South Florida/Caribbean Network

National Park Service
Inventory and Monitoring Division

Standard Operating Procedure RF04

Data Entry and Proofing—Version 2.0

Please cite this as:

South Florida/Caribbean Network. 2025. Data Entry and Proofing—Version 2.0. South Florida/Caribbean Network Standard Operating Procedure NPS/SFCN/SOP—RF04. South Florida/Caribbean Network, Palmetto Bay, Florida.

Purpose

This standard operating procedure (SOP) describes how to enter data into the RVC Data Entry Program, including the Sample, Substrate and Fish Species data and the Field/Boat Log. Steps also describe tier I and tier II data proofing and the verification process and structure of the Annual Excel RVC Database and Master RVC Analysis Ready Database.

Revision History Log

Revision Date	Author	Changes Made	Reason for Change	New Version #
6/18/2020	Mike Feeley	Replaced RVC.3.exe with a new data entry program.	Updated to web-based program.	2.0
6/18/2020	Mike Feeley	Deleted environmental log.	No longer used.	2.0
6/18/2020	Mike Feeley	Boat log entered within RVC Data Entry Program.	Update to program.	2.0
6/18/2020	Mike Feeley	Data merger and proofing scripts in R.	Converted merger and proofing processes from SAS program to R.	2.0
6/18/2020	Mike Feeley	Data uploaded to NOAA managed web-based data portal.	All RVC data is available for user download at a single location.	2.0

Contents

Page

Introduction 4

Data Entry 5

The RVC Data Entry Program 5

Directions for Use of the RVC Data Entry Program 5

Entering Sample Data 5

Entering Habitat Data 8

Entering Fish Species Data 10

Entering the Field/Boat Log 12

Editing Saved Data 14

Fish Surveys 14

Boat Log 15

Data QA/QC and Verification 15

Initial Diver Proofing 15

Participant-Level Proofing 16

Data Proofing Criteria 16

Proofing Report Procedure 16

Agency-Level Proofing 18

Preliminary Annual Master Spreadsheet File Verification 18

Final Verification 19

Data Portal 28

Data Archiving 28

Data Cataloguing 28

Literature Cited 29

Appendix A. Variables in Data Output Files and Data Verification Files 30

Appendix B. Final QAQC function list (R programming) 37

Appendix C. Data Quality Standards for A Cooperative Multiagency Reef Fish Monitoring Protocol for the Florida and U.S. Virgin Islands Coral Reef Ecosystems 69

Protocol Narrative Version—2.0 69

Introduction 69

Protocol Overview 69

Protocol Activities and Modules 70

Sampling Design 73

Data Quality Objectives 74

Literature Cited 78

Introduction

Management of the data is an important and multi-level process between the diver, agency data manager and the NOAA lead statistician. Steps that are taken to efficiently manage the data once they have been collected include data entry, proofing, and storage. The specific steps and those responsible for each step are outlined in Table 1 and explained in detail in the following sections of this SOP.

Table 1. Data management overview.

Steps	Responsible person	Forms/Programs used
(1) Field data collection	Boat Captain/Diver	Fish/Habitat Data Sheet Field/Boat Log sheet
(2) Enter/collate data	Diver	Fish/Habitat data entered into RVC Data Entry program
	Agency Data Manager	Boat Log entered into RVC Data Entry Program
(3) Cross check Boat Log and fish samples	Agency Data Manager	RVC Data Entry Program - Manager Dashboard
(4) Download proofing reports	Diver	RVC Data Entry Program
(5) Proof and edit data	Diver	Visual comparison of original data sheets to proofing report
(6) Create preliminary RVC Annual Master Excel file	Statistician (NOAA)	R
(7) QA/QC and verification procedures	Statistician (NOAA)	R
(8) Finalize RVC Annual Master analysis ready datasheet	Statistician (NOAA)	R
(9) Distribute RVC Annual Master analysis ready datasheet	Statistician (NOAA)	csv

Data Entry

The RVC Data Entry Program

Reef fish Visual Census (RVC) fish and habitat survey data are entered into a web-based database using the RVC Data Entry Program. This program was designed to standardize data entry and minimize the potential for error during the data entry process. Data are entered into the RVC Data Entry Program through three 'window' or data entry screens focusing on sample identifier information, habitat or substrate data, and fish species, size, and abundance data.

Directions for Use of the RVC Data Entry Program

To open the RVC Data Entry Program, begin by accessing the internet address for the Florida data entry application, https://grunt.sefsc.noaa.gov/rvc_data_entry/ or the United States Virgin Islands (USVI) data entry application, https://grunt.sefsc.noaa.gov/caribbean_data_entry/. On the home page (Figure 1), click on the login link in the upper left and sign in with the unique diver username and password that was issued to the individual diver.

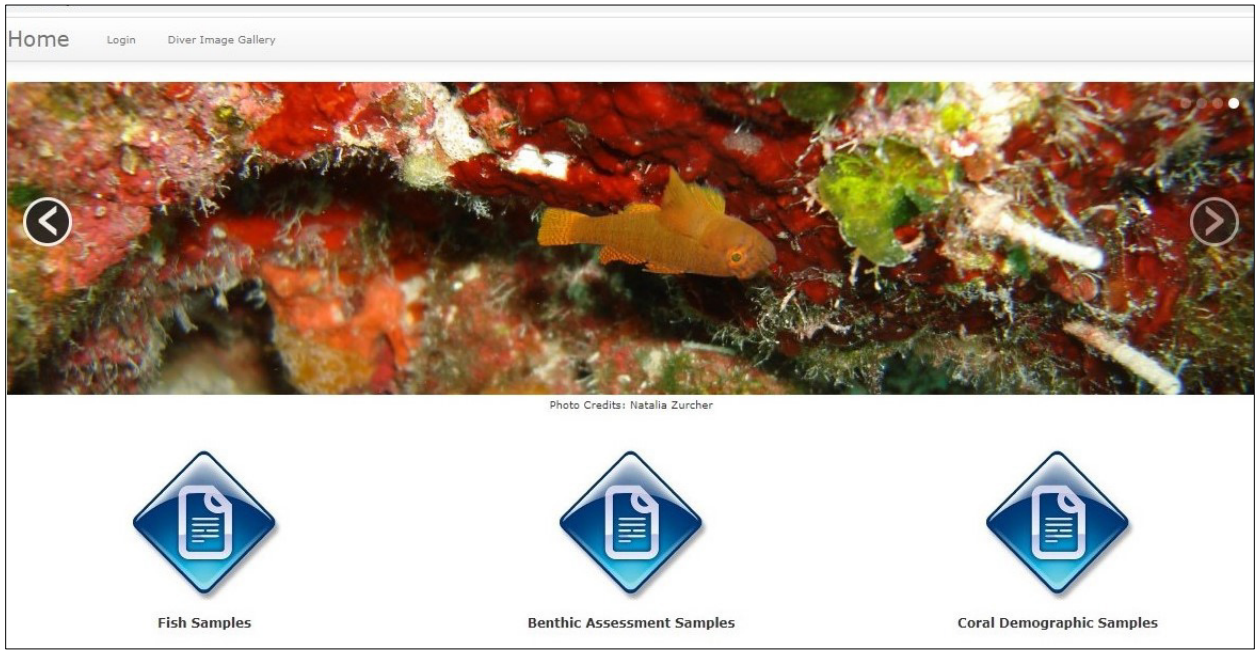


Figure 1. Home page for the RVC Data Entry Program.

Entering Sample Data

The first window of the home page contains icons for "Fish Samples," "Benthic Assessment

Samples," and "Coral Demographic Samples" (Figure 1). Click on the "Fish Samples" icon. This is where general information about the diver's fish samples are tabulated, and all samples that an individual diver has entered (per year or sampling time) will be available. The following steps will assist the user with RVC data entry.

- Click on the "New Fish Sample" icon. The Sample Section window for the new fish sample will appear (Figure 2).

- **Boatlog/Manager**. From the dropdown menu, select the agency/last name of the person identified by your agency or the mission as the agency data manager.

- **Sample Date**. This value will default to the current day identified on the computer of the user. Update this value (year-month-day) to the date of the survey if necessary. Failure to do this is a common mistake.

- **Diver** and **Buddy**. Enter the last name of the "Diver" and "Buddy" by clicking the drop-down arrow in each data field and choose the appropriate name from the given list of divers. The "Diver" and "Buddy" data fields must not be left empty.

- **Sample Type**. Choose the appropriate sample type using the drop-down arrows in the data entry field. For RVC surveys, the sample type will be "Bohn-Bannerot Point." If the diver is training, "Training – Observer" is selected from the dropdown list instead.

- **Habitat Type**. Choose the appropriate sample type using the drop-down menu in the data entry field.

- **Dive Begin Time, Dive End Time, Sample Begin Time, Sample End Time**. Enter the time when the dive started and ended, and the time when the sample started and ended in the appropriate entry fields. It is important that the date and times be entered in the proper, indicated format. The hour must be a number between 0 and 23, and the minutes must be a number between 0 and 59. The "Dive End Time" must be greater than the "Dive Start Time," and the "Sample End Time" must be greater than the "Sample Start Time." Additionally, the "Sample Start Time" and the "Sample End Time" must fall between the "Dive Start Time" and the "Dive End Time." Warning windows will pop up when very early start times and very late end times for the dive are entered.

- **Field ID**. The value of the "Field Number" will be provided by the mission coordinator / data manager. This number will consist of the four-digit primary sampling unit number, station sample number and a diver replicate letter designation. In the current sampling design, the station sample number will always be 1.

- **Dive Depth (feet [ft])**. Enter the maximum depth in feet reached on the dive. This entry must be an integer greater than zero.

- **Sample Depth (ft)**. Enter the depth in feet of the sampling station (bottom depth at center of sampling cylinder).

- **Visibility (meters [m])**. Enter the best estimate of the horizontal underwater visibility in measurement units of meters. This entry will be an integer greater than zero.

- **Water Temperature (°F)**. Enter the bottom temperature recorded by diver's dive computer in degrees Fahrenheit. This entry will be an integer greater than zero.

- **Current**. Enter an estimate of water current strength from the drop-down menu. (None, Moderate or Strong).

- **Fishing Gear**. Enter the type of fishing gear observed within the sample cylinder from the drop-down menu (None, Trap, Trap Debris, Fishing or Other). Possible types of fishing gear include lobster traps, trap line, and monofilament fishing line.

- **Comments**. Enter any important remarks regarding the dive and/or sample into the "Comments" entry field.

- **Go To Substrate**. Throughout the data entry process, dynamic evaluations of the data input into entry fields are performed by the program. Upon leaving an entry field, the program checks the validity of the data automatically. Prior to moving on to the next window, the information entered into the current window must be checked for errors once more by the program. If there is an error detected the cell will appear yellow. The correction must be made before the "Go To Substrate" icon will be active. If no errors are detected, the program will generate the "Master Sample number," and the Substrate section data entry window will open. The "Master Sample number" is the unique relational linking index with the Substrate data and Species data (see sections below). The "Master Sample Number" is composed of the following information: year+month+date+sample start time+diver number. The year, month, day, and sample start time are derived from entered data for a sample and the diver number is determined from the number (4 digit) assigned to each diver last name in the RVC data entry support file (diverlist.noaa).

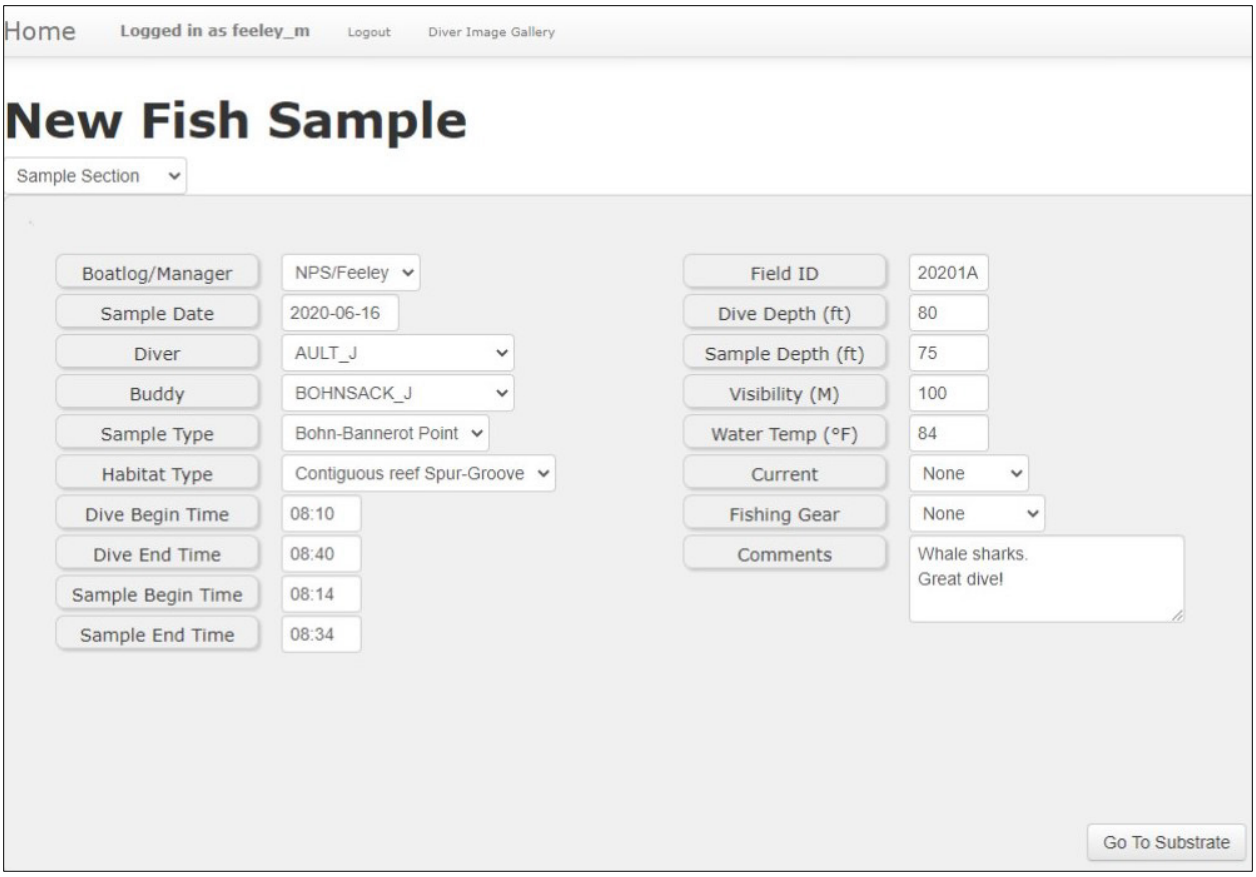


Figure 2. The Sample Section data entry window.

