NEON USER GUIDE TO LITTER CHEMICAL PROPERTIES AND LITTER STABLE ISOTOPES (DP1.10033.001)

PREPARED BY	ORGANIZATION
Samantha Weintraub-Leff	TOS

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	01/22/2018	Initial Release
B	08/01/2020	Added text on publication of litter chemistry and stable isotope data alongside litter mass fluxes plus discussion of flux-weighted chemistry inputs
C	04/25/2022	Updated section 4.3 Data Revision with latest information regarding data release.
D	12/08/2022	Clarifications throughout to reflect new site-level pooled samples for non leaves and needles functional groups
E	03/19/2024	Section 3.2, clarified bout selection guidance for different functional groups. Section 3.3, more details on interpretation of acid detergent lignin measurements and how the C:N ratio value is generated. Section 4.4, added two new dataQF. Minor text clarifications throughout.
F	02/18/2025	Section 3.10 Special Considerations, added information on possible delta 15N data impact from switching analytical labs in 2024. Added information about the new neonUtilities Python package.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

This document provides an overview of the data included in this NEON Level 1 data product, the quality controlled product generated from raw Level 0 data, and associated metadata. In the NEON data products framework, the raw data collected in the field - for example, the dry weights of litter functional groups from a single collection event, are considered the lowest level (Level 0). Raw data that have been quality checked via the steps detailed herein, as well as simple metrics that emerge from the raw data, are considered Level 1 data products.

The text herein provides a discussion of measurement theory and implementation, data product provenance, quality assurance and control methods used, and approximations and/or assumptions made during L1 data creation.

1.2 Scope

This document describes the steps needed to generate the chemistry and stable isotope components of the L1 data product Litterfall and fine woody debris production and chemistry (DP1.10033.001). This includes litter carbon (C) and nitrogen (N) concentrations and stable isotope ratios as well as litter 'lignin' as quantified by the acid detergent method. Prior to 2020, Litter chemical properties (NEON.DP1.10031) and Litter stable isotopes (NEON.DP1.10101) were delivered separate from litterfall biomass, but starting in August 2020 all three data types are delivered together. This document also provides details relevant to the publication of the data product via the NEON data portal, with additional detail available in the files NEON Data Variables for Litterfall and fine woody debris production and chemistry (DP1.10033.001) (AD[06]) and NEON Categorical Codes for Litterfall and fine woody debris production and chemistry (AD[07]), provided in the download package for this data product.

This document describes the process for ingesting and performing automated quality assurance and control procedures on C, N, and lignin concentration and stable isotope data measured in NEON litterfall samples. How the Level 0 data are processed is detailed in the files NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001) (AD[04]) and NEON Raw Data Validation for Plant lignin concentrations (DP0.10031.001) (AD[05]), provided in the download package for this data product. Please note that raw data products (denoted by 'DP0') do not always have the same numbers (e.g., '10033') as the corresponding L1 data product.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Associated Documents

AD[01]	NEON.DOC.000001	NEON Observatory Design (NOD) Requirements
AD[02]	NEON.DOC.000913	TOS Science Design for Spatial Sampling
AD[03]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products Catalog
AD[04]	Available with data download	NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001)
AD[05]	Available with data download	NEON Raw Data Validation for Plant lignin concentrations (DP0.10031.001)
AD[06]	Available with data download	NEON Data Variables for Litterfall and fine woody debris production and chemistry (DP1.10033.001)
AD[07]	Available with data download	NEON Categorical Codes for Litterfall and fine woody debris production and chemistry
AD[08]	NEON.DOC.001710	TOS Protocol and Procedure: Litterfall and Fine Woody Debris
AD[09]	NEON.DOC.000914	TOS Science Design for Plant Biomass, Productivity, and Leaf Area Index
AD[10]	NEON.DOC.000906	TOS Science Design for Terrestrial Biogeochemistry
AD[11]	NEON.DOC.000008	NEON Acronym List
AD[12]	NEON.DOC.000243	NEON Glossary of Terms
AD[13]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[14]	Available on NEON data portal	NEON Ingest Conversion Language Function Library
AD[15]	Available on NEON data portal	NEON Ingest Conversion Language

2.2 Acronyms

Acronym	Definition
δ13C	delta 13C, the stable carbon isotope ratio (13C:12C) in a sample compared to a reference material, reported in parts per thousand
δ15N	delta 15N, the stable nitrogen isotope ratio (15N:14N) in a sample compared to a reference material, reported in parts per thousand
C	Carbon
N	Nitrogen

