

August 10, 2001

\*\*\* *STI-SEFSC DRAFT 4* \*\*\*

## **ACCSP Data Reconciliation Project Requirement Specification Document**

### **1. General Information**

#### *Introduction*

This project is a data reconciliation programming project, funded through the Atlantic Coastal Cooperative Statistics Program (ACCSP). The application will compare data stored in the NMFS Southeast Science Center (SEFSC) to dealer trip data stored in the NMFS Northeast Science Center (NEFSC), and also to ACCSP tables from Florida and North Carolina. These procedures will be referred to as *SEFSC-external comparisons* in contrast to SEFSC-internal comparisons and/or links between the SEFSC dealer weighout (DLS) tables and pelagic logbook tables (PLL). The SEFSC HMS data is captured from logbooks, weighout sheets, and trip summaries. Trips are also randomly selected for observer coverage, but that data is not involved in this project. The customers are NMFS Northeast and Southeast Science Center analysts and database administrators as well as State database managers and State and NMFS fishery managers.

#### *Purpose & Scope*

The purpose of this project is to ensure that *vessel specific trip data* is accurate and complete (comprehensive) between State agencies, NMFS offices, and ACCSP. The ACCSP data warehouse captures data from State and Federal reporting systems. In the southeast, State partners are the primary source for trip level data that will be submitted to the ACCSP. In the northeast, NMFS submits trip landing data to the ACCSP for Federally regulated fisheries, while States will provide trip specific landings for State managed fisheries. The existence of redundant data collection programs requires the ability to compare and cross-check trip data so that duplicate records are not included in data warehouses. The reconciliation application (SEFSC-external comparisons) shall perform the following data comparisons:

- 1> NMFS Northeast data to NMFS Southeast data,
- 2> North Carolina State data in ACCSP to NMFS Southeast data, and
- 3> Florida State data in ACCSP to NMFS Southeast data.

The data comparisons and requirements described in this document will initially focus on trip level data for highly migratory species, primarily swordfish, tuna, wahoo and shark, recorded in the SEFSC DLS

system. A follow-up analysis comparing NMFS snapper/grouper logbook data with State trip ticket information will also be started with the limited funds provided as a second part of this ACCSP project. (*Refer to table 2 for a list of comparison species*)

The objective of this reconciliation project is to identify records that may be duplicated between State and Federal reporting systems, to identify trips reported in one system but missing from the backup system and to identify matched trips that have discrepancies in key variables such as species specific landed weight, State of landing, and/or gear. The system will support a reconciliation table as well as a meta data table and standardized reports which will be provided to the database administrators of each of the comparison files. Trips are defined as a vessel, landing date, and landing State combination.

The system will help identify when a fisher lands in North Carolina and sells part of his catch to a NC dealer and sells the second part of his catch to a South Carolina dealer (*split trip*). The dealer in NC may report the fishers entire catch, include the catch sold in SC, while at the same time the SC dealer reports his purchase under Federal reporting requirements. In order to capture these discrepancies and to provide tools to support investigating and reconciling trip records in the comparative databases, the system will need to include trip based screens for searching records sorted by different key variables as well as the backward linkages to the original comparative data sets (SEFSC, NEFSC, ACCSP-NC, ACCSP-FL). The backward linkages were originally thought to be appropriate trip numbers either explicitly coded in the comparison databases or based on a concatenation of variables. While the SEDLS does have a trip number variable, for the NEFSC ST1 staff thought that we could concatenate year, month, and document number. However, document number appears to be dealer-vessel trip specific rather than vessel and date specific, so split trips result in two different doc numbers for the same vessel-date unloading. Backward linkages may end up having to rely on the vessel, landing date, landing State combination.

## 2. System Operating Environment

### *Hardware and Software*

The application will be developed in a SUN environment. The web application server is a SUN 450, running Oracle Application Server version 4.0.8.1. The database server is a SUN machine, running Oracle enterprise database version 8i. The program will be a web-based application developed in Oracle forms and reports, version 6i for the web.

In production, the application will run on a SUN machine operating Oracle Application Server 4.0.7. This web server will reside in the NMFS Southeast Science Center. **The reconciled database will remain on the SUN machine, located in the ST/1 data center.** The reconciled database will have links to local databases residing at the Southeast Science Center and the Northeast Science Center.

A diagram showing the location of servers and data is attached as Appendix A.

### **3. Vessel -Trip Based Reporting Requirements**

Vessel owners and operators with NMFS HMS longline permits for swordfish, tuna, and sharks record daily logbook information on set catch (numbers) forms, trip summary records on a second form, and voluntary economic information on a third form. After receiving a copy of the unloading or weighout sheet, which lists the individual dressed weights of the landed catch, from the dealer, all forms are mailed to the NMFS Southeast Fisheries Science Center. This data has to be reported to NMFS within 7 days of unloading their catch.