Entering Habitat Data

The second window of the RVC Data Entry Program is titled "Substrate Section" (Figure 3). This is where the user enters information about the substrate of the sampled habitat. As data are entered, the program evaluates the data dynamically in a manner like that of the first window. The following steps will assist the user with data entry into the "Substrate Data" window.

- **Max Substrate Depth (ft)** and **Min Substrate Depth (ft)**. Enter the recorded maximum depth and minimum depth of the substrate slope within the sample cylinder in feet in the respective entry fields. The maximum depth must be a whole integer greater than zero, and the minimum depth must be less than or equal to the maximum depth.

- **Hard Substrate (m) and Soft Substrate (m)**. Enter the measured "Max Vertical Relief" of the hard substrate and soft substrate in meters in the respective entry fields. The relief of the hard and soft substrate may be a decimal value but must be greater than zero.

- **Surface Relief Coverage %**. The estimated percent of the intervals of vertical relief that comprise the surface of the sample area must be entered into the "Surface Relief Coverage %" section. This section is divided into the intervals of "*Hard*" vertical relief and "*Soft*" vertical relief. The data entry fields that are activated depend on the values entered in the

"Max Vertical Relief" section above. For example, if the value of the maximum vertical relief of the soft substrate equals zero (i.e. no soft substrate was found in the sample area), then the entry fields for the "*Soft*" vertical relief in the "Surface Relief Coverage %" will not be activated. The total percent coverage of both "*Hard*" and "*Soft*" vertical relief must equal 100%. Hardbottom relief may include coral (living or dead), bedrock and sand. Soft relief typically consists of biotic cover such as algae, seagrass, sponges and octocorals.

- **Abiotic Footprint %**. The percent of sand, hard-bottom, and rubble that comprises the "Abiotic Footprint" of the sample area must be entered into the fields labeled "Sand," "Hard-B," and "Rubble" respectively. The values entered into these fields must add to 100%. If the percent of "Sand" in the "Abiotic Footprint" is not equal to zero, then the "Biotic Cover Sand %" portion of the "Biotic Cover" section will be activated. Likewise, if "Hardbottom" in the "Abiotic Footprint" is not equal to zero, then the "Biotic Cover Hardbottom %" portion of the "Biotic Cover" section will be activated.

- **Biotic Cover Sand % and Biotic Cover Hardbottom %**. Values for the estimated percent of various categories of biota covering both "Sand" and "Hardbottom" are entered for each classification present. Any category that is not listed in this section, but was present in the sample area may be entered in one of the fields labeled "Other." All values entered in the "Sand" portion of the section must add to 100%, and all values entered in the "Hardbottom" portion of the section must add to 100%.

- **Go To Species**. As in the previous window, the information entered into the current window must be checked for errors by the program prior to moving on to the next data entry window. Errors will be indicated by a yellow highlighted box. If no errors are detected, select the "Go to Species" icon and the third and final data entry window will open.

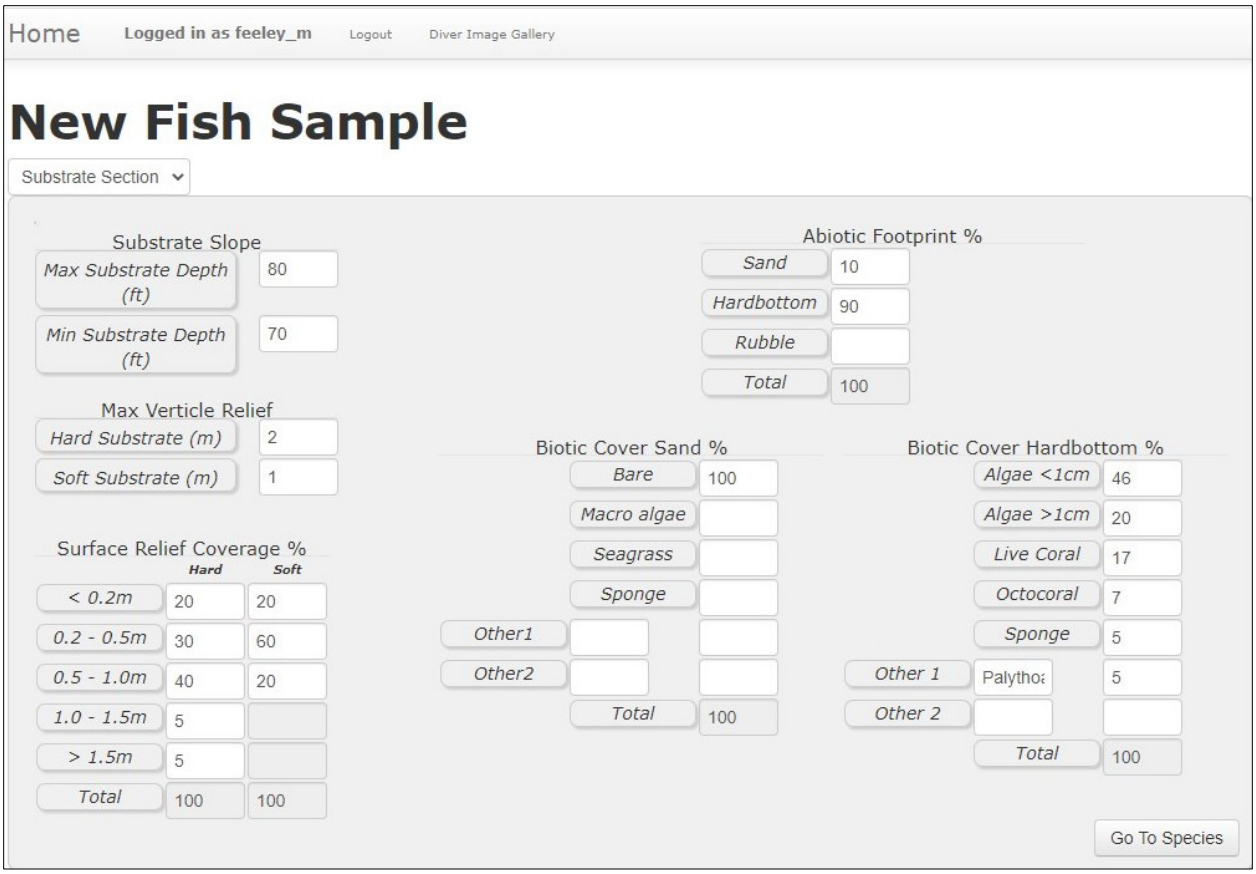


Figure 3. The Substrate Section data entry window.

Entering Fish Species Data

The third window of the RVC Data Entry Program is titled the "Species Section" (Figure 4). This is the data window where fish species that are observed and recorded in the sample area are entered. As in the previous two windows, dynamic data checking is also implemented by the program in this window. The following steps will assist the user with data entry of reef fish species.

To differentiate between species observed during the "First 5 Minutes" of the sampling period and species observed during the "5–10 Minutes" and "After 10 Minutes" periods, click the appropriate time period icon in the top left side of the window. Species observed in the first five minutes will be recorded in a white box, all species recorded after the first five minutes but before ten minutes have passed will be identified with a blue box, and all species recorded after 10 minutes will be noted with a green highlighted box.

- **Species**. Begin typing in the first species box using either the scientific abbreviation or common name of the fish. The program will search for and provide the possible corresponding common names and scientific names as you type in the drop-down menu. Note: scientific names may only be entered using the correct abbreviation for each species (i.e. the first 3 letters of the genus and the first 4 letters of the species). For example, CAR

BART_yellow jack. To begin entering records of observed species choose the appropriate species from the given list.

- **Num**. **Mode. Max. Min.** Next, enter the observed number of individuals of that species in the field labeled "Num." If "Num." is equal to one (i.e. only one individual of that species was observed), the estimated length is entered in the "Mode" field. If "Num." is equal to two or three, enter the estimated lengths in the "Min" and "Max" fields (2) or "Min," "Mode" and "Max" fields (3). When "Num." equals three, make sure the "Mode" field contains the size in between the "Min" and "Max." If "Num." is greater than three, enter the mode size of the individuals, and the minimum and maximum observed size in the "Mode," "Min," and "Max" field, respectively. The "Mode" value must be greater than or equal to the "Min" value and less than or equal to the "Max" value.

- If a species is observed that is not in the species list, first check with the data manager to confirm that the species is not on the list and has not been seen in a previous RVC survey. If it is confirmed that the species has not been seen before, enter data for this species using "*Otherx_new species x*" from the species drop down list. Identify the species in the "Comments" entry field of the "Sample Data" window. It is possible, however extremely unlikely, to add up to three new species for a site. Confirmation with the NOAA data manager or statistician must always occur prior to adding a new species.

- When the record for a particular species is complete, click the icon labeled "Add species," and the new data record will appear in the box below the entry fields. Data for additional species may be added in the same manner. To delete a record, simply click the "Remove Species" icon next to the data entry record for that species.

- The program is designed to allow the user to input multiple size classes within the same species. Therefore, multiple records for the same species may be created, but the "Min" and "Max" size values for each record must not overlap and must be within the same period.

- The program is also designed to allow the user to input multiple individual sizes of a selected species (commercially important snapper and grouper species). Individual sizes may not exceed three per line, may not overlap with a previous record, and the smallest size is entered in the "Min" field, the largest size is entered into the "Max" field and the in-between size is entered into the "Mode" field.

- If it is necessary to go back and edit the Substrate, Sample or Species sections select the appropriate section from the pull-down menu of the section icon in the upper left corner of the page. When all recorded data have been entered and no further editing is required, click the "Validate Species" icon. If no errors are detected a large green check will appear. If an error was detected in the species list a large red "X" will appear. Review and correct the error (generally missing values). Click the "Validate Species" icon again and the "Submit Sample" icon to enter and save the entire record.

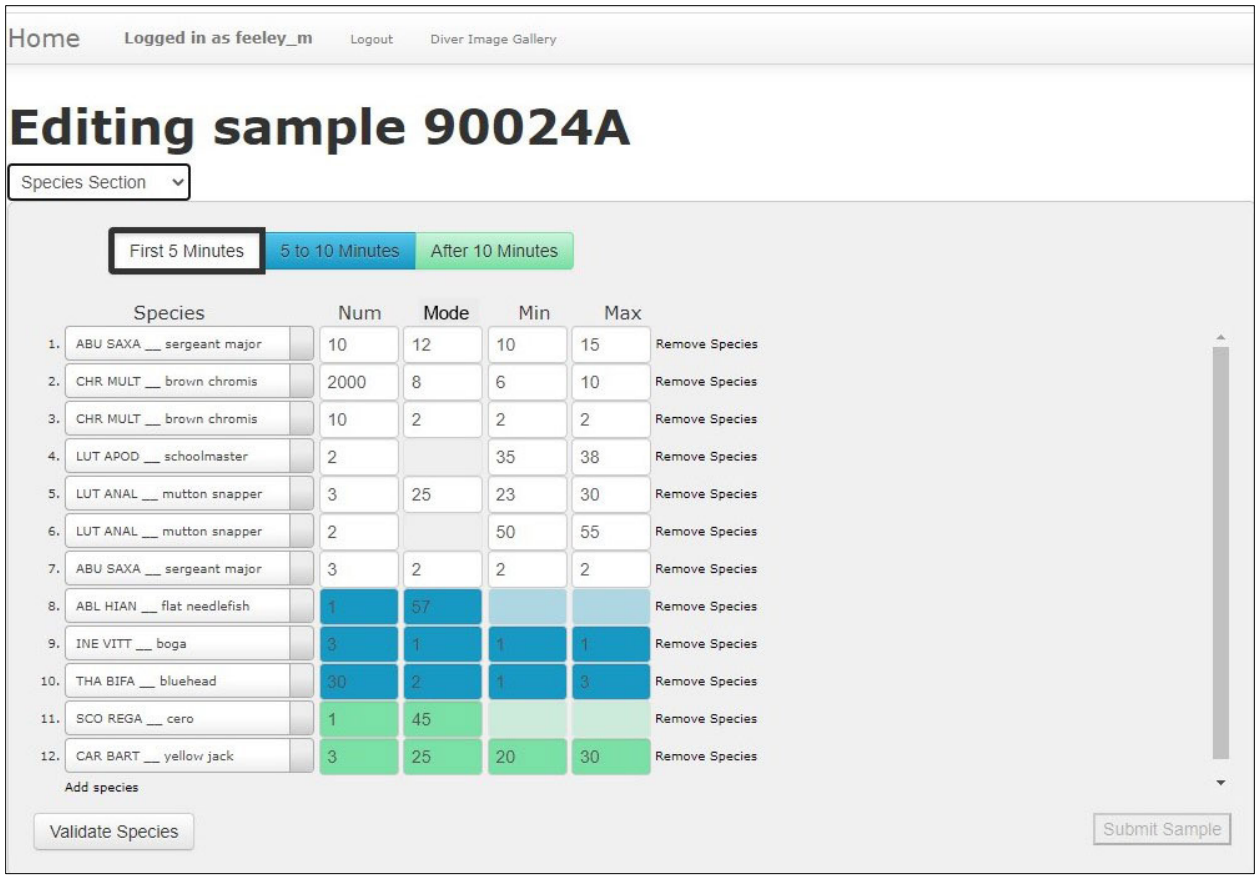


Figure 4. The Species Section data entry window.

Entering the Field/Boat Log

The agency data manager or his designee may enter the Field/Boat Log information. The agency data manager has access to a "Field/Boat Logs" icon on the home page of the RVC Data Entry Program. The following steps will assist the user with data entry of the Field/Boat Log (Figure 5).

