NEON USER GUIDE TO ROOT BIOMASS AND CHEMISTRY, PERIODIC (DP1.10067.001)

PREPARED BY	ORGANIZATION
Hannah Schartel	TOS
Samuel Simkin	TOS
Courtney Meier	TOS
Samantha Weintraub-Leff	TOS

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	04/18/2018	Initial Release
B	11/04/2019	Consolidation of 0-0.5 and 0.5-1 mm size classes, and discontinuation of sorting live vs. dead
C	08/28/2020	Details the combined delivery of root biomass, chemistry, and stable isotope data
D	03/02/2022	Added info for new dataQF value, gloves not worn while sorting roots; Updated section 4.3 Data Revision with latest information regarding data release
D.1	09/25/2023	Updated section 3.3 Sampling Design Changes with info about reducing dilution subsamples from 10 to 3
E	03/12/2024	Updated section 3.6 Spatial Resolution and Extent with information regarding new subplotID naming convention; Updated Figure 2 and Figure 3 to represent new subplotID naming convention; Updated section 3.3 Sampling Design Changes with serial dilution and sample freezing options; Updated 3.4 Theory of Laboratory Measurements with details on C:N ratio calculations; Updated 3.8.1 Product Instances for Missed or Incomplete Sampling to describe samplingImpractical

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 depth of a soil root coring sample 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 product, Root biomass and chemistry, periodic (DP1.10067.001), and associated metadata from input data. 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 Root biomass and chemistry, periodic (DP1.10067.001) (AD[06]) and NEON Categorical Codes for Root biomass and chemistry, periodic (DP1.10067.001) (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 the data collected in the field pertaining to TOS Protocol and Procedure: Plant Belowground Biomass Sampling (AD[08]). The raw data that are processed in this document are detailed in the files NEON Raw Data Validation for Root Sampling Tower Plots (DP0.10067.001) (AD[04]) and NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001) (AD[05]), 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.000914	TOS Science Design for Plant Biomass, Productivity, and Leaf Area Index
AD[04]	Available with data download	Validation csv
AD[05]	Available with data download	Validation csv
AD[06]	Available with data download	Variables csv
AD[07]	Available with data download	Categorical Codes csv
AD[08]	NEON.DOC.014038	TOS Protocol and Procedure: Plant Belowground Biomass Sampling
AD[09]	NEON.DOC.000008	NEON Acronym List
AD[10]	NEON.DOC.000243	NEON Glossary of Terms
AD[11]	NEON.DOC.002652	NEON Data Products Catalog
AD[12]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[13]	Available on NEON data portal	NEON Ingest Conversion Language Function Library
AD[14]	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

The Root biomass and chemistry, periodic (DP1.10067.001) data product enables estimation of the amount of plant belowground biomass <= 10 mm diameter as well as its carbon (C) and nitrogen (N) concentration and stable isotope values. Root biomass is measured once every 5 years at each NEON site. As of 2020, biomass, C and N concentrations, and stable isotopes are reported together, with DP1.10102.001 and DP1.10099.001 published as part of DP1.10067.001. Prior to this, the three types of measurements (biomass, chemistry, stable isotopes) were reported in separate data downloads.

Root biomass is measured within the same landsurface area from which NEON Tower eddy covariance data are derived. At many sites this will also be the dominant vegetation type(s). When possible, NEON uses a 3-inch outside diameter (6.65 cm inside diameter) soil corer for belowground biomass sampling. If coring is not possible, soil monoliths are collected instead and the area recorded. With either method, samples are collected to a maximum depth of 30 cm. This sampling depth is consistent with that used for soil biogeochemistry and microbe sampling. Within each 3 m x 0.5 m "cell" selected for belowground biomass sampling, up to two 30 cm max depth cores or monoliths are collected.