Other NMFS Southeast and Northeast permits have similar generic reporting requirements although the logbooks are different and specific to each permitted fishery. Northeast vessels primarily use Vessel Trip Report (VTR) forms, while many Southeast vessels report through Snapper-Grouper-Mackerel logbooks. VTR and Snapper-Grouper forms are handled by separate systems but both involve image scanning and data are entered into Oracle databases for compliance checks and subsequent tracking and verification of the landings records.

***The following description of the trip data collection processes in the NE and SE are provided so that system designers and developers understand the processes that result in redundant trip data requiring reconciliation.***

#### ***NMFS NE HMS Processes***

Northeast (Va. - Me.) seafood dealers with NMFS permits record their purchases either on weighout sheets or trip landing reports. These seafood dealers provide the weighout sheets for federally permitted vessels to NMFS port agents or in some cases directly to the SEFSC. These seafood dealers may also be required to report purchases from State permitted vessels to State offices where they either enter these records into State landings databases or the State provides these weighout sheets to NMFS port agents. State fishery reports are often entered as multi-trip totals into the NEFSC system. On an informal basis, some NMFS port agents send weighout copies to the NMFS Southeast Science Center (to Andy Bertolino).

NMFS Northeast port agents enter species and/or species market category summaries from the weighout sheets into the NMFS NERO-NEFSC CODES program. The data files produced by the CODES program are uploaded into an Oracle database at the NEFSC. Once the ACCSP system is operational, ACCSP partners (including the NMFS NEFSC) will upload required trip data from their Oracle database into the ACCSP Oracle data warehouse on a quarterly schedule. Additionally, the NEFSC periodically provides a copy of their trip landing records (CFDBS) to the NMFS SEFSC so that trip landings can be checked and international reporting obligations satisfied. Because of the wide-

ranging nature of federally managed fisheries and the diversity of vessels and seafood dealers, fishery managers have established redundant data collection systems to partially support verification of self-reported data.

Northeast seafood dealers with Bluefin tuna permits have to report daily landings of Bluefin Tuna by call-ins and must fax weekly reports to the NMFS Northeast Regional Office (NERO). **Monthly** the NERO Bluefin data is loaded into the NEFSC's Oracle database (CFDBS).

On a weekly basis, the NMFS Science and Technology Office (F/ST1) pulls landing data from the Northeast database. This data is stored in a database server residing in the ST/1 data center. The data are made available to the public through the ST/1 web site.

The NMFS NE HMS process is documented, schematically in Appendix B.

### ***NMFS SE HMS Processes***

Southeast seafood dealers (TX - NC) prepare weighout sheets and submit these sheets to both the NMFS SEFSC and when required to State trip ticket offices. The State offices, in turn, enter the data into State databases and provide the SEFSC with summary landing records (State, dealer, month, county, species totals with the weight reported as live weight equivalents) as input to the yearly Accumulated Landings file (**ALS**). The weighout sheet information is entered by NMFS personnel into the SEFSC's Oracle database, into a ***DLS\_DOMESTIC\_LONGLINE*** table (1998-present).

As mentioned previously, the NEFSC periodically provides a copy of their trip landing records (CFDBS) to the SEFSC for use in reconciling the trip landings. In addition, the NMFS Northeast port agents may send copies of weighout sheets to the SEFSC, on an informal basis.

According to ACCSP protocols, trip weighout information from Southeast States will be uploaded into the ACCSP Oracle database by the State offices, within 45 days for preliminary data and within 90 days for final records. North Carolina and Florida send ACCSP data by vessel trip, by dealer, date, county, species, and by market category.

On a weekly basis, the NMFS Science and Technology Office (F/ST1) pulls landing data from the SEFSC's **ALS** database. This data is stored in a database server residing in the ST/1 data center. The data are made available to the public through the ST/1 web site. ***During previous reconciliation activities SEFSC weighout trips that were identified as missing either from the NEFSC or SEFSC ALS databases have not always been added into the files that are accessed by the ST1 programs. This has resulted in landings discrepancies between published SEFSC reports and summaries available from the ST1 web site.***

The NMFS SE HMS process is documented, schematically in [Appendix C](#).

#### 4. Description of Comparative NMFS Databases.

##### **SEFSC DLS - Domestic Longline System** (Dealer weighout data - individual carcass weights)

The DLS system captures size-frequency data (dressed carcass weights) from swordfish and tuna dealers and also from cooperating owner-operators. This system is fairly self contained. The primary table is - *sefsc.dls\_domestic\_longline*. There also appears to be a trip data load routine that assigns system dates to records added to the DLS (*DLS\_LOAD\_HISTORY* table identified by Susan Molina). Other internal SEFSC programs and processes link, sometimes by hand, the DLS with the PLL so that the DLS Effort variables including days fished, sets fished, total hooks fished can be entered. ST1 staff will need to track the system load dates and the system change dates so that the comparison routines will be able to track database modifications. ST1 staff has developed a screen that allows you to explore the DLS table. A DLS screen displays header information and then individual dressed weights that can be scrolled through. A second, trip summary screen provides trip species total weights. He also provides the ability to submit a query in either the DLS screen or the trip summary screen for other trip codes.

