

Sierra Nevada Network Lake Monitoring Protocol

SOP 12. Data Management

Version 1.00, October 2007

Prepared by Andrea M. Heard

Revision History Log

Previous Version #	Revision Date	Author	Changes Made	Reason for Change	New Version #

SIEN Lake Monitoring Protocol

Overview

Data take on different forms during various phases of a project, and are maintained in different places as they are acquired, processed, documented, analyzed, reported, and distributed. The SIEN Data Management Plan models the recommended flow of data using a "data life cycle" model. We have adopted this model so it reflects the flow of data for the lake protocol (Figure SOP 12.1). Since data management actions occur throughout the monitoring process, procedures are located across several SOPs. This SOP walks you through the data management steps, from acquisition to dissemination. It points to specific SOPs and sections as appropriate. Significant portions of this SOP were adapted from the SIEN Data Management Plan (Cook and Lineback 2008).

1. *Acquire data:* Data are acquired in analog and digital form. Analog data include the field data sheets. Digital data include datalogger output files and laboratory results.
2. *Archive raw data:* Copies of all raw data files are archived intact. Digital files are copied to the digital library (the set of LAN folders created for the project); LAN folders are backed-up per the SIEN Data Management Plan (Cook and Lineback 2007). Hard copy forms are copied and placed in the archives; they may also be scanned and placed in the digital library. Archiving or scanning of hard copy data forms occurs at the end of the season as a means of retaining all marks and edits made during the verification and validation steps.
3. *Enter/import data:* Analog data are entered manually. Refer to SOP 13: Database User's Manual, Section 'Entering New Data' for data entry procedures. Digital data files are uploaded to the working database. Refer to SOP 13: Database User's Manual Section: Streamflow Data, for datalogger file importing procedures and Section 'Entering New Data-Lab Results' for importing laboratory results.
4. *Verify data*– The technician entering data follows *visual review after data entry* procedures to verify accurate transcription for 100% of the data. Corrections are made immediately and tracked in the editing log book. Refer to SOP 4: QAPP, Section D2.1. for verification procedures.
5. *Verify and validate:* At the completion of the field season, the protocol lead verifies data entry for 10% of the records. Data are processed to remove missing values and other flaws. Once verification is complete, it is documented in the database by checking the 'Verification Complete' box. The protocol lead also validates the data set through visual inspection and queries to capture missing data, out-of-range values, and logical errors. Refer to SOP 4: QAPP, Section 4.2.1. and 4.2.2. for verification procedures and SOP 13: Database User's Manual, Section: Verifying and Validating the Data.
6. *Documentation and certification:* Develop or update project metadata and certify the data set. Certification is a confirmation by the project leader that the data have passed all quality assurance requirements and are complete and documented. It also means that data and metadata are ready to be committed to the permanent database, uploaded to CEDEN and STORET, and disseminated. Certification is documented using the Project Data

SIEN Lake Monitoring Protocol

Certification form (attachment to this SOP).

7. *Commit data*: Certified data are committed from the working database to the permanent project database. This step might be skipped for short-term projects where there is no need to distinguish working data for the current season from the full set of certified project data. Refer to SOP #13: Database User's Manual, Section: Entry to Permanent, for data committing procedures.
8. *Archive versioned data set*: Copies of the certified data and metadata are placed in the digital library annually. This is accomplished by exporting data to a more software-independent format (e.g., ASCII text).
9. *Disseminate data to state and national databases*: Certified data and metadata, and digital image products are posted to state and national repositories to make them more broadly available to others.

Data are sent to NPS-WRD on an annual basis, following each field season. NPS-WRD performs qa/qc analyses, validates the data, and uploads it to EPA's STORET. Data are exported to NPSEDD Excel spreadsheets from the permanent database and sent to NPS-WRD (Refer to SOP 13: Database User's Manual, Section Exporting Data-STORET, for export procedures).

On an annual basis, data are also uploaded to the California Environmental Data Exchange Network. (*This export feature will be developed by the State of California.*)

Metadata are uploaded to NPS Data Store as needed.

10. *Reporting and analysis*: Certified data are used to generate data products, analyses, and reports, including annual summary reports and comprehensive status and trend reports. This is the protocol leads responsibility during the off-season months.
11. *Distribute information products*: Information products such as reports and maps are disseminated to the public through the SIEN website and NPS Focus, and catalogued in NatureBib as they are produced.
12. *Share data and information*: Data, metadata, reports, and other information products can be shared in a variety of ways: by FTP or mailing in response to specific requests, or by providing direct access to project records to park staff and cooperators.
13. *Edits*: All subsequent changes to certified data are documented in an edit log, which accompanies project data and metadata upon distribution. Significant edits will trigger reposting of the data and products to national databases and repositories.
14. *Store products*: Reports and other data products are stored according to format and likely demand, either in the digital library, on off-line media, or in the document archives.

SIEN Lake Monitoring Protocol

15. *Catalog project products*: Catalog products and all information associated with a project, including results of analyses and paths of dissemination, in SIEN's project tracking database.

Literature Cited

Cook, R.R. and P. Lineback. 2008. Sierra Nevada Network data management plan. Natural Resource Report NPS/NRPC/NRR--2008/070. National Park Service, Fort Collins, Colorado. <http://science.nature.nps.gov/im/units/sien/>.

