NEON USER GUIDE TO SOIL CHEMICAL PROPERTIES (DISTRIBUTED PERIODIC) (NEON.DP1.10078) AND SOIL STABLE ISOTOPES (DISTRIBUTED PERIODIC) (NEON.DP1.10100)

PREPARED BY	ORGANIZATION	DATE
Samantha Weintraub	FSU	01/22/2018

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	01/22/2018	Initial Release

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, soil temperature 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 L1 data products Soil chemical properties (Distributed periodic) (NEON.DP1.10078) and Soil stable isotopes (Distributed periodic) (NEON.DP1.10100), which include organic carbon and nitrogen concentrations and stable isotopes in surface soils (0-30 cm). This document also provides details relevant to the publication of the data products via the NEON data portal, with additional detail available in the files NEON Data Variables for Soil chemical properties (Distributed periodic) (NEON.DP1.10078) (AD[05]) and NEON Data Variables for Soil stable isotopes (Distributed periodic) (NEON.DP1.10100) (AD[06]), 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 organic carbon and nitrogen concentrations and stable isotope values measured in NEON soil samples. How the Level 0 data are processed is detailed in the file, NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (NEON.DP0.10103) (AD[04]), provided in the download package for this data product. Please note that raw data products (denoted by 'DP0') may 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]	NEON.DP0.10103.001 _dataValidation.csv	NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (NEON.DP0.10103)
AD[05]	NEON.DP1.10078.001 _variables.csv	NEON Data Variables for Soil chemical properties (Distributed periodic) (NEON.DP1.10078)
AD[06]	NEON.DP1.10100.001 _variables.csv	NEON Data Variables for Soil stable isotopes (Distributed periodic) (NEON.DP1.10100)
AD[07]	NEON.DOC.000906	TOS Science Design for Terrestrial Biogeochemistry
AD[08]	NEON.DOC.014048	TOS Protocol and Procedure: Soil Biogeochemical and Microbial Sampling
AD[09]	NEON.DOC.004130	TOS Standard Operating Procedure: Wetland Soil Sampling
AD[10]	NEON.DOC.000008	NEON Acronym List
AD[11]	NEON.DOC.000243	NEON Glossary of Terms
AD[12]	OS_Generic _Transitions.pdf	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[13]		NEON's Ingest Conversion Language (NICL) specifications

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

The Soil chemical properties (Distributed periodic) (NEON.DP1.10078) and Soil stable isotopes (Distributed periodic) (NEON.DP1.10100) data products provide data about soils collected using TOS Protocol and Procedure: Soil Biogeochemical and Microbial Sampling (AD[08]), or TOS Standard Operating Procedure: Wetland Soil Sampling (AD[09]) if the site is a wetland. These procedures implement the guidelines and requirements described in the TOS Science Design for Terrestrial Biogeochemistry (AD[07]). All laboratory data are reported at the spatial resolution of a single soil sample, which is collected from a unique x,y coordinate (+/- 0.5 meters) and horizon (mineral or organinc) within a NEON plot. The temporal resolution is that of a single collection date.

Measurements of soil carbon and nitrogen concentrations help to reveal drivers of variation in belowground element storage and stoichiometry at the plot, site, and continental scales. Measurements of stable isotopes can shed light on nutrient sources and transformations and serve as biogeochemical integrators of plant-soil-microbial dynamics. Repeat measurements of concentrations and stable isotopes will provide essential data for understanding change in soil biogeochemical dynamics over time.

3.1 Spatial Sampling Design

Briefly, soils are sampled at all terrestrial NEON sites from three pre-determined, randomly assigned x,y locations per 40 x 40 meter plot. Soils are sampled to a maximum depth of 30 cm and separated by horizon type (mineral vs organic). Ten plots per site are sampled, four within the Tower airshed and six others distributed across the landscape and located in dominant vegetation types. See AD[02] for further details on the NEON spatial design and the Soil physical properties (Distributed, periodic) (NEON.DP1.10086) User Guide (http://data.neonscience.org/api/v0/documents/NEON_soils_userGuide_vA) for more information on soil sampling specifically.