The following notes were provided by Susan Molina (SEFSC).

All DLS data are stored in *DLS.DLS\_DOMESTIC\_LOGLINE@SEFIN8*. All other DLS tables contain codes. Jean creates a file that is used to fill days fished, sets fished, total hooks fished, etc. The table these values are stored in is *DLS\_DOMESTIC\_EFFORT\_TEMP*, which is a temporary table. These values should be entered directly into the DLS system.

The report you mention is called *SUMMARY\_BY\_TRIP\_MASTER.rdf* and it is on the CD I sent you. It is accessed from the menu Reports/Species Summary option. There are a number of other validation routines that Jean has written in SAS that need to be incorporated into the FLS system. The new client/server version of DLS is currently being tested by Amy.

##### ***SEFSC - DLS Variables***

SEFSC VESSEL_ID	LOG_DATE ( landing date)	STATE_CODE
PORT_CODE	DEALER_CODE	SE GEAR_CODE
DATA_SOURCE	LOCATION_CODE	TOTAL_HOOKS_FISHED
DAYS_OF_FISHING	TOTAL_NUMBER_OF_SETS	SE SPECIES_CODE
WEIGHT 1 to WEIGHT 32	TRIP_NUMBER	SOURCE_FILE

BATCH\_NUMBER

ACTION\_CODE

CUTOFF\_CODE

The SE TRIP\_NUMBER variable links trip records from 1996 to 2001 in the weighout, logbook, trip summary and observer databases.

ST1 staff needs information on variables in *DLS\_LOAD\_HISTORY*.

### **NEFSC - CFDBS Trip Landing Records**

The NERO-NEFSC CODES program captures trip based landings primarily from Federally permitted dealers operating in a variety of both Federal and State managed fisheries from Virginia to Maine. Data are routinely submitted by port agents to the database managers at the NEFSC. The NEFSC has periodically provided a copy of a portion of their trip landing records (CFDETS) to the SEFSC for use in reconciling a selected list of highly migratory species, primarily swordfish, sharks, and tuna against the SEFSC trip weighout (size frequency) data.

ST1 Staff has knowledge about the CFDETS tables that are used to generate the tables that are set up for both Jean and Daryl to use.

***SEFSC, NEFSC, and ST1 staff will have to agree on a final list of selected species keeping in mind Jean's priorities as well as the interests of other NMFS scientists involved with HMS research or management. This needs to be done ASAP.***

## **5. Overview of System Approach**

ST1 recommends a trip summary view, *se\_ne\_recon\_v*, using variables from the *dls\_longline*, *dls\_species\_code*, *dls\_states*, *nehms*, *cfspp* and *sppname* tables. From these, tables are created on shark: *dls\_longline*, *dls\_species\_code*, *dls\_states*, *nehms*, *ne\_species* and *ne\_species\_nomk*. To link vessel identification numbers in the southeast to hull numbers in the northeast, a table on shark, *se\_cg9899*, is used. Species from each region are linked via a nmfs species code. The trip summary records in the view are created by summarizing by trips, species, vessels and landing dates. From the view, a temporary table, *se\_ne\_recon\_t*, is created to merge matched records from both regions according to landing dates, NMFS HQ species codes, no market size categories from the northeast and vessel identification numbers. The reconciled table, *recon\_se\_ne*, is then populated from matched records in the *se\_ne\_recon\_t* table using PL/SQL to test for matching species code, landing date and possible matching weights. Other fields in the reconcile table that are updated include data source, match result and reconciled date. The reconciled table can be edited, if necessary, through an Oracle data entry form and if additional research is needed to resolve discrepancies or editing of any records

to reflect appropriate trip matches, the data entry form will be used. Subsequent data loads would only look at new records rather than any records already reconciled. The reconciliation table will allow Andy, Jean, NMFS or state scientist responsible for reconciling trip reports, to perform comparisons and archive resulting annual species totals reported in national fisheries report to ICCAT and other international fisheries organizations. Other NMFS scientist will be able to access the reconciliation database to review Bluefin and shark landings. NEFSC and headquarter offices would be able to utilize this table to support us landings on the web pages. This design will address long-standing problems with NMFS web pages disagreeing with ICCAT National Reports. The approach would also be adaptable for other species and fisheries provided that appropriate species, vessel and gear translation tables are available. This later aspect is critical for satisfying the existing ACCSP grant which included an HMS component and a snapper-grouper add-on.

Tables used from northeast database

NEHMS1998

Gear	gear(s) used on a trip
Hullnum	vessel number
Nespp4	1 <sup>st</sup> 3 digits of the species number is used to designate no market category.
Port	1 <sup>st</sup> 2 digits of the port number is used to designate state code.
Year  month  day	3 fields concatenated together makes up landed date.
Spplnlbb	used to sum weights together per landed date, species and vessel

CFSPP - species table

nespp4	4 digit species code
nespp3	3 digit, no market size code
nmfs_cd	HQ nmfs species code
name	associated species name

SPPNAME - no market size table

nespp3	no market size code
name	associated species name

Tables used from southeast database.

