

SOP 11: Data Quality Assurance and Quality Control

Version 1.21 (March 2020)

Overview

Careful, accurate data collection and stewardship are critical to the success of the vegetation monitoring program. This SOP describes data quality assurance and quality control (QA/QC) procedures for the vegetation monitoring process. Ensuring data quality must be a priority during training, and throughout the data collection, entry, verification, and validation process.

Field teams will typically consist of a field crew leader, who is the lead botanist, and one or two additional crew members. Ideally all members of the field team will have read all protocol SOPs and attended annual training at the beginning of each field season. However, CUPN invites and encourages park personnel to participate in field data collection as much or as little as their schedule allows. This assistance and interaction in the field is always appreciated! With that said however, it is important to assign tasks based on a team members previous training and experience with this protocol. It has been a common practice in the past to assign minimally trained or experienced crew members the task of data recorder. However, in an effort to evaluate and potentially reduce data recording errors, this is discouraged beginning with the 2020 field season. Instead various tasks have been added to the field SOPs for this protocol that can be assigned to “temporary” assistants with limited training and/or experience. Examples include verifying tag tree information and laying out 10m² quadrats ahead of the botanist.

The crew leader (i.e., lead vegetation ecologist) will:

1. collect data according to established methods
2. possess adequate plant id skills for completing assigned tasks
3. understand how to correctly complete data forms in the field
4. collect and process unidentified plants and/or plants new to the park
5. oversee all activities of crew members

Crew members will:

1. correctly and legibly record all data values on the data sheets
2. complete all tasks assigned by the crew leader
3. adhere to any additional instruction provided by the crew leader in the field
4. when in doubt, about plant id or methods, ask questions

Data Collection

Field data is collected by entering data onto a series of paper datasheets (refer to Appendix D). Field data may in the future be collected using handheld devices or field laptops.

Regardless of the data collection method, all data values shall be called out by the “measurer” and repeated by the recorder when one crew member is measuring and one crew member is recording. The measurer shall clearly call out the value record. The recorder shall clearly repeat

the value aloud as it is recorded, to ensure that the value was heard and recorded accurately. If the measurer does not hear a response from the recorder, the measurer shall repeat calling out the measurement until a response is obtained. If the measurer hears a different value from the value measured, the measurer must immediately correct the value.

Datasheets

Carry datasheets within an enclosed clipboard to keep them clean, dry and secure in one location. **Print legibly** all information on datasheets using pencil.

Once data collection is complete and before leaving the plot, crew members must review all sections of the paper datasheets to ensure all fields are complete and no obvious errors exist. Each data sheet should be reviewed by someone other than the person who originally recorded the data on that sheet. Once it has been determined that the data sheet is complete, the reviewer will sign his or her initials in the “Reviewed by” box in the header for each SOP on the field data sheets. Common data omissions to keep an eye out for include:

1. Decay Class for standing snags on the Tree sheet
2. Incomplete plot diagram
3. Any data associated with the last few trees or herbaceous plants in the plot
4. Failure to map all trees >20cm DBH
5. Coarse woody debris data for both transects
6. Lack of marks indicating which species were collected
7. USNVC classification information
8. Incomplete plot establishment information including detailed directions to plot and total time on plot
9. Failure to record a presence value in the herb (SOP 7) list after having recorded it in tree or seedling data

Field data forms will be maintained by CUPN and the lead botanist on the team will be able to request digital scans of datasheets if needed to aid in plant specimen review. Once all specimens have been reviewed, the lead botanist communicates with CUPN to update the original datasheets, while photocopies/digital scans may remain with the botanist (e.g., NatureServe offices).

Datasheet Review and Midseason Training

Upon return to the office and prior to scanning, datasheets are reviewed again by experienced staff for data collection errors and/or omissions. Substantive or consistent issues are to be communicated back to the field crew members as soon as practicable. Additionally, an interim field season training or check-in will be scheduled to discuss among other topics, an overview of field data collection issues. This session, which will most likely be conducted via teleconference, should at minimum include the protocol lead, data entry staff and those involved in data collection.

Data Entry

The goal of data entry is to transcribe field observations into a computer database with 100% accuracy. Recognizing however, that errors are unavoidable verification steps are to be

incorporated. As a general rule, the following guidelines should be incorporated for data entry for all CUPN projects:

1. To the extent possible and practicable, data entry will be automated
2. Data will be entered as soon as reasonably possible following collection
3. Data entry will be completed by someone familiar with data collection
4. Data will be entered into pre-designed database forms that resemble field sheets and maximize error control.

When the person doing data entry is uncertain about writing, she/he will highlight that information and follow-up with the original data collection team to get clarification on what they intended.

Currently data are entered into the Vegetation Monitoring Database from the paper field forms. **A user manual is currently in development for the database and should be reviewed before attempting data entry. In addition, the data manager should also be contacted prior to initiating data entry to ensure database versions, plant lists, etc. are up to date and to determine what level of data entry training is in order.**

The database incorporates many functions and utilities to help reduce data entry errors such as look-up fields and appropriate upper and lower bounds for fields. However care should be taken to avoid any potential entry errors. If there are any uncertainties about an entry (e.g., data are illegible or inconsistent with the database), do not guess. Instead, mark the data sheet with a Post-it Note. Describe the illegible or inconsistent entry on the Post-it Note with an arrow pointing to the questionable item. In addition, utilize the data entry notes fields in the database, as these will be reviewed during the data validation process.