For each sample, NEON sorts roots and root fragments >= 1 cm length by hand, then dries them to estimate biomass in each category. Prior to 2019, roots were sorted into two rootStatus classes (live or dead) and 4 sizeCategory bins based on root diameter: <= 0.5 mm, 0.5 - 1 mm, 1 - 2 mm, and 2 mm - 10 mm. However, in spring 2019, input from the community including the Terrestrial Plant Productivity and Biomass Technical Working Group as well as NEON staff lead to a simplification of the design. From 2019-05-06 onward, roots are *not* sort by live/dead and the smallest size categories are lumped, resulting in only 3 sizeCategory bins: <= 1 mm, 1 - 2 mm, and 2 mm - 10 mm. See section 3.3 for more detail on these changes.

When sufficient mass is available, dried, ground root samples, either from the live (pre 2019-05-06) or mixed (2019-05-06 onward) categories, are pooled by cell and size cateogry, then analyzed for C and N concentrations and stable isotopes by an external laboratory. Additional pooled sample material is committed to the NEON Biorepository and available upon request.

Root fragments < 1 cm length may comprise a significant portion of total plant belowground biomass, and for 20 randomly selected cores or monoliths per sampling event per site, NEON employs a dilution sampling technique to quantify total root fragment mass (Koteen and Baldocchi 2013). This yields root fragment biomass per core, not separated by siteCategory or rootStatus.

For more details on the sampling protocol, see the TOS Protocol and Procedure: Plant Belowground Biomass Sampling (AD[08]).

3.1 Spatial Sampling Design

At each terrestrial NEON site, roots are sampled from all base plots that fall in the tower airshed ('Tower Plots', Figure 1), with the goal of estimating plant belowground biomass (<= 10 mm diameter) within the same landsurface area where NEON Tower eddy covariance data are derived. This equates to either 20 or 30 plots depending on the stature of site vegetation. Belowground biomass sampling does not occur in Distributed base plots.

