

National Park Service
Inventory and Monitoring Division



Data Management—Version 1.0

Bracewell, J., W. Granger 2017. Data Management—Version 1.0. Gulf Coast Network Standard Operating Procedure NPS/GULN/SOP—SHR08. Gulf Coast Network, Lafayette, Louisiana.

Proper data management and standardization are essential to the effective utilization of information gained through long-term ecological monitoring activities. This standard operating procedure (SOP) provides guidelines to the GULN data manager and to the GULN project leader for the development, storage, and distribution of monitoring data associated with the network's shoreline position monitoring protocol. The SOP describes how the many pieces of information generated at various points in the data-gathering, analysis, and reporting processes are to be organized, stored, and disseminated.

A formal information management plan for all GULN monitoring protocols is available at <https://irma.nps.gov/DataStore/Reference/Profile/2194250> (Granger 2007) and should be consulted for more detailed procedural and policy information.

[illegible]

Contents

	Page
Summary.....	1
Revision Log.....	1
Shoreline Data Set and Products.....	3
File Management and Documentation	4
Generalized Shoreline, Baseline and Transects.....	4
Raw GPS Data.....	5
Corrected GPS Data	5
Shoreline Shapefiles.....	6
Field Data Forms	6
Digital Photos	6
Distance Matrices	7
Reports.....	7
Survey Planning Documents and Data	7
Protocol Implementation Plan and Related SOPs	8
Database Administration and Maintenance	8
Data Entry QA/QC	8
Database Backup	8
Database Revision Control	9
Archiving	9
Data Distribution.....	11
Literature Cited	12

Shoreline Data Set and Products

The shoreline monitoring program produces a number of electronic files that includes formal written reports, GPS and GIS spatial data sets, a Microsoft Access database, and paper field forms. The large number of files in various formats will require conscientious and formal attention to their management.

In addition to the collection and initial processing of shoreline positional data, periodic analysis of shoreline change using the United States Geological Survey (USGS) Digital Shoreline Analysis System (DSAS) will be conducted. This analysis involves the calculation of distances between multiple shoreline features and an established baseline along pre-defined azimuths, resulting in a matrix of distance measurements that can be used for subsequent statistical trend analyses.

Finally, outside investigators may be contracted to produce long-term (10-year) reports summarizing and interpreting the results from the above analyses. Copies of these reports will be delivered to the GULN project leader and stored primarily in Adobe .pdf format.

Several products generated from the GULN shoreline monitoring protocol will be maintained as long-term records and have specific storage and archiving requirements. These products include:

- Original hard-copies or scanned versions of all field data sheets (.pdf);
- All photographs taken during survey (.jpg);
- Raw GPS data files (e.g., .ssf format for Trimble) downloaded from the GPS unit;
- Differentially-corrected GPS data files (e.g., .cor for Trimble units);
- Exported shapefile for editing (.shp);
- A cleaned, exported ESRI point shapefile (.shp) (if applicable);
- A continuous shoreline line-feature shapefile (.shp);
- Generalized shoreline, baseline and trendline (Personal Geodatabase, .mdb);
- Exported shoreline distance table (.csv);
- Matrices of shoreline change distances (.xlsx);
- Shoreline monitoring database (.mdb);
- Metadata file for the above geospatial data which complies with Federal Geographic Data Committee (FGDC) standards (.xml format); and
- Project Reports (.pdf).

File Management and Documentation

It is the responsibility of the GULN project leader to assemble and maintain the various components of the shoreline data set described above. The shoreline monitoring protocol generates a number of raw data products and reports that are not suitable for direct storage within the database, including uncorrected (.ssf) and corrected (.cor) GPS data files, as well as exported ESRI shapefiles (.shp). In order to ensure proper functioning of the shoreline change monitoring database and to generally ease the data retrieval process, the Database User's Guide (Bracewell 2017f) identifies a coherent, consistent directory structure to store data files (Figure 1).