SIEN Lake Monitoring Protocol

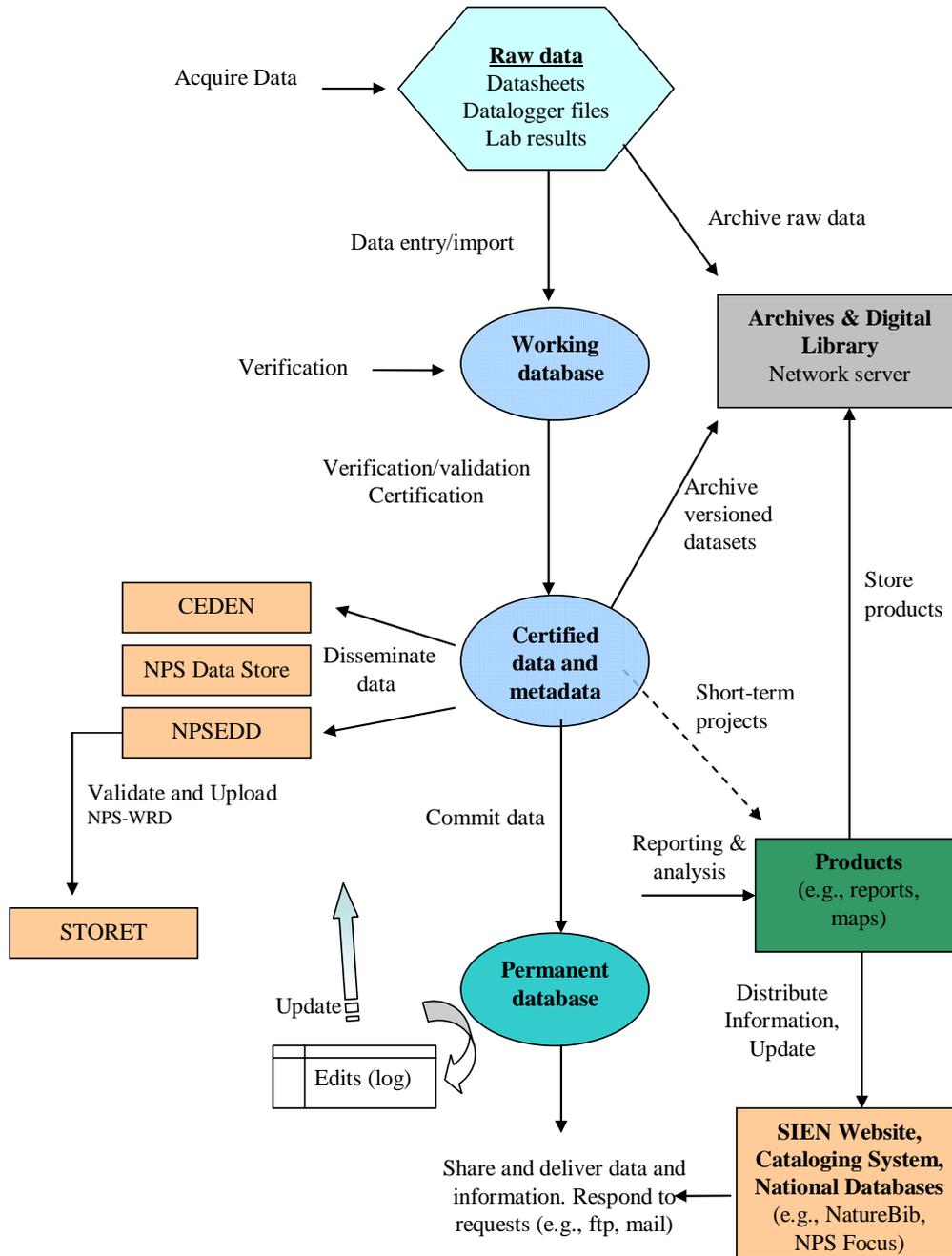


Figure SOP 12.1. Data flow.

SIEN Lake Monitoring Protocol

Project Data Certification Form

Data certification is a benchmark in the project information management process and indicates that: 1) the data are complete for the period of record; 2) they have undergone and passed specified quality assurance checks; and 3) that they are appropriately documented and in a condition for archiving, posting and distribution as appropriate. Certification does not guarantee that the data are completely free of errors or inconsistencies which may or may not have been detected during quality assurance reviews.

1) Certification date: _____

2) Certified by: _____

Title: _____

Affiliation: _____

3) Project code: _____

Project title: _____

4) Range of dates for certified data: _____

5) Description and scope of data being certified:

6) List the parks covered in the certified data set, and provide any park-specific details about this certification.

Park	Details

7) This certification refers to data in accompanying files. Check all that apply, and indicate file names to the right:

_____ Database file(s):

_____ Spatial data theme(s):

_____ Geodatabase file(s):

_____ Other (specify):

SIEN Lake Monitoring Protocol

_____ Certified data are already in the master version of a park, Network or NPS database. Please indicate the database system(s):

8) Is there any sensitive information in the certified data which may put resources at greater risk if released to the public (e.g., spotted owl nest sites, cave locations, rare plant locations)?

_____ No _____ Yes

Details:

9) Description of data processing and quality assurance measures. (Note: These can be cut and pasted from appropriate sections of the protocol.)

10) Results and summary of quality assurance reviews, including details on steps taken to rectify problems encountered during data processing and quality reviews.