

Standard Operating Procedure (SOP) 6: Raw HEX Files (OC_B Creation) version 3

Glacier Bay National Park and Preserve oceanographic monitoring protocol

Change History

Previous Version #	Revision Date	Revised By	Changes	Justification	New Version #
2	February 2020	B. Johnson	Replace actions referencing integration website with Data Store. Changes references to “certified” with “accepted”. Rename GLBA server path to \\nps\akrdfs\GLBA\Science\Data\Oceanography.	Integration website was deprecated. Terminology guidance changed by CO. The GLBA file server share had been renamed since the prior SOP version.	3
1	September 2018	B. Johnson	Added Data Steward role.	To lessen workload on PI.	2
-		B. Johnson	Initial version	For original protocol	1

Please cite this Standard Operating Procedure as:

Johnson, W. 2020 Standard Operating Procedure (SOP) 6: Raw HEX Files (OC_B Creation) version 3 Glacier Bay National Park and Preserve oceanographic monitoring protocol. National Park Service, Juneau, Alaska. Available online at <https://irma.nps.gov/DataStore/Reference/Profile/2272144>

Summary

During a cruise, at the end of each day data from each cast recorded on the CTD must be moved to a laptop computer file. The laptop files then undergo basic quality assurance. When the laptop files are judged usable, the CTD memory is erased to make room for the next collection of data.

At the end of each cruise, the Project Leader copies the collected laptop files to a specific location on the park’s network file server.

At the end of each season, the Data Steward zips the network file copies into one file for the year. He submits the one file to the Data Manager as OC_B. Validation and quality assurance iterations are performed until the Project Leader and Data Steward are satisfied the data are as accurate and complete as is practical. The Project Leader accepts the deliverable and the Data Manager arranges for repository and dissemination.

Standard Operating Procedure (SOP) 6: Raw HEX Files (OC_B Creation) version 3
Glacier Bay National Park and Preserve oceanographic monitoring protocol

SOP 6 requires users to work in the internal NPS network under valid Active Directory accounts. Access questions should be directed to the Data Manager.

Detailed steps

A. Project Leader tasks

1. During the field season, the data capture procedures described in SOP 5 must be successfully completed. They provide the required basis for performing this SOP.
2. After the season is over and all captured data have been placed in the proper network directory, notify the Data Steward that OC_B is ready to process.

B. Data Steward tasks

1. Verify all of the year's cruises are present as hex files in
\\nps\akrdfs\GLBA\Science\Data\Oceanography\Data\YYYY\HEX_raw\
 1. If any hex files are missing and cannot be located, this should be marked on the original field log sheet before it is scanned and preserved. That information will be needed when building the OC_M quality assessment and OC_K annual report at the end of the cruise season.
 2. If multiple CTDs were used in a cruise year, then quality control tasks will require the ...\\HEX_raw\ folder to have each CTD hex file set be moved into it for processing and then moved out. When generating OC_B, the combined set of all hex files for the cruise year must be copied into ...\\HEX_raw\.
2. Zip all the hex files in ...\\HEX_raw\ into a single file stored in that same folder using Windows Explorer and Winzip or similar application.
 1. The filename form should be HEX_YYYY_a.ZIP, where YYYY is the cruise year and "a" is A for the first submission, B for the second, etc.
3. Submit the deliverable file by sending an email notification to the Data Manager for validation.
 1. Do not attach the ZIP file to the email, as it may be too large to be delivered.
 2. Specify in the message body that it is deliverable OC_B, as defined in the appropriate protocol package.
 3. Specify the fully qualified filename that the ZIP resides in.
 4. Do not alter the deliverable file after the submission email has been sent; that could break the validation process.

C. Data Manager tasks

1. On receipt of the submission, assign the next formal submission number to this file using the "Update Submission Log" tool on the Data Management website at http://sean-dm.glba.nps.gov/0_submission_update.aspx.
2. Complete Submission_Log details up through the Submission_Date column, entering the submission scope as the four-digit cruise year.
3. If this is the first OC_B submission for the cruise year, update the deliverable tracking spreadsheet in Development with the date of first submission under OC_B.
4. Copy the zip file into the staging area for validation at
\\sean.glba.nps.gov\data\SEAN_Data\Staging\OC\OC_B.
 1. Put it in its own subdirectory whose name is the same as the submission number.

