NEON ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD): SOIL CHEMICAL PROPERTIES: MEGAPIT

PREPARED BY	ORGANIZATION	DATE
Edward Ayres	FIU	06/02/2015
Josh Roberti	FIU	06/02/2015
Janae Csavina	CVAL	06/02/2015

APPROVALS	ORGANIZATION	APPROVAL DATE
Andrea Thorpe	SCI	11/13/2015
Dave Tazik	SCI	11/02/2015
Hanne Buur	PSE	10/20/2015

RELEASED BY	ORGANIZATION	RELEASE DATE
Jennifer DeNicholas	CM	12/08/2015

See configuration management system for approval history.

Change Record

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	12/08/2015	ECO-03235	Initial release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

Various chemical properties will be measured on soil samples collected by soil horizon from the Terrestrial Instrument System (TIS) soil pit, also known as the Megapit. This represents a one-time sampling effort, normally during site construction, at all NEON core and relocatable sites. This document details the processes necessary to convert "raw" measurements into meaningful scientific units and their associated uncertainties are described. The soil properties data will be generated by an external laboratory.

1.1 Purpose

This document details the algorithms used for creating the NEON Level 1 data products from Level 0 data, and ancillary data as defined in this document (such as calibration data), obtained from analyses at an external laboratory. It includes a detailed discussion of measurement theory and implementation, appropriate theoretical background, data product provenance, quality assurance and control methods used, approximations and/or assumptions made, and a detailed exposition of uncertainty resulting in a cumulative reported uncertainty for this product.

1.2 Scope

The theoretical background and entire algorithmic process used to derive Level 1 TIS Soil Properties data product from Level 0 data are describe in this document. The scope of the document is limited to soil chemical properties measured on soil samples collected from the TIS Soil Pit (Megapit). This document assumes the soil samples will be analyzed in accordance with the USDA NRCS Soil Survey Laboratory Methods Manual (Soil Survey Staff 2004). This document does not provide computational implementation details, except for cases where these stem directly from algorithmic choices explained here.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Applicable Documents

AD[01]	NEON.DOC.000001 NEON Observatory Design (NOD) Requirements
AD[02]	NEON.DOC.005003 NEON Scientific Data Products Catalog
AD[03]	NEON.DOC.005004 NEON Level 1-3 Data Products Catalog
AD[04]	NEON.DOC.005005 NEON Level 0 Data Product Catalog
AD[05]	NEON.DOC.000782 ATBD QA/QC Data Consistency
AD[06]	NEON.DOC.011081 ATBD QA/QC plausibility tests
AD[07]	NEON.DOC. 002840 NEON Data Publication Workbook for Soil Chemical Properties (Megapit)
AD[08]	NEON.DOC.000746 Evaluating Uncertainty (CVAL)
AD[12]	NEON.DOC.000784 ATBD Profile Development
AD[13]	NEON.DOC.000785 TIS Level 1 Data Products Uncertainty Budget Estimation Plan
AD[14]	NEON.DOC.000751 CVAL Transfer of standard procedure
AD[15]	NEON.DOC.000747 Uncertainty Assessment Methodologies (CVAL)
AD[17]	NEON.DOC.001307 TIS Soil Pit Sampling Protocol

2.2 Reference Documents

RD[01]	NEON.DOC.000008	NEON Acronym List
RD[02}	NEON.DOC.000243	NEON Glossary of Terms
RD[03]		
RD[04]		

2.3 Acronyms

Acronym	Explanation
ATBD	Algorithm Theoretical Basis Document
CI	Cyberinfrastructure Product Team
CVAL	NEON Calibration, Validation, and Audit Laboratory
DAS	Data Acquisition System
DP	Data Product
FIU	Fundamental Instrument Unit Product Team
GRAPE	Grouped Remote Analog Peripheral Equipment
L0	Level 0
L1	Level 1
UQ	Unquantifiable uncertainty
USDA	United States Department of Agriculture
NRCS	Natural Resource Conservation Service

2.4 Verb Convention

"Shall" is used whenever a specification expresses a provision that is binding. The verbs "should" and "may" express non-mandatory provisions. "Will" is used to express a declaration of purpose on the part of the design activity.

