Vegetation Community Monitoring Protocol for the Heartland Inventory and Monitoring Network

Standard Operating Procedure 12: Double sampling

Version 1.00 (2019)

Revision History Log:

| Previous Version # | Revision Date | Author | Changes Made | Reason for Change | New Version # |
| --- | --- | --- | --- | --- | --- |
|  |  | S.A. Leis | New SOP | Improve data quality | 1.00 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Introduction

Error introduced by different observers identifying species can be difficult to measure. Never-the-less, doing so can provide insight into data trends over time and serve as a tool to improve team performance. The process of measuring the rate of agreement among observers improves quality in the field and in post-collection processing and reporting opportunities.

Measuring error rates is intended to evaluate team performance (sampling error), not individuals. The information is intended to provide transparent information about error rates used for continuous improvement rather than simply evaluation. The double sampling procedure described here used to measure this error can also help observers in the field to strive for accuracy.

Goals of this procedure are based on three assumptions. 1. We assume that two botanists work together to complete observations at each monitoring site. Each botanist samples one of the two transects. 2. Botanists work collaboratively for the first site at a new park or plant community. Because of this effort to calibrate and establish basic flora knowledge, the first site at a park or plant community should not be included in the double sampling effort. 3. Double sampling should not influence the way plots are sampled in terms of normal levels of collaboration.

Equipment

* See prior SOPs for equipment lists for sampling gear. In particular, SOP 1-Preparations and Equipment Setup Prior to the Field Season.

Additional equipment

* Blank or duplicate datasheets
* Pre-randomized list of sites for double sampling
* Randomization app

Goals

Accuracy goals were developed based on existing research and guidance.

1. At least 10% of monitoring plots within a reference frame for a given sampling event should be re-sampled for quality assurance purposes. Additionally, a sufficient number of plots should be resampled for calculation of means and variances (N ≥ 3 plots). As a result, at small parks 1 or 2 plots in every site (excluding the first site) may be re-sampled.
2. Agreement of species-level observations: ≥ 80% (averaged across sites) between observers during a sample event. This would translate into a rate of 20% pseudoturnover (i.e., differences in species lists due to observer error). In the literature, pseudoturnover typically ranges between 10-30% for inter-observer error and < 15% for intra-observer error (Morrison 2016).
3. Agreement of cover measurements: 95% of observations within 1 class. In the literature, between one-third and one-half of cover class estimates typically disagree, although most differ by only one cover class (Morrison 2016). The initial observation is the authority.
4. Agreement of seedling counts should be within 5% among observers (USDI National Park Service 2003, page 102).

Procedures

Determining sample numbers

Two plots at each of the sites targeted for resampling will be sampled after the initial traditional monitoring is completed. Each botanist will sample their designated transect (five plots). Then botanists will swap transects and sample one plot. For example, Botanist X will sample transect A and one additional plot from transect B. Botanist Y will sample transect B and one additional plot from transect A. The result is sampling of the 10 plots and two additional plots are double sampled.

For small parks or communities (i.e., 2-5 sites): all but the first site will be targeted for double sampling. For larger parks or communities (>5 sites), the number of sites to resample can be based on the 10% of plots goal; see Table 12.0 for target numbers specific to each park. Resampling two plots at each target site is 20% of the site. For TAPR where there are 30 sites to monitor, 29 would be eligible for monitoring. A goal of double sampling 15 sites (2 plots X 15 sites = 30 plots) would meet requirements. However, if time will not allow for that many sites, aim to complete double sampling on a minimum of 10 sites (2 plots X 10 sites = 20 plots.

Table 12.0. Double sampling targets 10% of plots by park. Some park targets are calculated by plant community.

| Park | Community | Sites Installed | Sites to Resample |
| --- | --- | --- | --- |
| EFMO | Goat Prairie | 8 | \*Plot target: 8 |
| EFMO | Forest | 18 | 9 |
| GWCA | Prairie | 7 | 4 |
| HEHO | Prairie | 6 | 3 |
| HOME | Prairie | 7 | 4 |
| HOME | Forest | 3 | 2 |
| HOSP | Forest | 7 | 4 |
| LIBO | Forest | 4 | 3 |
| PERI | Forest | 7 | 4 |
| PIPE | Prairie | 13 | \*Plot target: 13 |
| TAPR | Prairie | 30 | 15 |
| WICR | Prairie | 6 | 3 |
| WICR | Forest | 4 | 3 |

\*Because of sites with only 5 plots, number of sites is adjusted to meet plot target.

Identifying sites to double sample

For small parks, the first site to sample (therefore, is not double sampled) is often chosen for logistical reasons. The project manager should take care to avoid choosing this site with bias to avoid double sampling. At larger parks, where not all sites will be sampled, apply the following process.

