 1 lbs or more of the chosen species, year, region combination
Days/Trips	Days/Trips is the average number of Days per trip which landed 1 lbs or more of the chosen species, year, region combination
Crew/Trips	Crew/Trips is the average number of Crew per trip which landed 1 lbs or more of the chosen species, year, region combination
Average Landings (lbs/trips)	Average Landings is the average reported weight from fishing trips having landed 1 lbs or more of the chosen species, year, region combination
Nominal Revenue (\$/yr)	Nominal Revenue is the dockside revenue (ex-vessel price) from fishing trips having landed 1 lbs or more of the chosen species, year, region combination
Other Species Landings (lbs gw)	Other Species Landings is the other jointly caught species on a per trip bases. Other Species Landings jointly caught from the same fishing trips having landed 1 lbs or more of the chosen species, year, region combination
Other Jointly Caught Species Revenue (\$/yr)	Other Jointly Caught Species Revenue is the dockside revenue (ex-vessel price) from the other species from the same fishing trips having landed 1 lbs or more of the chosen species, year, region combination
Other Trips (cnt) w/o Species	Other fishing trips, for the same vessels, where trips did not catch the species of interest. Represents counts of the number of trips taken by vessels that caught at least 1 lbs of the named species but did not land the named species on that trip. The values are from vessel participating in the fishery but not catch the named species for these trips.

Other Days (cnt) w/o Species	Sum of Days Away, for the same vessels, where trips did not catch the species of interest. Represents counts of the number of trips taken by vessels that caught at least 1 lbs of the named species but did not land the named species on that trip. The values are from vessel participating in the fishery but not catch the named species for these trips.
Misc Species Landings (lbs gw/yr) w/o Species	Reported weight of other species for the same vessels where trips did not catch the species of interest. Represents counts of the number of trips taken by vessels that caught at least 1 lbs of the named species but did not land the named species on that trip. The values are from vessel participating in the fishery but not catch the named species for these trips.
Other Nominal Revenue w/o species (\$/yr)	Reported revenue of Other Species, for the same fishing trips, where trips did not catch the species of interest. Represents counts of the number of trips taken by vessels that caught at least 1 lbs of the named species but did not land the named species on that trip. The values are from vessel participating in the fishery but not catch the named species for these trips.
Other Region All Species Landings	Represents all species landings in the other region for the vessels which caught at least 1 lbs of the named species in the chosen region.
Other Region All Species Revenue	Represents all species revenue in the other region for the vessels which caught at least 1 lbs of the named species in the chosen region.
Other Region All Species Trips (cnt)	Represents the count of all trip in the other region for the vessels which caught at least 1 lbs of the named species in the chosen region.
Other Region All Species Vessels (cnt)	Represents the count of vessels with fishing activity in the other region for the vessels which caught at least 1 lbs of the named species in the chosen region.
Other Region All Species Species (cnt)	Represents the count of all species landed in the other region for the vessels which caught at least 1 lbs of the named species in the chosen region.
Nominal Revenue Adjustment (\$/yr)	Represents the Nominal Revenue adjusted by the chosen inflation adjustment. See Nominal Revenue.
Jointly Caught Species Revenue Adjustment (\$/yr)	Represents the Jointly Caught Species Revenue adjusted by the chosen inflation adjustment. See Jointly Caught Species Revenue.
Other Nominal Revenue w/o species Adjustment (\$/yr)	Represents the Other Nominal Revenue w/o species adjusted by the chosen inflation adjustment. See Other Nominal Revenue w/o species.
Other Region All Species Revenue Adjustment (\$/yr)	Represents the Other Region All Species Revenue adjusted by the chosen inflation adjustment. See Other Region All Species Revenue.

Trip Limit Results

The Trip Limit results show what percentage of the fleet was affected at each trip limit level. The Trip limit revenue statistics is the total number of vessels that landed over each trip limit by any amount and that also derived 50% of their revenue from the selected species.

Column Name	Description
Year	From year(s) chosen in filter criteria
Region	The region(s) chosen from the filter
Count Trips	Count of fishing trips that had at least 50% of their revenue from the named species.
Count Trips > TL1 Count Trips > TL2 Count Trips > TL3 Count Trips > TL4	Count of fishing trips that had the named species of their trip landings fell below the trip limit threshold.
Count Vessels	Count of vessels that had at least 50% of their revenue from the named species.
Count Vessels > TL1 Count Vessels > TL1 Count Vessels > TL1 Count Vessels > TL1	Count of vessels that had the named species of their trip landings fell below the trip limit threshold.
Count Trips Revenue %	Count of fishing trips that had revenue equal or greater than the user provided revenue % from the named species.
Count Trips Revenue % > TL1 Count Trips Revenue % > TL2 Count Trips Revenue % > TL3 Count Trips Revenue % > TL4	Count of fishing trips that had revenue equal or greater than the "user provided revenue % parameter" from the named species of their trip landings fell below the trip limit threshold.
Count Vessels Revenue %	Count of vessels that had revenue equal or greater than the user provided revenue % from the named species.
Count Vessels Revenue % > TL1 Count Vessels Revenue % > TL2 Count Vessels Revenue % > TL3 Count Vessels Revenue % > TL4	Count of vessels that had revenue equal or greater than the user provided revenue % from the named species of their trip landings fell below the trip limit threshold.

"project narrative with figures and appendix", page 48

Graphs

The Graphs displayed by the Web-Base Query Tool represent revenue and weight by years. The Revenue by Weight includes other trips revenue, other species revenue, and reported revenue in a bar chart for each year chosen in the parameter filters. The Weight by Year shows similar information for landed weight amounts.