3 DATA PRODUCT DESCRIPTION

3.1 Variables Reported

This ATBD describes the steps needed to generate the L1 data products: Soil Physical Properties: Megapit (NEON.DOM.SIT.DP1.00097). Subproducts for this data product are listed below (Table 1). Detailed lists of the associated subproducts and metadata products are provided separately, along with example data in publication-ready spreadsheets in the NEON Data Publication Workbook for Soil Chemical Properties (Megapit) (AD[07]).

Table 3-1. List of subproducts produced in this ATBD for the data product, Soil Physical Properties: Megapit (NEON.DOM.SIT.DP1.00097). The list is not exhaustive and a variety of supporting data will also be made available.

Number	**Field Name**	**Description**
NEON.DOM.SITE.DP1.00097.001.01271.003.001.001	cTot	Total carbon of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01272.003.001.001	nTot	Total nitrogen of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01273.003.001.001	sTot	Total sulfur of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01274.003.001.001	eoc	Estimated organic carbon of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01238.003.001.001	alMjelm	Total dissolution aluminum from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01239.003.001.001	caMjelm	Total dissolution calcium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01240.003.001.001	feMjelm	Total dissolution iron from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01241.003.001.001	kMjelm	Total dissolution potassium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01242.003.001.001	mgMjelm	Total dissolution magnesium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01243.003.001.001	mnMjelm	Total dissolution manganese from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01244.003.001.001	naMjelm	Total dissolution sodium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01245.003.001.001	pMjelm	Total dissolution phosphorous from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01246.003.001.001	siMjelm	Total dissolution silicon from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01247.003.001.001	srMjelm	Total dissolution strontium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01248.003.001.001	tiMjelm	Total dissolution titanium from the <2 mm fraction

NEON.DOM.SITE.DP1.00097.001.01249.003.001.001	zrMjelm	Total dissolution zirconium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01250.003.001.001	phCacl2	pH of the <2 mm fraction in CaCl2
NEON.DOM.SITE.DP1.00097.001.01251.003.001.001	phH2o	pH of the <2 mm fraction in water
NEON.DOM.SITE.DP1.00097.001.01252.003.001.001	ec12pre	1:2 Electrical conductivity of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01224.003.001.001	gypsum	Gypsum content of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01225.003.001.001	caco3	Carbonate content of the <2 mm fraction experssed as calcium carbonate
NEON.DOM.SITE.DP1.00097.001.01226.003.001.001	caNh4d	Positive charge of ammonium acetate extractable Calcium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01227.003.001.001	kNh4d	Positive charge of ammonium acetate extractable potassium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01228.003.001.001	mgNh4d	Positive charge of ammonium acetate extractable magnesium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01229.003.001.001	naNh4d	Positive charge of ammonium acetate extractable sodium from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01230.003.001.001	cecdNh4	Positive charge of ammonium acetate cation exchange capacity (CEC) of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01231.003.001.001	alSatCecd33	Aluminum saturation of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01232.003.001.001	baseSumCecd10	Positive charge of sum of Ammonium acetate extractable bases from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01233.003.001.001	bsesatCecd10	Base saturation of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01234.003.001.001	ececCecd33	Positive charge of effective cation exchange capacity (CEC) of the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01235.003.001.001	alKcl	Positive charge of KCl extractable aluminum from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01236.003.001.001	feKcl	KCl extractable iron from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01237.003.001.001	mnKcl	KCl extractable manganese from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01253.003.001.001	bSatx	Negative charge of boron in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01254.003.001.001	brSatx	Negative charge of bromine in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01255.003.001.001	caSatx	Positive charge of calcium in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01256.003.001.001	clSatx	Negative charge of chlorine in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.0009	co3Satx	Negative charge of carbonate in the saturation extract