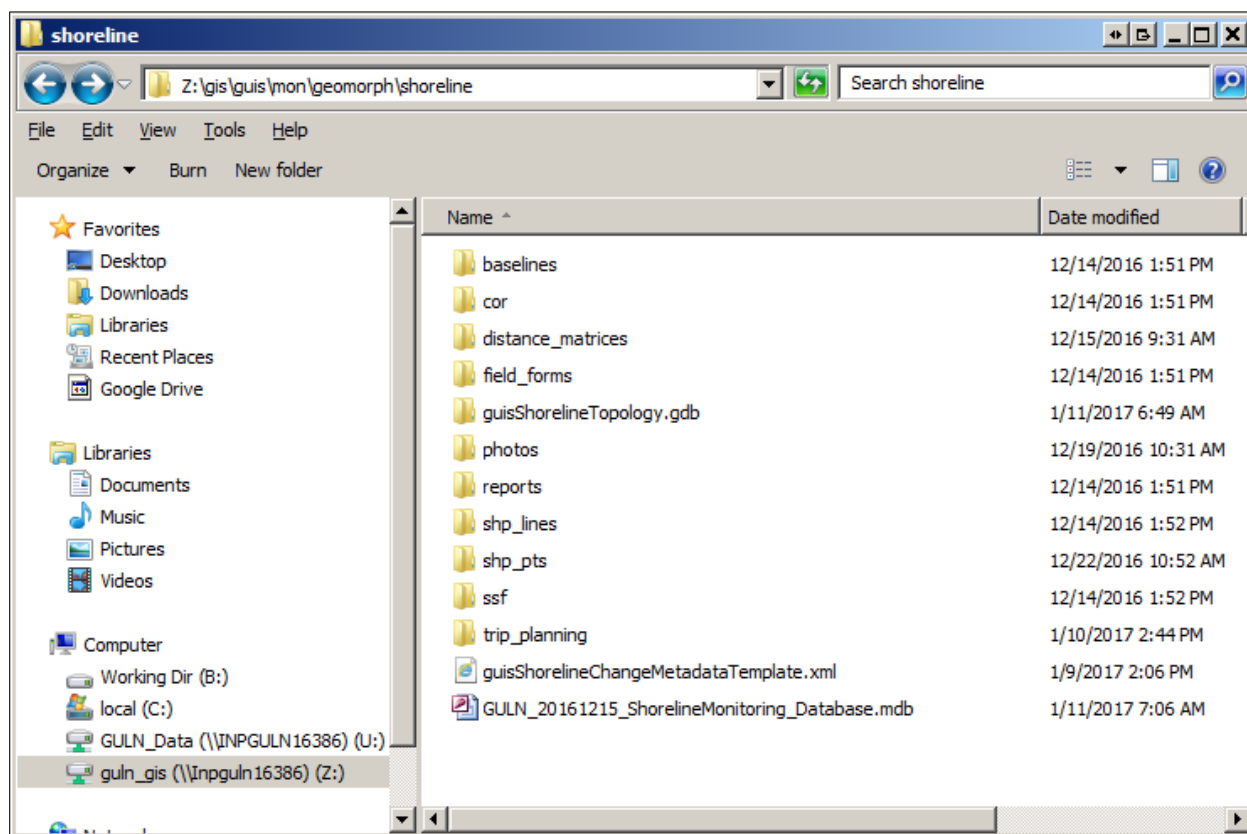


Figure 1. Top Level Shoreline Monitoring File Structure.

In order for hyperlinks within the database to work properly, the main folder 'shoreline' (Figure 1) should be placed at the same directory level (i.e., within the same folder) as the Microsoft Access monitoring database itself.

This directory structure, along with naming conventions for the various file types, is outlined in detail in Section 1.1 of the Database User's Guide (Bracewell 2017f).

