Standard Operating Procedure (SOP) #17

Data Entry and Verification

Version 1.01 (March 18, 2022)

Change History

New Version #	Revision Date	Author	Changes Made	Reason for Change	Previous Version #
1.01	3/18/2022	Kelly Kozar, Kim Weisenborn	The database application is now on SQL Server, located in Seattle, and requires difference instructions for accessing the database. Therefore, information has been revised to refer to the FTPC Monitoring Database User Guide on SharePoint instead.	SQL server is now located in Seattle and is accessed using a different method. Instructions for installing the database are already in the user guide, so no need to duplicate.	1.0

Only changes in this specific SOP will be logged here. Version numbers increase incrementally by hundredths (e.g., version 1.01, version 1.02) for minor changes. Major revisions should be designated with the next whole number (e.g., version 2.0, 3.0, 4.0). Record the previous version number, date of revision, author of the revision, changes made, and reason for the change along with the new version number.

Purpose

This SOP document describes the general procedures for entry and verification of field data in the working project database for the Pacific Island Inventory and Monitoring Network (PACN I&M) Focal Terrestrial Plant Communities (FTPC) Monitoring Protocol. Refer to protocol sections Overview of Database Design (section 4.3 of the protocol narrative), and Data Entry and Processing (4.4) for related guidance and a clarification of the distinction between the working database and the master database.

FTPC Monitoring Database

The FTPC Monitoring Database resides within a remote SQL Server database and is stored on a server in Seattle. The enterprise SQL Server database allows multiple users to simultaneously access the database to enter, edit, or view data. Security is greatly enhanced when using SQL Server, backups are created automatically, and transactions are logged, which allows a database to be "rolled back" to a point in time. Users will use a Microsoft® Access front-end database file that is linked to SQL Server to enter, edit, and view records.

Since the FTPC SQL Server database is stored on a remote server in Seattle, it must be accessed remotely using the Remote Desktop Connection application. A shortcut is stored on the PACN I&M SharePoint site[1] in the FTPC Monitoring project workspace and will need to be downloaded to the user's Desktop to enter, edit, or view FTPC data.

Refer to the FTPC Database User Guide for detailed instructions on how to download, install, and operate the SQL Server database application.

Database Instructions

The first action to be taken is to make sure the user's local workspace is set up properly so that the shortcut to the FTPC SQL Server can be downloaded from SharePoint to the user's computer in order to access the FTPC Monitoring Database. Additionally, a OneDrive shortcut to the database images folder will need to be created in order to link photos to the database. See SOP #14 Workspace Setup and Project Records Management for more information on how to setup the local workstation and create a OneDrive shortcut.

See the FTPC Monitoring Database User Guide for information on downloading the Remote Desktop Connection from SharePoint.

Data Entry

Refer to the FTPC Monitoring Database User Guide for detailed instructions on how to operate the database application. The user guide provides step-by-step database instructions accompanied by screen-capture images. Ultimately, it is the PACN Botanist's and PACN Data Manager's shared responsibility to ensure that all data entry staff understand how to enter data and follow all applicable SOPs. Data entry technicians are responsible for becoming familiar with the field datasheets, the database software and structure, import file procedures, and any standard codes for data entry.

The following are general guidelines to keep in mind:

1. Data entry will be conducted by the PACN Botanist or another trained biological science technician at their duty station.

2. Data entry should occur as soon after data collection as possible so that field crews keep current with data entry tasks and catch any errors or problems as close to the time of data collection as possible.

3. Data entry should be completed by the person who collected the data or someone who is familiar with the project and data. The primary goal of data entry is to transcribe the data from paper and photographic records into the computer with 100% accuracy.

4. If any sensitive data is collected (e.g., endangered plant locations), then make a note of this on the original data form and let the PACN Data Manager know to flag the data as sensitive.

---

[1] PACN I&M SharePoint site, https://doimspp.sharepoint.com/sites/nps-PWR-PACNIM (accessed 21 April 2022).

5. Each data entry form is patterned after the layout of the field form and has built-in quality assurance components such as pick lists and validation rules to test for missing data or illogical combinations. The database allows the users to view only the pre-built forms and does not permit users to view the raw data tables and other database objects as a way of ensuring the maximum level of quality assurance.

6. As data are being entered or imported, the person entering the data should visually review each data form to make sure that the data on screen match the field forms. This should be done for each record prior to moving to the next form for data entry.

7. At regular intervals and at the end of the field season the PACN Botanist should inspect the data that have been entered to check for completeness and perhaps catch avoidable errors.

8. Update the FTPC Database Log Microsoft® Excel spreadsheet. Include the date and name of the technician entering data. Use the data log to keep notes on data entry, including suggestions for updates, errors encountered, and who entered or proofread the data.