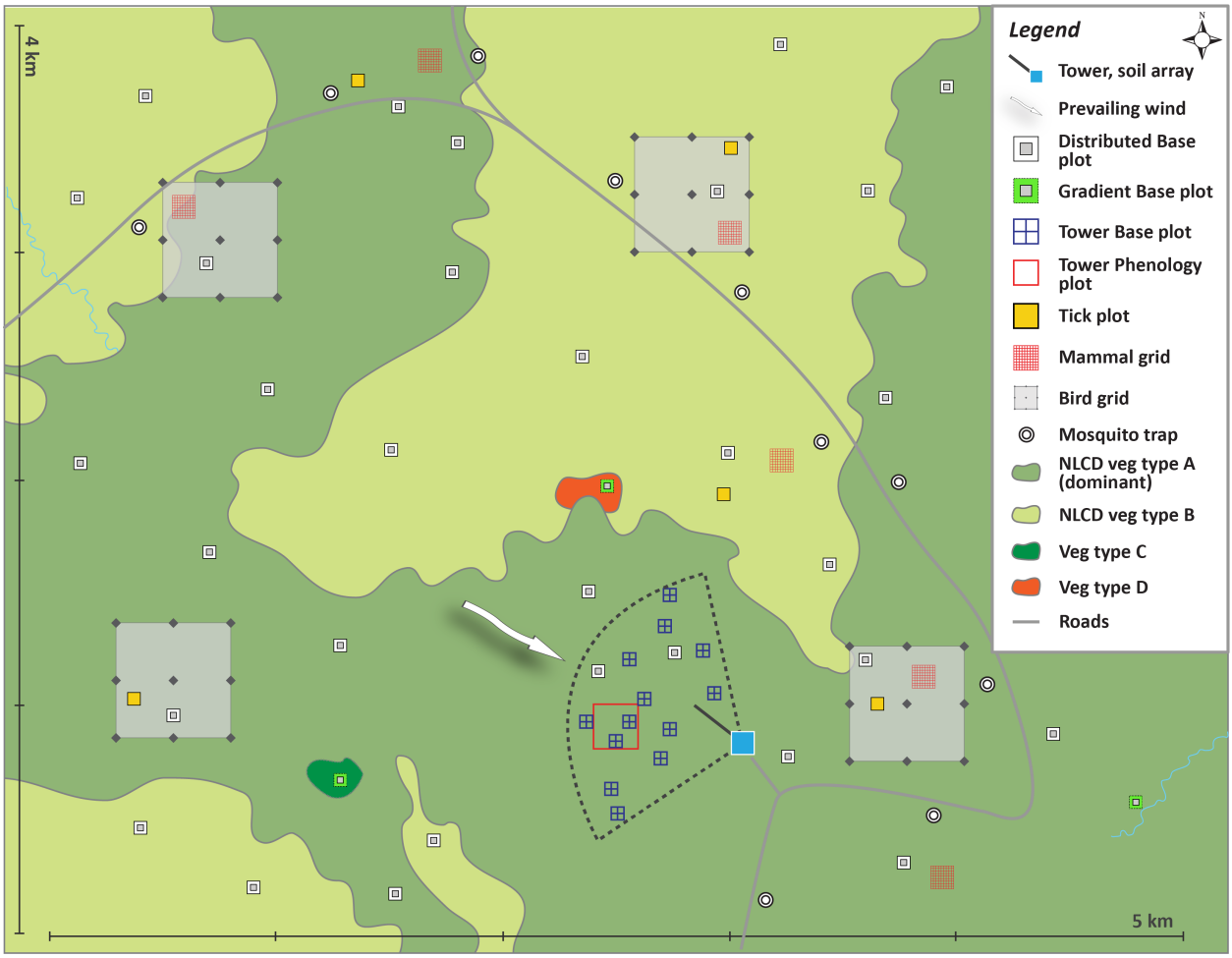

Figure 1: Representation of a NEON site with Tower and Distributed Plots; note that the actual Tower Plot number sampled (n=20 or 30) is not accurately represented in the schematic.

In each Tower plot, one or two cells are randomly chosen for root coring, depending on plot size, and in a given root sampling year, the same cells are also used for aboveground Herbaceous Biomass sampling (DP1.10023.001). Sampling cell locations are randomized in advance, and once sampled, cells are removed from consideration for future sampling. See AD[03] for further details.

In 20m x 20m Tower Plots in short-stature vegetation, up to two soil cores are sampled from one cell per bout. In 40m x 40m Tower Plots installed in tall-stature vegetation, soil core sampling occurs in two randomly selected 20m x 20m subplots (out of four), and up to two soil cores are sampled from one cell per subplot per bout. This strategy means that at sites with thirty 20m x 20m Tower Plots, there will be a maximum of n=60 soil cores collected per bout, and at sites with twenty 40m x 40m Tower Plots, there will be a maximum of n=80 soil cores collected per bout (Figure 2).