Dls\_Longline

trip number	
gear_code	gear used for a trip
vessel_id	vessel identification number
species_code	species of fish caught on trip
state_code	state of landing
log_date	date of landing
weight1-32	individual weights summed per trip, landing date and species

Dls\_Species\_Code

Dls_Species_Code	southeast species code
Nmfs_Species_Code	associated nmfs hq species code
Common_Name	species name

Se\_Cg9899

Dls_Id	southeast vessel identification
Cg_Nmbr	northeast hullnum and coast guard vessel number



Reconciliation tables stored on shark

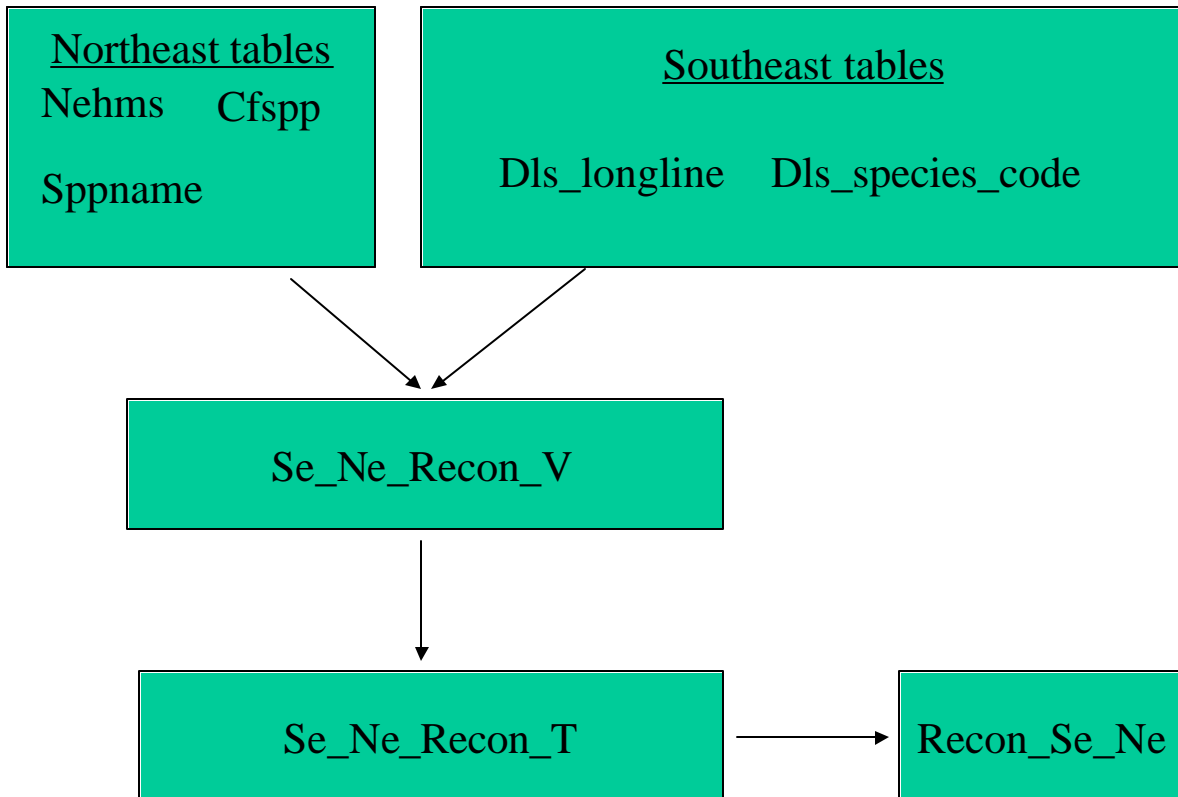
Recon\_Se\_Ne

Data_Source	1 = Se, 2 = Ne, 3 = ACCSP
Se_Vessel_Id	southeast vessel identification number
Ne_Hullnum	northeast vessel identification number
State_Code	2 character state code
Landing_Date	date of trip landings
Gear_Code	southeast gear code used on a trip
Species_Code	species code
Species_Name	species name
Pounds	summed species weights
Se_Trip_Number	southeast trip number
ACCSP_Rec_Num	ACCSP record number
Reconciled_Date	date record was entered in system as reconciled
Match_Results	1 = exact weight match, 2 = record matched, weight was different. Greater of weight values entered in table, 3 = no match

The reconciliation log file, *Recon\_Meta*, will reside on shark and store any modifications to a reconciled record. Users will have the capability to view any changes.

Recon\_Meta

trip_number	southeast unique trip_number
se_vessel_id	southeast boat numbers
ne_hullnum	northeast hullnum
field_name	name of field being modified
old_value	value of a field before modification
modified_value	new value of a record after modification
status_type	whether record was changed or deleted
change_date	date of modification
user_name	name of person modifying record
comments	field for comments on a record



## 6. Programming Specifics for HMS Reconciliation

> All weight data will appear as “landed pounds” (or dressed weight). This is the same weight stored in ACCSP.