3.2 Temporal Sampling Design

Soils are analyzed for chemistry and stable isotope composition every 5 years at each site, using samples collected during the period of peak greenness. Soil sampling for other parameters occurs more frequently, but does not result in chemical or stable isotopic measurements. When soils are analyzed for carbon and nitrogen concentrations and stable isotopes, an air-dried subsample is also created for archive purposes. This archive sample will be available to the community upon request.

3.3 Theory of Laboratory Measurements

Concentrations of total soil organic carbon and total nitrogen are commonly measured via combustion and elemental analysis (EA). If stable isotope data are also desired, as is the case with NEON samples, isotope ratio mass spectrometry (IRMS) can be coupled to elemental analysis, yielding simultaneous concentration and stable isotope measurements.

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, δ15N values are normalized to atmospheric N2 and δ13C values are normalized to Vienna Pee Dee Belemite.

Standard operating procedures for laboratories performing soil chemical and stable isotope analyses can be found in the NEON Data Portal document library (http://data.neonscience.org/documents), in the External Lab Protocols section.

3.4 Variables Reported

All variables reported from the laboratory (L0 data) are listed in the file, NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (NEON.DP0.10103) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the files NEON Data Variables for Soil chemical properties (Distributed periodic) (NEON.DP1.10078) (AD[05]) and NEON Data Variables for Soil stable isotopes (Distributed periodic) (NEON.DP1.10100) (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 GEOID09 for its reference gravitational ellipsoid. 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 soil chemistry and stable isotope data are reported is a unique x,y sampling location and horizon within a NEON plot.

cnSampleID (unique ID given to the individual soil sample analyzed for carbon and nitogen) → **plotID** (ID of plot within site) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain.

The basic spatial data included in the data downloaded include domain, site, and plot. However, by joining chemistry and stable isotope data tables to the field metadata present in the Soil physical properties (Distributed, periodic) (NEON.DP1.10086) data tables and following the instructions contained in the User Guide for that product, sample easting and northing for each unique soil sample location can be determined. Shapefiles of all NEON Terrestrial Observation System plot locations can be found in the Document Library: http://data.neonscience.org/documents.

3.6 Temporal Resolution and Extent

The finest resolution at which temporal data are reported is the **collectDate**. All samples associated with a sampling event have collect dates within a ~14-day window. The total number of sampling events will be one, once every 5 years during the period of historic peak greenness. The NEON Data Portal currently provides data in monthly files for query and download efficiency. Code to stack files across months is available here: https://github.com/NEONScience/NEON-utilities.

3.7 Associated Data Streams

In order to make full utility of the data contained in Soil chemical properties (Distributed periodic) (NEON.DP1.10078) and Soil stable isotopes (Distributed periodic) (NEON.DP1.10100), users will most likely wish to download the field and laboratory metadata for the parent soil samples. These can be found in the Soil physical properties (Distributed periodic) (NEON.DP1.10086) data product. **cnSampleID** is the variable name needed to join chemistry and isotope tables to NEON field and laboratory tables.

Soils subsampled for carbon and nitrogen concentrations and stable isotopes are also analyzed to produce several microbial data products. As well, these soils are analyzed with an *in-situ* incubation to determine net nitrogen transformation rates. A list of relevant data products and the variable names needed to join and link tables across them are provided in the User Guide for Soil physical properties (Distributed periodic) (http://data.neonscience.org/api/v0/documents/NEON_soils_userGuide_vA).

3.8 Product Instances

A maximum of 10 plots will be sampled per site. For each soil horizon present (maximum of 2, organic and mineral), 3 samples will be collected per plot. Thus on average, this will result in 30-60 unique soil samples per year per site sampled for carbon and nitrogen. Approximately 8-10 terrestrial NEON sites per year will be sampled for carbon and nitrogen. Thus, we expect 240-600 unique data records per year.

3.9 Data Relationships

TOS Protocol and Procedure: Soil Biogeochemical and Microbial Sampling dictates that each x,y location sampled for soil yields a unique **sampleID** per horizon per collectDate (day of year, local time) in sls_soilCoreCollection in the Soil physical properties (Distributed periodic) (NEON.DP1.10086) data product. This sample is then subsampled for chemical and stable isotopic analyses, yielding a corresponding **cnSampleID** in the sls_bgcSubsampling table in that same data product. Each of these child records may then appear from zero to four times in the sls_soilStableIsotopes and sls_soilChemistry tables. Most **cnSampleID**s will appear once, but some may appear more than once if analytical replicates were conducted or if samples were acified to remove carbonate (more in *3.10 Special Considerations*). 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*.