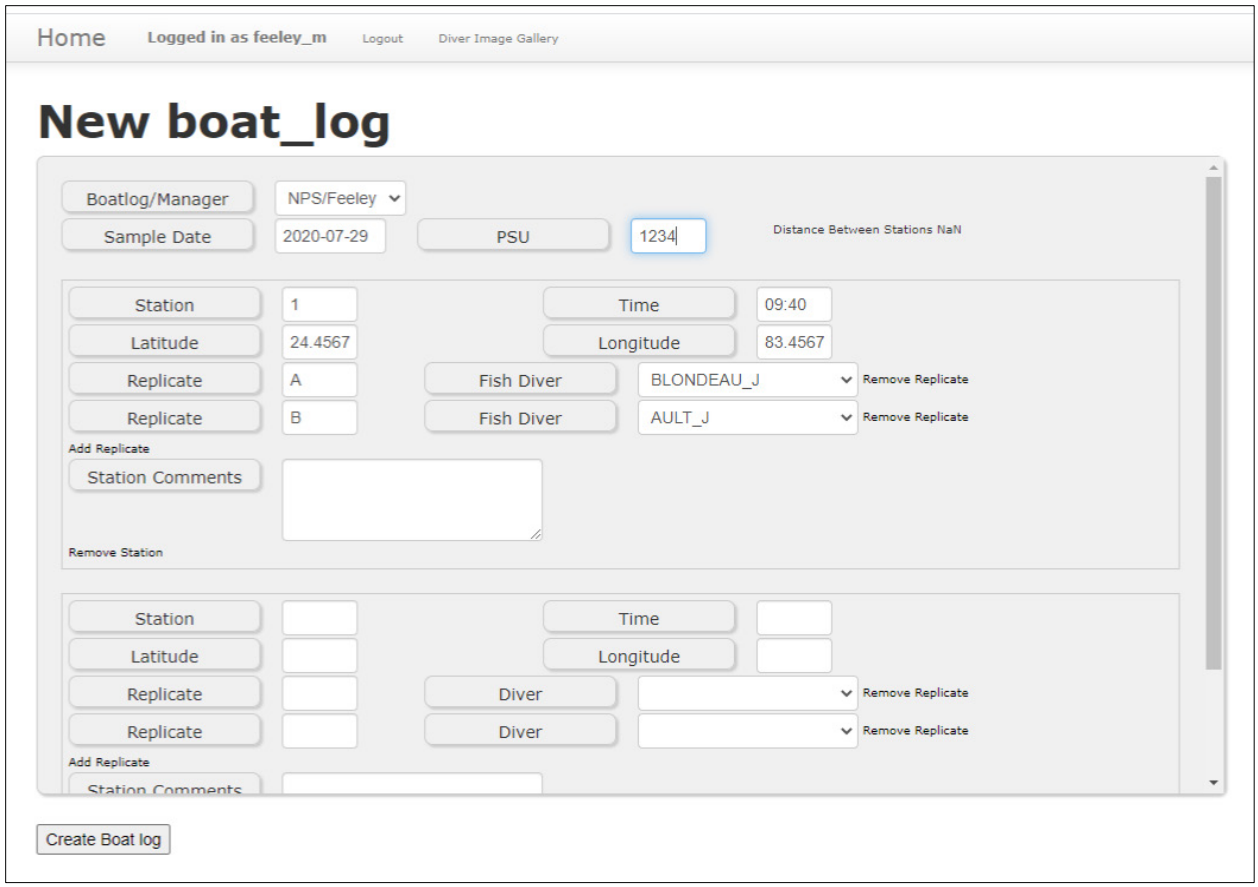


Figure 5. The Boat log data entry window.

- **New Boat Log**. Click on the "New Boat Log" button.
- **Boatlog/Manager**. From the dropdown menu select the last name of the person identified by your agency or the mission as the agency data manager.
- **Sample Date**. This value will default to the current day identified on the computer of the user. Update this value (year-month-day) to the date of the survey if necessary.
- **PSU**. The 4-digit value "Field Number" for the primary sampling unit (PSU) will be provided by the mission coordinator/data manager.
- **Station**. The designation of each "team" number of replicate fish divers in the primary sampling unit. In the current sampling design, this will always be 1.
- **Time**. The estimated time of the beginning of dive. It is important that the time be entered in the proper indicated format. The hour must be a number between 0 and 23, and the minutes must be a number between 0 and 59.

- **Latitude and Longitude**. The coordinates of the team's location are recorded in decimal degrees. It is important that this information is entered in the proper indicated format (xx.xxxxxx°), with longitudinal values expressed as a negative number. There are several sources of this information. The primary source is the information collected from the data cloud generated by the flag mounted GPS unit carried by the divers and attached to the dive flag. If this device fails, the position of surface vessel when divers enter the water is also recorded as a waypoint on a boat mounted / handheld GPS unit and/or is recorded on the Field/Boat Log.

- **Replicate** (or Team letter designation). Letter code identifying the type of survey data being collected by the diver within their dive team. Codes are assigned to diver positions within the team and type of data collected; therefore, diver team codes could change by station as divers potentially rotate. A two-diver fish team consists of a Diver A and Diver B. Additional replicate information may be added by clicking the "Add Replicate" link if additional divers were present (extra fish, LPI or Demo divers).

- **Diver**. Enter the last name of the Divers by clicking the drop-down arrow in each data field and choose the appropriate name from the given list of divers.

- **Station comments**. Any additional information judged to be important by the surface data collector (e.g., waypoint, notation of a specific camera used by a diver) may also be entered.

- A station may be removed by clicking the "Remove Station" link or added by clicking the "Add Station" link. Note: The second station must be removed if only one was sampled (as in the current sampling design).

- *Note:* Distance Between Stations (station 1 vs. station 2) is calculated and displayed as a quick check of invalid or incorrectly entered values.

- **Create Boat Log**. Select the "Create Boat log" button to save the boat log information.

Editing Saved Data

Fish Surveys

First click on the "Fish Samples" icon to manage your fish surveys. All records are listed by Diver, Field ID, Date and Start Time. The program defaults to listing all surveys by Primary Sample Unit, but any heading may be selected to sort records in ascending order by that heading. The user may also search for a record by entering key information under a heading. Once the desired survey is located, highlight the sample and click that survey row. The next window that appears is the sample information page with three icons: "Edit Fish Sample," "Delete Fish Sample" and "Fish Sample List." Click these icons to edit the survey, delete the survey record or return to the master sample list of surveys. To edit a previously saved data record, the user must click the "Edit Fish Sample" icon. When this function is chosen, the program first opens the "Sample" section sheet for the survey, but

the user can navigate to the desired section by selecting it from the Section pull-down menu – either the Sample, Substrate or Species. The user must scroll through the available records to find the record to be edited. When the user selects a particular record, the data files corresponding to that entry may be replaced or corrected. At this point the user can change the previously entered data in each of the three windows. When the editing process is completed, the user must navigate to the Species Section, "Validate Species" and "Submit Sample" again. If the user exits the program without submitting the file, then the data in the edited record *will be lost*.

Boat Log

To manage the boat log records the agency data manager must click on the "Boat log" icon. All records are listed in a table format by Primary Sample Unit, Date, Time, and Comments. The program defaults to listing all surveys by Primary Sample Unit, but any heading may be selected to sort records in ascending order by that heading. The agency data manager may also search for a record by entering key information under a heading. Once the desired survey is located, highlight the sample and click that survey row. The first window that next appears is the sample information page with three icons "Edit Boat log," "Delete Boat log" and "Boat log List." Click these icons to edit the boat log record, delete the boat log record, or return to the master boat log list. To edit a previously saved data record, the user must click the "Edit Boat log" icon. When this function is chosen, the program opens the boat log input page. The user must scroll through the available records in order to find the record to be edited. When the user selects a particular record, the data files corresponding to that entry may be replaced or corrected. At this point the user is able to change the previously entered data. When the editing process is completed, the user must select "Create Boat log" to save the record again. If the user exits the program without submitting the file, then the data in the edited record *will be lost.*

Data QA/QC and Verification

Data proofing and verification are important steps in a monitoring program and there are a series of tasks that must be followed to accurately proof, correct, and verify the data before analysis. If errors are not caught during this phase, they will likely go undetected, introducing error into the database. The most crucial data proofing is carried out by each diver immediately following each dive and at the time the diver enters his data in the RVC Data Entry Program. There is much more to proofing than simply verifying that the values for each variable were entered correctly. The proofers must be aware of the procedures and guidelines related to the data and recognize situations where the variables entered do not seem reasonable.

Initial Diver Proofing

These are the most crucial proofing steps to ensure the accuracy of the data. Initial data proofing is carried out by each diver immediately following each dive and involves reviewing the data sheet for errors, missing data, and legibility, as well as discussing subjective data collection variables (e.g., habitat type and bottom current) with their dive buddy.

Following this, the diver is responsible for providing the agencies with the most correct data. The diver will make a conscious effort to enter the data into the RVC Data Entry Program as accurately and as soon as possible. This includes counting the number of species on the field data sheet and comparing it with what was entered into the program immediately after entering the data, as well as going back through the three windows and checking for any errors prior to saving the record.

Participant-Level Proofing

Data Proofing Criteria

Each agency will be responsible for maintaining the integrity of data collected and entered by their participants. All data collected by any participant and entered into the RVC database is proofed after it is entered. Initial proofing consists of comparing data collected on underwater datasheets to that entered into the database, which is printed out as a proofing report. If there is any inconsistency between data recorded on the data sheet and that entered into the system, that sample is recorded as an error. The number of samples recorded without error divided by the number of samples proofed is the accuracy for that participant. A minimum accuracy level of 90% is maintained for any participant. Each agency will follow their own steps to assure that all participants maintain this level of accuracy.

Proofing Report Procedure

Proofing reports can be created for a specific diver after data for that diver have been entered. The RVC Data Entry Program can generate proofing reports in the common species name style (Figure 6), or the RVC genus/species code style (Figure 7) depending on the diver's preference and mode of recording of fish names on the underwater data sheet. The proofing report is designed to look like the underwater field sheet to ease the proofing process. After the diver has logged in, select the "Fish Samples" icon. Select either the "Download Proofing Report (Species Code)" or "Download Proofing Report (Common Name)" button. After a proofing report has been downloaded and printed, the information on the proofing report is compared to the data on corresponding underwater data sheets as per the data proofing criteria described previously. Each diver can accomplish necessary corrections or changes discovered during the proofing report process using the "Edit" function of the RVC Data Entry Program.

Diver	FEELEY_M	Date	2019-06-10	Field Number	90024A	Boatlog/Manger:	NPS/Feeley
Buddy	RICE_PATRICK	Sample Start	15:01	Habitat	Contiguous reef Spur-Groove	Visibility	40
Dive Start	15:00	Sample End	15:29	Fish Gear	none	Water Temp	83.0
Dive End	15:30	Max Depth	99	Stn Depth	50	Current	none

Substrate Slope	
Max Depth:	75
Min Depth:	30
Max Vertical Relief	
Hard Relief:	1.2
Soft Relief:	0.8

Comments:

Surface Relief %		
	Hard	Avg. Soft
<0.2	20	30
0.2-0.5	20	40
0.5-1.0	40	30
1.0-1.5	20	
>1.5		

Abiotic Footprint	
Sand	
Hard-B	100
Rubble	

Biotic Cover %			
Sand		Hardbottom	
Bare		Algae(<1)	17
Macro Algae		Algae(>1)	10
Seagrass		Live Coral	45
Sponge		Octocoral	10
1:		Sponge	8
2:		1: cyanobacter	10
		2:	

	Period	Species	N/Mod-Min-Max
1	1	sergeant major	10 / 12 - 10 - 15
2	1	brown chromis	2000 / 8 - 6 - 10
3	1	brown chromis	10 / 2 - 2 - 2
4	1	schoolmaster	2 / - 35 - 38
5	1	mutton snapper	3 / 25 - 23 - 30
6	1	mutton snapper	2 / - 50 - 55
7	1	sergeant major	3 / 2 - 2 - 2
8	2	flat needlefish	1 / 57 - -
9	2	boga	3 / 1 - 1 - 1
10	2	bluehead	30 / 2 - 1 - 3
11	3	cero	1 / 45 - -
12	3	yellow jack	3 / 25 - 20 - 30

Figure 6. Example proofing report sheet in common species names style.

Diver	FEELEY_M	Date	2019-06-10	Field Number	90024A	Boatlog/Manger:	NPS/Feeley
Buddy	RICE_PATRICK	Sample Start	15:01	Habitat	Contiguous reef Spur-Groove	Visibility	40
Dive Start	15:00	Sample End	15:29	Fish Gear	none	Water Temp	83.0
Dive End	15:30	Max Depth	99	Stn Depth	50	Current	none

Substrate Slope	
Max Depth:	75
Min Depth:	30
Max Vertical Relief	
Hard Relief:	1.2
Soft Relief:	0.8

Comments:

Surface Relief %		
	Hard	Avg. Soft
<0.2	20	30
0.2-0.5	20	40
0.5-1.0	40	30
1.0-1.5	20	
>1.5		

Abiotic Footprint	
Sand	
Hard-B	100
Rubble	

Biotic Cover %			
Sand		Hardbottom	
Bare		Algae(<1)	17
Macro Algae		Algae(>1)	10
Seagrass		Live Coral	45
Sponge		Octocoral	10
1:		Sponge	8
2:		1: cyanobacter	10
		2:	

	Period	Species	N/Mod-Min-Max
1	1	ABU SAXA	10 / 12 - 10 - 15
2	1	CHR MULT	2000 / 8 - 6 - 10
3	1	CHR MULT	10 / 2 - 2 - 2
4	1	LUT APOD	2 / - 35 - 38
5	1	LUT ANAL	3 / 25 - 23 - 30
6	1	LUT ANAL	2 / - 50 - 55
7	1	ABU SAXA	3 / 2 - 2 - 2
8	2	ABL HIAN	1 / 57 - -
9	2	INE VITT	3 / 1 - 1 - 1
10	2	THA BIFA	30 / 2 - 1 - 3
11	3	SCO REGA	1 / 45 - -
12	3	CAR BART	3 / 25 - 20 - 30

Figure 7. Example proofing report sheet in RVC genus/species code style.

Agency-Level Proofing

Preliminary Annual Master Spreadsheet File Verification

Once divers proof all their entered data for an annual sampling cycle, verification that divers have entered all RVC counts in the RVC data entry system can occur. This initial check by the agency data manager ensures all the samples collected have been entered into the system and incorporates a few checks on fields that are prone to errors.

The agency data manager and the NOAA statistician accomplish this task via a series of record checks of the diver entered data against the Boat/Field Log data in the web-based data entry system. The agency data manager has access to the “Manager Dashboard” icon on the RVC Data Entry Program homepage. After clicking on the dashboard icon, a tally sheet of all agency diver’s Boat Log Counts and Fish Counts is presented (along with possible LPI Counts and Demographic Counts). The sum of the sample counts should equal the sum of the Boat Log Counts for each diver. By selecting

the “Crosscheck Boat log and Samples” icon, the agency data manager can review and cross check the data with tables that list records by Diver, Date and Field ID. One table lists all boat logs that do not have a matching diver sample and the second table lists all surveys that have been entered but do not have a matching Boat Log entry. If missing diver samples are discovered during this check, divers will need to be contacted to resolve the missing data. The resolution may include a revision to Boat/Field Log or entry of missing RVC sample data. In addition to the missing data check, some preliminary checks of certain data elements such as Field Numbers, outliers in lengths or numbers of fishes, sampling dates, etc. is performed. The data proofing procedures described previously should reduce or eliminate occurrences of questionable data elements. However, should questionable data elements be discovered during this preliminary check, individual divers might need to be contacted for resolution. Once there are no longer any ‘mismatches’ between diver samples and boat logs in the “Crosscheck Boatlog Sample” tables, the agency level QA/QC is completed and ready for next level.