3 DATA PRODUCT DESCRIPTION

Litter chemistry and stable isotope data provide measurements of litterfall samples collected using TOS Protocol and Procedure: Litterfall and Fine Woody Debris (AD[08]). Litterfall is defined as material dropped from the forest canopy with a butt end diameter < 2 cm and a length < 50 cm, collected in elevated 0.5 m^2 PVC traps. Fine woody debris is also collected using ground traps during execution of TOS Protocol and Procedure: Litterfall and Fine Woody Debris (AD[08]), but is not measured for chemistry or isotopes so will not be discussed further here. Following field collection, litter is sorted into 8 functional groups, with mass and chemistry reported for each of them separately. Litterfall sampling and analyses implement the guidelines and requirements described in the TOS Science Design for Plant Biomass, Productivity, and Leaf Area Index (AD[09]) and TOS Science Design for Terrestrial Biogeochemistry (AD[10]). Chemical and isotopic data are reported at the spatial resolution of a plot or site, depending on the functional group. The temporal resolution is that of a single collection event.

Measurements of litterfall carbon, nitrogen and lignin concentrations and C and N stable isotopes help to reveal drivers of variation in aboveground net primary productivity, stoichiometry, foliar resorption, and rates of decomposition at the plot, site, and continental scales. They also provide essential data for understanding change in ecosystem biogeochemical dynamics over time.

3.1 Spatial Sampling Design

Litterfall sampling is executed at each terrestrial NEON site that contains woody vegetation > 2 m tall. Sampling occurs only in base plots that fall in the tower airshed, in 10-30 Tower plots per site depending on the height of woody vegetation, plot sizes, and the degree of spatial heterogeneity in litterfall flux. For more information on the spatial design and how it has changed over time, see the related Data Product User Guide for Litterfall and fine woody debris production and chemistry. For larger Tower plots (40 x 40 m), two elevated litter traps are deployed and material is collected and weighed separately, whereas smaller Tower plots (20 x 20 m) only have one trap. The chemistry of leaf and needle materials are measured at the plot scale, using a pooled sample for larger plots. Starting in 2022 and onward, the other functional groups are measured using a site-level pooled sample, combining material across plots prior to grinding, subsampling, and chemical analyses.

As much as possible, sampling occurs in the same locations over the lifetime of the Observatory. However, over time some sampling locations may become impossible to sample, due to disturbance or other local changes. When this occurs, the location and its location ID are retired. A location may also shift to slightly different coordinates. Refer to the locations endpoint of the NEON API for details about locations that have been moved or retired: https://data.neonscience.org/data-api/endpoints/locations/

3.2 Temporal Sampling Design

Following field collection, litter material is sorted, oven-dried at 65C, and weighed. Litterfall sampling schedules vary according to site vegetation type, with frequent sampling (1 x every 2 weeks) in deciduous sites during senescence and less frequent collections the rest of the year, compared to year-round sampling (1 x every 1-2 months) at evergreen sites. However, only material from a single bout per functional group is analyzed for chemistry and stable isotopes, and this occurs once every five years.

Selection of bouts for chemistry and stable isotopes varies by functional group. For leaves and needles, a bout that has high mass and trapping duration < 1.5 months is selected; these tend to be in the fall. For the other functional groups, whichever bout has the highest mass for that group is selected, regardless of trapping duration. For woody material and twigs/branches, this often ends up being a spring collection, while flowers, seeds, and other materials peak at variable times of year. Pooling and subsampling for chemistry can occur many months after samples are collected since it is often necessary to wait until end of the sampling season to select the most appropriate bout for each functional group. See TOS Protocol and Procedure: Litterfall and Fine Woody Debris (AD[08]) and TOS Science Design for Plant Biomass, Productivity, and Leaf Area Index (AD[09]) for more details as well the related Data Product User Guide for Litterfall and fine woody debris production and chemistry for additional info.

3.3 Theory of Laboratory Measurements

Concentrations and stable isotope ratios of carbon and nitrogen are measured simultaneously using elemental analysis coupled to isotope ratio mass spectrometry (EA-IRMS). Percent data are reported after rounding to 2 decimal places, whereas the ratio of C:N is generally reported using non-rounded values. This can lead to subtle differences in the C:N ratio in the data compared to a user generated value. In some cases, in order to get good N data, the CO_2 peak must be trapped, requiring two analytical runs to get both C and N values. This is noted in the data, and for such cases no C:N ratio is provided.

Isotopes are measured as the abundance ratio of a heavy, rare isotope (H) to a light, more common isotope (L), normalized by those same ratios in a standard reference material.

$$\delta = [(R_{sample}/R_{standard} - 1)] \times 1000$$

where R = H/L. For all NEON stable isotopic data, $\delta^{15}N$ values are expressed on the atmospheric N_2 scale and $\delta^{13}C$ values are expressed on the Vienna Pee Dee Belemite scale.