Standard Operating Procedure (SOP) 6: Raw HEX Files (OC_B Creation) version 3
Glacier Bay National Park and Preserve oceanographic monitoring protocol

2. Ensure its name meets the standards defined in the data dictionary (SOP 19).
5. Unzip the file into its component HEX files in the staging area.
6. Invoke validation of the HEX files according to current criteria by using the web program at http://sean-dm.glba.nps.gov/OC_DM_validate_OCB.aspx.
 1. See the data dictionary (SOP 19) for detailed validation criteria.
 2. A findings file is produced in the same directory as the data detailing validation faults.
 3. Validation status is automatically recorded in the Submission_Log.
7. If submission fails mandatory criteria, reply with a “failure email” that includes:
 1. The submission number assigned
 2. The deliverable ID
 3. The protocol package ID
 4. A specific list of all the mandatory criteria failed: attach the “findings.htm” file that was automatically created in the staging subdirectory.
8. If submission passes mandatory criteria, reply with a “success email” that includes:
 1. The submission number assigned
 2. The deliverable ID
 3. The protocol package ID
 4. Documentation listing any specific optional criteria failed (attach findings file)
 5. Request to certify deliverable as complete, or withdraw

D. Data Steward tasks

1. On receipt of a failure email:
 1. Make corrections so the deliverable meets mandatory criteria.
 1. Print out the findings from the email if desired for ready reference.
 2. Expect most of the errors to be related to manually entered header information in the individual files.
 3. It may be convenient to have the scanned field log sheets (OC_H) accessible in order to help resolve exceptions.
 4. Edit each original HEX file using WordPad.
 1. Be aware these are the original source records and editing must be done with care to avoid introducing new errors.
 2. When a file is corrected, save it on top of the existing file.
 5. If a file cannot be made to comply with mandatory criteria (e.g., required heading information is unknown), then rename the file with a “.BAD” extension to exclude it from the processing stream. If possible, also note the specific problem on the OC_H log form, which will inform the OC_M process.
 2. In case of editing accident, files may be restored to original state from the staging copy of previous submission available from the Data Manager.
 3. Restart the process from the beginning by initiating a new submission with the corrected files.
2. On receipt of a success email, confer with Project Leader regarding any failed optional criteria:

Standard Operating Procedure (SOP) 6: Raw HEX Files (OC_B Creation) version 3
Glacier Bay National Park and Preserve oceanographic monitoring protocol

1. If these are acceptable reply to both project leader and data manager with an “acceptance email” stating the deliverable is accepted and may be disseminated.
2. If these are unacceptable:
 1. Reply to both project leader and data manager with a “withdrawal email,” stating the deliverable is withdrawn.
 2. Take remedial action to obtain a corrected deliverable.
 3. Restart the process from the beginning.

E. Data manager tasks

1. On receipt of a withdrawal email:
 1. Mark the withdrawal in the Submission_Log’s Status column using the web tool.
 2. Terminate the process.
2. On receipt of an acceptance email:
 1. If sensitive information is in the deliverable:
 1. Products containing sensitive information cannot be disseminated.
 2. Flag the Submission_Log’s Status column as Sensitive using the web tool.
 3. Leave the file in the staging directory – do not upload it to Data Store.
 4. Notify the project leader and data steward the file(s) were sequestered due to sensitivity but may be retrieved by the data manager on request.
 5. Terminate the process.
 2. Set the acceptance ‘C’ in the Status column of Submission_Log using the web tool. Also update Date Certified and Certified By. Mark as decertified any obsolete prior accepted submission for this submission unit.
 3. Update the deliverable tracking spreadsheet with the date of acceptance for OC_B.
 4. Update Data Store with the new file
 1. Open the Data Store reference for OC_B (currently generic dataset 2257925).
 2. In the generic dataset reference upload the ZIP digital file. Reorder the uploads so they are presented by descending year. If this file covers the same submission period as an already present file, add the word DEPRECATED to the older file but do not delete it.
 3. Update the generic dataset reference to indicate a new Date of Issue and Content End Date.
 5. Verify deliverable(s) are accessible from Data Store and the CMS website.
 6. Update the deliverable tracking spreadsheet with dates for the Repository and Dissemination milestones.
 7. Be sure to retain the individual files in the staging area.
 1. They will serve as file backup for the original CTD data.
 2. The original CTD files on the park server are subject to editing in the validation process, so the staged versions are important as backup.