Final Verification

Below is the flow of the final verification procedures and creation of the analysis ready (AR) dataset.

1. **Import Boatlogs**
    a. Check for duplicates (fieldid)



2. **Import Samples**
    a. Merge with boatlog (by = YEAR, MONTH, DAY, fieldid, DIVER_NR)
    b. Check for mismatches
    c. Check for duplicates



3. **Check for gross errors in Lat/Lon** (missing negative longitude values)
    a. Check distance from planned



4. **Output actual coordinates for GIS plotting and spatial join of MAPGRID attributes**
    a. Check again for errors in lat/lon
    b. Export boatlog (with map attributes)



5. **Import substrate initial**

a. Check for outliers (SUBSTRATE_MAX_DEPTH, SUBSTRTATE_MIN_DEPTH, HARD_VERTICAL_RELIEF, SOFT_VERTICAL_RELIEF)

b. Run code to correct relief coverages

c. Export to sample_corrected



6. **Import substrate corrected, sample corrected, boatlog corrected**

    a. Merge sample and boatlog and check for mismatches

    b. Merge sample and substrate and check for mismatches

    c. Check for duplicates in master sample codes

    d. Within station (buddy pair) substrate variable checks

        i. UNDERWATER_VISIBILITY, SUBSTRATE_MAX_DEPTH, SLOPE, HARD_VERTICAL_RELIEF, SAND_PERCENTAGE, HARD_VERTICAL categories,

    e. Export verify_2b_hab_region_year

        i. This is used in the habitat/strata verification process



7. **Import species corrected**

    a. Run species before 5 code to export verify_3_sppbef5_region_year

        i. This creates a field for number of unique species seen in the first 5 minutes and is used in the verification process



8. **Import species corrected**

    a. Merge with sample to check for mismatches

    b. Check for any null SPECIES_CD, SAMPLE_TYPE_NR or SAMPLE_TYPE_NR > 1

    c. Check for duplicate species entries (by = MSN, SP_CD, NUM, MODE, MIN, MAX)

    d. Create species histograms and look through every species for outliers in NUMBER_OF_INDIVIDUAL, and sizes

e. Rerun check for duplicates in case you introduced them when fixing outliers



9. **Create verify_4_psu_region_year file**

    a. Imports boatlog, sample, verify_2b_hab_region_year and verify_3_sppbef5_region_year

    b. Exports verify_4_psu_region_year



10. **Create primary_sample_unit_info, station_info, sample_info**

    a. Check that each PSU has unique (HABITAT_CD, DEPTH_CLASS, SUBREGION, ADMIN)

    b. Export *info files for AR dataset creation



11. **Create AR dataset**



12. **Test AR dataset**

During most of the QAQC process, data fields are being checked for outliers and agreement between buddy pairs. When discrepancies are found, the original underwater datasheet is used to verify and make appropriate edits. Prior to creation of the 'Analysis Ready' dataset, a PSU verification is made to ensure SSU's are in the same stratum (for 2-stage designs) or that SSU and mapped stratum agree. This is a crucial step in the process to ensure that a sample is analyzed in the correct stratum and that final metric estimates (i.e. density, occurrence) are correct. To complete this step, every PSU is looked at individually with all SSU data tabulated. Any habitat, depth, rugosity, discrepancies between stations and map data are scrutinized. Diver data, GIS map data (habitat, bathymetry, aerial imagery, etc.) and site photos are used to determine that each PSU, and all station level data within, are in agreement. There are times where the paired SSU's or the individual surveys do not agree and based on the information are obviously not in the same stratum and grid cell. In these cases, the SSU's are split into two separate PSU's for analysis. In a single-stage design, only one survey would represent an SSU (instead of two diver surveys). An NPS Data Quality Standards document is included in at the end of this SOP.

Once the PSU verification process is complete and all changes have been documented in the PSU corrected file, the 'Analysis Ready' dataset is created. This dataset is a flat .csv file organized at the SSU level. A record of the original sheets [sample, substrate and species) is still preserved. Each individual species recorded goes through a discrete triangular distribution algorithm to distribute the

number (NUM) seen into five (5) distinct length sizes based on the "MIN," "MAX," and "MODE" values entered by the diver (Ault et al. 1998; Meester et al. 1999; Serafy et al. 2003; Serafy et al. 2007; Faunce and Serafy 2007; Jones et al. 2010). Table 2 provides a small example of length data from the RVC data entry program. The method uses a smoothing function based on the assumption that these data have a mode and upper and lower bounds based on the diver collected data (Jones et al. 2010). Five new length variables are assigned a length value and corresponding number of individual fish as *L1-A1, L2-A2….L5-A5,* the triangular distribution algorithm (see Table 3), which creates the resulting *L-A* dataset (Table 4). The 5 *L-A* variables are combined into one length and one number variable and transposed (Table 5). A discrete size distribution is created, rather than continuous, which more accurately reflects the size structure of schooling or aggregated conspecific fish (Hoare et al. 2005). All species are combined for each SSU (A and B), and transposed, including species that are absent ".", then transposed back to the original table format with 1 row per species per SSU with missing values (Table 6–8). Next, at each SSU, every species not seen is assigned a NUM = 0. The diver-pair species and length data are then averaged for each species-length combination to get final SSU-level species information (Table 9). The relative frequency of the average catch determines the final number variable for each size used in the analysis ready data set. This makes for a rather large flat file, but ensures correct metric estimates (i.e. density, occurrence) at any level.

Table 2. Example of length and number variables from the RVC data entry program. The diver has recorded 24 Yellowtail Snapper with a mode (moL) length of 13 cm, minimum (minL) of 11 centimeters (4.3 in) and maximum (maxL) of 16 centimeters (6.3 in); two additional fish are also recorded at 28 and 33 centimeters (11 and 13 in).

species	number	moL	minL	maxL
yellowtail snapper	24	13	11	16
black grouper	1	67	67	67
mutton snapper	2	34	34	52
stoplight parrotfish	3	12	17	39
yellowtail snapper	2	28	28	33

Table 3. A discrete triangular distribution algorithm distributes the number seen into five (5) distinct length sizes based on the "MIN" (minL), "MAX" (maxL), and "MODE" (moL) values entered by the diver. Five new length variables are assigned a length value and corresponding number of individual fish as *L1-A1, L2-A2….L5-A5,* where n = number.

Number	**L1**	**A1**	**L2**	**A2**	**L3**	**A3**	**L4**	**A4**	**L5**	**A5**
1	–	–	–	–	moL	1	–	–	–	–
2	minL	1	–	–	–	–	–	–	maxL	1
3	minL	1	–	–	moL	1	–	–	maxL	1
4	minL	1	–	–	moL	2	–	–	maxL	1
5 - 99	minL	1	$\frac{moL\ -minL}{2}$	$\frac{number\ -(A1+A2+A3)}{2}$	moL	0.5*number	$\frac{maxL\ -moL}{2}$	A2	maxL	1
≥ 100	minL	0.01*number	$\frac{moL\ -minL}{2}$	$\frac{number\ -(A1+A2+A3)}{2}$	moL	0.5*number	$\frac{maxL\ -moL}{2}$	A2	maxL	0.01*number

Table 4 Five new length variables are assigned a discrete length value and assigned a corresponding number of individual fish by the algorithm *L1-A1, L2-A2, …. L5-A5* which creates the *L-A* dataset.

species	L1	A1	L2	A2	L3	A3	L4	A4	L5	A5
yellowtail snapper	11	1	12	5	13	12	14.5	5	16	1
black grouper	–	–	–	–	67	1	–	–	–	–
mutton snapper	34	1	–	–	–	–	–	–	52	1
stoplight parrotfish	12	1	–	–	17	1	–	–	39	1
yellowtail snapper	28	1	–	–	–	–	–	–	33	1

Table 5. Transposed data based on the 5 length and number value (*L-A)* variables combined into one length and one number variable.

species	length	number
yellowtail snapper	11	1
yellowtail snapper	12	5
yellowtail snapper	13	12
yellowtail snapper	14.5	5
yellowtail snapper	16	1
yellowtail snapper	28	1
yellowtail snapper	33	1
black grouper	67	1
mutton snapper	34	1
mutton snapper	52	1
stoplight parrotfish	12	1
stoplight parrotfish	17	1
stoplight parrotfish	39	1

Table 6. All recorded species are combined for each SSU (A and B).

–	atoll_cd	zone_depth	PSU	SSU	num
CHMI	FFS	FRF_M	FFS-203	A	2.5
BOBI	FFS	FRF_M	FFS-203	A	1.0
NAUN	FFS	FRF_M	FFS-203	A	1.5
CHMI	FFS	FRF_M	FFS-203	B	7.0
CHPE	FFS	FRF_M	FFS-203	B	0.5
STFA	FFS	FRF_M	FFS-203	B	4.5

Table 7. All species data is then transposed, and species that are absent "–" are added per secondary sampling unit (SSU).

atoll_cd	zone_depth	PSU	SSU	CHMI	BOBI	NAUN	CHPE	STFA
FFS	FRF_M	FFS-203	A	2.5	1.0	1.5	–	–
FFS	FRF_M	FFS-203	B	7.0	–	–	0.5	4.5

Table 8. The data is transposed back to the original table format with 1 row per species per SSU.

species	atoll_cd	zone_depth	PSU	SSU	num
CHMI	FFS	FRF_M	FFS-203	A	2.5
BOBI	FFS	FRF_M	FFS-203	A	1.0
NAUN	FFS	FRF_M	FFS-203	A	1.5
CHPE	FFS	FRF_M	FFS-203	A	–
STFA	FFS	FRF_M	FFS-203	A	–
CHMI	FFS	FRF_M	FFS-203	B	7.0
BOBI	FFS	FRF_M	FFS-203	B	–
NAUN	FFS	FRF_M	FFS-203	B	–
CHPE	FFS	FRF_M	FFS-203	B	0.5
STFA	FFS	FRF_M	FFS-203	B	4.5

Table 9. Buddy pair averaging procedure for numbers-at-length by second stage unit (SSU). In this example for Yellowtail Snapper with 2 replicates (A and B), diver A recorded 24 fish and diver B recorded 20 fish.

station	rep_cd	species	length	rep_number	rel_frq	av_catch	station_number
1	A	yellowtail snapper	11	1	0.02273	22	0.5
1	A	yellowtail snapper	12	5	0.11364	22	2.5
1	A	yellowtail snapper	13	12	0.27273	22	6
1	A	yellowtail snapper	14.5	5	0.11364	22	2.5
1	A	yellowtail snapper	16	1	0.02273	22	0.5
1	A	yellowtail snapper	28	1	0.02273	22	0.5
1	A	yellowtail snapper	33	1	0.02273	22	0.5
1	B	yellowtail snapper	14	1	0.02273	22	0.5
1	B	yellowtail snapper	15.5	4	0.09091	22	2
1	B	yellowtail snapper	17	10	0.22727	22	5
1	B	yellowtail snapper	18.5	4	0.09091	22	2
1	B	yellowtail snapper	20	1	0.02273	22	0.5
–	–	–	Sum =	44	–	–	–

Data Portal

After final verification, the 'Analysis Ready' dataset and diver-level sample, species, substrate and boat log information are available for distribution. Data variables for all output files are defined in Appendix A. The South Florida data are available annually beginning in 1999 (biennially, beginning 2014) and the USVI point count data beginning in 2017 (biennially). In addition, stratum data on the number of primary sample units per stratum (NTOT) for each year and region is available to calculate important weighting factors for domain-wide estimates. Taxonomic and life history data for species observed in the Reef Visual Census are also on file. This includes scientific and common names, species code (first three letters of the generic name, and first four letters of the specific name). It also includes life history data for select species including: LM - median length at maturity (mm), LC - minimum length at capture (mm), WLEN_A - linear coefficient of the allometric growth equation in g/mm, WLEN_B - exponential coefficient of the allometric growth equation (Stevens et al. 2019). Analysis-ready datasets and data attributes for the Florida Keys, Tortugas, St. Thomas/St. John and St. Croix regions are shared on-line at the South Florida National Coral Reef Monitoring Program data portal managed by NOAA: https://grunt.sefsc.noaa.gov/rvc_analysis20/.

Data Archiving

In addition to the SEFSC data portal where 'Analysis Ready,' stratum and taxonomic datasets can be downloaded for each available year, all data are archived at NOAA's National Centers for Environmental Information (NCEI). The NCEI provides public access of all data for each sampling year and region in a citable accession package that includes:

1. 'Analysis Ready,' stratum, and taxonomic datasets (for immediate analysis)
2. Diver level 'Raw' data, including, sample, substrate, species, stratum NTOT, boatlog and PSU files
3. Data documentation, including data dictionaries and protocols
4. GIS shapefiles

Data Cataloguing

All data are catalogued, and backups created such that several independent copies of the data are held in different physical locations but where they are accessible to all groups involved. These locations will include UM-RSMAS, NPS, NOAA, and FWRI. The following cataloguing guidelines will be followed:

- All agency data entry and proofing will be completed by December 1 each sampling year.
- After collation at NOAA, all raw data will be available to all agencies.
- Data are made available to all agencies as soon as annual QA/QC is completed.

- All groups agree on data release to outside data requests prior to actual release.
- Data collected the previous year will generally be available for download on the NOAA data portal by August 1 of the following year. For example, 2025 surveys will be posted to the portal by August 1, 2025.

Literature Cited

Ault J. S., J. A. Bohnsack and G. A. Meester. 1998. A retrospective (1979–1996) multispecies assessment of coral reef fish stocks in the Florida Keys. Fishery Bulletin 96:395–414.

Faunce C. and J. E. Serafy. 2007. Nearshore habitat use by gray snapper (*Lutjanus griseus)* and bluestriped grunt (*Haemulon sciurus)*: environmental gradients and ontogeneticshifts. Bulletin of Marine Science 80:473–495.