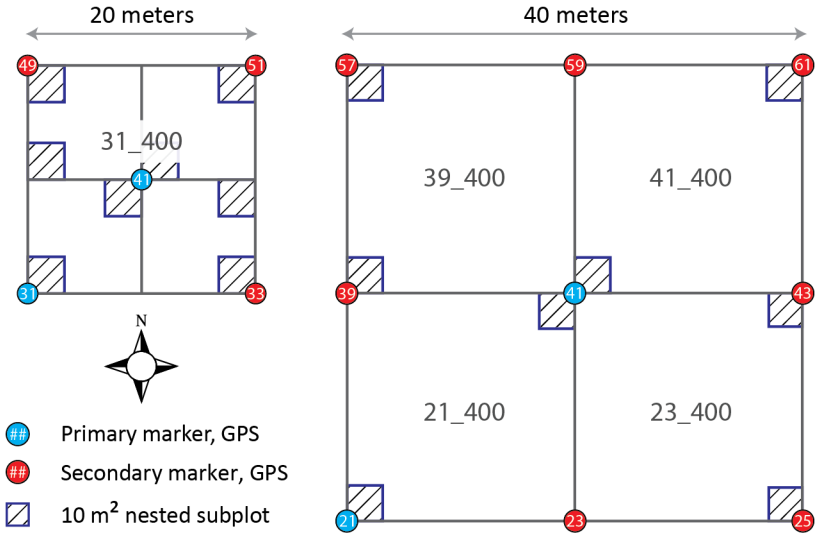

Figure 2: Illustration of two NEON plot sizes used for belowground biomass soil core sampling. Grey text indicates subplotIDs (XX_YYY format); in each 40 m x 40 m plot two 400 m2 subplots are randomly selected for plant belowground biomass sampling. Sampling cells are not provided for and cores are not collected from nested subplots <= 10 m2 (blue hashed squares). Blue and red circles represent plot markers, and the white numbers are pointIDs.

Prior to sampling, crowns, corms, rhizomes, and other perennial belowground parts that are not roots are removed from the top 3 cm of soil and discarded. In some ecosystems, these non-root belowground plant parts may constitute a significant portion of the belowground biomass; however, the NEON protocol is focused on measuring fine root biomass. A core or monolith is taken to 30 cm maximum depth from both the northern and southern end of the sampling cell (Figure 3).

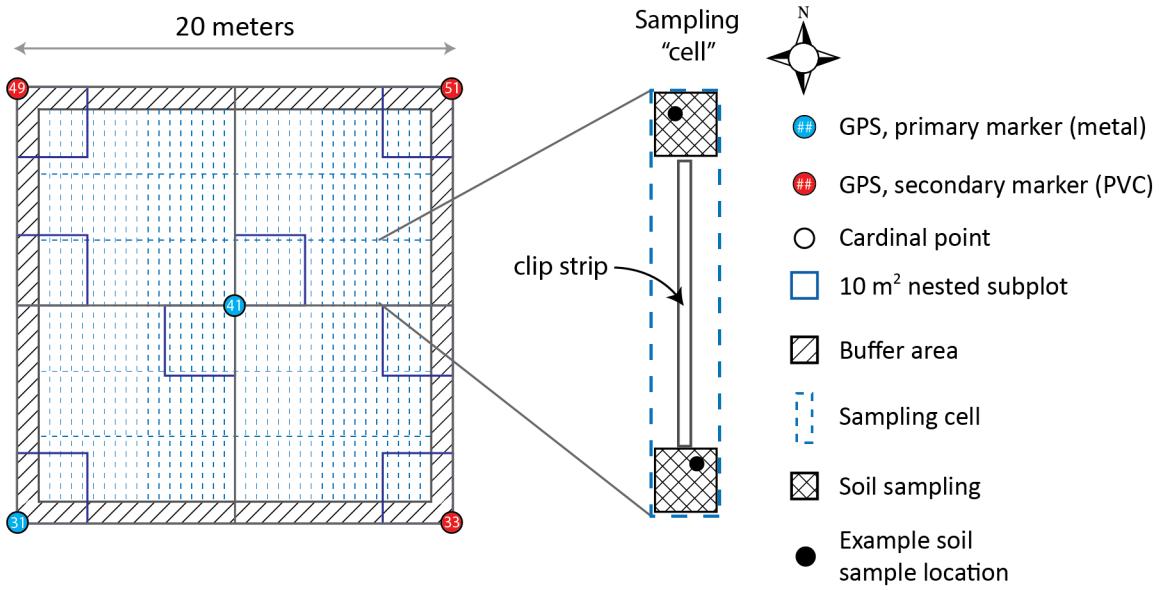

Figure 3: Example layout of NEON Tower Plot showing the locations of 3 m x 0.5 m cells used for belowground biomass soil core sampling (left); sampling cells that overlap 10 m2 nested subplots are not sampled and other nested subplots are omitted for clarity. Within a cell selected for soil core sampling, one core is collected from each of the 0.5 m x 0.5 m areas to the North and South of the clip-strip used for Herbaceous Biomass sampling (right).

