NEON USER GUIDE TO ROOT BIOMASS AND CHEMISTRY, MEGAPIT (DP1.10066.001)

PREPARED BY	ORGANIZATION
Sam Simkin	TOS
Courtney Meier	TOS
Cody Flagg	TOS
Samantha Weintraub	TOS

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	07/05/2017	Initial Release
B	07/01/2020	Details the combined delivery of root biomass, chemistry, and stable isotope data

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 root samples 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, Megapit, and associated metadata from input data. 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 Root biomass and chemistry, Megapit (DP1.10066.001) (AD[06]) and NEON Categorical Codes for Root biomass and chemistry, Megapit (DP1.10066.001) (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: Soil Pit Sampling for Plant Belowground Biomass (AD[08]). The raw data that are processed in this document are detailed in the files NEON Raw Data Validation for Root sampling (Megapit) (DP0.10066.001) (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	Variables csv
AD[05]	Available with data download	Variables csv
AD[06]	Available with data download	Validation csv
AD[07]	Available with data download	Categorical Codes csv
AD[08]	NEON.DOC.001708	TOS Protocol and Procedure: Soil Pit Sampling for Plant Belowground Biomass
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
CRB	Coarse root biomass
FRB	Fine root biomass
N	Nitrogen
TOS	Terrestrial Observation System
TIS	Terrestrial Instrument System

3 DATA PRODUCT DESCRIPTION

The Root biomass and chemistry, Megapit data product includes one-time measurements of the standing stock of fine and coarse root biomass as well as root carbon (C) and nitrogen (N) concentrations and stable isotopes in TIS soil Megapits. Megapits were established in each NEON Tower airshed, in a location expected to be representative of NEON sensor-based soil plots. For more detail about TIS Megapits, refer to the Data Product User Guide for Soil physical and chemical properties, Megapit (DP1.00096.001). As of 2020, root biomass, C and N concentrations, and stable isotopes are reported together, with records from DP1.10099.001 and DP1.10102.001 published as part of this product (DP1.10066.001). Prior to this, the three types of measurements (biomass, chemistry, stable isotopes) were reported in separate data downloads, and Megapit root chemistry and isotopes were reported together with chemistry and isotopes of distributed root samples (now in DP1.10067.001 Root biomass and chemistry, periodic).

Soil samples for roots are collected incrementally within each Megapit, from the soil surface to the bottom of the pit, typically 2 m deep. Three vertical sampling profiles are collected per Megapit. Biomass is separated into coarse root biomass (CRB) and fine root biomass (FRB), and further into living and dead fractions. The cutoff between CRB and FRB is generally 2 mm, but in ~ 30% of sites sampled to date the cutoff was 4 mm (this is recorded in the data). After drying and weighing, all samples with sufficient mass are sent for C and N concentration and stable isotope analyses via an external laboratory. Prior to September 2017, nitrile gloves were not worn during laboratory processing of Megapit root samples, though forceps were generally used for the separating and sorting steps. For additional details on the sampling protocol, see the TOS Protocol and Procedure: Soil Pit Sampling for Plant Belowground Biomass (AD[08]).

Measurements of root biomass and chemistry may help to reveal drivers of variation in belowground carbon allocation, nutrient cycling, soil organic matter, and microbial dynamics at the plot, site, and continental scales.

3.1 Spatial Sampling Design

The Root biomass and chemistry, Megapit data product is available for the majority (44 of 47) terrestrial NEON sites. The soil Megapit location was chosen to be representative of the sensor-based soil plots based on soil type, vegetation and topography, as well as being accessible by a backhoe and outside the main measurement zone of other sensors. The Megapit was usually within a few hundred meters of the sensor-based soil plots and NEON tower. Within each soil Megapit three profiles were used for root sampling, with the orientation of the pit profiles as follows:

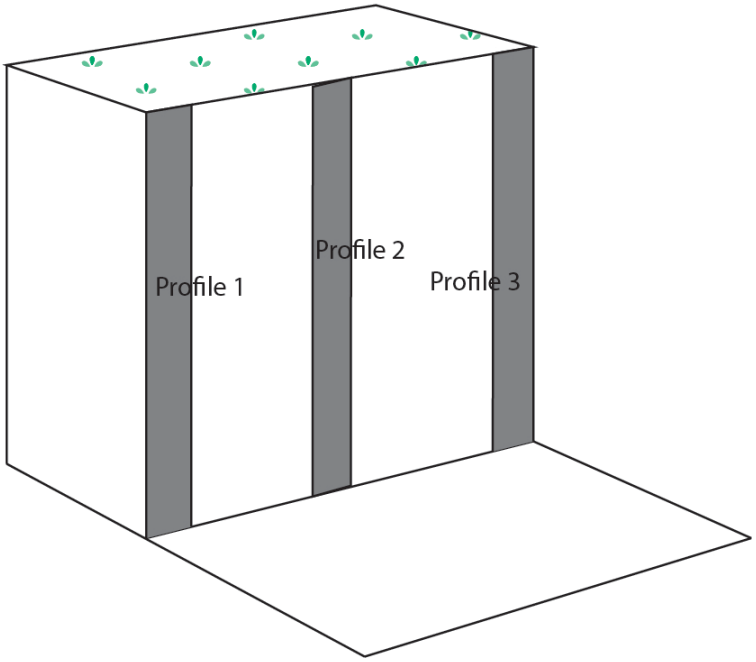

Figure 1: Representation of a NEON megapit with three pit profiles

Soil samples were collected at 10 cm depth increments to the first 100 cm below the surface, then 20 cm depth increments thereafter to 200 cm or bedrock.