Generalized Shoreline, Baseline and Transects

For each monitoring location, investigators will have developed a generalized shoreline trend line, against which all subsequent GPS'ed shorelines will be compared during shoreline change analyses.

Along with these baselines is a suite of permanent transects used to measure distances between the shoreline position and the baseline at approximately 10-meter (32.8-ft) intervals. (These analyses and the procedures used to generate the baseline and transect datasets are outlined in detail in GULN SOP SHR07—*Change Calculation, Data Analysis, and Reporting*; Bracewell 2017d). Each baseline and transect data set is stored as a park, or park unit specific, feature class within a DSAS configured ESRI Personal Geodatabase. Each feature class includes tabular attributes required by the DSAS application for analysis. Distance tables produced in DSAS are exported and appended to a park's Shoreline Distance Matrix.

Raw GPS Data

The original GPS data points are an integral component of the shoreline data set, allowing for the re-creation and replication of any subsequent calculations or derivations. The raw data file directly downloaded from the GPS data collector contains information beyond the horizontal point position, including precision and accuracy of the position, time and date of collection, equipment used, Positional Dilution of Precision (PDOP), number of satellites, and signal-to-noise ratio. If Trimble GPS data collectors such as the ProXR or GeoXT are used, these data will be in the proprietary .ssf file format. Different file formats may be generated if a different sub-meter accuracy GPS unit is utilized.

If there is any question whether the GPS data collection process was compromised, or accuracy thresholds exceeded (see GULN SOP SHR06—*Initial Post-Survey Processing*; Bracewell 2017c), the GULN project leader should refer to these files to determine that GPS unit settings for data collection have been correctly applied. Because these files are typically in non-tabular and proprietary formats, they are linked in the database rather than entered as text. Step-by-step procedures for linking these files are contained in the Database User's Guide, Section 5.1 (Bracewell 2017f).

Corrected GPS Data

Post-processed data products, described in GULN SOP SHR06 (Bracewell 2017c), consist of differentially-corrected versions of the raw GPS data file (e.g., a .cor file, in the case of using Trimble units). An accompanying log file is created along with the .cor file and contains references to the CORS beacon base files used for correction, as well as estimates of the accuracy of the positions. This log file is bundled with the .cor file (i.e., placed alongside the .cor file in the same directory location). Like the raw GPS file, the .cor file is in a non-tabular and proprietary format and is linked to the database as opposed to entered as text (see Section 5.1 of Database User's Guide; Bracewell 2017f).

It is recommended the GULN project leader inspect the .cor and log file in the unlikely event a substantial number of positions were not able to be corrected. In such cases, correction data from one or more nearby beacons should be downloaded and the original .ssf file re-processed to yield additional corrected positions.

Shoreline Shapefiles

In addition to the .ssf and .cor file, GULN SOP SHR06 (Bracewell 2017c) outlines the procedures for correcting the GPS files and creating a cleaned shoreline line-feature shapefile. This line data set represents the final GIS product on which all subsequent trend-analyses are based. Required attributes for these shapefiles are specified in GULN SOP SHR06 (Bracewell 2017c).

The ESRI line shapefile, similar to the raw GPS .ssf and corrected .cor files, is a non-tabular, proprietary data format and is likewise linked to the shoreline monitoring database rather than entered as text. The steps are outlined in the Database User's Guide, Section 5.1 (Bracewell 2017f). The GULN project leader should inspect this final product for completeness against known, trusted GIS base data sets (e.g., the shoreline trendlines or contemporary remotely sensed imagery or elevation described in GULN SOP SHR06; Bracewell 2017c), and ensure that an FGDC-compliant metadata file accompanies the line-feature shapefile. The metadata record is created in ArcCatalog and exported as a standalone .xml file. Metadata in .xml format is viewed in a number of web browsers and metadata editing software packages. A shoreline project metadata template is available at "...\\mon\\geomorph\\shoreline\\guisShorelineChangeMetadataTemplate.xml."