Prior to going to the field, determine the number of sites to be resampled based on the above guidance. Conduct this process during the winter months so that you will not remember which sites are to be resampled. A random process should be used to identify which sites will be resampled. <https://www.random.org/> provides tools that can support this work. There you can enter the site numbers and how many sites you want randomly chosen. Print the list for field use but place it in an envelope or sleeve with a sheet of colored paper over the list. Once the sites for double sampling are identified, put the list aside and do not consult it.

In the field, charge a scribe with managing the list of sites to double-samples. This will prevent the botanists from seeing the list and biasing the monitoring. At the completion of sampling plots at a site, the randomly generated list of double sample sites is consulted. If the site in on the resample list, next determine which plots will be resampled. Using a cell phone app such as *Random* enter the numbers 1-5, e.g., 1,2,3,4,5. Ener 1 for the length of characters desired and “,” for the type of character separator. Then click the Generate button. The app can then generate a random number in the given range. The generated number will refer to the order of the five plots along the transect tapelines i.e., if a 3 is generated the 3rd plot would be sampled (on the A line, the plot at 30 m and on the B line, the plot at 20 m would be sampled).

Datasheets

For small parks, two copies of all datasheets are taken into the field. Once the double-sample plot is chosen, choose the correct blank datasheets for each plot. For larger parks, generate a binder that has the prior species lists for each monitoring site and a set of blank datasheets corresponding to the number of plots to be resampled. Once a double sample plot has been identified in the field, the correct species list and a blank datasheet will be given each botanist-scribe team.

Sampling

Double sampling of the plots adds roughly 15 minutes to a site visit. Collect the species data, ground cover, and regeneration data as you would for the initial pass of sampling. Botanists and scribes should remain in the same teams as during the initial pass of sampling. Place sampling frames as you would in the initial pass, without trying to relocate the previous team’s position. Sampling frame placement is part of the error we want to capture.

Although two botanists at each site are preferred, there may be occasions for fewer or more botanists to be present. If only one botanist is available for the sampling, only double sample one plot from each site. When choosing the plot to double sample, number plots 1-10 starting with 10A through 50A and continuing with 0B through 40B. Randomize which of the 10 plots will re-read. If there are more than two botanists, rotate botanists at each site so that everyone has equal chances of leading identification at a site that gets double sampled.

Typically, botanists confer about unusual or unknown species as they read plots. The botanist and scribe also work as a team to identify and estimate species. During double sampling however, the teams may continue to consult among themselves, but should not include the other botanist in their deliberations.

Initially, we thought that teams could compare notes in the field and improve throughout a monitoring event. The time investment and tools to do so are not readily available in the field. As a result, we do not present a formal procedure for using the data in the field.

Data Management

A separate database mirroring the current VegMon database was constructed to house the double sampling data. QA\_VegMon11.6\_Shell.accdb houses this data. It is important make sure that updates to the species look-up table are mirrored in both databases. This can be done by making edits in each database simultaneously or by importing the table from the main database to QA\_VegMonXX.X\_Shell.accdb as needed, most likely at the completion of a project or data entry. Data can be exported from the main database and the quality assurance database for comparison and analysis.

Data Analysis

Species and cover data

Analysis is conducted at the plot scale (10 m2). In a spreadsheet format, paste the original data (plot identifier, species name, cover class) and double-sampled data for each plot in side-by-side columns. Insert blank rows as needed to match up the species lists. Next, identify species in the following categories using a color-coding, tallying, or formula system: overlooking (one botanist saw the plant but the other did not), specificity (one botanist lumped a species to genus while the other was more specific), and misidentification (these can sometimes be harder to spot in the data) (Figure 12.0). Misidentifications include species where published spatial distributions do not include the park and species are unlikely to occur there. Misidentifications may be easier to designate in the case that one botanist lists a species with more appropriate distribution and ecology than the other botanist. Known botanical errors based on field experience with the team, errors resulting from misspellings, and reevaluation of unknowns may also be classified as misidentification. Sum the number of species each botanist recorded and how many species were identified in the error categories. Average the results across all the plots for a monitoring event.