Concentrations of 'lignin' are determined using the acid detergent lignin method. This technique builds on the more common acid detergent fiber analysis, but includes a sulfuric acid digest plus ashing in order to separate lignin from cellulose. Lignin is thus defined operationally as the acid-insoluble residue, and concentration estimates of both lignin and cellulose are provided. Note that lignin values from the acid detergent procedure may differ when compared to direct molecular measurements of lignin.

NEON partners with external laboratories to conduct these chemical and stable isotope analyses. Standard operating procedures (SOPs) from these labs can be found in the NEON Data Portal document library (http://data.neonscience.org/documents), in the External Lab Protocols > Terrestrial Biogeochemistry section.

Some labs analyze a percentage of samples in duplicate in order to monitor internal consistency and repeatability. In these cases, NEON passes along replicate analyses in case the uncertainty information is of interest. Before using the data, end users may wish to remove or average the information from analytical replicates, by taking 'mean' or 'first' of all measurements of a given analyte associated with the same sampleID.

3.4 Variables Reported

All variables reported from the laboratory (L0 data) are listed in the files NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001) (AD[04]) and NEON Raw Data Validation for Plant lignin concentrations (DP0.10031.001) (AD[05]). All variables reported in the published data (L1 data) are also provided separately in the file NEON Data Variables for Litterfall and fine woody debris production and chemistry (DP1.10033.001) (AD[06]).

Field names have been standardized with Darwin Core terms (http://rs.tdwg.org/dwc/; accessed 16 February 2014), the Global Biodiversity Information Facility vocabularies (http://rs.gbif.org/vocabulary/gbif/; accessed 16 February 2014), and the VegCore data dictionary (https://projects.nceas.ucsb.edu/nceas/projects/bien/wiki/VegCore; accessed 16 February 2014), where applicable. NEON TOS spatial data employs the World Geodetic System 1984 (WGS84) for its fundamental reference datum and Geoid12A geoid model for its vertical reference surface. Latitudes and longitudes are denoted in decimal notation to six decimal places, with longitudes indicated as negative west of the Greenwich meridian.

Some variables described in this document may be for NEON internal use only and will not appear in downloaded data.

3.5 Spatial Resolution and Extent

The finest spatial resolution at which litter chemical and stable isotope data are reported is a plot, since traps from the same plot are pooled for leaf and needle samples prior to analyses.

plotID (ID of plot within site) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain).

However, by joining litter chemistry and stable isotope data to field collection metadata for associated samples and following the instructions contained in the related Data Product User Guide for Litterfall and fine woody debris production and chemistry, more precise geolocations for the sampled traps can be calculated. Shapefiles of all NEON Terrestrial Observation System sampling locations can be found on the NEON science webpage at https://www.neonscience.org/data-samples/data/spatial-data-maps.

Note that chemistry and stable isotope data for the six functional groups other than leaves and needles, namely twigs/branches, woody material, seeds, flowers, other, and mixed, are reported at the spatial resolution of site, since all material from the highest-mass bout is pooled across plots.

3.6 Temporal Resolution and Extent

The finest resolution at which temporal data are reported is the range between **setDate** and **collectDate** for samples that contributed to the pooled chemistry material. The date used in querying files is the **collectDate**. The NEON Data Portal currently provides data in monthly files for query and download efficiency. Queries including any part of a month will return data from the entire month. For code resources to work with these files, see Data Relationships (3.9).

3.7 Associated Data Streams

massSampleMixtureID, **cnSampleID** and **ligninSampleID** are linking variables that can be used to join pooled chemistry data to mass flux data from contributing samples from individual traps. Note that the

identifiers for leaves and needles are different compared to the other functional groups since the former contain the plotID while the latter do not. **massSampleMixtureID** is a list and so before doing any joins, it will be necessary to separate out all the samples in that list, then arrange in a long-format dataframe.

Users may wish to link litter chemistry measurements to other biogeochemical pools and fluxes measured at the site during the same year (soil chemistry, foliar chemistry, etc). In some cases, measurements will be coincident at the plot level, and different measurement streams can be joined using the **plotID** variable. In other cases, data will not overlap in space and users may need to take site-level means and join using **siteID**, or use other spatially explicit approaches to align the data.

3.8 Product Instances

Once every five years, one litterfall sampling bout per functional group is analyzed for chemistry and stable isotopes per site. For each site, 10 to 30 plots are sampled, and one or two functional groups are analyzed per plot plus six additional site-level samples for the non-leaf and needle functional groups. In an average year, 5-6 sites conduct litterfall biogeochemistry sampling. Thus, we expect between 80-396 data records per year Observatory-wide that will result in litterfall chemistry and stable isotope data.