Hoare D. J., J. Krause, N. Peuhkuri and J.-G. J. Godin. 2005. Body size and shoaling in fish. Journal of Fish Biology 57: 1351 – 1366.

Jones, D. L., J. F. Walter, E. N. Brooks and J. E. Serafy. 2010. Connectivity through ontogeny: fish population linkages among mangrove and coral reef habitats.
Marine Ecology Progress Series 401:245-258.

Meester G. A., J. A. Ault, J. A. Bohnsack. 1999. Visual censusing and the extraction of average length as an indicator of stock health. Naturalista Siciliano 23:205–222.

Serafy J. E., C. H. Faunce and J. J. Lorenz. 2003. Mangrove shoreline fishes of Biscayne Bay, Florida. Bulletin of Marine Science 72:161–180.

Serafy J. E., M. Valle, C. H. Faunce, J. Luo. 2007. Species-specific patterns of fish abundance and size along a subtropical mangrove shoreline: an application of the delta approach. Bulletin of Marine Science 80:609–624.

Stevens, M. H., S. G. Smith and J. S. Ault. 2019. Life history demographic parameter synthesis for exploited Florida and Caribbean coral reef fishes. Fish and Fisheries 00:1–22. DOI: 10.1111/faf.12405

Appendix A. Variables in Data Output Files and Data Verification Files

This appendix includes metadata tables that describe the variables included in RVC data (sample, substrate, species, and boatlog) output files and the variables and R/excel names included in the data verification files. Tables 10 to 16

Table 10. Variables in the sample output file (_Raw_Sample.csv). Diver level sample/site information collected using the RVC methodology.

Variable	Description
MASTER_SAMPLE_CD	unique sample identifier consisting of year, month, day, sample time, diver number
DIVER_NR	primary diver number
BUDDY_NR	diver number of primary diver's buddy
SAMPLE_TYPE_NR	all sample types are 1, indicating Bohnsack-Bannerot stationary point count
Divhab	integer representing habitat name
Habname	habitat name
YEAR	year
MONTH	month
DAY	day
Divbegh	dive begin hour
Divbegm	dive begin minute
Divendh	dive end hour
Divendm	dive end minute
SAMPLE_START_TIME_HOURS	sample start hour
SAMPLE_START_TIME_MINUTES	sample start minute
SAMPLE_END_TIME_HOURS	sample end hour
SAMPLE_END_TIME_MINUTES	sample end minute
Divdepth	max dive depth
SAMPLE_DEPTH	Sample depth, in feet
FISHING_GEAR	field indicating fishing gear or debris present
Fieldidi	ID of sample consisting of primary sample unit, station number and replicate
UNDERWATER_VISIBILITY	Visibility underwater in meters
WATER_TEMP	underwater temperature in f
CURRENT	underwater current low, medium, high
SAMPLE_DESCR	diver notes

Table 11. Variables in the boat log output file (_Boatlog.csv). Sample level information includes the primary sample unit (PSU) number, date, time, latitude, longitude and divers involved.

Variable	Description
Field id	ID of sample consisting of primary sample unit, station number and replicate
YEAR	year
MONTH	month
DAY	day
PRIMARY_SAMPLE_UNIT	A code indicating the primary sample unit in which a sample was collected
STATION_NR	station number
REPLICATE_CD	replicate code
Logbegh	boat dive log begin hour
Logbegm	boat dive log begin minute
DIVER_NR	primary diver number
LAT_DEGREES	latitude
LON_DEGREES	longitude
LOG_COMMENTS	boat dive log comments
REGION	A code indicating the region in which a sample was taken.
GRID_ID	A number indicating gridded shapefile number
DEPTH	Maximum depth, in meters, of secondary sampling unit.
HABITAT	A code indicating the habitat type–
DEPTH_STRAT	A code indicating less than or greater than 12 meters depths– Caribbean and GOM regions only–
SUB_REGION	Code indicating a geographical area in Region
ADMIN	Code indicating a named managed area
STRAT	A code indicating the stratum in which a sample was taken– Differs by region.

Table 12. Variables in the substrate output file (_Raw_Substrate.csv). Diver level substrate and benthic information collected using the RVC methodology.

Variable	Description
MASTER_SAMPLE_CD	unique sample identifier consisting of year, month, day, sample time, diver number
SUBSTRATE_MAX_DEPTH	max depth (f) in diver cylinder
SUBSTRATE_MIN_DEPTH	min depth (f) in diver cylinder
HARD_VERTICAL_RELIEF	max height of hard substrate (m)
SOFT_VERTICAL_RELIEF	max height of soft substrate (m)
h_relcat_0	proportion of hard substrate with height from 0-20 cm
h_relcat_1	proportion of hard substrate with height from >20 - 50 cm
h_relcat_2	proportion of hard substrate with height from >50 - 100 cm
h_relcat_3	proportion of hard substrate with height from >100 - 150 cm
h_relcat_4	proportion of hard substrate with height from >150 cm
s_relcat_0	proportion of soft substrate with height from 0-20 cm
s_relcat_1	proportion of soft substrate with height from >20 - 50 cm
s_relcat_2	proportion of soft substrate with height from >50 - 100 cm
s_relcat_3	proportion of soft substrate with height from >100 - 150 cm
s_relcat_4	proportion of soft substrate with height from >150 cm
SAND_PERCENTAGE	abiotic cover sand percentage
HARDB_PERCENTAGE	abiotic cover hardbottom percentage
RUBBLE_PERCENTAGE	abiotic cover rubble percentage
SAND_BARE	biotic cover percentage in sand that is bare
SAND_MACRO_ALGAE	biotic cover percentage in sand that has macro algae
SAND_SEAGRASS	biotic cover percentage in sand that is seagrass
SAND_SPONGE	biotic cover percentage in sand that is sponges
s_pcov_oth1_lab	diver labeled other category
s_pcov_oth2_lab	diver labeled other category
s_pcov_oth1	biotic cover percentage in sand that is labeled other_1
s_pcov_oth2	biotic cover percentage in sand that is labeled other_2
HARDBOTTOM_ALGAL_TURF	biotic cover percentage on hardbottom that has turf algae

Variable	Description
HARDBOTTOM_MACRO_ALGAE	biotic cover percentage on hardbottom that has macro algae
HARDBOTTOM_LIVE_CORAL	biotic cover percentage on hardbottom that has live coral
HARDBOTTOM_OCTOCORAL	biotic cover percentage on hardbottom that has octocoral
HARDBOTTOM_SPONGE	biotic cover percentage on hardbottom that is sponges
h_pcov_oth1_lab	diver labeled other category
h_pcov_oth2_lab	diver labeled other category
h_pcov_oth1	biotic cover percentage on hardbottom that is labeled other_1
h_pcov_oth2	biotic cover percentage on hardbottom that is labeled other_2

Table 13. Variables in the species output file (rvc_species.xls). Diver level species information collected using the RVC methodology.

Variable	Description
MASTER_SAMPLE_CD	unique sample identifier consisting of year, month, day, sample time, diver number
SPECIES_CD	A code indicating the species for a sample– Consists of the first three letters of the genus and the first four of the species name–
NUMBER_OF_INDIVIDUALS	The number of individuals of a given species and length observed in a secondary sampling unit–
AVERAGE_LENGTH	average length of most of the species from number of individuals (cm)
MIN_LENGTH	minimum length of species (cm)
MAX_LENGTH	maximum length of species (cm)
TIME_SEEN	A number indicating when, during sampling, an individual was observed.
instance_field	legacy field referencing order of species entered into database
instance_key	legacy field referencing time period that species entered into database

Table 14. Variables in the verification file used (verify_2.xls).

R/Excel name	Variable	Description
Field id	–	ID of sample consisting of primary sample unit, station number and replicate
stn	STATION_CD	–
Rep	REPLICATE_CD	–
Year	YEAR	Year
Month	MONTH	Month
Day	DAY	Day
Divnum	DIVER_NR	Primary diver number
pr_num	–	–
pr_mod	–	–
Hab	HABITAT_NR	–
Divhab	–	–
Maphab	–	–
sampdepth	SAMPLE_DEPTH	Sample depth, in feet
Mapgrid	MAPGRID_CD	–
Region	REGION_NR	–
Subreg	SUBREGION_NR	–
Mpa	MPA_NR	–
pr_unit	PRIMARY_SAMPLE_UNIT	A code indicating the primary sample unit in which a sample was collected

Table 15. Variables in the map grid ground truthing file (mapgrid_groundtruth.txt) used in verification.

R/Excel name	Variable	Description
Mapgrid	MAPGRID_CD	–
Year	YEAR	–
Month	MONTH	–
Day	DAY	–
Maphab	–	–
Obshab	HABITAT_NR	–
hab_chng	–	–
Depcat	–	–
Obsdepcat	–	–
dep_chng	–	–

Table 16. Variables used in the regional species master file (species_master_region.xls) used in verification.

R/Excel name	Variable	Description
Species	SPECIES_CD	A code indicating the species for a sample– Consists of the first three letters of the genus and the first four of the species name–
Latin	SCIENTIFIC_NAME	–
Common	COMMON_NAME	–
Family	FAMILY_NAME	–
com_fam	COMMON_FAMILY	–
Specnum	SPECIES_NR	–
Famnum	FAMILY_NR	–
max_size	MAX_SIZE	–
Spectype	SPECIES_TYPE	–
Bbtrawl	–	–
nps_creel	–	–
rvc_FL	–	–

Appendix B. Final QAQC function list (R programming)