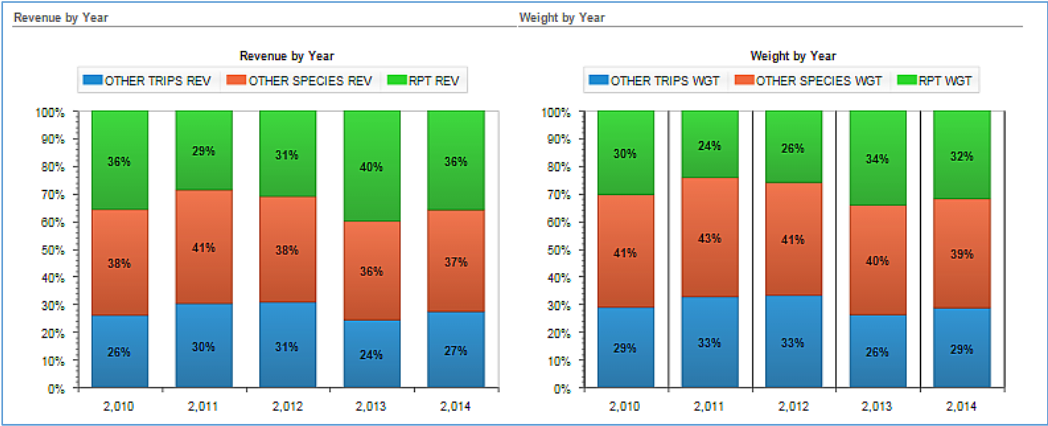


Figure 15: Revenue by Year and Weight by Year Graphs

Appendix

New Desired Functionality

On May 12, 2015 the development team conducted a live demonstration of the Web-based Query Tool. After the demo the participants held brainstorming sessions and put together the following list of new desired functionality. Some items have been completed and delivered while other may be pending or require historic data maintenance (cleanup)

#	Function	Description	Status
1	Cross region functionality	Ability to capture a vessel's activity in another region. For example, if the query is for vessels that harvest/land king mackerel in the Gulf, their other activity output (trips w/o Gulf king mackerel) will capture fishing/landing that may have exclusively occurred in the South Atlantic. This is particularly relevant for vessels in some fisheries.	Done
2	Custom species grouping	Allow users to select more than one species and produce estimates for vessels that landed any of them.	Done
3	Dealer queries	If we link it to the 'raw' ALS data (vs. what Larry currently works up), then multiple dealer queries should be supported. like find all dealers that purchase a species/group, then show where they are, how much, how much of other species, etc. This would be an expansion of the "number of dealers" request.	In Process [Requires historic data cleanup of dealer entities]
4	Number of dealers	Provide a count of dealers located within different units of geography	In Process
5	Aggregate data at community (city) level	Provide option to select data presented at the community (city) level including a drop down list similar to the county list.	Pending

6	Trip limit modification (change to existing calculation)	It may be more useful to show the total number of vessels that landed over each trip limit, rather than showing the number of vessels in each trip limit range (i.e. those that landed over one limit but not the next). The idea is we want to show what percentage of the fleet will be affected at each trip limit level, so the way it's currently set up we would have to add them together. Granted vessels that are over the limit by a lot would be expected to be more affected by a trip limit change than those that are only over by a little, but I don't think we will typically try and interpret the data at that level, rather we just want to identify those likely affected, and drill down deeper in the RFA as needed, probably with custom SAS code. This would also apply to the Trip limit revenue statistics (i.e. total number of vessels that landed over each trip limit by any amount and that also derived 50% of their revenue from the selected species). Also, as a side note, the description of these variables in the website documentation for the Vessels TL1 through TL5 says these vessels earned a plurality of their revenues from this species, but I think that only applies to the Revenue TL1 through TL5 variables.	Done
7	Cross gear functionality	Ability to capture a vessel's activity with other gears. For example, if the query is for vessels that harvest/land black sea bass in the South Atlantic with pot gear, their other activity output (trips w/o South Atlantic black sea bass harvested with pots) will capture fishing/landing of any/all other species with any other gears.	Pending
8	Breakdown by gear	Provide option to pull statistics by species and gear type. Gear types could be broad categories (i.e. trawl nets, gill nets, longlines, etc.)	Done
9	Breakdown by permit type	Retrieve species-specific and/or general revenue statistics for permit types. Note that this should include the identification/inclusion of vessels without a/the federal permit. This will/could be significant with the inclusion of state trip ticket data in the accessible domain of data.	Pending
10	Search by vessel or permit number	Useful to identify total landings of an owner's fleet	Pending
11	Inclusion of US Caribbean query capability	Similar to capabilities for other regions, as modified by data content differences.	Pending

"project narrative with figures and appendix", page 51

Economic Web-Based Query Tool

FY15 Final Report [Draft]

12	Breakdown by revenue category	Pull statistics by total annual vessel-level revenue ranges for selected species (i.e. Total \$ = 0 to 50K, 50K to 100K, 100K to 200K, 200K to 300K, 300K to 500K, 500K to 700K, 700K to 1mill, and 1mill+). This could help get at industry structure. Open to suggestions on ranges.	Pending
13	Breakdown by vessel length category	Pull statistics by vessel length using vessel length ranges (i.e. 0 to 30 ft, 30 to 50 ft, 50 to 75 ft, 75+ feet). Ranges could vary by species groups based on vessel size distribution (i.e. reef fish vessels vs. shrimpers). Open to suggestions on ranges.	Pending
14	Inflation adjustment	Provide option to select and apply a price deflator to revenue estimates for desired dollar year, including Consumer Price Index (CPI), Producer Price Index (PPI), and Gross Domestic Product (GDP) deflators	Done [Being reviewed by users]
15	Reverse query capability	Instead of identifying all vessels that harvest a species/group, query for all vessels that do not.	Pending