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

Belowground biomass soil samples are collected once every five years per site. It is theoretically desirable to perform belowground biomass sampling when the root crop is at peak biomass. However, peak belowground biomass does not necessarily correspond with peak aboveground biomass, and in some ecosystems, does not vary in a predictable manner within a growing season from year to year (Milchunas and Lauenroth 2001). Combined with the fact that belowground biomass timecourse data are unavailable for the majority of NEON sites, the timing of belowground biomass soil core sampling is guided by these two factors, listed in order of importance:

1) Date of peak biomass herbaceous clip harvest: Perform belowground biomass soil coring either immediately before, during, or immediately after the herbaceous clip harvest associated with the greatest aboveground peak biomass.

2) Soil hardness: At some sites, peak herbaceous biomass occurs during hot, dry parts of the year when soils are extremely hard and veritably impenetrable due to high clay content (e.g. D10 CPER).

At sites where these conditions occur, the timing of soil core sampling may be moved to earlier in the growing season when soil moisture is more conducive to core sampling.

In addition to the above primary timing criteria, sampling must be timed to avoid standing water in potential soil sampling locations. If a plot is partially submerged but still accessible for terrestrial sampling, cells that contain standing water must be rejected for root sampling, and a new sampling cell location must be chosen.

3.3 Sampling Design Changes

There have been several changes to the root biomass sampling protocol over time. The rationale for these changes is outlined below.

3.3.1 Live-Dead Sorting

There is a high degree of certainty for the total root biomass in a given size class. However, the discrimination between live and dead root biomass via visual sorting has significant uncertainty if an independent method, such as staining, is not used to confirm live and dead categories. Consequently, NEON discontinued sorting root biomass into live and dead categories (recorded in field “rootStatus”) on 2019-05-06. Before 2019-05-06, roots were sorted to ‘live’ and ‘dead’ status with the inherent confidence issues described above; after 2019-05-06, all roots have rootStatus = ‘mixed’ and the mass values reported are the total fine root biomass from both live and dead roots with no sorting attempted. Mixed samples are used for chemical and stable isotopic analyses.

3.3.2 Size Category sorting

Sorting of root cores is very time and labor intenstive. In order to save resources while still providing ecologically valuable data, NEON decided to consolidate the 0-0.5 and 0.5-1 mm size classes into a single 0-1 mm size class on 2019-05-06. This results in substantial labor savings but still preserves the functionally important separation of smaller ‘absorptive’ roots from larger ‘transport’ roots (see McCormack et al. 2015). A 1 mm cutoff for sorting the finest roots yields valuable information about the biomass of roots within these two functional classes and is broadly useful across both forested and grassland ecosystems.

3.3.3 Number of dilution subsamples

We reduced the number of dilution subsamples per root core in table bbc_dilution from 10 to 3 at all terrestrial sites. Analyses showed that total fine root biomass (the sum of fragments quantified in table bbc_dilution and non-fragments quantified in table bbc_rootmass) is adequately quantified if the number of dilution subsamples is reduced from 10 to 3, and therefore this reduction from 10 to 3 subsamples was implemented to save time and reduce labor costs, effective 2023-08-02.

3.3.4 Option to freeze samples prior to processing

Beginning in 2024, there is an option to freeze cores after field collection when processing cannot be initiated within 72 hours. Core samples are placed in the freezer as soon as possible after collection in the

field (and always within 72 hours) and can be stored at either -20C or -80C. Additionally, there is an option to freeze the residual fraction within 72 hours of field collection and process the dilution sample at a later date. Sample storage is recorded in field "samplePrepMethod" in the bbc_rootmass table and field "dilutionSamplePrepMethod" in the bbc_dilution table.

3.3.5 Option for serial dilution

Beginning in 2024, there is an option to perform a serial dilution when the residual fraction is relatively massive and large volumes of distilled water (>2 L) are required to create a dilution sample such that each dilution subsample can be sorted within the timing guidelines. Effective dilution sample volume may now range up to 13,500 mL. The dilution method is recorded in field "dilutionMethod" in the bbc_dilution table.

3.4 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.

