

SOP 11: Data Quality Assignment (OC_M Creation)

Version 1

Revision History Log:

Prev. Version #	Revision Date	Author	Changes Made	New Version #
[none]	Oct. 2010	S. Danielson, W. Johnson, L. Sharman, G. Eckert, B. Moynahan	Initial version	1

Summary

At the end of the survey year, a review is done to determine if recalibration data indicate sensor errors over the season are significant. The original log sheet images are reviewed for annotations regarding exceptions. The raw and processed vertical profile data plots are reviewed both separately and jointly for data errors.

The Project Leader notes questionable and unusable casts for the year on the data quality report spreadsheets. The Data Manager, after validating the spreadsheet forms, updates appropriate columns in the cumulate database to reflect exceptions to quality.

Detailed Steps

1. Project Leader Tasks

1. Obtain a blank Data Quality Report form from the “toolbox” link on the SEAN web site.

1. Its form is an XLSX spreadsheet.
2. Ensure there are nine tabs, one for each cruise of the year.
3. Head the top of each form with cruise-specific information.
4. Record results of the following analyses on the appropriate sheet.
 1. *Use the order shown.*
 2. *If a cast/station is marked BAD at a point in the process for any reason, no further quality review for that cast is performed.*
 3. *If a cast/station is marked QUESTIONABLE, review continues in order to determine if it should be further marked down to BAD.*

2. Review the latest certified calibration certificates.

1. Examine the certified images on the SEAN website of the pre- and post-season calibration documents (deliverable OC_G).
2. Following the guidance in SOP 9, note any data quality exceptions on the data quality report sheets.
3. The scope of this finding may cover the entire season if a serious drift was present.

3. Review all certified field log sheet images for the year on the web site (deliverable OC_H).

1. Individual cast exceptions may be noted in the COMMENTS field.

2. Overall cruise exceptions and problems may be noted in the NOTES field.
3. Be sure to check the backside image of each sheet for further details.
4. Mark technically flawed casts as BAD on the quality spreadsheet.
 1. *Use the codes listed at the bottom of the sheet, which follow the standard used by the Ocean Data View system.*
 2. *Include a brief explanatory comment. Each comment will be embedded in the historic database to guide future researchers when they select data to work with.*
5. Mark/comment casts as questionable if one or two sensors appear faulty for some or all of a cast.
 1. *It isn't possible to provide a quality rating at the CTD/dump/cast/individual-sensor level, so any problem has to be recorded for the cast as a whole.*
6. Check appropriate boxes on the Data Quality Sheet to mark any failed or disconnected sensors for particular casts or for the dump as a whole.
 1. *This will cause the OC_D database to be null in the corresponding cells, removing the wild readings typical of broken sensors.*
4. **Review the raw in-season profile plots made for individual casts made for QA purposes during the season (SOP 7).**
 1. Mark/comment as questionable casts having one or two improbable plot lines.
 2. Mark/comment as bad casts displaying three or more improbable plot lines.
5. **Review the processed post-season profile plots made for individual casts made for QA purposes during the season (SOP 8).**
 1. Mark/comment as bad casts for which nothing is plotted
 1. *The data processing tasks discard observations that are widely outside acceptable limits – in a seriously failed cast it is possible all points get deleted.*
 2. Compare processed post-season plots against raw in-season plots.
 1. *If strong discrepancies appear, investigate possible causes.*
 2. *The Project Leader's judgment, following currently accepted practices in the field, will determine whether to mark casts as questionable or bad.*
6. **Submit the deliverable file via email attachment to the Data Manager for validation.**
7. **Specify in the message body it is deliverable OC_M, as defined in protocol OC-2010.1.**

2. Data Manager Tasks

1. **On receipt of the submission, assign the next formal Submission Number to this file, as found in the master Submission_Log table.**
 1. Use the "Update Submission Log" web tool at http://165.83.57.239/0_submission_update.aspx.
 2. Complete Submission_Log details up through the Submission_Date column.
2. **Copy the file into the staging area for validation at \\npglbafs03\data\SEAN_Data\Staging\OC\OC_M.**
 1. Ensure its name meets the standards defined in Appendix J.
3. **Validate the XLSX file according to current criteria.**
4. **Record validation summary data in the Submission_Log using the web tool.**
5. **If submission fails mandatory criteria, reply with a "failure email" that includes:**
 1. The submission number assigned
 2. The deliverable ID
 3. The protocol ID
 4. Documentation listing all the specific mandatory criteria failed

6. If submission passes mandatory criteria, reply with a “success email” that includes:

1. The submission number assigned
2. The deliverable ID
3. The protocol ID
4. Documentation listing any specific optional criteria failed
5. Request to certify deliverable as complete

3. Project Leader Tasks

1. On receipt of a failure email:

1. Make corrections so the deliverable meets mandatory criteria.
2. Make another submission with the corrected deliverable candidate.

2. On receipt of a success email, review any failed optional criteria:

1. If these are acceptable:
 1. Reply with a “certification email” stating the deliverable is certified and may be disseminated.
 2. Store log sheets in an annual paper file kept by the Project Leader.
2. If these are unacceptable:
 1. Reply 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.

4. Data Manager Tasks

1. On receipt of a withdrawal email:

1. Mark the withdrawal in the Submission_Log’s Status column with the web tool.
2. Terminate the process.

2. On receipt of a certification email:

1. Due to the nature of these data NO sensitive information is in the deliverable.
2. Copy the submitted single spreadsheet file to test environment at AUXREP\OC\OC_M\yyyy.XLSX.
3. Create web page links to the new file.
4. Update the test database contents.
 1. Use the OC_M update tool on the Data Management web site at http://165.83.57.239/OC_DM_update_OCM.aspx.
5. Mark the certification in the Status column in Submission_Log using the web tool
6. Propagate from test to production environment.

3. Update the annual deliverable tracking spreadsheet with the date of completion for OC_M.

4. Update the scope of the formal metadata so it includes the new date coverage.