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

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

bgc_CNiso_externalSummary.csv - > One record expected per **laboratoryName** x **analyte** x **sampleType** x **labSpecificStartDate** combination. Can use corresponding variables in isotope and chemistry tables to associate sample data with relevant uncertainty values.

3.10 Special Considerations

Using the coupled EA-IRMS system, dryland sites that contain soil carbonates must be analyzed twice - once using an acidified subsample where inorganic carbon has been removed (and values for organic carbon are reliable), and again with a non-acidified subsample, which is needed to produce reliable nitrogen values. In such dryland sites, **cnSampleID**s will appear up to four times in any given sls_soilStableIsotopes or sls_soilChemistry data table - once for nitrogen-only measurements with **acidTreatment** = N, once for carbon-only measurments with **acidTreatment** = Y, and once or twice more if analytical replicates are conducted for either subsample type.

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 document NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (NEON.DP0.10103), provided with every download of this data product. Contained within this file is a field named 'entryValidationRulesParser', which describes the validation rules for external labs that submit spreadsheets to the NEON database. Data entry constraints are described in Nicl syntax in the validation file provided with every data download, and the Nicl language is described in NEON's Ingest Conversion Language (NICL) specifications (AD[13]).

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 also not subject to the 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[12]).

4.3 Data Revision

All data are provisional until a numbered version is released; the first release of a static version of NEON data, annotated with a globally unique identifier, is planned to take place in 2020. 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 Change Log section of the data product readme, provided with every data download, 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. See below for an explanation of **dataQF** codes specific to these two data products.

fieldName	value	definition
dataQF	legacyData	Data collected using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow

4.5 Analytical Facility Data Quality

Analytical labs that generate soil chemical and stable isotope data calibrate each run of NEON samples with primary reference materials, and include secondary reference materials analyzed as unknowns alongside NEON samples in order to gauge run acceptability. Labs communicate run-level issues with the accuracy of secondary reference materials, as well as record-level issues with samples or measurements, using the suite of quality flags described below. In general, an entry of 0 in a quality flag field means there is no issue to report.

In addition, long-term analytical precision and accuracy of secondary reference material analyses are reported for each lab to allow users to interpret and analyze soil chemistry and stable isotope data in the context of their uncertainty ranges. The data table bgc_CNiso_externalSummary, which is available in the data product expanded package, contains 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 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.

sls_soilChemistry

fieldName	value	definition
cnPercentQF	0	Indicator for both percentNitrogen and percentCarbon in range
cnPercentQF	1	Indicator for percentNitrogen out of range
cnPercentQF	2	Indicator for percentCarbon out of range
cnPercentQF	3	Indicator for both percentNitrogen and percentCarbon out of range
percentAccuracyQF	0	Observed values of run QA material fall within acceptance criteria for percent concentration
percentAccuracyQF	1	Observed values of run QA material do not meet acceptance criteria for nitrogenPercent
percentAccuracyQF	2	Observed values of run QA material do not meet acceptance criteria for carbonPercent
percentAccuracyQF	3	Observed values of run QA material do not meet acceptance criteria for nitrogenPercent or carbonPercent

sls_soilStableIsotopes

fieldName	value	definition
cnIsotopeQF	0	Indicator for both d15N and d13C in range
cnIsotopeQF	1	Indicator for d15N out of range
cnIsotopeQF	2	Indicator for d13C out of range
cnIsotopeQF	3	Indicator for both d15N and d13C out of range
isotopeAccuracyQF	0	Observed values of run QA material fall within acceptance criteria for stable isotopes
isotopeAccuracyQF	1	Observed values of run QA material do not meet acceptance criteria for d15N
isotopeAccuracyQF	2	Observed values of run QA material do not meet acceptance criteria for d13C
isotopeAccuracyQF	3	Observed values of run QA material do not meet acceptance criteria for d15N or d13C