3.2 Temporal Sampling Design

Root samples from soil Megapits are sampled once during construction of a NEON terrestrial site. They represent a point in time.

3.3 Variables Reported

All variables reported from the field or laboratory technician (L0 data) are listed in the files NEON Raw Data Validation for Root sampling (Megapit) (DP0.10066.001) (AD[04]) and NEON Raw Data Validation for Carbon and nitrogen concentrations and stable isotopes in plants and soil (DP0.10103.001) (AD[05]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Root biomass and chemistry, Megapit (DP1.10066.001) (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 five 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.

3.4 Theory of Laboratory Measurements

Concentrations of C and N in plant material are commonly measured via combustion and elemental analysis (EA). If stable isotope data are also desired, as is the case with NEON Megapit root samples, isotope ratio mass spectrometry (IRMS) can be coupled to elemental analysis, yielding simultaneous concentration and stable isotope data.

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. In these cases, 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 Spatial Resolution and Extent

The finest spatial resolution at which Root biomass and chemistry, Megapit are tracked is per sampling depth increment within a pit profile (three per Megapit) within the soil Megapit (one per NEON terrestrial site). Separate data tables are generated per pit profile and per sampling depth increment. There are up to up to 15 depth increments per pit, 10 within the top 100 cm and 5 within the 100-200 cm depth. As such, the spatial hierarchy is as follows:

sampleID (unique ID given to the individual root sample per sizeCategory and sampleStatus) → **depthIncrementID** (ID of the depth increment within a profile) → **pitProfileID** (ID of the profile within a pit) → **pitID** (ID of pit within a site) → **siteID** (ID of NEON site) → **domainID** (ID of NEON domain).

The basic spatial data included in the data downloaded include the latitude, longitude, and elevation of the Megapit where sampling occurred + associated uncertainty due to GPS error.

3.6 Temporal Resolution and Extent

The finest temporal resolution that Root biomass and chemistry, Megapit data are reported is the **collectDate**, the date when the sample was collected. The Root biomass and chemistry, Megapit data product is generated once for each soil pit profile and sampling depth increment at each NEON core and relocatable site. Data represent a point in time.

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 resources to work with these files, see Data Relationships (3.9).

3.7 Associated Data Streams

The Root biomass and chemistry, Megapit data product is directly linked by the variable **pitID** to the Soil physical and chemical properties, Megapit (DP1.00096.001) data product, as these samples are collected from the same soil Megapit over the same time period. However, linking by depth increment requires some discretion as sub-sampling for soil chemical and physical properties is carried out per soil horizon while roots are sampled at fixed depth intervals. Megapit root data are also linked to other data generated from samples collected from the same Megapits, such as the soil water content sensor calibration equation and data generated from analysis of Megapit Soil Archive samples by external users.

Those interested in comparing the deep but single timepoint, single location data from Root biomass and chemistry, Megapit to more spatially distributed, recurring root observations from surface soils only could download the Root biomass and chemistry, periodic (DP1.10067.001) data product and compare the two datasets within a given **siteID**.

3.8 Product Instances

There is one Megapit at each NEON terrestrial core and relocatable site. There are three profiles per pit and 15 depth increments per profile, except for when it is impossible to dig to the full prescribed depth of 200 cm. Roots are sorted into two size catagories and both live and dead fractions. Between 15-128 Megapit root records are recorded per site and nearly 3,000 records across all 44 current terrestrial NEON sites where Megapit root sampling has occurred.

3.9 Data Relationships

Following TOS Protocol and Procedure: Soil Pit Sampling for Plant Belowground Biomass (AD[08]), each **pitID** should appear three times in the mpr_perpitprofile table, one for each **pitProfileID**. Each **pitProfileID** will have up to 15 records in the mpr_perdepthincrement table, one per unique **depthIncrementID**. Each **depthIncrementID** may then appear up to 4 times in mpr_perrootsample, one for each size category and live/dead status, yielding unique **sampleIDs**. Each **sampleID** is expected to appear up to two times in mpr_carbonNitrogen in the **cnSampleID** field. Most **cnSampleIDs** will appear once, but some may appear twice if analytical replicates were conducted or not at all if there was insufficient mass for analysis. 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*.

mpr_perpitprofile.csv - > Three records expected per **pitID**, each generating a unique **pitProfileID** .

mpr_perdepthincrement.csv - > Fifteen records expected per **pitID** by **pitProfileID** combination, each generating a unique **depthIncrementID**.

mpr_perrootsample.csv - > Up to four records expected per **depthIncrementID**, since there are two different root status values (live or dead) and two different size classes (CRB, > 2 or 4 mm and FRB, < 2 or 4 mm). Each generates a unique **sampleID** used for chemistry and stable isotope analyses.

mpr_carbonNitrogen.csv - > One record expected per **cnSampleID** x **analyticalRepNumber** combination. **cnSampleID** is the same as the **sampleID** in the mpr_perrootsample table.

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.

If any sample material remains after laboratory analyses, it is retained for archive in the NEON Biorepository (**archiveID**).

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.

4 DATA QUALITY

4.1 Data Entry Constraint and Validation

A 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 document NEON Raw Data Validation for Root sampling (Megapit) (DP0.10066.001), 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[12]).

4.2 Automated Data Processing Steps

Following data entry 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[11]).

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. At present, there are no known dataQF entries for this data product.

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 mpr_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, Megapit (DP1.10066.001) (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.