Finally, the completed line shapefile and its metadata will be imported into the DSAS Geodatabase. This can be done in ArcCatalog using the following steps:

1. Navigate to the Shoreline\\baselines folder outlined above.
2. Right-click on the appropriate geodatabase and choose "Import -> Feature class (single)."
3. In the Input Features field browse to the line shapefile.
4. In the Output Feature Class field repeat the name of the shapefile, minus the .shp file extension.
5. Click OK.

Field Data Forms

The Field Data Form completed in conjunction with the GPS position data provides information essential to the usability and verification of the data set. If not already converted, the GULN project leader scans hard-copy field data forms and stores them as Adobe .pdf files. These files are linked to the shoreline monitoring database according to the steps outlined in the Database User's Guide, Section 5.1 (Bracewell 2017f).

In addition to the electronic versions, all original paper Field Data Sheets are housed with GULN central files at GULN headquarters. The central file system is maintained by the GULN data manager.

Digital Photos

Digital photos are taken intermittently throughout the survey and serve to verify and interpret information, such as a specific benchmark location or the shoreline feature being surveyed. Photos

are downloaded and managed with the shoreline project as defined in GULN SOP SHR05—*Conducting the GPS shoreline survey* (Bracewell 2017b).

Distance Matrices

Shoreline position is collected over many years and is processed relative to a pre-determined baseline shoreline feature generated for each park or park survey unit. Distances from the baseline to each shoreline are calculated at 10-meter (32.8-ft) intervals along the shoreline at set azimuths (“Procedural Background” step in GULN SOP SHR07; Bracewell 2017d). The DSAS program provides tools for exporting these distance calculations as a matrix in comma separated value format (.csv), which itself can be assimilated into a Microsoft Access data table.

As periodic analyses are conducted and annual and long-term reports are generated (see GULN SOP SHR07; Bracewell 2017d), additional shoreline distances, recorded in the .csv export, are appended to a park, or park unit’s shoreline distance matrix (Figure 2). The GULN project leader maintains exported .csv files, and incorporates the table into the shoreline monitoring database each time a new shoreline is entered. Procedures are outlined in Section 7 of the Database User’s Guide (Bracewell 2017f).

A		B	C	D	E	G
1		Distance from Baseline				
2		Meters				
3		Transect ID	4/15/00	4/15/01	4/15/02	4/15/03
4	8	-246.6	-250.7	-287.5		
5	9	-239.1	-228.7	-261.9		
6	10	-225.8	-218.3	-239.9	-283.4	
7	11	-211.9	-203.9	-226.3	-280.7	
8	12	-217.9	-208.2	-227.6	-261.3	
9	13	-236.4	-219.2	-221.6	-254.6	
10	14	-246.4	-230.1	-237.1	-262.3	
11	15	-253.9	-238.5	-253.3	-273.3	

Figure 2. Example shoreline distance matrix in Microsoft Excel spreadsheet format.

Reports

The final products included in the shoreline monitoring database are the regular reports produced in GULN SOP SHR07 (Bracewell 2017d). These reports are completed for each park unit. There are two types, a biennial report summarizing the survey event, and a long-term trend report that summarizes and analyzes the trends of changes in shoreline position after five survey events. Reports should be stored as Adobe .pdf files and linked to the monitoring database as described in Section 8 of the Database User’s Guide (Bracewell 2017f).

Survey Planning Documents and Data

Several documents and data sets are used in planning for a survey trip. These documents help to describe survey logistics and the conditional state of survey areas. Tide predictions, PDOP plans and weather forecasts are always accessed for trip planning. Other relevant planning elements include

correspondence with the park about a current project or access restriction that may impact the survey. These ancillary documents will be exported and stored locally (see GULN SOP SHR03—*Survey timing and GPS mission planning*; Bracewell 2017a).