> To compare the DLS records of individual carcass weights to trip ticket weights from another source, the individual weights will be summed by species and trip number from the NMFS Southeast database (*dls\_domestic\_longline*) and stored in a single view created from an SQL query that summarizes data from each database and merges the two queries into 1 data set. This view, ***RECON\_SE\_NE\_V***, is stored on the *Shark ST server*. The SE species trip totals are going to be **compared** to the ACCSP or NEFSC trip records by **vessel\_id, landing\_date, state\_code, and species weight**.

(NOTE - Gear has been dropped as a matching characteristic because of conversion problems and one to many relationships - gear codes from each comparative database will be tracked.

>**Three external-comparison outcomes are possible during the first stage**: matched trips that agree (vessel, date, State and Species are primary match fields); matched trips with discrepancies in species weight variables; and unmatched trips that exist in one or the other of the comparison data sets.

In the existing SEFSC reconciliation process, trip records are merged in a SAS table with summary records that list the gear, SEFSC trip number, SEFSC vessel number, NEFSC vessel number, landing date (Month - Day), landing State, and the landed weights for 11 HMS species from both sources, and the differences in those weights (**Example - Table 1**). ***The general business rule is that the larger species weight value for each matched trip is accepted. A question has been raised about flagging differences that exceed 25% of SEFSC totals.*** Un-matched trips are listed in vessel ID, date, and State order and will have blank fields for the comparison database.

> Reconciliation Scenarios.

1. Trips that match exactly with no discrepancies in trip header variables (vessel ID, landing date, State, species) or for any of the species weight totals are coded as reconciled (**match\_result\_code = 1**). The reconciled table will include the SEFSC trip number and the associated NE Vessel number, NE landed date, and NE State. Original ideas was to track NE Doc number (concatenating landed year, month and Doc number for unique receipt identification) but we have found that there can be two document numbers for a single vessel-date, state trip record.

2. Trips that match but that have discrepancies in species weight totals will be entered into the reconciliation table with (**match\_result\_code = 2**), identifying that **the larger species weight value was selected as the preferred total weight and the source responsible for that value**. A meta-data table will record the comparison results which is not stored in the reconciliation table. Both the meta-data and reconciled tables will be available to the data providers.
3. Un-matched trips with **match\_result\_code = 3** from either SEFSC or NEFSC records will be added to the reconciled table with the appropriate source variables indicating data source (e.g. SEFSC trips would have null values for NE document number). A data entry form will allow users to query the reconciliation database and review un-matched trips. Trips that didn't match were either missing from the comparative source or the matching variables (vessel id, date, and State) differed. The operator who is using the utility will be able to sort, view, and edit records in the reconciliation table. If the operator finds records from the comparative table that he or she determines should be matched, an editing screen will allow the appropriate matching variables to be added to the retained record while the deleted record and the variables that were changed to allow a match will be saved to the meta-data file. A comment field allows the user to explain why a record was deleted. This comment is added to the meta-data table. NE and SE database managers are given access to the meta-data file to allow them to review changes made by staff personnel responsible for reconciling trip records. SEFSC staff will work with the ST1 developer during system testing to refine search options.

> Although the initial priority is to develop and test the system for the 11 HMS species that are currently compared, SEFSC, NEFSC, and ST1 staff should review the list of species in *Table 2* that reflect all of the species recorded in the DLS system. By including Bluefin tuna and more shark species, the proposed system will satisfy the reconciliation requirements of other NMFS and State scientists and managers. The reconciliation data entry form is able to query and track the reconciliation status for all selected species records. This allows reconciliation of many discrepancies relating to differences in species identifications for sharks and several other species caught in HMS fisheries. Several of these species are already of management interest and landings data quality problems have limited progress (e.g. porbeagle landings) in both stock assessments and management. In either the SEFSC or external files there may be trips that land only dolphin, wahoo, and/or king mackerel so these records must be tracked since they will need to be compared against other State and Federal logbook systems.

> The system provides editing capabilities for SEFSC staff to, not only the edit records in the reconciliation table, but the *dls\_domestic\_longline* tables.

> The proposed system may eventually integrate other SEFSC routines that provide effort and location information which is added to the *DLS* table from data contained in the logbook PLL trip summary table. Dealer and location data from the trip summaries may help in the reconciliation process. Fishing effort variables in particular are often taken from owner/operator daily logbook catch reports. For Atlantic HMS the merging of logbook effort and trip landings files has traditionally occurred once a year prior to international reporting deadlines and annual stock assessments. ***This component will become a joint effort between ST1, SEFSC, and NEFSC IT staff -- NOTE this is where ST1 needs assistance so that Joan can get the effort and area data she needs and so that Susan can assist Jean in tying the logbook and weighout data together.***

> The SEFSC trip number and the associated NEFSC vessel, State, and date fields will be included in the ST1 reconciliation table and allow the NEFSC to query information from the SEFSC tables on **fishing effort and area fished**. A potential problem may result if a split trip with two NE dealer reports is matched against a single SEFSC unloading receipt. In the reconciliation table only 1 trip record would be retained.