Data Verification

Quality control of data is a critical step in data management. Verification of data (ensuring data on the field datasheets match data entered into the database) is the responsibility of the PACN Botanist and designees. See recommendations outlined in Chapter 4 of the FTPC Monitoring Protocol. Any modifications to the FTPC Monitoring Database will be described in the edit log and the functionality of the verification routines will be explained in detail in the FTPC Monitoring Database User Guide.

Data verification checks that the digitized data match the source data. The following guidelines regarding data verification should be followed:

1. The PACN Botanist is responsible for specifying in the project protocol one or more of the data verification methods available and ensuring proper execution. At the discretion of the PACN Botanist, additional verification methods may be applied.

2. Data verification is carried out by staff thoroughly familiar with data collection and entry.

3. All records (100%) should be verified against original source data using the method below.

   a. Visual review after data entry: Upon completion of data entry, all records are printed and compared with the original values from the hard copy. Errors are clearly marked and corrected in the database as soon after data entry as possible. Reliability increases if someone other than the person keying the data performs the review. Alternatively, two technicians (one reading from the original data and one checking the entered data) can perform this review.

4. A subset of randomly selected records (10%) should be reviewed after initial verification by the PACN Botanist. If errors are found, the entire data set should be verified again.

5. A record of the verification process for each data set, including number of iterations and results, will be prepared by the PACN Botanist as part of formal metadata generation.

6. Spatial data collected as part of the project will be viewed in GIS and visually inspected for accuracy (e.g., points located outside park boundaries, upland locations occurring in water).

Database Log

The FTPC Database Log is a Microsoft Excel® spreadsheet that is used to keep track of processed FTPC monitoring data for each plot within a sampling year. It is located in the "Database_log" folder in the FTPC Monitoring project workspace in SharePoint and should be filled out as data processing occurs (Figure SOP 17.1). Processed data is composed of scanning datasheets, processing images; and entering, verifying, and certifying data. For questions about how to fill in the FTPC Database Log, please contact the PACN Botanist, Jacob Gross (jacob_gross@nps.gov) or PACN Lead Biological Science Technician, Kathryn Akamine (kathryn_akamine@nps.gov).
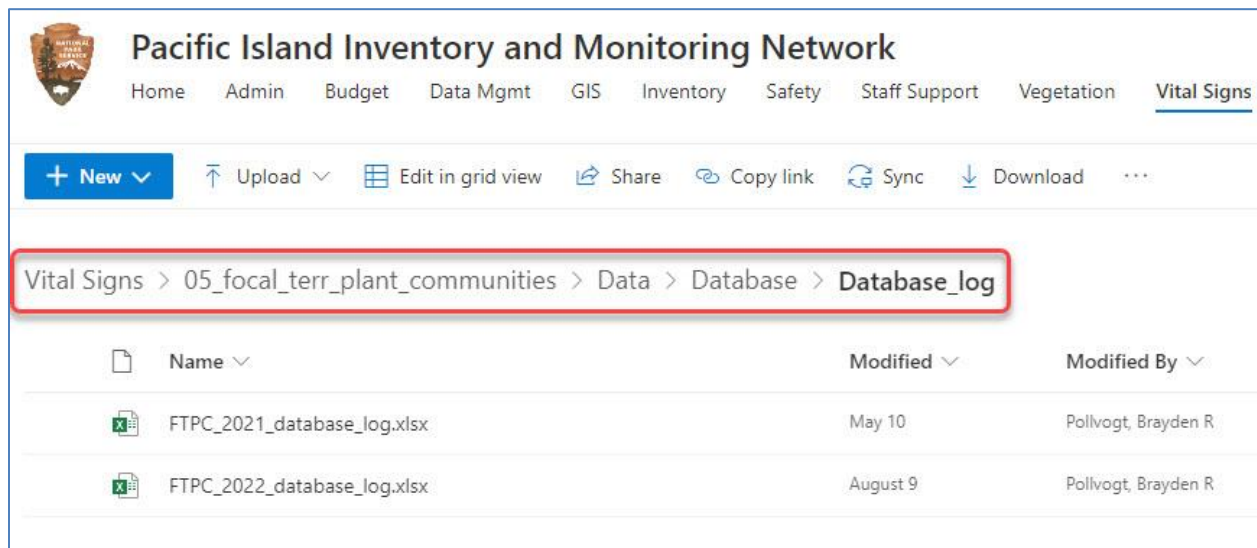
Figure SOP 17.1. Folder location in SharePoint for the FTPC Database Log Microsoft Excel® spreadsheet for keeping track of data entry and verification.