7.001.01257.003.001.001		from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01258.003.001.001	ecSatp	Electrical conductivity of the saturated paste from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01259.003.001.001	flSatx	Negative charge of florine in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01260.003.001.001	h2oSatx	Water content of the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01261.003.001.001	hco3Sx	Negative charge of bicarbonate in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01262.003.001.001	kSatx	Positive charge of potassium in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01263.003.001.001	mgSatx	Positive charge of magnesium in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01264.003.001.001	naSatx	Positive charge of sodium in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01265.003.001.001	no2Satx	Negative charge of nitrite in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01266.003.001.001	no3Satx	Negative charge of nitrate in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01267.003.001.001	pSatx	Negative charge of phosphate in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01268.003.001.001	phSp	pH of the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01269.003.001.001	resist	Resistivity of the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01270.003.001.001	so4Satx	Negative charge of sulfate in the saturation extract from the <2 mm fraction
NEON.DOM.SITE.DP1.00097.001.01289.004.001.001	archiveID	An identifier for the archive sample

3.2 Product Instances

Soil Chemical Properties: Megapit data products will be produced for each soil horizon identified in the TIS soil pit (Megapit), unless this is not appropriate because of methodological constraints (see Theory of Measurement/Observation section). There is one Megapit at each NEON terrestrial core and relocatable site. The number of soil horizons identified at each pit will vary among the sites.

3.3 Temporal Resolution and Extent

The finest temporal resolution that Soil Chemical Properties: Megapit data will be tracked is at the level of a day. The Soil Chemical Properties: Megapit data products will be generated once for each soil horizon at each NEON core and relocatable site. They represent a point in time, although many of the soil properties typically change fairly slowly (e.g., over decades).

3.4 Spatial Resolution and Extent

The finest spatial resolution at which Soil Chemical Properties: Megapit will be tracked is per sample from a defined depth increment within a soil horizon identified within a megapit soil pit (one megapit soil pit per NEON terrestrial core and relocatable site, number of soil horizons per megapit varies among sites). The Soil Chemical Properties: Megapit will be generated for a single horizontal location at each NEON terrestrial core and relocatable site (i.e., the location of the megapit). Separate data products will be generated for each soil horizon identified in the Megapit; therefore, the data products will be generated at several depth increments with the number of depth increments depending on the number of soil horizons. The data products are assumed to be representative of the entire depth increment of the soil horizon from which the soil sample was collected. The data products represent a point in horizontal space. Overall, this results in a spatial hierarchy of:

spatial hierarchy =

biogeoID (Identifier for the biogeochemistry sample) (biogeoTopDepth, biogeoBottomDepth)→

horizonName (Soil horizon name) →

pitID (ID of megapit soil pit within site) (*decimalLatitude, decimalLongitude*) →

siteID (ID of NEON site) → *domainID* (ID of a NEON domain)

3.5 Associated Data Streams

The Soil Chemical Properties: Megapit data product is directly linked to the TIS Soil Physical Properties: Megapit (NEON.DOM.SIT.DP1.00096) data products, as well as other data generated from samples collected from the same soil pit, such as the soil water content sensor calibration equation and data generated from analysis of Megapit Soil Archive samples by archive users. The Soil Chemical Properties: Megapit data product is derived from the same parent sample as the Soil Physical Properties derived data product. The parent sample of the Soil Chemical Properties: Megapit and Soil Physical Properties: Megapit (excluding the bulk density, soil taxonomy, and horizon name subproducts) is subsampled from the same parent sample as the Soil Archive: Megapit sample. The Soil Physical Properties: Megapit bulk density subproduct is derived from a soil sample collected from the same soil pit and soil horizon as the Soil Chemical Properties: Megapit sample and the sample that produces the remaining Soil Physical Properties: Megapit data product subproducts. The soil taxonomy and horizons subproducts, and the soil water content sensor calibration equation, relate to the soil pit where the samples for Soil Chemical Properties: Megapit and Soil Physical Properties: Megapit data products were collected.

4 SCIENTIFIC CONTEXT

4.1 Theory of Measurement/Observation

The methods used to quantify the data products included in this ATBD are described in the USDA NRCS Soil Survey Laboratory Methods Manual (Soil Survey Staff 2004). Method codes from the Soil Survey Laboratory Methods Manual for each data product are listed in Table 3 in the Appendix. See the USDA NRCS Soil Survey Laboratory Methods Manual for a description of each method.

Because the primary users of the data products included in this ATBD are expected to be US-based ecologists, soil scientists, and other environmental science researchers, the USDA NRCS soil taxonomic classification system is used since it is expected to be the classification system that is most familiar to the users.