Standard operating procedures for laboratories performing root 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 > Terrestrial Biogeochemistry section. Many labs that work with NEON analyze a percentage of samples in duplicate in order to monitor internal consistency and repeatability. NEON passes along replicate analyses because the uncertainty information may be of interest. However, end users will likely wish to take 'mean' or 'first' of these replicate measurements before proceeding with data analysis.

3.5 Variables Reported

All variables reported from the field or laboratory (L0 data) are listed in the files NEON Raw Data Validation for Root Sampling Tower Plots (DP0.10067.001) (AD[04]) and NEON Raw Data Validation for Root Sampling Tower Plots (DP0.10067.001) (AD[05]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Root biomass and chemistry, periodic (DP1.10067.001) (AD[06]).

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.

3.6 Spatial Resolution and Extent

The finest spatial resolution at which Root biomass and chemistry, periodic data will be tracked is per soil sample within a cell (two per cell) within the designated subplots of Tower Plots at each NEON terrestrial site. The number of soil samples per NEON site is up to 60 (when 20m x 20m plots are sampled) or 80 (when 40m x 40m plots are sampled).

The naming convention for subplots within Base plots consists of the identity of the plot point in the southwest corner of the subplot and the scale or size of the subplot. For example, subplot '21_400' is located with point 21 in the southwest corner and is 400 m2 (20 m x 20 m, Figure 2) and subplot '41_100' is located such that point 41 is in the southwest corner and is 100 m2 (10 m x 10 m, Figure 2). Subplots within Base plots in data releases prior to the 2024 data release – Release 2024 – follow a slightly different naming convention. Previously, subplots of 100 m2 or 400 m2 were identified only by the identity of the point in the southwest corner of the subplot (e.g., '21' or '41'). The difference is that the updated subplotID contains the scale or area of the subplot in the string (e.g., '21_400' or '41_100').

Spatial hierarchy:

Tower Plots (>= 1600 m^2): sampleID (ID of individual core or monolith within a clip cell) → clipID → subplotID → plotID → siteID → domainID

Tower Plots (400 m^2): sampleID (ID of individual core or monolith within a clip cell) → clipID → plotID → siteID → domainID

The basic spatial data included in the data download include the latitude, longitude, and elevation of the *centroid* of the plot where sampling occurred, plus associated uncertainty due to GPS error and plot width. Shapefiles of all NEON Terrestrial Observation System sampling locations can be found in the Document Library: http://data.neonscience.org/documents.

There are two options to derive a more precise location:

- Use the getLocTOS() function from the geoNEON package, available here: https://github.com/NEONScience/NEON-geolocation/tree/master/geoNEON, or
- Follow these steps to perform the same calculation:
    1. Construct the named location of the southwest corner pointID of each subplot for records in cfc_fieldData by concatenating the fields for namedLocation and the first two characters of the subplotID. For example, a sample collected in subplotID '31_400' of namedLocation 'WOOD_002.basePlot.bbc' has a pointID named location of 'WOOD_002.basePlot.bbc.31'.
    2. Use the API (http://data.neonscience.org/data-api; e.g. http://data.neonscience.org/api/v0/locations/WOOD_002.basePlot.bbc.31) to query for easting ("locationUtmEasting"), northing

("locationUtmNorthing"), coordinateUncertainty ("Value for Coordinate uncertainty"), and utmZone ("locationUtmZone") for each pointID named location as inputs to the next step.

3. Use the clip cell lookup table, available here: http://data.neonscience.org/api/v0/documents/clipCellNumber_lookup (clicking on link will initiate download), to find the offsets for each clipCellNumber and pointID.
4. Use these offsets to adjust the easting and northing values downloaded in step 2, using equations (1) and (2):

$$Easting = easting.pointID + offsetEasting \quad (1)$$

and

$$Northing = northing.pointID + offsetNorthing \quad (2)$$

5. Increase coordinateUncertainty by an appropriate amount to account for error introduced by navigating within plots (suggested 0.5 m). Keep in mind that the calculated value is the center of the clip strip.

