

Information on Data Collection and Organization from the SGS-LTER

This data package was produced by researchers working on the Shortgrass Steppe Long Term Ecological Research Project. This project was supported by National Science Foundation from 1982-2014. This data package includes one or more tab-delimited data tables, tab-delimited files that denote header definitions and data types for each column, and detailed metadata within an Ecological Metadata Language document (i.e. XML). Example image files of plots, digital datasheets, or schematics of the experimental design may also be included when applicable. Background information on the SGS-LTER project is contained in related series of objects within the Digital Collections of Colorado and the Colorado State University archives. Together data packages and other background information, and items such as images, proposals, and reports contribute to a comprehensive SGS-LTER collection.

The data tables and associated EML documents represent components of the LTER data package, which may be discovered and accessed through secondary repositories serving specific ecosystem science domains (e.g. PASTA (LTER Network Repository), DataONE, or The Knowledge Network for BioComplexity).

The following information is copied from the SGS-LTER field protocols to provide specific details on how these data were collected.

A General Guide for Producing Lab Nitrogen Files from CHN Data

I. Steps for making the initial “CHN” file for lab use:

- A. Enter Sample ID's in an “ID file”, including lines inserted to identify Blanks and Standards.
- B. Match up “.csv” file containing corresponding CHN results from the LECO to the “ID file” (or manually add CHN results to the “ID file”). This forms one complete basic file that includes all of the preliminary lab information needed to form the next two files (one with final lab data and one with condensed data results to send on to the scientist). Save this preliminary file in the CHN directory for the appropriate experiment.

II. Steps for making the “LAB” file to assess the quality of the LECO run:

- A. Convert “Run Dates” to “Run Numbers”.
- B. Code Blanks and Standards as “1A, 1B, 1C..., 2A, 2B, 2C...etc.”, if it hasn't already been done.
- C. Add a column down the left edge of the file and fill with consecutive numbers for the

length of the file. This is so that the data may always be sorted back to the original running order.

D. Sort/move Blanks and Standards to the end of the spreadsheet file; then, sort experiment samples back to running order.

E. Graph out all data points for complete set of data (excluding Blanks and Standards) to check for outlying values and other anomalies.

F. Blanks:

1. Make sure that the Blanks at the end of the file are in order by run and within runs.
2. Examine values for all Blanks individually, looking for trends within runs to determine if a rerun should be done for this reason.
3. If a Blank was run immediately after the analyzer was opened up by the technician, that Blank value may need to be discarded. It will most likely be an unusual Blank reading. The same may be true for the first Blank run after the analyzer has been idle for a period of time.

G. Standards:

1. Make sure that the Standards are in order by run and within runs.
2. Calculate Percent Error for each individual Standard value. (Should be within + or - 5% of Expected Value)...

$$\% \text{ Error} = \frac{(\text{Actual Value} - \text{Expected Value})}{\text{Expected Value}} * 100$$

...or, use the + and - 5% limits to determine if Standard values are within acceptable limits for the run.

3. Use graphs of Standard values as tools for comparison, if that is helpful to you.
4. If Standard results don't look good for a run or runs, either rerun the affected samples or see the scientist or Dan Reuss at NREL for directions on how to proceed.

H. Duplicates:

1. Sort samples by Field ID to match Duplicates. (This also gives an opportunity to check that samples for all Field ID's have been analyzed when the entire set of samples have been run).
2. Make a separate chart within the spreadsheet file for Dupes; usually this chart is located at the top of the spreadsheet file to the right of the main body of information. Copy Duplicate Sets from the main file to the Dupe Chart; do not remove the Duplicate Sets from the main file.

3. For each set of Duplicates calculate Duplicate Set Means, Variances and Variances as Percent of Duplicate Set Mean.

$$\text{Variance} = (\text{Value 1} - \text{Value 2}) / ((\text{Value 1} + \text{Value 2}) / 2)$$

4. Dupe Set Variance as a Percent of Dupe Set Mean would normally be 10% or less. In fact our values for DSV as % of DSM are usually well below 10%. A value much higher than 10% indicates the need for a rerun of the sample. Also, check vial labels for both duplicate samples to make sure they truly are duplicates. Sometimes errors may be made in data entry of vial ID's.
5. If all looks well with the values obtained, insert the Dupe Mean Values for both Carbon and Nitrogen into the main spreadsheet file and remove extra lines for the Duplicate samples.

III. Steps for making the "FINAL" file for the scientist to use and the data manager to archive:

- A. See the scientist to find out how the file should be formatted.
- B. Save a new copy of the file. Then remove "extra" info. including Blanks, Standards and the Dupe Chart from the new file copy.
- C. If any sample is missing data that will not be attainable (for example, if there is not enough sample to analyze), insert an "M" in the % Nitrogen column, but leave in the Sample ID to make sure that sample is accounted for in the Nitrogen file.
- D. If any sample is missing data because a rerun is necessary, let the scientist know about it and send rerun information to the scientist in a separate file.
- E. Put notes to the scientist at the very beginning of the file, such as, "This data is for the 2003 XXXXX SGS LTER Exp.; the data is complete and ready to use (or incomplete and not ready to use)". Add a sheet to the file for Metadata if necessary.
- F. E-mail copies of the file to the scientist and the data manager.
- G. Make a lab file folder for the printed data and update the Lab Notebook for the experiment, and make sure that all final lab files are saved on ascalon.

Mixed Grass Standard

N = 2.53%
(+ or - 5% = 2.40 - 2.65)

C = 43.78%
(+ or - 5% = 41.59 - 45.97)

Meadow Hay Standard

N = 1.765%
(+ or - 5% = 1.677 - 1.853)

C = 44.615%
(+ or - 5% = 42.38 - 46.85)

Mixed Grass Standard

N = 2.53%
(+ or - 5% = 2.40 - 2.65)

C = 43.78%
(+ or - 5% = 41.59 - 45.97)

Meadow Hay Standard

N = 1.765%
(+ or - 5% = 1.677 - 1.853)

C = 44.615%
(+ or - 5% = 42.38 - 46.85)