A brief description of the methods is included below (see the USDA NRCS Soil Survey Laboratory Methods Manual (Soil Survey Staff 2004) for details). All analyses for the data products included in this ATBD are performed on <2 mm fraction.

The total concentration of C, N, and S are assessed using an elemental analyzer. This analysis is performed on all samples.

The total concentration of other elements (Al, Ca, Fe, K, Mg, Mn, Na, P, Si, Sr, Ti, Zr) are assessed via hydrofluoric acid, nitric acid, and hydrochloric acid digestion followed by analysis on an inductively coupled plasma atomic emission spectrometer (ICP-AES). This analysis is performed on all samples.

Soil pH is measured in soil-deionized water (1:1) and soil-salt solution (1:2 $CaCl_2$) using an electrode. This analysis was performed on all samples.

Electrical conductivity is determined by mixing soil with deionized water (1:2 weight basis), leaving the sample to equilibrate for several hours, and measured using a conductivity meter. This analysis is performed on all samples except some samples high in plant material.

Gypsum content is measured by dissolving it on water, extracting a subsample of the liquid and adding acetone to precipitate out the gypsum. The gypsum is then dissolved in water again and the concentration is determined by measuring the electrical conductivity of the solution. This analysis is performed on samples that had an electrical conductivity of ≥ 0.5 dS m^{-1}.

Carbonate content is measured by determining the amount of CO_2 released following the addition of hydrochloric acid (HCl) to the soil. Carbonate content was expressed as calcium carbonate. This analysis is performed on samples that had a pH measured in $CaCl_2 \geq 6.95$ and/or effervescence of greater than "none" in the NRCS soil description.

Ammonium acetate extractable Ca, K, Mg, Na, the sum of extractable bases, and ammonium acetate cation exchange capacity (cecdNh4) are measured by leaching the soil with ammonium acetate buffered at a pH of 7 to saturate the cation exchange sites in the soil. The leachate is collected and analyzed for Ca, K, Mg, and Na content using an atomic absorption spectrophotometer. The sum of extractable bases represents the sum of Ca, K, Mg, and Na content in the leachate. The ammonium acetate cation exchange capacity (cecdNh4) is determined by rinsing the ammonium acetate saturated soil with ethanol to remove excess ammonium, extracting the remaining ammonium with potassium chloride solution, and determining the concentration of ammonium via steam distillation and titration. Ammonium acetate extractable Ca, K, Mg, Na and ammonium acetate cation exchange capacity are measured on all samples. However, for samples that required grinding to pass through an 80-mesh sieve in order to perform the analysis (usually organic soil samples) the sum of extractable bases is not calculated as it may not be representative of field condition (NRCS Kellogg Soil Survey Laboratory comment).

Potassium chloride (KCl) extractable Al and Mn are measured by leaching the soil with potassium chloride solution, adding hydrochloric acid to the leachate, and analyzing it on an inductively coupled plasma atomic emission spectrophotometer (ICP-AES). This analysis ia performed on samples that had a pH measured in $CaCl_2$ ≤5.05.

Effective cation exchange capacity is calculated by adding the sum of ammonium acetate extractable bases to the potassium chloride extractable Al. This calculation is performed on samples that are not ground to pass through an 80-mesh sieve.

Base saturation is determined by dividing the sum of ammonium acetate extractable bases by the ammonium acetate cation exchange capacity (cecdNh4) and multiplying by 100 to express it as a percentage. This calculation is performed on samples that are not ground to pass through an 80-mesh sieve.

Aluminum saturation is calculated by dividing potassium chloride extractable Al by the effective cation exchange capacity and multiplying by 100 to express it as a percentage. This calculation is performed on samples where both these measurements were available.

Electrical conductivity of saturated paste and water content of saturated paste are determined by adding reverse osmosis water to soil to achieve a saturated paste. A saturated paste has the following subjective characteristics: 1) glistens as it reflects light; 2) flows slightly when the container it is in is tilted; and 3) slides freely and cleanly from a spatula (except soils with high clay content). The paste is covered and left overnight and the characteristics are rechecked and additional water is added if necessary. Since the amount of water that needs to be added to achieve a saturated paste is somewhat subjective, the water content of the paste is determined by weighing a subsample followed by oven-dying and reweighing, and reported as a percentage of soil dry weight. The electrical conductivity of the saturated paste is determined using a conductivity meter. This analysis is performed on samples that had an electrical conductivity of ≥0.25 dS m^{-1}.