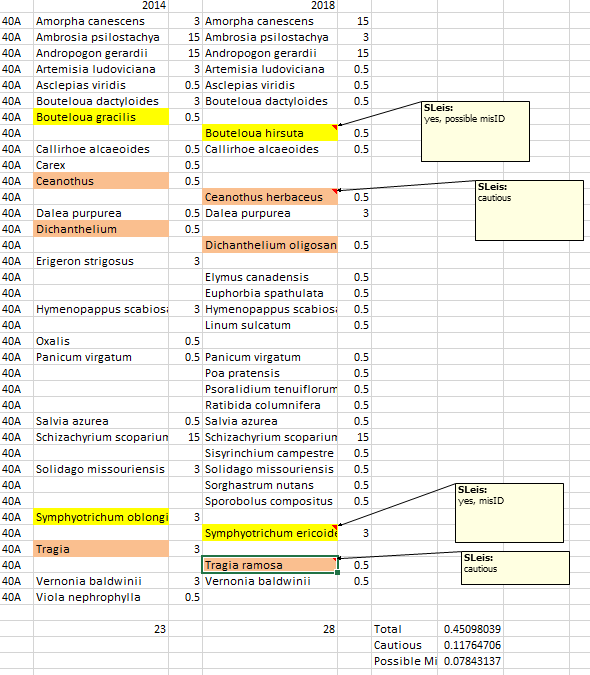


Figure 12.0. Example of error assignments for a double sampled plot. Species are aligned then assigned an error type if appropriate.

Cover class comparison

To compare cover classes, use a function such as ABS to subtract the two observations. If the observations are equal, the value returned will be 0. If cover values are not equal, the value will reflect the number of classes different. Then the number of species with equal cover, cover within 1 class may be summed. The percentage of species with equal cover classes may then be calculated. Overlooked species should not be included in the calculations (Figure 12.1).

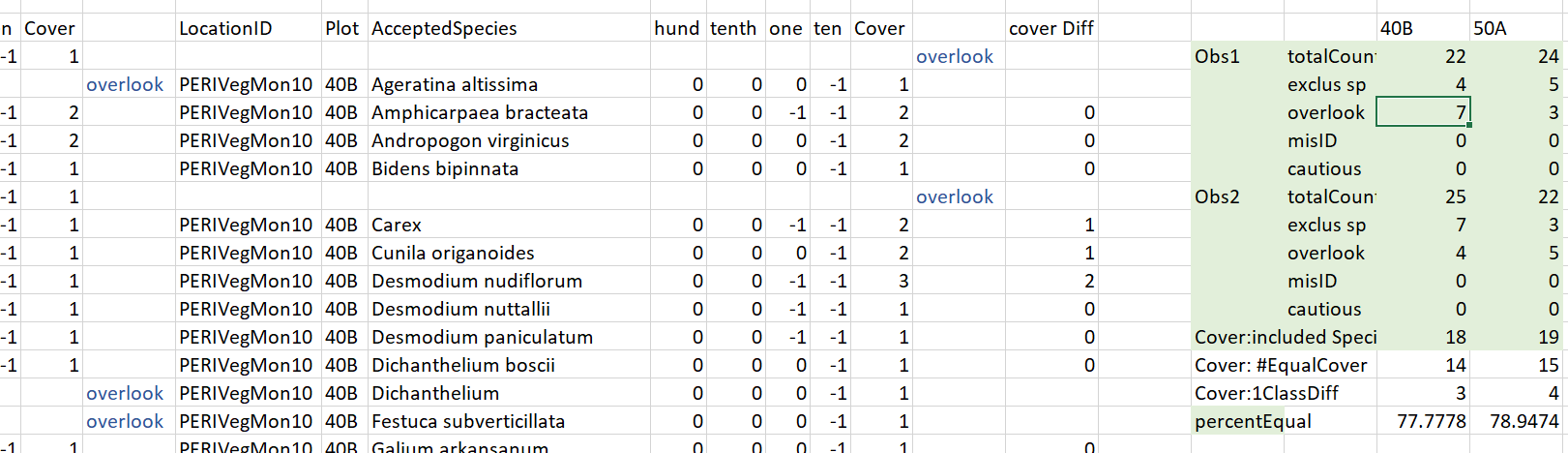


Figure 12.1. Comparing cover classes of two observations for the same plot. The two columns of cover data are differenced in the cover Diff column. The number of equal observations and observations within 1 class are summed and a percentage calculated on the right in a summary area.

Seedling Counts

Seedling counts will be assessed by comparing plots where double sampling took place. Counts from both assessments on a plot will be differenced and if values are outside the tolerance, retraining on technique and identification may be necessary.

Outcomes

If sampling error is within prescribed levels, sampling teams should continue to be vigilant and maintain skills from season to season. If sampling error is above prescribed levels, sampling teams should increase efforts to study and calibrate in preparation for future monitoring events. The calculated error rates should be reported along with other analyses in trend-type reports described in SOP 16-Reporting. Although this does not allow us to create corrections for metrics, it will aid in the interpretation of trends from year to year.

Literature Cited

Morrison, L. W. 2016. Observer error in vegetation surveys: a review. Journal of Plant Ecology 9:367-379.

USDI National Park Service. 2003. Fire monitoring handbook. National Interagency Fire Center, Boise, ID.