3.9 Data Relationships

Each record in the ltr_chemistrySubsampling table may appear from one to four times in the ltr_litterCarbonNitrogen table, depending on whether C and N were run separately (see Section 3.3) and if analytical replicates were conducted. Each record in the ltr_chemistrySubsampling table may appear from zero to two times in the ltr_litterLignin table, depending on whether there is enough mass for the analysis and if analytical replicates are conducted. The associated mass flux data appear in other tables that are described in the related Data Product User Guide. Duplicates and/or missing data may exist where protocol and/or data entry abberations have occurred; *users should check data carefully for anomalies before joining tables*.

ltr_litterCarbonNitrogen.csv - > One record expected per **cnSampleID** x **analyticalRepNumber** x **co2Trapped** combination

ltr_litterLignin.csv - > One record expected per **ligninSampleID** x **analyticalRepNumber** combination

bgc_CNiso_externalSummary.csv - > One record expected per **analyte** x **sampleType** x **laboratoryName** x **qaReportingStartDate** combination, used to associate sample data with relevant C and N uncertainty values.

lig_externalSummary.csv <- One record expected per **analyte** x **sampleType** x **laboratoryName** x **qaReportingStartDate** combination, used to associate sample data with relevant lignin uncertainty values.

Identifiers and barcodes will be generated for each sample. An archive sample will be created whenever there is sufficient material; this leads to creation of an **archiveSampleID** and **archiveSampleCode** in the ltr_chemistrySubsampling table. Any excess physical sample material will be discarded following analyses.

Data downloaded from the NEON Data Portal are provided in separate data files for each site and month requested. The neonUtilities package in R and the neonutilities package in Python contain functions to

merge these files across sites and months into a single file for each table. The neonUtilities R package is available from the Comprehensive R Archive Network (CRAN; https://cran.r-project.org/web/packages/neonUtilities/index.html) and can be installed using the install.packages() function in R. The neonutilities package in Python is available on the Python Package Index (PyPi; https://pypi.org/project/neonutilities/) and can be installed using pip. For instructions on using the package in either language to merge NEON data files, see the Download and Explore NEON Data tutorial on the NEON website: https://www.neonscience.org/download-explore-neon-data.

3.10 Special Considerations

3.10.1 Calculating mass-weighted chemistry fluxes

Prior to 2022, litter material from only two of the eight functional groups (leaves and needles) was analyzed for chemistry. While these materials do comprise the majority of the litter mass flux, other components can make a significant contribution in certain sites. Users will need to make assumptions in order to calculate flux-weighted C, N, and lignin values for these data - either applying the chemistry values from leaves and needles to the rest of the material, or using values from the literature.

3.10.2 2020 Chemistry Prototype

In 2020, two sites (D01 - Bartlett Experimental Forest (BART) and D05 - Treehaven (TREE)) prototyped a different method for litter sample pooling and chemical analysis. For the entire season, these sites pooled all functional groups at the level of the site (e.g., combining across plots), for all bouts in the year, and sent all functional groups for chemistry and stable isotope analyses. This was in addition to creating the plot-level samples for leaves and needles from a single fall bout according to the standard workflow. The results were used to inform the design change for litter chemistry that was adopted in 2022. The data along with a presentation compiling some of the results are available as a prototype dataset on the NEON data portal: https://data.neonscience.org/prototype-datasets/a6d42f4b-8d76-f192-9bbd-872b3e264437.

3.10.3 Sample drying issue, C-N data

Due to a miscommunication, prior to 2020-08-14 litterfall samples analyzed for carbon (C) and nitrogen (N) concentrations and stable isotopes were not re-dried prior to weighing and analysis at the external lab. All NEON litterfall samples are dried at 65C in the domain labs, but they are sometimes then stored in paper bags or coin envelopes for weeks to months before being ground, transferred to vials, and shipped. During this time they may accumulate moisture, especially in humid areas.

Subsequent testing revealed that %C data measured prior to 2020-08-14 are likely underestimated by 1.5-2.5% due to this lack of re-drying prior to analysis. As litterfall samples tend to have high %C (30% - 55%), this bias may have only minor impacts on many analyses, but is something for users to keep in mind. For the other parameters (%N, C:N, d15N, d13C), testing suggests there were no detectable differences between re-dried samples and originals. All affected records have been flagged, see section 4.4 for more details. All samples collected after 2020-08-14 are re-dried prior to external analysis.