Protocol Implementation Plan and Related SOPs

Because the Protocol Implementation Plan (PIP) and related SOPs are considered living documents that may periodically change and influence the interpretation of survey results, maintenance and access to the complete set of versioned PIPs, SOPs and appendices is required. These documents, and their versioned copies, reside on the GULN Data Server in the shoreline monitoring directory (Figure 3), as well as on the National Park Service's Integrated Reference Management System (IRMA) under the Shoreline Monitoring Project (<https://irma.nps.gov/DataStore/Reference/Profile/2237508>). GULN SOP SHR09 – *Revising the Protocol* (Bracewell 2017e) includes detailed information about tracking protocol revisions.

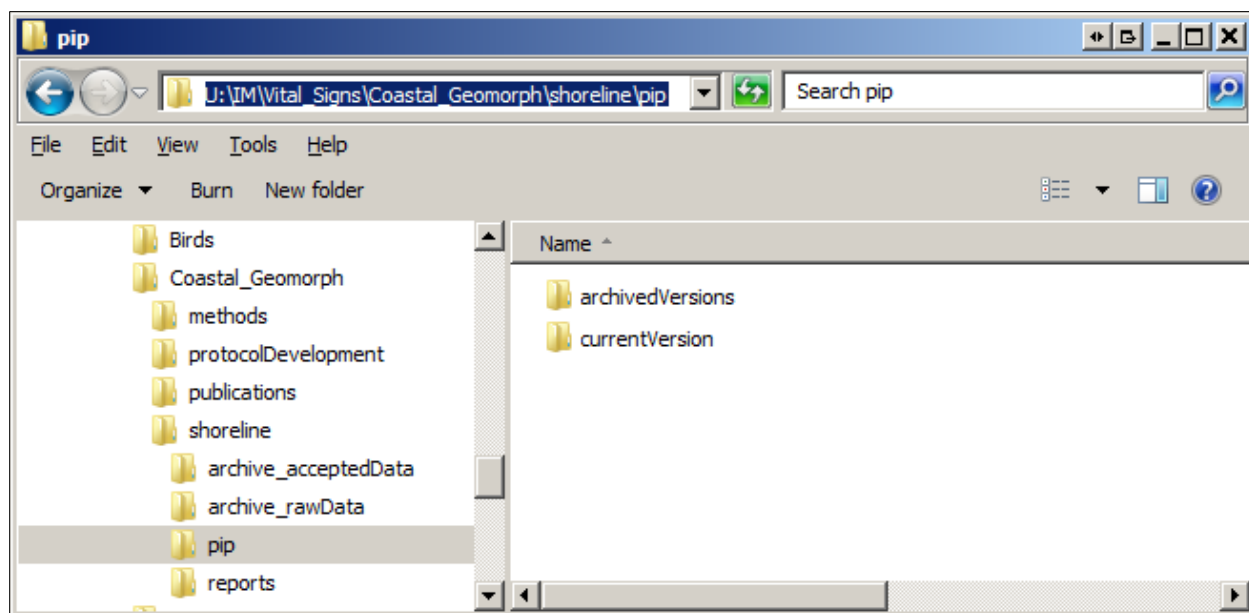


Figure 3. Local directory for PIP and SOPs.

Database Administration and Maintenance

Data Entry QA/QC

The GULN project leader is responsible for entering all pertinent data into the shoreline monitoring database in a timely manner. To minimize transcription errors, each line of data is checked against the original field data sheets by a second person. If no staff member is available, the project leader should check 100% of the data entry the following day.

Database Backup

The database and project directory reside on the GULN server, which is backed up using a tape library and external hard drive. Differential tape backups occur daily with a full back up scheduled

for the last day of each month. Both full and differential backups are retained for three months. A complete replacement backup of the server is copied once a month to an external hard drive.