This appendix includes the R programing code used to complete the final data merger and verification process and creation of the analysis ready data sheets– This code was developed by Jerry Ault and Steve Smith (University of Miami) and Jeremiah Blondeau (NOAA – Southeast Fisheries Science Center)– Individual R-code functions that support the verification process are highlighted–

```
createCorrectedSubstrate <- function(region, year) {

  a <- hard_relief_correction(substrate = substrate_initial_data(region, year))

  write_xlsx(a, path = paste(getwd(), "/", region, "_", year, "_substrate_corrected–xlsx", sep = ""))

}


}


hard_relief_correction <- function(substrate) {

 a <-  substrate %>%

       mutate_if(is–numeric, ~replace(–, is–na(–), 0)) %>%

       filter(!h_relcat_0 < SAND_PERCENTAGE)

 b <-  substrate %>%

       mutate_if(is–numeric, ~replace(–, is–na(–), 0)) %>%

       filter(h_relcat_0 < SAND_PERCENTAGE) %>%

       mutate(h_rel_sum = h_relcat_0 + h_relcat_1 + h_relcat_2 + h_relcat_3 + h_relcat_4, h_rel_tot = (100 -
SAND_PERCENTAGE)) %>%

       mutate(h_relcat_0 = round(((h_relcat_0/h_rel_sum)*h_rel_tot) + SAND_PERCENTAGE), 0) %>%

       mutate(h_relcat_1 = round(((h_relcat_1/h_rel_sum)*h_rel_tot)), 0) %>%

       mutate(h_relcat_2 = round(((h_relcat_2/h_rel_sum)*h_rel_tot)),0) %>%

       mutate(h_relcat_3 = round(((h_relcat_3/h_rel_sum)*h_rel_tot)), 0) %>%

       mutate(h_relcat_4 = round(((h_relcat_4/h_rel_sum)*h_rel_tot)), 0) %>%

       select(-c(h_rel_sum, h_rel_tot, "0"))

  c <- bind_rows(a,b)

 return(c)

 }

 }
```

```
exportActualCoords <- function(region, year) {
 a = read_jumplog(region, year) %>%
     select( YEAR,  MONTH,  DAY,  siteid,  STATION_NR, REPLICATE_CD,  logbegh, logbegm, DIVER_NR,  lat,  lon,
     LOG_COMMENTS ) %>%
  mutate(
    lat = format(lat, nsmall = 5),
    lon = format(lon, nsmall = 5),
    LOG_COMMENTS = replace_na(LOG_COMMENTS, "")
  ) %>%
  ) %>%
  filter(REPLICATE_CD == "A" | REPLICATE_CD == "B")
  write–csv(a, paste(region, year, "actual_latlon–csv", sep = "_"))
 }


 }

exportHabitat <- function(region, year) {
 a = corrected_data(region, year)



 b = merge(a$sample,
         a$substrate,
         by = c("MASTER_SAMPLE_CD"),
         all = T) %>%
  select(
    fieldid,
    siteid–x,
    STATION_NR–x,
    REPLICATE_CD–x,
    HARD_VERTICAL_RELIEF,
    slope,
```

```
    wt_rug,
    SAND_PERCENTAGE,
    MASTER_SAMPLE_CD
   )
   )
 write–csv(b, file = paste("verify_2b_hab_", region, year, "–csv", sep = ""))
}

}

exportSppBefFive <- function(region, year) {
 a = sppbeffive(region, year)
 write–csv(a, paste("verify_3_sppbef5_", region, year, "–csv", sep = ""))
 }

 }

exportVerifiedHab <- function(region, year) {
 a = corrected_data(region, year)
 b = read–csv("verify_2b_hab_prico2019–csv")
 c = read–csv("verify_3_sppbef5_prico2019–csv")
 d = merge(
  a$boatlog,
  a$sample,
  by = c("fieldid", "DIVER_NR", "YEAR", "MONTH", "DAY"),
  all = T
 ) %>%
 ) %>%
  select(  DIVER_NR, divhab, habname, YEAR,  MONTH, DAY, SAMPLE_DEPTH, fieldid, GRID_ID, HABITAT,
           DEPTH, SUBREGION, ADMIN, MASTER_SAMPLE_CD ) %>%
  merge(–, b, by = "MASTER_SAMPLE_CD", all = T) %>%
```

```
  merge(–, c, by = "MASTER_SAMPLE_CD", all = T) %>%
  select( fieldid–x, STATION_NR–x,  REPLICATE_CD–x, YEAR, MONTH,  DAY, DIVER_NR, siteid–x, habname,
          divhab,HABITAT, SAMPLE_DEPTH, HARD_VERTICAL_RELIEF, slope, wt_rug, SAND_PERCENTAGE,
          sppbeffive, GRID_ID, DEPTH, SUBREGION, ADMIN ) %>%
  rename(
   fieldid = fieldid–x,
   STATION_NR = STATION_NR–x,
   REPLICATE_CD = REPLICATE_CD–x,
   siteid = siteid–x
  )
  )
  write_xlsx(d, paste("verify_4_psu_", region, "_", year, "–xlsx", sep = ""))
 }

 }

substrate_initial_data <- function(region, year) {
 dir <- paste(getwd(), "/", region, "_", year, "_", sep = "")
 substrate <-
  read_fish_substrate(
   paste(dir, "samples–xlsx", sep = ""),
   paste(region, "_", year, "_", "substrate_initial", sep = "")
   return(substrate)
 }

 }

corrected_data <- function(region, year) {
 dir <- paste(getwd(), "/", region, "_", year, "_", sep = "")
```

```
boatlog <-
 read_boatlog(
  paste(dir, "boatlogs–xlsx", sep = ""),
  paste(region, "_", year, "_", "boatlog_corrected", sep = "")
 )
 )
 sample <-
 read_fish_sample(
  paste(dir, "samples–xlsx", sep = ""),
  paste(region, "_", year, "_", "sample_corrected", sep = "")
 )
 )
 substrate <-
 read_fish_substrate(
  paste(dir, "samples–xlsx", sep = ""),
  paste(region, "_", year, "_", "substrate_corrected", sep = "")
 )
 )
sub_merge <-  substrate %>%
 left_join(select(
  sample,
  "MASTER_SAMPLE_CD",
  "siteid",
  "STATION_NR",
  "REPLICATE_CD")) %>%
 mutate(wt_rug = ((h_relcat_0 / 100) * –1) +
           ((h_relcat_1 / 100) * –35) +
           ((h_relcat_2 / 100) * –75) +
           ((h_relcat_3 / 100) * 1–25) +
           ((h_relcat_4 / 100) * (( HARD_VERTICAL_RELIEF + 1–5 ) / 2))) %>%
 mutate(slope = SUBSTRATE_MAX_DEPTH - SUBSTRATE_MIN_DEPTH)
```

```
  species <-
    read_fish_species(
      paste(dir, "samples–xlsx", sep = ""),
      paste(region, "_", year, "_", "species_corrected", sep = "")
    )
    )
  sp_merge <-  species %>%
    left_join(select( sample, "MASTER_SAMPLE_CD", "siteid", "STATION_NR", "REPLICATE_CD" ))



  return(list(
    boatlog = boatlog,
    sample = sample,
    substrate = sub_merge,
    species = sp_merge
  ))
}



  ))
}


read_jumplog <- function(region, year) {
  dir <- paste(getwd(), "/", region, "_", year, "_", sep = "")
  x <-
    read_xlsx(
      path = paste(dir, "jumplogs–xlsx", sep = ""),
      sheet = paste(region, "_", year, "_", "jumplog_corrected", sep = "")
```

```
  )
  )
 x$date <- as–Date(x$date, format = "%Y–%m–%d")
 x <- parse_fieldid(x)
 x <- split_date(x)
 x$logbegh <- substr(x$time, 1, 2)
 x$logbegm <- substr(x$time, 4, 6)
 return(x)
}

}

read_boatlog <- function(file, worksheet, ——) {
 x <- read_xlsx(file, sheet = worksheet)
 x$date <-
  as–Date(paste(x$YEAR, x$MONTH, x$DAY, sep = "–"), format = "%Y–%m–%d")
 return(x)
}

}

read_fish_sample <- function(file, worksheet, ——) {
 x <- read_xlsx(file, worksheet)
 x$date <-
  as–Date(paste(x$YEAR, x$MONTH, x$DAY, sep = "–"), format = "%Y–%m–%d")
 x <- parse_fieldid(x)
 return(x)
}

}
```

```
read_fish_substrate <- function(file, worksheet, ——) {

 x <- read_xlsx(file, sheet = worksheet)
 return(x)
}

}

read_fish_species <- function(file, worksheet, ——) {
 x <- read_xlsx(file, sheet = worksheet)

 return(x)
}

}

check_jump_duplicates <- function(region, year) {
 jump = read_jumplog(region, year)
 dup_fieldid = find_duplicate(jump, "fieldid")
 dup_date_time_diver = jump %>%
  group_by(date, time, DIVER_NR) %>%
  count() %>%
  filter(n > 1)
  dups = list(fieldid = dup_fieldid, group = dup_date_time_diver)
 return(dups)
}

}

check_jump_sample <- function(region, year) {
 a = jumplog_v_sample(region, year)
```

```
  return(a)
 }

 }

parse_fieldid <- function(DT) {
 DT <-
  mutate(
   DT,
   siteid = substr(fieldid, 1, 4),
   STATION_NR = substr(fieldid, 5, 5),
   REPLICATE_CD = substr(fieldid, 6, 6)
  )
}

  )
}

split_date <- function(DT) {
 DT <- mutate(
  DT,
  YEAR = year(date),
  MONTH = month(date),
  DAY = day(date)
 )
}

 )
}

check_substrate_variables <- function(region, year) {
```

```
 s = initial_data(region, year)
  for (v in list(
  "SUBSTRATE_MAX_DEPTH",
  "SUBSTRATE_MIN_DEPTH",
  "HARD_VERTICAL_RELIEF",
  "SOFT_VERTICAL_RELIEF"
 )) {
 )) {
  print(variable_histogram(a, v))
 }
}

 }
}

variable_histogram <- function(df, variable) {
 ggplot(df, aes_string(variable)) +
  geom_histogram(binwidth = 1,
                 color = "black",
                 fill = "grey") +
  theme_minimal()
 }

 }

between_diver_check <- function(sample, field) {
 a <- sample %>%
  select(siteid, REPLICATE_CD, field) %>%
  spread(REPLICATE_CD, field) %>%
  mutate(diff = abs(A - B)) %>%
  arrange(desc(diff))
```

```
  return(a)
}

}

sppbeffive <- function(region, year) {
 s = corrected_data(region, year)
  a <-  s$species %>%
  filter(TIME_SEEN == 1) %>%
  group_by(MASTER_SAMPLE_CD) %>%
  summarise(sppbeffive = n_distinct(SPECIES_CD))
  return(a)
 }

 }

lookupSpecies = function(spp) {
 sub = speciesMasterTable[speciesMasterTable$SPP_CD == spp, ]
 sub
}

}

plot_species_histograms = function(data, species) {
 sub = subset(data, SPECIES_CD == species)
 len = data–frame(c(sub$AVERAGE_LENGTH, sub$MIN_LENGTH, sub$MAX_LENGTH))
 names(len) = "length"
 t = paste("Max Length = ", lookupSpecies(species)$max_size)
 text_position = grobTree(textGrob( t,  x = 0–05,  y = –95, hjust = 0, gp = gpar( col = "black", fontsize = 10,
                  fontface = "bold" )
 ))
```

```
 ))
 a = ggplot(sub, aes(x = NUMBER_OF_INDIVIDUALS)) +
  geom_histogram( binwidth = 1, boundary = -0–5,color = "black",  fill = "white" ) +
  ggtitle(species)
 b = ggplot(len, aes(x = length)) +
  geom_histogram( binwidth = 1,  boundary = -0–5, color = "black",  fill = "white"    ) +
 ggtitle(lookupSpecies(species)$common) + annotation_custom(text_position)
 e = grid–arrange(a, b, ncol = 2, nrow = 2)
 print(e)
}

}

PrintSpeciesToPdf = function(region, year) {
 dataCorrected = corrected_data(region, year)
 speciesMasterTable = read_xlsx(paste(getwd(), "/species–xlsx", sep = ""), sheet = "species")
  sppList = sort(unique(dataCorrected$species$SPECIES_CD))
  for (i in sppList) {
  if (!i %in% speciesMasterTable$SPECIES_CD) {
    stop(c(i, " not in species master table"))
  }
 }


  }
 }

 pdf(paste(region, year, "_species_hist–pdf", sep = ""), onefile = TRUE)
 for (i in sppList) {
  plot_species_histograms(dataCorrected$species, i)
 }
 }
```

```
 dev–off()
}

}

find_duplicate <- function(a, b) {
 x <- data–frame(table(a[b]))
 return(x[x$Freq > 1,])
}

}

find_duplicate_fish <- function(region, year) {
 s = corrected_data(region, year)
 a = s$species %>%
  select( MASTER_SAMPLE_CD, SPECIES_CD, NUMBER_OF_INDIVIDUALS, AVERAGE_LENGTH, MIN_LENGTH,
    MAX_LENGTH )
 return(a[duplicated(a), ]) }


jumplog_v_sample <- function(region, year) {
 jumplog = read_jumplog(region, year) %>%
  filter(REPLICATE_CD == "A" | REPLICATE_CD == "B")
  dataCorrected = corrected_data(region, year)
  a = merge(
  jumplog,
  dataCorrected$sample,
  by = c("date", "fieldid", "DIVER_NR"),

 ) %>%
  all = T
```

```
  ) %>%
    select(date, fieldid, DIVER_NAME, MASTER_SAMPLE_CD) %>%
    filter_all(any_vars(is–na(–)))
    return(a)

}


}

boatlog_vs_sample <- function(region, year) {
  dataCorrected <- corrected_data(region, year)


  a = merge(
    dataCorrected$boatlog,
    dataCorrected$sample,
    by = c("fieldid", "YEAR", "MONTH", "DAY"),

  ) %>%
    all = T
  ) %>%
    select(fieldid, YEAR, MONTH, DAY, SUBREGION, MASTER_SAMPLE_CD) %>%
    filter_all(any_vars(is–na(–)))


  return(a)
}

}
```

```
sample_vs_substrate <- function(region, year) {
  dataCorrected <- corrected_data(region, year)


  a = merge(
    dataCorrected$sample,
    dataCorrected$substrate,
    by = c("MASTER_SAMPLE_CD"),

  ) %>%
    all = T
  ) %>%
    select(MASTER_SAMPLE_CD, SAMPLE_DEPTH, SAND_PERCENTAGE) %>%
    filter_all(any_vars(is–na(–)))



  return(a)
}

}

sample_vs_species <- function(region, year) {
  dataCorrected <- corrected_data(region, year)
    a = merge(
    dataCorrected$sample,
    dataCorrected$species,
    by = c("MASTER_SAMPLE_CD"),
    all = T
  ) %>%

  ) %>%
```

```
  select(MASTER_SAMPLE_CD, SAMPLE_DEPTH, SPECIES_CD) %>%
  filter_all(any_vars(is–na(–)))
   return(a)
}

}

arTest <- function(region, year) {
 ar <- read–csv(paste(region, "_", year, "_dat1–csv", sep = ""))
  b <- corrected_data(region, year)
  ar_sub <- ar %>%
  filter(NUM > 0) %>%
  group_by(PRIMARY_SAMPLE_UNIT, STATION_NR, SPECIES_CD) %>%
  summarise(ar_num = sum(NUM))


 rep_cnt <- b$species %>%
  group_by(siteid, STATION_NR) %>%
  summarise(rep_count = n_distinct(REPLICATE_CD))


 sp_num <- b$species %>%
  group_by(siteid, STATION_NR, SPECIES_CD) %>%
  summarise(sp_num = sum(NUMBER_OF_INDIVIDUALS)) %>%
  merge(rep_cnt, by = c("siteid", "STATION_NR")) %>%
  rename(PRIMARY_SAMPLE_UNIT = siteid)


 t <   merge(ar_sub,
       sp_num,
       by = c("PRIMARY_SAMPLE_UNIT", "STATION_NR", "SPECIES_CD")) %>%
```

```
  mutate(is_equal = as–logical(ar_num * rep_count == sp_num))



  result <- t %>%
  filter(is_equal == F)
  return(result)
}

}

dataMerge <- function(region, year) {
 a <- corrected_data(region, year)



 b = a$boatlog
 s = a$sample %>%
  select(-c("date", "siteid", "STATION_NR", "REPLICATE_CD"))



 h = read_xlsx(paste(getwd(), "/verify_4_psu_", region, "_", year, "–xlsx", sep = ""),
           sheet = "verify_psu_import") %>%
  select(-c("STATION_NR", "REPLICATE_CD"))



 bs = merge(b,
         s,
         by = c("fieldid", "DIVER_NR", "YEAR", "MONTH", "DAY"),
         all = T)



 bsh = merge(bs, h, by = c("YEAR", "fieldid"), all = T)
```

```
    return(bsh)
  }

  }

psuChecks <- function(region, year) {
  d = dataMerge(region, year)


  for (i in list("HABITAT_CD", "DEPTH", "SUBREGION", "ADMIN")) {
    a = d %>%
      group_by(PRIMARY_SAMPLE_UNIT, UQ(as–name(i))) %>%
      tally() %>%
      filter(n != 2)
    print(a)
  }
  }

  }
  }

createAR <- function(region, year) {
  x = dataMerge(region, year)
  dataCorrected = corrected_data(region, year)
  speciesMasterTable = read_xlsx(paste(getwd(), "/species–xlsx", sep = ""), sheet = "species")


  a <-
    merge(dataCorrected$species,
          speciesMasterTable,
          by = 'SPECIES_CD',
```

```
      all = T) %>%
  select(-c("siteid", "STATION_NR", "REPLICATE_CD"))




b <- a %>%
  filter(is–na(MASTER_SAMPLE_CD)) %>%
  mutate(
    MASTER_SAMPLE_CD = replace(MASTER_SAMPLE_CD, values = "A2019071808190307"),
    NUMBER_OF_INDIVIDUALS = replace(NUMBER_OF_INDIVIDUALS, values = 0),
    AVERAGE_LENGTH = replace(AVERAGE_LENGTH, values = 0),
    TIME_SEEN = replace(TIME_SEEN, values = 1)
  )


  )

c <- a %>%
  filter(!is–na(MASTER_SAMPLE_CD))




d <- rbind(b, c) %>%
  select(
    MASTER_SAMPLE_CD,
    SPECIES_CD,
    NUMBER_OF_INDIVIDUALS,
    AVERAGE_LENGTH,
    MIN_LENGTH,
    MAX_LENGTH,
```

```
    TIME_SEEN
  )




  )



c1 <- d %>%
  filter(NUMBER_OF_INDIVIDUALS == 0) %>%
  mutate(
    L1 = NA,
    A1 = NA,
    L2 = NA,
    A2 = NA,
    L3 = AVERAGE_LENGTH,
    A3 = NUMBER_OF_INDIVIDUALS,
    L4 = NA,
    A4 = NA,
    L5 = NA,
    A5 = NA
  )


  )


c2 <- d %>%
  filter(NUMBER_OF_INDIVIDUALS == 1) %>%
  mutate(
    L1 = NA,
    A1 = NA,
```

```
  L2 = NA,
  A2 = NA,
  L3 = AVERAGE_LENGTH,
  A3 = NUMBER_OF_INDIVIDUALS,
  L4 = NA,
  A4 = NA,
  L5 = NA,
  A5 = NA
 )

 )

c3 <- d %>%
 filter(NUMBER_OF_INDIVIDUALS == 2) %>%
 mutate(
  L1 = MIN_LENGTH,
  A1 = 1,
  L2 = NA,
  A2 = NA,
  L3 = NA,
  A3 = NA,
  L4 = NA,
  A4 = NA,
  L5 = MAX_LENGTH,
  A5 = 1
 )

 )

c4 <- d %>%
 filter(NUMBER_OF_INDIVIDUALS == 3) %>%
```

```
  mutate(
    L1 = MIN_LENGTH,
    A1 = 1,
    L2 = NA,
    A2 = NA,
    L3 = AVERAGE_LENGTH,
    A3 = 1,
    L4 = NA,
    A4 = NA,
    L5 = MAX_LENGTH,
    A5 = 1
  )

  )

c5 <- d %>%
  filter(NUMBER_OF_INDIVIDUALS == 4) %>%
  mutate(
    L1 = MIN_LENGTH,
    A1 = 1,
    L2 = NA,
    A2 = NA,
    L3 = AVERAGE_LENGTH,
    A3 = 2,
    L4 = NA,
    A4 = NA,
    L5 = MAX_LENGTH,
    A5 = 1
  )

  )
```

```
c6 <- d %>%
 filter(NUMBER_OF_INDIVIDUALS >= 5 &
        NUMBER_OF_INDIVIDUALS <= 99) %>%
 mutate(
  L1 = MIN_LENGTH,
  A1 = 1,
  L3 = AVERAGE_LENGTH,
  A3 = (–5 * NUMBER_OF_INDIVIDUALS),
  L5 = MAX_LENGTH,
  A5 = 1,
  L2 = (L3 + L1) / 2,
  A2 = (NUMBER_OF_INDIVIDUALS - (A1 + A3 + A5)) / 2,
  L4 = (L3 + L5) / 2,
  A4 = A2
 )

 )

c7 <- d %>%
 filter(NUMBER_OF_INDIVIDUALS > 99) %>%
 mutate(
  L1 = MIN_LENGTH,
  A1 = (–01 * NUMBER_OF_INDIVIDUALS),
  L3 = AVERAGE_LENGTH,
  A3 = (–5 * NUMBER_OF_INDIVIDUALS),
  L5 = MAX_LENGTH,
  A5 = (–01 * NUMBER_OF_INDIVIDUALS),
  L2 = (L3 + L1) / 2,
  A2 = (NUMBER_OF_INDIVIDUALS - (A1 + A3 + A5)) / 2,
  L4 = (L3 + L5) / 2,
```

```
   A4 = A2
  )

  )

c8 <- rbind(c1, c2, c3, c4, c5, c6, c7)




d1 <- c8 %>%
  filter(!is–na(L1)) %>%
  mutate(len = L1, abun = A1) %>%
  select(MASTER_SAMPLE_CD, SPECIES_CD, len, abun, TIME_SEEN)


d2 <- c8 %>%
  filter(!is–na(L2)) %>%
  mutate(len = L2, abun = A2) %>%
  select(MASTER_SAMPLE_CD, SPECIES_CD, len, abun, TIME_SEEN)


d3 <- c8 %>%
  filter(!is–na(L3)) %>%
  mutate(len = L3, abun = A3) %>%
  select(MASTER_SAMPLE_CD, SPECIES_CD, len, abun, TIME_SEEN)


d4 <- c8 %>%
  filter(!is–na(L4)) %>%
  mutate(len = L4, abun = A4) %>%
```

```
  select(MASTER_SAMPLE_CD, SPECIES_CD, len, abun, TIME_SEEN)



d5 <- c8 %>%
  filter(!is–na(L5)) %>%
  mutate(len = L5, abun = A5) %>%
  select(MASTER_SAMPLE_CD, SPECIES_CD, len, abun, TIME_SEEN)



d6 <- rbind(d1, d2, d3, d4, d5) %>%
  arrange(MASTER_SAMPLE_CD)




# Merge with averaged sample information



e <-  dataCorrected$sample %>%
  select(
    MASTER_SAMPLE_CD,
    YEAR,
    MONTH,
    DAY,
    siteid,
    STATION_NR,
    REPLICATE_CD,
    SAMPLE_DEPTH,
    UNDERWATER_VISIBILITY
  ) %>%
  ) %>%
```

```
  group_by(YEAR, siteid, STATION_NR) %>%
  mutate(depth = (mean(SAMPLE_DEPTH)) * –3048,
         vis = mean(UNDERWATER_VISIBILITY)) %>%
  select(-c(SAMPLE_DEPTH, UNDERWATER_VISIBILITY))


f <- merge(e, d6, by = "MASTER_SAMPLE_CD") %>%
  select(-MASTER_SAMPLE_CD)




# Add in zero records


g <-  f %>%
  group_by(siteid, STATION_NR, REPLICATE_CD, SPECIES_CD) %>%
  summarise(abun = sum(abun))


# Transpose 1


t1 <- g %>%
  spread(–, key = SPECIES_CD, value = abun)


#  Transpose 2


# Bring back in variables
```

```
t2a <- f %>%
  group_by(siteid, STATION_NR, REPLICATE_CD) %>%
  summarise(
    vis = min(vis),
    depth = min(depth),
    YEAR = min(YEAR),
    MONTH = min(MONTH),
    DAY = min(DAY)
  )

  )

t2b <- merge(t2a, t1)


t2 <-
  gather(
    data = t2b,
    key = SPECIES_CD,
    value = abun,
    -siteid,
    -STATION_NR,
    -REPLICATE_CD,
    -YEAR,
    -MONTH,
    -DAY,
    -vis,
    -depth
  ) %>%
```

```
  ) %>%
  mutate(abun = replace_na(abun, 0))




t3 <- t2 %>%
  mutate(name1 = SPECIES_CD, name2 = SPECIES_CD) %>%
  group_by(name1, name2) %>%
  summarise(abund = n())



t4 <- t2 %>%
  filter(abun == 0) %>%
  mutate(len = 0, TIME_SEEN = 1)




t5 <- rbind(f, t4) %>%
  arrange(SPECIES_CD, siteid)



t6 <- t5 %>%
  group_by(YEAR, MONTH, DAY, siteid, STATION_NR) %>%
  summarise(freq = n_distinct(REPLICATE_CD)) %>%
  merge(–, t5, by = c("YEAR", "MONTH", "DAY", "siteid", "STATION_NR"))
```

```
t7 <- t6 %>%
 group_by(YEAR,
          MONTH,
          DAY,
          siteid,
          STATION_NR,
          depth,
          vis,
          SPECIES_CD,
          len) %>%
 summarise(
  freq = max(freq),
  abun = format(round((sum(
   abun
  ) / freq), 4)),
  TIME_SEEN = mean(TIME_SEEN)
 ) %>%
 ) %>%
 arrange(SPECIES_CD, siteid) %>%
 select(-freq)
```

```
s <- x %>%
 select(
  YEAR,
```

```
  MONTH,
  DAY,
  siteid,
  STATION_NR,
  LAT_DEGREES,
  LON_DEGREES,
  GRID_ID,
  SUBREGION,
  ADMIN,
  HABITAT_CD,
  DEPTH
 ) %>%
 ) %>%
 mutate(
  LAT_DEGREES = round(LAT_DEGREES, 5),
  LON_DEGREES = round(LON_DEGREES, 5),
  DEPTH_CLASS = DEPTH
 ) %>%
 ) %>%
 select(-DEPTH)



s1 <-
 merge(t7, s, by = c("YEAR", "MONTH", "DAY", "siteid", "STATION_NR")) %>%
 distinct(–)



s2 <- speciesMasterTable %>%
 select(SPECIES_NR, SPECIES_CD)
```

```
s3 <- merge(s1, s2, by = "SPECIES_CD")



dat1 <- s3 %>%
 rename(
  PRIMARY_SAMPLE_UNIT = siteid,
  DEPTH = depth,
  UNDERWATER_VISIBILITY = vis,
  LEN = len,
  NUM = abun
 ) %>%
 ) %>%
 select(
  YEAR,
  MONTH,
  DAY,
  PRIMARY_SAMPLE_UNIT,
  STATION_NR,
  LAT_DEGREES,
  LON_DEGREES,
  DEPTH,
  UNDERWATER_VISIBILITY,
  GRID_ID,
  HABITAT_CD,
  DEPTH_CLASS,
  SUBREGION,
  ADMIN,
  SPECIES_NR,
  SPECIES_CD,
  LEN,
  NUM,
```

```
  TIME_SEEN
) %>%
) %>%
arrange(PRIMARY_SAMPLE_UNIT, STATION_NR, SPECIES_NR, LEN)
return(dat1)}
```