3.7 Temporal Resolution and Extent

The finest temporal resolution that Root biomass and chemistry, periodic data are tracked is a collect date. Root biomass and chemistry, periodic data are collected once every five years at each NEON core and relocatable site. The NEON Data Portal currently provides data in monthly files for query and download efficiency.

3.8 Associated Data Streams

Root biomass and chemistry, periodic data are spatially linked to the Herbaceous clip harvest (DP1.10023.001) data product, as these samples are collected from the same sampling cell in a given Tower Plot in the same year, ideally with the Root biomass and chemistry, periodic sampling occurring either shortly before or after the Herbaceous clip harvest sampling.

Additional plant belowground biomass data are available in the Root biomass and chemistry, Megapit data product (DP1.10066.001). This data product is generated one time per site during initial NEON site characterization and involves root sampling in one large pit in the NEON tower vicinity in fixed intervals to 2 m depth.

3.9 Product Instances

At each terrestrial NEON site, one or two cells are sampled from 20-30 plots, with each cell having up to 2 root samples. For each sample collected, roots are sorted into up to eight (pre 2019-05-06) or three (2019-05-06 and onward) size category by live/dead status combinations. As such, each NEON site will

generate a maximum of 480-640 (pre 2019-05-06) or 160-240 (2019-05-06 and onward) unique root mass observations. For chemistry and isotope analyses, biomass is pooled by size category across samples within a cell, and only live or mixed samples are analyzed. As such, a maximum of 120-160 (pre 2019-05-06) or 90-120 (2019-05-06 and onward) chemistry and isotope data points are expected per site per sampling year. Because root samples are collected every 5 y from a given site, approximately 8-11 terrestrial NEON sites per year will be sampled for belowground biomass.

3.9.1 Product Instances for Missed or Incomplete Sampling

Missed or incomplete sampling bouts are recorded using the **samplingImpractical** field located in the bbc_percore data table. Any value for samplingImpractical that is not 'OK' indicates that sampling did not occur. A record will be created for every expected sampling location (i.e., soil core) that was unable to be sampled. No **sampleID** will be generated. If sampling was attempted but unsuccessful due to variables impacting the sampling area, **samplingImpractial** will equal 'obstruction' to indicate that a sample was not produced. The deprecated field **rootSamplingPossible** was previously used to report what is now 'obstruction' and all prior data has been corrected to reflect this change.

3.10 Data Relationships

The Root biomass and chemistry, periodic data product is comprised of samples collected within the same cell IDs as those used to generate the Herbaceous Clip Harvest data product (DP1.10023.001). In addition, TOS Protocol and Procedure: Plant Belowground Biomass Sampling (AD[08]) dictates that each core or monolith collected is associated with a unique **sampleID**. The roots in each unique sizeCategory x rootStatus combination are then assigned a unique **subsampleID**. Live (pre-2019) or mixed/unsorted (2019 and onward) subsamples are then pooled for the two cores collected from the same cell, creating a **poolSampleID**, and pooled root samples are prepared for chemical and isotopic analyses, yielding a **cnSampleID**. Each **cnSampleID** may appear from one to four times in the bbc_carbonNitrogen table. Most will appear once, but some will appear more than that if analytical replicates are conducted and/or CO_2 trapping is required to get good N data. 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.

bbc_percore.csv → Two records (i.e., unique **sampleIDs**) expected per **clipID**.

bbc_rootmass.csv → Pre-May 2019 data: up to eight child records (one per **subsampleID**; i.e., 4 sizeCategories x 2 rootStatus classes) are expected per **sampleID**. May 2019 and onwards: up to three child records (one per **subsampleID**; i.e., 3 sizeCategories) are expected per **sampleID**.

bbc_dilution.csv → One **dilutionSampleID** expected per each of n=20 randomly selected **sampleIDs**, and 10 child (pre-August 2023) or 3 child (August 2023 and onwards) **dilutionSubsampleID** records expected per **dilutionSampleID**.

bbc_chemistryPooling.csv → Up to four (pre-May 2019) or three (May 2019 and onwards) records (**cnSampleID**) per **clipID**.

bbc_rootChemistry.csv → One record expected per **cnSampleID** x **analyticalRepNumber** x **co2Trapped** combination

bgc_CNiso_externalSummary.csv → One record expected per **analyte** x **sampleType** x **laboratoryName** x **qaReportingStartDate** combination, used to associate sample data with relevant uncertainty values. Available in the expanded data package.