After each plot is entered, the person entering the data should print their initials and date in the lower right corner of the field data sheets, page 1 (Refer to Appendix D, Field Data Sheets).

Data Verification

Following data entry, network staff will implement procedures to ensure data are accurately and completely entered into the database. After each section of data is verified, the person doing the verification will add their initials and date in the lower right corner of the field data sheets, page 1. If data verification finds significant problems with the data entry for a particular plot or park, this will trigger a full review of all data entered for that plot or park. In addition, steps should be reviewed by the data manager and/or project leader to correct deficiencies and avoid a repeat of future errors.

Data Validation

Following data verification, summarize and review the data to identify any potential mistakes. Verification can include:

1. Run summary queries, and tallies on the data
2. Identify out-of-range values, including new species
3. Evaluate outliers
4. Compare plot location coordinates in database with coordinates in GIS files from GPS unit.

Observer Bias

To determine whether the plots are being measured consistently by field crews, randomly selected plots will be re-sampled during each field season. Before the beginning of each field season, the GIS specialist will randomly choose two plots in each of three previously selected park units (5% of all plots sampled in any given year) to be re-sampled that field season. Logistics will be worked out such that a different field crew (from the one that did the original plot sampling) will be selected for plot re-sampling. Each year, a new selection process will take place so that the same plots will not be used every year for data quality assessment. This data will be analyzed and compared with the data collected by the primary data collection team to make sure that measurement techniques yield consistent results. If significant discrepancies are noted between the two teams that have measured the same plots, a re-evaluation of sampling techniques and training will be undertaken and adjustments will be made in future years to try to address the discrepancy.

To control trampling impacts, re-sampling should not occur on plots which are particularly sensitive (e.g, high moss cover, steep unstable slopes), and no plot should be measured more than three times in four years.

If consistent observer biases are seen, there are a few options:

1. Mid-season training if the problems are caught early
2. Renewed focus on training for the measures that were off in future years
3. “Disqualification” of some data if it is inconsistent between observers

References

- Chapman, A.D. 2005. *Principles of Data Quality*, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen. <http://www2.gbif.org/DataQuality.pdf>
- Perles, S., J. Finley, and M. Marshall. 2010. Vegetation and soil monitoring protocol for the Eastern Rivers and Mountains Network, Version 2. Natural Resource Report NPS/ERMN/NRR-2010/183. National Park Service. Fort Collins, Colorado.
- Wagner G. 1999. Basic approaches and methods for quality assurance and quality control in sample collection and storage for environmental monitoring. University of the Saarland, Centre of Environmental Research, Institute of Biogeography, 66041, Saarbrücken, Germany. [http://dx.doi.org/10.1016/0048-9697\(95\)04830-8](http://dx.doi.org/10.1016/0048-9697(95)04830-8)

Revision History

Version numbers are incremented by decimals where minor grammatical changes or textual clarifications are incremented by hundredths (e.g., 1.01, 1.02, 1.03, . . . etc.) and moderate changes, such as equipment updates or changes in methodology that do not impact analysis are incremented by tenths (e.g., 1.1, 1.2, 1.3 . . . etc.). Whole number increments (e.g., 1.0, 2.0, 3.0, . . . etc) are reserved for major revisions. Of course, differentiating between what is moderate vs. major may not always be clear, but it is anticipated that with a little “common sense” and discussion protocol leads can adequately document and track revisions through time to the benefit of future data users. It should be noted that prior to 2020 all revisions were incremented by tenths.

Revision History Log

Version #	Date	Revised by	Changes
1.1	01/15/12	B. Moore and R. White	- Additional details about field team make up and roles were added, as well as encouragement to involve park staff in sampling - Made reference to protocol's Appendix D in this SOP and added copy of field datasheet and field datasheet descriptions to Appendix D. Previously Appendix D was a placeholder - Updated list of common data omissions - Updated directions that datasheets should remain with individual identifying any collected specimens (plot botanist) until this identification is complete. Originals are returned to CUPN headquarters. Ideally this is done withing 3 weeks of plot work - Changed references to protocol database from future tense to present. Protocol database was not complete when written protocol initially approved - Made reference to the use of "Data Entry Notes" field in the protocol database. Data entry personnel can leave notes for future data processing steps (e.g., verification, validation...) - Removed "review 10% of all entered data" from verification section (all entries should be reviewed) - Other minor editorial changes
1.2	9/27/12	B. Moore and R. White	- Updated list of common data omissions - Instructed data entry personnel to follow-up with field crew regarding any illegible handwriting
1.21	3/25/2020	B. Moore	- Added text discouraging use of minimally trained or experienced crew members as data recorders - Added plan to have a midseason training or recalibration where data collection errors will be discussed with field crew members - Moved Revision History to end of SOP and added more detailed list of changes made to date