Appendix C. Data Quality Standards for A Cooperative Multiagency Reef Fish Monitoring Protocol for the Florida and U.S. Virgin Islands Coral Reef Ecosystems

Protocol Narrative Version—2.0

(Note: This appendix is designed to eventually become a stand-alone National Park Service published document, a Quality Assurance Plan.)

Introduction

The purpose of this report is to document the standards, used by the South Florida/Caribbean Network, for activities related to the collection, processing, storage, analysis, and publication of monitoring data as described in *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025).The policies and procedures documented in this quality assurance plan complement the quality assurance plans for other monitoring activities conducted by the South Florida/Caribbean Network and supplement National Inventory & Monitoring Division Quality Management Plan (in development). The plan also serves as a guide for all South Florida/Caribbean Network personnel and partners who are involved in protocol/program activities and as a resource for identifying memoranda, publications, and other literature that describe associated techniques and requirements in more detail.

Protocol Overview

Reef fish populations are of great economic, cultural and ecological significance throughout the Caribbean and are essential components of coral reef ecosystems in U.S. North Atlantic territorial waters including Florida, United States Virgin Islands (USVI), Puerto Rico and Flower Garden Banks. Recent stressors on these populations include severe quality degradation of essential habitats, such as widespread coral bleaching and disease outbreaks, as well as significant increases in recreational and commercial fishing efforts. Reef fish populations have historically been difficult to assess due to a lack of traditional fisheries dependent data and limitations on interpreting catch data at a fine spatial scale. The monitoring methodologies described in *A Cooperative Multiagency Reef Fish Monitoring Protocol for the Florida and U.S. Virgin Islands Coral Reef Ecosystems, Protocol Narrative Version—2.0* (Feeley et al. 2025) are a culmination of over 40 years of counting fish and survey design work by dedicated fisheries scientists at multiple agencies and universities. The data collected according to these procedures are useful for understanding the ecological effect of natural and man-made stressors as well as management actions to promote sustainability of reef fish populations and communities.

This protocol is a revised and unified version of two previous published reef fish protocols (Brandt et al. 2009; Bryan et al. 2013). Brandt et al. (2009) was originally developed as a collaborative, multi-agency effort to monitor reef fish populations in the Florida Keys and the Tortugas. The Bryan et al. (2013) protocol was later created to monitor and assess reef fish communities and populations within the U.S. Virgin Islands and specific shelf sub regions (e.g., the islands of St. Croix, St. John and St. Thomas). These earlier versions shared a common survey design but relied on different fish count methods (point count and belt transect count) and site selection protocols. *A Cooperative Multiagency Reef Fish Monitoring Protocol for the Florida and U.S. Virgin Islands Coral Reef Ecosystems, Protocol Narrative Version—2.0* (Feeley et al. 2025) represents the first time long term reef fish fishery-independent monitoring methods have been consolidated across all U.S. territorial interests in the Caribbean and Florida. This protocol focuses on shallow coastal shelf areas within and near U.S. National Park Service (NPS) parks and is the product of nearly four decades of independent localized reef fish population monitoring programs. A complementary NOAA Technical Memorandum is planned for the Caribbean and Atlantic waters outside National Park Service unit areas of interest including SE coastal Florida, Puerto Rico and the Flower Garden Banks in the Gulf of Mexico. The statistical sampling design and survey methods described here are also the basis for a large-scale diver visual survey monitoring program that surveys coral reef fish assemblages and habitats of most of the U.S.-affiliated tropical Pacific Ocean islands. Currently these regions include the main and Northwestern Hawaiian Islands and affiliated geopolitical regions of American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the Pacific Remote Islands Areas.

Protocol Activities and Modules

Data are collected or derived as a part of A *Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025) in 16 different activities or modules (Table 1). To efficiently sample within a heterogeneous environment, the sampling effort is optimized using a single-stage stratified random sampling design. Before sampling occurs, the sampling domains are partitioned into 50 × 50 meter (164 × 164 feet [ft]) grid cells, or primary units, which are each assigned a stratum designation based on habitat type, geographic sub-region, management type (open vs. closed to extraction), depth and rugosity. The number of primary units to be sampled in each stratum type is based on a Neyman allocation scheme, which accounts for the stratum's size and standard deviations for focal fish species densities in each region. Once the optimal number of primary units per stratum is chosen based on a target precision and the capabilities of the agencies, primary units to be sampled are randomly selected from a list of all primary units for each stratum. Within each selected primary unit, one smaller second-stage unit is selected for survey.

The second-stage unit consists of a buddy pair of divers who each perform a Reef Visual Census (RVC) timed point count within a 15-meter (49.2 ft) diameter cylinder. A stationary point count method was chosen as the *in situ* visual sampling method of choice because the complexity and spatial extent of habitats often precludes the use of linear transects and it is a better method for exploited fish species (snappers and groupers) that may exhibit diver avoidance. The RVC method allows divers to simultaneously collect data on the relative density and size distributions of the entire fish community, as well as information on benthic habitat features. Atlantic and Caribbean sampling

occurs biennially in the summer months, generally between May and September. The Florida Keys and Tortugas and the Flower Gardens are surveyed the same year and are alternated biennially with surveys in the U.S. Virgin Islands and Puerto Rico every other year. Qualified and trained personnel participate from all participating agencies.

The goals of the surveys are to support multispecies stock assessments and determine the effectiveness of different marine management zones in the Florida Keys and USVI regions (Smith et al. 2011). The measurable objective is to track the status and trends over time in:

- Relative density, abundance, and proportion of occurrence of target species
- Size-structure (size frequency distribution) of target species
- Spatial distribution of target species
- Average size in the exploited phase of target species
- Species richness

These changes are evaluated each year with respect to the habitat features, the physical environment and potential management actions that may have affected the dynamics of the populations.

Table 1. Protocol activity matrix for *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025).