> NMFS NEFSC will perform **monthly** refreshes of the CFDETS table from the Bluefin Tuna table.

> Application shall have capability to run data reconciliation at any time frame (quarterly is suggested initially) determined by the person responsible for data reconciliation. Reconciliation would be facilitated if the responsible operator ran the routines whenever significant additions were made to the *DLS* tables. Timely attempts to identify missing trips can assist port agents and other NMFS staff responsible for quota monitoring. Discrepancies will be easier to resolve if the time between the event and efforts to resolve the differences are short.

### ***Concerns relating to data processing from multiple comparative sources.***

> There will often be multiple records in ACCSP and NEFSC for the same species on a trip since fish species can fall into different market categories and damaged carcasses (chunks) can be coded separately. As the system is adapted for use with different Federal and State databases, documentation on coding practices will be critical.

> ST1 staff will work with SEFSC and NEFSC staff to develop and maintain a cross reference table with ITIS species codes (the ACCSP standard), NMFS HQ species codes, SEFSC, NEFSC and State codes necessary to support these applications.

> A hull number + permit number, (hull # may be the same as state registration or coast guard number), uniquely identifies a vessel. NMFS HQ to create a cross reference table that reconciles the SEFSC vessel numbers, NEFSC vessel numbers, Coast Guard numbers, HMS permit numbers, and subsequent State lists as they become available through the ACCSP (Liz, John, Jean, and Andy).

> Dealer cross-reference lists may also be required to facilitate reconciliation.

### ***Backend Processes***

There are a number of tables involved in creating the reconciliation between the SE and NE data sets.

- > SE - DLS\_longline table
- > SE - DLS\_Species\_Code
- > SE - DLS\_Gears
- > SE - DLS\_States
  
- > NE - NEHMS1998 table (data subset of CFDETS)
- > NE - NEHMS1999 table (data subset of CFDETS)
- > NE - NEHMS2000 table (data subset of CFDETS)
- > NE - NEHMS2001 table (data subset of CFDETS)
- > NE - CFSPP
- > NE - SPPNAME

The NE DBA has provided ST1 staff with read access to a copy of CFDETS (*Daryl please resolve with Joan*). This copy will be updated on a quarterly or bi-monthly basis.

Data from the CFDETS table will be pulled (selected) for specific highly migratory species and stored on the HQ database server, SHARK. This table on SHARK will be called NEHMS. The NEHMS table will contain the summarized version of the CFDETS (sums of no market category).

Data from the SE DLS\_domestic\_longline table will be pulled by trip\_number and stored on the HQ database server, SHARK. This table on SHARK will be called DLS\_longline. When data is pulled from the SE DLS\_longline table, it will sum the individual weights of a species for a specific trip number. The records from these two trip views (SE & NE) will be compared and stored in the reconciled table on SHARK, called RECON\_NE\_SE. The record comparing process is as follows.

STEPS:

1) The records from each table will be matched by the following fields:

RECON_NE_SE	SE field name	NE field name
vessel_number	vessel_id	hull_number
state_code	state_code	substr (port_code)
log_date	log_date	<u>datelanded</u> - year  month  day
species_code	species_code	substr(nespp4,1,3)
total_weight	sum(weights)	sum(spplndlb)

2) IF processing -

Scenario A

If the records match exactly,

then the record is added to the recon\_ne\_se\_v table,  
the update\_date field is updated with the system date,  
the data\_source field is updated with a 1 (1=SE as a default),  
the trip\_num field is updated with the SE trip number,  
the NE vessel, NE State, and NE date variables are retained as tracking  
variables since NE doc\_num is not unique for a vessel trip,  
and the match field is updated with a 1 (identifying an exact match).

Scenario B

If the total weight does not match but all other field match,

then the record is added to the recon\_ne\_se\_v table with the larger weight of  
the two records,  
the update\_date field is updated with the system date,  
the data\_source field is updated with a 1 or 2 (1=SE, 2=NE) to identify  
the source of the larger value entered in the weight field,  
the trip\_num field is updated with the SE trip number,  
the NE vessel, NE State, and NE date variables are retained as tracking  
variables since NE doc\_num is not unique for a vessel trip,  
the match field is updated with a 2,  
and the meta\_table is updated with the two records and a field showing the  
difference between the weights.

Scenario C



If the record is found in one of the data sets and not in the other data set  
(ie. found in SE and not NE or vice versa)

then the record is added to the recon\_ne\_se table,  
the update\_date field is updated with the system date,  
the data\_source field is updated with the a 1 or 2 (1=SE, 2=NE),  
the trip\_num field is updated with the SE trip number,  
the NE vessel, NE State, and NE date variables are retained as tracking  
variables since NE doc\_num is not unique for a vessel trip,

and the match field is updated with a 3 (identifying only one source).