If any sample material remains after subsampling for laboratory analyses, root material is retained for the Biorepository, and assigned a **bgcArchiveID**.

Data downloaded from the NEON Data Portal are provided in separate data files for each site and month requested. The neonUtilities R package contains functions to merge these files across sites and months into a single file for each table described above. The neonUtilities 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. For instructions on using neonUtilities 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.

3.11 Special Considerations

3.11.1 Sample drying issue

Due to a miscommunication, prior to 2019-10-30 root 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 root 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 2019-10-30 are likely underestimated by 1.5-2.5% due to this lack of re-drying prior to analysis. As root 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 2019-10-30 are re-dried prior to external analysis.

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 applications and database ingest are described in the documents NEON Raw Data Validation for Root Sampling Tower Plots (DP0.10067.001) and NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001), provided with every download of this data product. 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]).

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. 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 catch-all quality flag for known issues that apply to a record. Values are added by NEON Science upon data review. Below are specific **dataQF** codes relevant to this data product.

fieldName	value	definition
dataQF	legacyData	Data collected during early operations, did not include the full suite of quality control features
dataQF	dryingProtocolError	Samples were not re-dried prior to external lab analysis, weight percent C values are likely underestimated by 1.5-2.5 percent
dataQF	gloves not worn while sorting roots	Gloves not worn while sorting roots, could have a small but unknown impact on root chemistry or isotope values

Additionally, several other data quality/sample integrity fields have been added to the root biomass and chemistry tables over time in order to communicate anomalous sample conditions or method deviations. These include variables such as **biophysicalCriteria**, **sampleCondition**, **subsampleCondition**, and others. Definitions for the categorical codes used for these fields are included in the file NEON Categorical Codes for Root biomass and chemistry, periodic (DP1.10067.001) (AD[07]), provided in the download package for this data product. Fields have been added over time and entries may be missing in older data.

Records of land management activities, disturbances, and other incidents of ecological note that may have a potential impact are found in the Site Management and Event Reporting data product (DP1.10111.001).

4.5 Analytical Facility Data Quality

Analytical labs that generate root chemistry and stable isotope data include standards or secondary reference materials run as unknowns alongside NEON samples to gauge run acceptability. Labs communicate batch-level issues with the accuracy of check-standards or secondary reference materials, as well as record-level issues with samples or measurements, in the bbc_carbonNitrogen table using a suite of quality flags. Definitons for the categorical codes used for these QF fields are included in the file NEON Categorical Codes for Root biomass and chemistry, periodic (DP1.10067.001) (AD[07]). 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 root 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 > Terrestrial Biogeochemistry section.

5 REFERENCES

Koteen, L. E., and D. D. Baldocchi. 2013. A randomization method for efficiently and accurately processing fine roots, and separating them from debris, in the laboratory. Plant and Soil 363:383-398.

McCormack, M. L, I. A. Dickie, D. M. Eissenstat, T. J. Fahey, C. W. Fernandez, D. Guo, H-S. Helmisaari, E. A. Hobbie, C. M. Iversen, R. B. Jackson, J. Leppalammi-Kujansuu, R. J. Norby, R. P. Phillips, K. S. Pregitzer, S. G. Pritchard, B. Rewald, and M. Zadworny, 2015. Redefining fine roots improves understanding of below-ground contributions to terrestrial biosphere processes. New Phytologist 207: 505-518.

Milchunas, D. G., and W. K. Lauenroth. 2001. Belowground primary production by carbon isotope decay and longterm root biomass dynamics. Ecosystems 4:139-150.