Category	Activity	Description
Site Recon, Establishment, Maintenance	1. Site Reconnaissance	Shapefiles of sampling frame, habitats, bathymetry, strata, number of total cells available within each stratum, and number of cells sampled each year are documented (see SOP RF01 *Statistical Sampling Design—Version 2.0* [SFCN 2025a]).
	2. Site Establishment	Sites are identified in a random sampling allocation draw and surveyed in the field. A dive flag and float are temporarily tied to the hard bottom at the site marking the center point of the established station. Sites are visited once (see SOP RF01 and SOP RF03 *Field Data Collection—Version 2.0* [SFCN 2025b]).
Field Observations	3. Reef fish Visual Point Counts	Circular plot observations of reef fish species, including estimate of fork length (size) and number of fish for each species within a 7.5 radius cylinder (see SOP RF03).
	4. Substrate and benthic characteristics	Substrate slope, maximum vertical relief, vertical surface relief coverage, abiotic footprint, dominant biotic coverage on sand and hardbottom (see SOP RF03).
	5. Boat log	Dive-of-day, Field ID, Station, Estimated Time, Dive team, location, GPS/Flag #, and surface-data-collector comments. (see SOP RF03)
Sensor Data	6. GPS location	A GPS attached to a dive flag/float and reel is used to record the position of the paired circular visual assessment cylinders. Logging interval (tracks) set to one point every 10 seconds. (see SOP RF03)
	7. Bottom temperature	Bottom water temperature measurement taken once at depth by the diver. (see SOP RF03)
	8. Habitat photos (4)	Digital photos taken in four directions from the center of a cylinder to characterize habitat type. (see SOP RF03)
	9. Depth	Survey depth is measured by a digital depth sensor (dive computer). (see SOP RF03)
Derived Data	10. Relative density	Relative density of reef fish species or reef fish groups per secondary sampling unit, stratum, and domain. (see SOP RF05 *Data Analyses and Reporting—Version 1.1* [SFCN 2025c])
	11. Relative proportion of occurrence	Relative proportion of occurrence of reef fish species or reef fish groups per secondary sampling unit, stratum, and domain. (see SOP RF05)
	12. Average size	Average size of species in the exploited phase of the stock, average size of all individuals or life stage (adult or juvenile). (see SOP RF05)
	13. Species richness	Species richness of reef fish species or reef fish groups per secondary sampling unit, stratum, and domain. (see SOP RF05)
Quality Control Data	14. Repeated measurement of reef fish visual point count and substrate metrics	Two separate point counts by two observers are collected at a station within the same primary sampling unit and averaged (see SOP RF03 and SOP RF04 *Data Entry and Proofing—Version 2.0* (SFCN 2025d) Experienced and inexperienced divers are paired together and if observations among observers do not meet with acceptable variability the data will be flagged and disregarded if necessary.

Category	Activity	Description
	15. Repeated measurement of survey location	The diver held GPS provides the primary source of location information of the survey. However, waypoint or location coordinates of the divers are also recorded by the support vessel. (see SOP RF03)
	16. Standardized training	Pre-season training with fish models of known size and species are used to evaluate proficiency test scores and test observer reliability. (see SOP RF02 *Participant Training—Version 1.1* [SFCN 2025e)

Sampling Design

Information regarding the sampling design is provided in Table 2.

Table 2. Activity-level sample design matrix for *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025).

Category	Activity	Sampling Design	Revisit Design
Site Recon, Establishment, Maintenance	2. Site Establishment	Probability, with gridded sample frames for St. Thomas/St. John, St. Croix, Florida Keys and the Tortugas. Randomized stratified survey design with sites (primary sampling units) allocated across strata according to Neyman allocation. Secondary sampling units (selected for hardbottom) are targeted within primary sampling units.	Sites randomized every 2 years in each region (South Florida and USVI)
Field Observations	3. Reef fish Visual Point Counts	Targeted, timed estimate of all reef fish within 7.5-meter (24.6 ft) radius cylinder	Sites randomized every 2 years in each region (South Florida and USVI)
	4. Substrate and benthic characteristics	Targeted, quantitative estimate of substrate and benthic characteristics	Sites randomized every 2 years in each region (South Florida and USVI)
	5. Boat log	Targeted, record of sampling event	Sites randomized every 2 years in each region (South Florida and USVI)
Sensor Data	6. GPS location	Targeted, with Activity 3	Continuous with Activity 3
	7. Bottom temperature measurement	Targeted, with Activity 3	Once per survey Activity 3
	8. Habitat photos (4)	Targeted, with Activity 3	Once per survey Activity 3
	9. Depth	Targeted, with Activity 3	Once per survey Activity 3

Data Quality Objectives

Data quality values and standards for implementation are provided in Table 3 through Table 6.

Table 3. Data Quality Values (DQVs) for *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025).

Type	Data Quality Value	Definition	Protocol Considerations
Intrinsic Data Quality	Accuracy	Measurements reflect the true value of the parameter being observed. This applies to measures (length, width, position) or classes (species, types, or categories). Includes components of precision and bias	Precision and potential observer bias are balanced by collecting large samples sizes (n = 200–400) across regions and various habitats using numerous observers. Data collectors are paired to match experienced divers with less experienced personnel on paired reef fish point counts that are averaged together to create one count per station. See SOP RF02 (SFCN 2025e), SOP RF03 (SFCN 2025b) and SOP RF05 (SFCN 2025c) for more details. Note: Trainee data are not included in survey analyses.
	Representativeness	Measurements represent conditions at the time of sampling. Combined with accuracy, leads to repeatable data collection.	Parameters and visual assessment methods (RVC Reef fish Visual Census) were chosen based on published and established protocol by Bohnsack and Bannerot (1986) that uses a standardized timed survey of reef fish on hardbottom habitat. Counts, identification and size measurements are the best direct indicators of fisheries independent data. Primary sampling units are chosen using a stratified random survey design based on depth, habitat or habitat zone, rugosity and management zone. Secondary sampling units are targeted to occur on hardbottom habitat.
Contextual Data Quality	Comparability	The degree to which data can be compared among sample locations, data sources, or periods of time.	Reef fish surveys are completed at the same time of year from May to October across the Florida and USVI regions to ensure data are comparable over time and allow for trend analyses. An effort is made to sample each region at approximately the same time each year. Data are collected and shared by a cooperative interagency effort. Members of all agencies meet prior to the field season to refresh and review necessary training, methodology, standards, and sampling design.

Type	Data Quality Value	Definition	Protocol Considerations
	Timeliness / Currency	How recent the data need to be to be considered valid for their intended use. Data represents conditions and/or is available and in a format for use at the appropriate time in the decision-making process.	Data processing occurs in a timely fashion based on certification requirements and intended use of data (generally available to the public in 1 year). Where certified data cannot be provided for resource management decisions in a timely manner, provisional use and dissemination of data is appropriate within the agencies.
	Completeness	All data/measures required to evaluate accuracy representativeness are present; incomplete data sets (either at a location, across sampling locations, or over time) lose utility or relevance. Data records contain values as planned across the period of record.	Methods, sampling plans, and analyses are designed and implemented such that they result in a complete dataset across space and the planned period of record. An extensive QA/QC and validation process takes place annually to create an analysis-ready data set including sample, substrate, species and NTOT (survey) data.
Representational Data Quality	Consistent Representation	Use of standard definitions when describing data quality or resource quality based on data	Data are defined as “provisional”, “approved” and “certified”. The data go through an extensive proofing and validation process at the diver level and at the data manager level. Derivative data are compared to sustainable fisheries metrics. Fishery regulations (bag limits and legal size of capture) and park management (take or no-take reserves) vary by park and by region. This is particularly relevant when reporting data in comparison to management, ecological, or regulatory thresholds.
Data Accessibility	Secure	Access to data, products, and systems limited to appropriate audiences.	Data are publicly available once they have completed the QA/QC verification process and a copy is maintained by each agency and is publicly available at the NOAA-SEFSC data portal and NOAA’s Coral Reef Information System (CoRIS) Geoportal. Sensitive species location data, and location of transient fish spawning aggregations are protected at the park level. Protected data may be released in limited cases only with park Superintendent approval following park procedures.

Table 4. Measurement Quality Objectives for *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025).

Activity	**Measure / Quality Indicator**	**Quality Objective**
Site Recon	Location Accuracy (PSU)	Sample location within +/- 25 m of preseleted centroid waypoint (within 50 × 50 meter [164 × 164 ft] grid)
Site Establishment	Location Accuracy	+/- 8 meters (26.2 ft) horizontal measured by GPS unit under normal conditions). Additionally, actual GPS position will swing with prevailing wind and current above the physical survey location.
Reef Fish Visual Point Counts	Taxonomic Resolution	Identifications made to the species level
	Identification Accuracy	95% accuracy for visual observations
	Identification Completeness	98% of observations made to species level
	Size	+/- 5 centimeters (2 inches [in]) fish < 50 centimeters (19.7 in) fork length; +/- 10 centimeters (3.9 in) ≥ 50 centimeters fork length
	Distance accuracy	+/- 0.5 meters (1.6 ft)
Substrate and Benthic Characteristics	Taxonomic Resolution	Identifications made to the functional group (turf algae, macroalgae, sponge, stony coral, octocoral, seagrass, etc.)
	Identification Accuracy	95% accuracy
	Habitat type	100% accurate agreement between observers
	Abiotic footprint	+/- 5% for each value (sand, rubble, hardbottom)
	Surface relief coverage	+/- 5% for each category
	Maximum vertical relief	+/- 3 meters (9.8 ft)
	Visibility	+/- 3 meters
	Biotic cover	+/- 5% for each category
	Slope	+/- 1 meter (3 ft)
	Depth	+/- 0.3 m (1 ft)
	Current	100% agreement between observers (none, moderate, high)
GPS Location	Location accuracy	+/- 3 meters 95% of the time
	Location position precision	+/- 5 meters (16.4 ft; unit can swing around anchoring point)
Bottom Temperature	Accuracy (temperature)	+/- 1° Fahrenheit
Habitat Photos	Image resolution	10–12 megapixel
	Image format	JPG w/ embedded w metadata

Table 5. Taxonomic standards to be used in *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystem, Narrative Version—2.0* (Feeley et al. 2025).

Options	Activity	Standard to be Used	Reference / Authority
Collection of Observational Data	3. Reef fish Visual Point Counts	Scientific Publication	Humann and DeLoach (2014)
	4. Substrate and benthic characteristics	Scientific Publication	Littler et al. (1989) Humann and DeLoach (2013) Feeley et al. (2025)
Reporting of Observational Data	3. Reef fish Visual Point Counts	Scientific Publication	Humann and DeLoach (2014)
	4. Substrate and benthic characteristics	Scientific Publication	Littler et al. (1989) Humann and DeLoach (2013) Feeley et al. (2025)

Table 6. Data Protection standards for *A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, Narrative Version—2.0* (Feeley et al. 2025). With the exceptions noted, all data collected are to be made publicly available in a timely fashion.

Information Type	Type of Data	Level of Protection	Rules for Dissemination
Resource Data and Information	Fish spawning aggregation locations	Operationally Protected	Location limited to the park level all parks.
	Epinephelus striatus, Nassau Grouper locations	Legally Protected	Commercially valuable and federally threatened; location information redacted from dissemination.
	Epinephelus itajara, Goliath Grouper locations	Legally Protected	Commercially valuable, federally threatened and territorially endangered. Locations of observations redacted from dissemination in USVI parks.
Personally Identifiable Information	Non-NPS Staff Information	Legally Protected	All but first name and last initial redacted from public release

Literature Cited

Bohnsack, J. A. and S. P. Bannerot. 1986. A stationary visual census technique for quantitatively assessing community structure of coral reef fishes. NOAA Tech. Report NMFS 41. 15 p.

Feeley, M. W., M. E. Brandt, D. R. Bryan, N. Zurcher, A. Acosta, J. S. Ault, J. Blondeau, J. A. Bohnsack, L. J. W. Grove, J. Keller, V. McDonough, L. Richter, S. G. Smith, D. W. Swanson. 2025. A Cooperative Multiagency Reef Fish Monitoring Protocol for Florida and U.S. Virgin Islands Coral Reef Ecosystems, v. 2.00. Natural Resource Report NPS/SR—2025/XXX. National Park Service, Fort Collins, Colorado. Available at: DataStore - Project for A Cooperative Multiagency Reef Fish Monitoring Protocol for the Florida and US Virgin Islands Coral Reef Ecosystems Protocol Narrative Version—2.0. (last accessed May 2025).

Humann, P. and N. DeLoach. 2013. Reef Coral Identification: Florida, Caribbean, Bahamas 3rd Edition. New World Publications, Jacksonville, FL. 276 p.

Humann, P. and N. DeLoach. 2014. Reef Fish Identification: Florida, Caribbean, Bahamas, 4th Edition. New World Publications, Jacksonville, FL. 548 p.

Littler, D. S., M. L. Littler, K. E. Bucher, and J. N. Norris. 1989. Marine Plants of the Caribbean. A Field Guide from Florida to Brazil. Smithsonian, Washington, D.C. 272 p.

South Florida/Caribbean Network. 2025a. Statistical Sampling Design—Version 2.0. South Florida/Caribbean Network Standard Operating Procedure NPS/SFCN/SOP—RF01. South Florida/Caribbean Network, Palmetto Bay, Florida. Available at: irma.nps.gov/DataStore/Reference/Profile/2309696 (last accessed April 2025).

South Florida/Caribbean Network. 2025b. Field Data Collection—Version 2.0. South Florida/Caribbean Network Standard Operating Procedure NPS/SFCN/SOP—RF03. South Florida/Caribbean Network, Palmetto Bay, Florida. Available at: irma.nps.gov/DataStore/Reference/Profile/2309698 (last accessed April 2025).

South Florida/Caribbean Network. 2025c. Data Entry and Proofing—Version 2.0. South Florida/Caribbean Network Standard Operating Procedure NPS/SFCN/SOP—RF04. South Florida/Caribbean Network, Palmetto Bay, Florida. Available at: irma.nps.gov/DataStore/Reference/Profile/2309699 (last accessed April 2025).

South Florida/Caribbean Network. 2025d. Data Analyses and Reporting—Version 1.1. South Florida/Caribbean Network Standard Operating Procedure NPS/SFCN/SOP—RF05. South Florida/Caribbean Network, Palmetto Bay, Florida. Available at: irma.nps.gov/DataStore/Reference/Profile/2309701 (last accessed April 2025).

South Florida/Caribbean Network. 2025e. Participant Training—Version 1.1. South Florida/Caribbean Network Standard Operating Procedure NPS/SFCN/SOP—RF02. South

Florida/Caribbean Network, Palmetto Bay, Florida. Available at: irma.nps.gov/DataStore/Reference/Profile/2309697 (last accessed April 2025).