Database Revision Control

Because the shoreline monitoring database follows the NPS I&M Division Natural Resource Database Template v.3 (NRDTV.3), a table describing database revision history (tbl_Db_Revisions) and linking to the database metadata (tbl_Db_Meta) is an included core table. Revisions are performed periodically by or in conjunction with GULN staff as the need arises and should be documented fully in the above revisions table, as well as in the database metadata. The GULN project leader is also responsible for reflecting pertinent changes to the shoreline monitoring database in the accompanying User's Guide (Bracewell 2017f), as well as this document.

Archiving

Many files are produced in planning and conducting the shoreline monitoring protocol. It is imperative these files are managed and preserved in succinct packages in the course of long-term monitoring so they will be readily available and easily comprehensible. The types of files generated for this project include both raw data and accepted data. Raw data are those in their original format that have not undergone editing or quality control procedures. Accepted data have undergone quality QA/QC measures as prescribed by the protocol and related SOPs. Formal NPS I&M definitions for these data processing levels can be found in the NPS I&M draft quality management plan (Devivo J, Unpublished).

For this project, raw data includes:

1. Uncorrected GPS/GNSS files,
2. corrected GPS/GNSS files,
3. the exported shapefile, and
4. the edited shapefile.

Accepted data includes:

1. Renamed photos,
2. verified field forms,
3. the QA/QCed database,
4. the appended DSAS geodatabase,
5. the exported distance table, and
6. the master distance matrix.

In addition to raw and accepted data, the PIP, related SOPs and appendices, biennial reports, and trend reports are archived on an annual basis in their native formats. Archiving occurs locally on the GULN Data Server and remotely under the Gulf Coast Network Shoreline Monitoring project on IRMA. This IRMA project serves both as an off-site storage container and data delivery mechanism. Table 1 lists the shoreline monitoring archive products and the related reference type used when creating a new IRMA reference. Raw files are packaged together into one zip file; accepted files will be packaged into another zip file. Local archiving occurs on the GULN Data Server in the shoreline monitoring directory at U:\IM\Vital_Signs\Coastal_Geomorph\shoreline using the basic structure outlined in Figure 4. Remote archiving will occur on The Gulf Coast Network Shoreline Monitoring Project, available at <https://irma.nps.gov/DataStore/Reference/Profile/2237508> (National Park Service 2017).

Table 1. IRMA reference type and related material.

IRMA Data Store Reference Type	GULN Shoreline Monitoring Files
Protocol Narrative	Current and Preceding PIP versions
Standard Operating Procedure	Current and Preceding versions of SOPs and appendices
Generic Dataset (1)	Raw Data
Generic Dataset (2)	Accepted Data
Published Report (Data Series)	Biennial Report
Published Report	Trend Report