This backend processing will be executed quarterly. There are three data checks for updating the recon\_ne\_se\_v table. These data checks occur after the “IF” testing as stated above in the scenarios A-C, and becomes a nested “IF” under the “THEN” statements.

- 1) If a trip number is not equal to any trip number in the recon\_ne\_se\_v table,  
then the record is appended to reconcile table.
- 2) If a trip number in the reconcile table is equal to an existing trip number in the recon\_ne\_se\_v table and the match field is equal to a 1 or 2,  
then the record is already reconciled
- 3) If a trip number is equal to an existing trip number in the recon\_ne\_se\_v table and the match field is equal to a 3,  
then the record overwrites the existing record in the recon\_ne\_se\_v table in case any modifications were done at the remote site, since the last data load.

The meta-data (*recon\_meta*) table are readable by both NE and SE data managers.

The meta table will contain both the old and new/modified values, and the date of when any changes occurred. The meta table description needs to be finalized. Both the SE and NE will have read permissions to the meta table.

Key Oracle tables and associated access level:

- SEFIN8, read access to DLS\_domestic\_longline table
- SEFIN8, read access to DLS\_species\_code table
- HQ vessel cross-reference table (Se\_Cg9899)
- HQ dealer cross-reference table (dls codes - SERO & NERO permit numbers)
- HQ SHARK, read access to reconciled table
- HQ SHARK, read access to meta data table
- NE, read access to CFDETS\_copy table
- ACCSP, read access to Landings table for NC, FL, and NMFS NE
- ACCSP, read access to Species table for NC, FL, and NMFS NE
- ACCSP, read access to Dealer\_Reports table
- ACCSP, read access to Vessels table

**HQ DLS\_Longline** (Table copied from southeast database to Shark)

Column Name	Type	Description
Vessel_Id	Number(5)	Southeast Boat identification number
Log_Date	Number(8)	Trip landing date
State_Code	Varchar2(2)	State where trip was landed
Dealer_Code	Number(3)	
Gear_Code	Varchar2(2)	Gear used on trip
Species_Code	Varchar2(3)	Species caught
Weight1...32	Number(3)	Weight of individual species caught
Trip_Number	Varchar2(10)	Unique identifier of the time a boat leaves shore until its return

--	--	--

**\*Note\*** Columns used in retrieval and reporting of southeast data. To get total weight per species, the 32 weight fields are summed.

### NEHMS

Column Name	Type	Description
Year	Varchar2(4)	Year of landing as recorded on weighout sheets
Month	Varchar2(2)	Month of landing as recorded on weighout sheets
Day	Varchar2(2)	Day of landing as recorded on weighout sheets
Port	Varchar2(6)	Use a substr of 1 <sup>st</sup> two digits to get the state of landing
Hullnum	Varchar2(10)	Northeast recorded vessel number
Nespp4	Varchar2(4)	Northeast 4 digit species code
SppIndlb	Number(10)	Landed weight of species

**\*Note\*** The 1<sup>st</sup> three digits of the species code is the no market size category of a species. This is what is used to compare landed weight in the northeast to the landed weight in the southeast.

### DLS\_Species\_Code

Column Name	Type	Description
DLS_Species_Code	Varchar2(3)	Species code linked to the dls_longline table to retrieve species names
Common_Name	Varchar2(50)	Name of Species
Nmfs_Species_Code	Varchar2(4)	Nmfs species code as taken from the species_cross_reference table

--	--	--

**\*Note\*** NMFS\_Species\_Code used to link southeast and northeast data. Will ultimately build one table with NE, SE, HQ, and ITIS species codes

**lpritch.Se\_Cg9899**

Column Name	Type	Description
Cg_Nmbr	Number(7)	Coast guard register boat number
Dls_Id	Number(5)	Southeast boat number
NEFSC Vessel ID		NEFSC vessel number
State_ID		State Registration number
CG_Name		

**\*Note\*** This table is used to link vessels in the southeast and northeast through coast guard vessel identification numbers

**DLS\_States**

Column Name	Type	Description
DLS_State_Code	Varchar2(2)	Southeast defined state code
State_Label	Varchar2(30)	State name
Nmfs_State_Code	Varchar2(2)	Nmfs state code equivalent

**\*Note\*** NMFS\_State\_Code is used to link southeast state coding to the northeast state coding.

### Recon\_Se\_Ne

Column Name	Type	Description
Data Source	Number	1=SEFSC, 2=NEFSC, 3=ACCSP-NC, 4=ACCSP-FL
SE Vessel_Id	Number(5)	Southeast Boat identification number
Landing_Date	Varchar2(8)	Trip landing date
State_Code	Varchar2(2)	State where trip was landed
Pounds	Number(6)	Species weight
SE Trip_Number	Varchar2(10)	Unique identifier of the time a boat leaves shore until its return
Match_Result_code		1=Exact match, 2=Match larger weight, 3 = unmatched
Reconciled_Date		Date of reconciliation
Ne_Hullnum	<b>varchar2(10)</b>	<b>Ne vessel number</b>

**Table 1.** Example of existing program output for SEFSC comparison of SEFSC and NEFSC trip records. The first trip was from the NEFSC file and did not have a match in the SEFSC. It missed a region designation and SEFSC trip and vessel number. The second trip was a match between the SEFSC and NEFSC and a 99 pound discrepancy for swordfish. The third and fourth trip appear to be a match but the trip record from the SEFSC seems to be divided into two areas, Grand Banks and Northeast Coastal, and it is matched against a single NEFSC record.