3.10.4 Implications of lab change in 2024 for C-N data

After working with the same lab for many years to analyze carbon and nitrogen concentrations and stable isotope, results of a competitive process resulted in selection of a new lab to conduct these analyses starting in 2024. In order to assess the possible data implications of this lab switch, a set of samples was analyzed by both labs. While there were significant differences in observed chemistry values between the two labs according to paired t-tests, the differences were subtle and did not exceed the long-term analytical uncertainty except for delta 15N. For this analyte, values were higher (on the order of 0.5 per mill) when analyzed by the new lab compared to the previous lab; users should interpret the data in this context. For a more detailed look at the lab comparison results, see this document.

4 DATA QUALITY

4.1 Data Entry Constraint and Validation

Many quality control measures are implemented at the point of data entry within a mobile data entry application or web user interface (UI). For example, data formats are constrained and data values controlled through the provision of list-of-value options, which reduce the number of processing steps necessary to prepare the raw data for publication. An additional set of constraints are implemented during the process of ingest into the NEON database. The product-specific data constraint and validation requirements built into data entry and database ingest are described in the documents NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001) and NEON Raw Data Validation for Plant lignin concentrations (DP0.10031.001), provided with every download of this data product. Contained within this file is a field named 'entryValidationRulesParser', which describes syntactically the validation rules for each field that is performed upon ingest of the data into the NEON Cyberinfrastructure. Rules are based on a standardized data validation language (Nicl) internal to NEON. Please see AD[14] and AD[15] for more information about the Nicl language.

Note that field data collected prior to 2017 were processed using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow. Moreover, external laboratory data were not subject to same full suite of quality controls.

4.2 Automated Data Processing Steps

Following data entry into a mobile application or web user interface, the steps used to process the data through to publication on the NEON Data Portal are detailed in the NEON Algorithm Theoretical Basis Document: OS Generic Transitions (AD[13]).

4.3 Data Revision

All data are provisional until a numbered version is released. Annually, NEON releases a static version of all or almost all data products, annotated with digital object identifiers (DOIs). The first data Release was made in 2021. During the provisional period, QA/QC is an active process, as opposed to a discrete activity performed once, and records are updated on a rolling basis as a result of scheduled tests or feedback from data users. The Issue Log section of the data product landing page contains a history of major

known errors and revisions.

4.4 Quality Flagging

The **dataQF** field in each record is a quality flag for known issues applying to the record, added by NEON Science upon data review. The list of **dataQF** values for litterfall chemistry is given below. Note that whenever possible, bouts where gloves were worn are selected for chemistry and stable isotopes, but in some cases mass values do not support this.

tableName	fieldName	value	definition
litterCarbonNitrogen	dataQF	dryingProtocolError	Samples were not re-dried prior to external lab analysis, weight percent C values are likely under-estimated by 1.5-2.5 percent
litterCarbonNitrogen	dataQF	Gloves not worn during lab sorting	Gloves were not worn when sorting litter to functional group in the lab. Contamination is possible and may affect chemistry values (especially nitrogen)
litterCarbonNitrogen	dataQF	Gloves not worn for field collection or lab sorting	Gloves were not worn when collecting litter in the field or sorting litter to functional group in the lab. Contamination is possible and may affect chemistry values (especially nitrogen)

4.5 Analytical Facility Data Quality

For the C and N data, primary reference materials are used to calibrate each EA-IRMS run and secondary reference materials are analyzed in order to gauge run acceptability. For the lignin data, no standards are used for calibration (since the technique is based purely on differences in mass), but a set of in-house standards with known lignin concentration are included in each run to gauge run acceptability. Labs communicate run-level issues with the accuracy of secondary reference materials or standards, as well as record-level issues with samples or measurements, using a suite of quality flags. Definitions for the categorical codes used for these QF fields are included in the file NEON Categorical Codes for Litterfall and fine woody debris production and chemistry (AD[07]). In general, an entry of OK in a quality flag field means there is no issue to report. Fields have been added over time and entries may be missing in older data.

In addition, long-term analytical precision and accuracy of check-standard or secondary reference material analyses are reported per lab to allow users to interpret and analyze chemistry and stable isotope data in the context of their uncertainty ranges. The data tables bgc_CNiso_externalSummary and lig_externalSummary, which are available in the data product expanded package, contain the long-term precision and accuracy of lab analyses.

For further information about individual laboratory QA procedures, refer to the lab-specific SOPs found in the NEON Data Portal document library (http://data.neonscience.org/documents), External Lab Protocols > Terrestrial Biogeochemistry section. NEON's Calibration/Validation department has regular

procedures for auditing the quality assurance of external laboratories and their reports are also available to data users.

5 REFERENCES