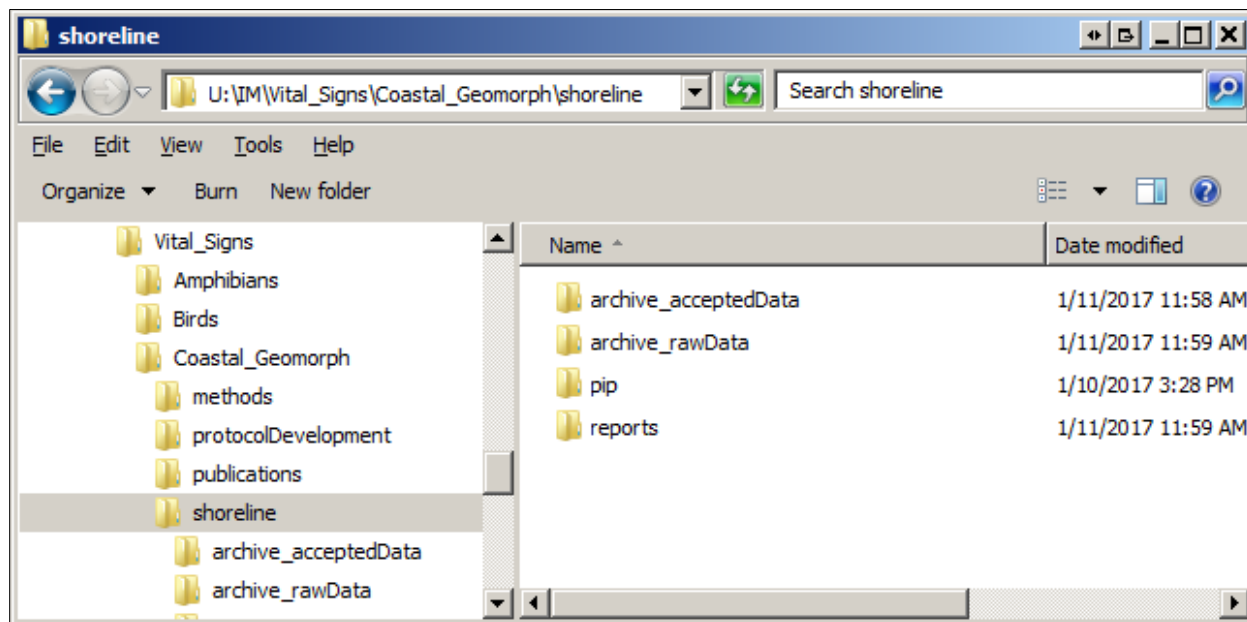


Figure 4. Directory structure for local archive.

Data Distribution

Access to GULN data products will be facilitated via IRMA which allows users to browse, search, and download a variety of NPS-related data and documents. The core elements of this project, as described in the Archiving section of this document, will be organized and published on IRMA under the Gulf Coast Network Shoreline Monitoring Project and can be found at <https://irma.nps.gov/DataStore/Reference/Profile/2237508> (National Park Service 2017).

Literature Cited

- Bracewell, J. 2017a. Survey timing and GPS mission planning—Version 1.0. Gulf Coast Network Standard Operating Procedure NPS/GULN/SOP—SHR03. Gulf Coast Network, Lafayette, Louisiana. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2242635> (last accessed August 2017).
- Bracewell, J. 2017b. Conducting the GPS shoreline survey—Version 1.0. Gulf Coast Network Standard Operating Procedure NPS/GULN/SOP—SHR05. Gulf Coast Network, Lafayette, Louisiana. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2242637> (last accessed August 2017).
- Bracewell, J. 2017c. Initial post-survey processing—Version 1.0. Gulf Coast Network Standard Operating Procedure NPS/GULN/SOP—SHR06. Gulf Coast Network, Lafayette, Louisiana. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2242638> (last accessed August 2017).
- Bracewell, J. 2017d. Change calculation, data analysis, and reporting—Version 1.0. Gulf Coast Network Standard Operating Procedure NPS/GULN/SOP—SHR07. Gulf Coast Network, Lafayette, Louisiana. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2242639> (last accessed August 2017).
- Bracewell, J. 2017e. Revising the protocol—Version 1.0. Gulf Coast Network Standard Operating Procedure NPS/GULN/SOP—SHR09. Gulf Coast Network, Lafayette, Louisiana. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2242641> (last accessed August 2017).
- Bracewell J. 2017f. GULN Shoreline Change Monitoring Database User's Guide. Lafayette, LA. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2238356> (last accessed February 2017).
- Devivo, J. Unpublished. DRAFT Quality Management Plan for the National Park Service Inventory & Monitoring Program. Natural Resource Report NPS/2017/ NRR—2017/XXX. National Park Service, Fort Collins, Colorado.
- Granger, W. 2007. Gulf Coast Network Data Management Plan Version 1.2. Natural Resource Report NPS/GULN/NRR-2007/00X. National Park Service, Fort Collins, Colorado.
- National Park Service. 2017. Integrated Resource Management Applications Gulf Coast Network Shoreline Monitoring Project. Available at: <https://irma.nps.gov/DataStore/Reference/Profile/2237508> (last accessed February 2017).