2000 Landings data from SEFSC and NEFSC by id and date

Region	Gear		Trip				Vessel				Id	Month	Day	State
LL				.						1022773	10	2	MA	
	swo	yft	bet	skj	alb	bon	blf	lta	wah	dol	oth			
wo	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ne	472	0	0	0	0	0	0	0	0	0	0	0	0	0
dif	.	.	.	.	.	.	.	.	.	.	.	.	.	.
est	472	0	0	0	0	0	0	0	0	0	0	0	0	0
GB LL	20034954				3026					1022773	10	3	MA MA	
	swo													
wo	40867	0	6812	0	998	0	0	0	0	0	0	0	0	0
ne	40966	0	6812	0	998	0	0	0	0	0	0	0	0	0
dif	99	0	0	0	0	0	0	0	0	0	0	0	0	0
est	40966	0	6812	0	998	0	0	0	0	0	0	0	0	0
GB LL	20041243				3026					1022773	11	15	MA MA	
	swo	yft	bet		alb									
wo	6745	47	673	0	22	0	0	0	0	0	0	0	0	0
ne	11772	380	1261	0	157	0	0	0	0	0	0	0	0	0
dif	5027	333	588	0	135	0	0	0	0	0	0	0	0	0
est	11772	380	1261	0	157	0	0	0	0	0	0	0	0	0
NE LL	20041243				3026					1022773	11	15	MA MA	
	swo	yft	bet		alb									
wo	5108	333	588	0	135	0	0	0	0	0	0	0	0	0
ne	11772	380	1261	0	157	0	0	0	0	0	0	0	0	0
dif	6664	47	673	0	22	0	0	0	0	0	0	0	0	0
est	11772	380	1261	0	157	0	0	0	0	0	0	0	0	0

Table 2. List of SEFSC DLS species codes.

Existing SEFSC Comparison species:

SE SPECIES	Name and Priority	SE-DLS species codes		
SWO	Swordfish - 1	SWO - SWX - SWD		
YFT	Yellowfin Tuna - 1	YFT - YFX - YFD		
BET	Bigeeye Tuna - 1	BET - BEX - BED		
SKJ	Skipjack Tuna - 1	SKJ - SKD		
ALB	Albacore Tuna - 1	ALB - ALX - ALD		
BON	Bonita - 1	BON - BOD		
BLF	Blackfin Tuna - 1	BLF - BLD		
LTA	Little Tunny - 1	LTA - LTD		
WAH	Wahoo - 1	WAH - WAD		
DOL	Dolphin Fish - 1	DOL - DOD		

\* If acronym ends in X - weight is chunk weight - If acronym ends in D - weight is lump sum of a group of fish

#### Other SEFSC DLS species

SE SPECIES	Name and Priority	SE-DLS species codes		
OTH	Other Tuna -2	OTH - OTX - OTD		
BFT	Atl.Bluefin Tuna - 2	BFT - BFX - BFD		
BUM	Blue Marlin	BUM -BUD		
WHM	White Marlin	WHM - WHD		
SAI	Sailfish	SAI - SAD		
SPF	Spearfish	SPF - SPD		
BIL	Billfish Unid	BIL - BID		
SHK	Sharks Unid. - 2	SHK - SHD		
BSH	Blue Sharks - 2	BSH - BSD		
MAK	Mako shark - 2	MAK - MAD		
SFM	Shortfin mako - 2	SFM - SFD		
LFM	Longfin Mako - 2	LFM - LFD		

THR	Thresher Shark - 2	THR - THD		
POB	Porbeagle Shark - 2	POB - POD		
BTP	Blacktip shark - 2	BTP - BTD		
BRN	Brown	BRN - BRD		
SBR	Sandbar shark - 2	SBR - SBD		
DSK	Dusky Shark - 2	DSK - DSD		
SLK	Silky Shark - 2	SLK - SLD		
HMH	Hammerhead Sharks	HMH - HMD		
WTS	Whitetip Shark - 2	WTS - WTD		
TGR	Tiger Shark	TGR - TGD		
BBL	Bull Shark	BBL - BED		
LMN	Lemon Shark	LMN - LMD		
NGT	Night Shark	NGT		
FIN	Shark Fins	FIN		
BHF	Bluefish	BHF - BHD		
KNF	Kingfish	KNF - KND		
AMJ	Amberjack	AMJ - AMD		
TLF	Tilefish	TLF - TLD		
OLF	Oilfish	OLF - OLD		