Ca, K, Mg, and Na in the saturation extract are determined by extracting the water from a saturated paste under a vacuum. An ionization suppressant (lanthanum oxide) is added to the extract and Ca, K, Mg, and Na concentration is measured using an atomic absorption spectrophotometer. This analysis is performed on samples that had an electrical conductivity of ≥0.25 dS m^{-1}.

Bromide, chloride, fluoride, nitrite, nitrate, phosphate, and sulfate in the saturation extract are determined by extracting the water from a saturated paste under a vacuum. The extract is diluted with reverse osmosis water and injected into an ion chromatograph to separate the anions. A conductivity detector is used to identify the anions and measure their concentration. This analysis is performed on samples that had an electrical conductivity of ≥0.25 dS m^{-1}.

Carbonate and bicarbonate in the saturation extract are determined by extracting the water from a saturated paste under a vacuum. A subsample of the extract is diluted with reverse osmosis water and the concentration of carbonate and bicarbonate are determined via titration. This analysis is performed on samples that had an electrical conductivity of ≥0.25 dS m^{-1}.

The pH of the saturated extract is determined on samples that had an electrical conductivity of ≥0.25 dS m^{-1}.

Estimated organic content is calculated by subtracting the carbon content of calcium carbonate (% gravimetric) measured in the sample from the total carbon content (% gravimetric). The carbon content of the calcium carbonate is determined by dividing weight of calcium carbonate (% gravimetric) by the molecular mass of calcium carbonate (100.087), and multiplying the result by the molecular mass of carbon (12.011). This calculation is performed on all samples.

4.2 Theory of Algorithm

4.2.1 Summary of algorithm for the mgc_permegapit data

1. Generate uid

4.2.2 Summary of algorithm for the mgc_perhorizon data

1. Generate uid

4.2.3 Summary of algorithm for the mgc_perbiogeosample data

1. Generate uid
2. Calculate **cTot, nTot**, **sTot**, **eoc**, **gypsum**, and **caco3** using algorithm described below

3. Calculate **alMjelm, caMjelm, feMjelm, kMjelm, mgMjelm, mnMjelm, naMjelm, pMjelm, siMjelm, srMjelm, tiMjelm, zrMjelm, phCacl2, phH2o, ec12pre, caNh4d, kNh4d, mgNh4d, naNh4d, cecdNh4, alSatCecd33, baseSumCecd10, bsesatCecd10, ececCecd33, alKcl, feKcl, mnKcl, bSatx, brSatx, caSatx, clSatx, co3Satx, ecSatp, flSatx, h2oSatx, hco3Sx, kSatx, mgSatx, naSatx, no2Satx, no3Satx, pSatx, phSp, resist,** and **so4Satx** using algorithm described below

4.2.4 Summary of algorithm for mgc_perarchivesample data

1. Generate uid

5 ALGORITHM IMPLEMENTATION

Unless otherwise specified below, all variables that appear in tables mgc_permegapit_pub, mgc_perhorizon_pub, mgc_perbiogeosample_pub, and mgc_perarchivesample_pub can be passed directly from the L0 or CI datastore to the L1 variable with the same name.

5.1 Run the following processing steps for NEON Raw Data Ingest Workbook for Soil Chemical Properties (Megapit) (mgc_permegapit)

None

5.2 Run the following processing steps for NEON Raw Data Ingest Workbook for Soil Chemical Properties (Megapit) (mgc_perhorizon)

None

5.3 Run the following processing steps for NEON Raw Data Ingest Workbook for Soil Chemical Properties (Megapit) (mgc_perbiogeosample)

1. Calculate **cTot**
    - Do not calculate **cTot** if **bulkDensSampleType** = Audit or **biogeoSampleType** = Audit for the same **pitID** and **horizonName**.
    - If **biogeoHorizonProportion** = 1 for the same **pitID** and **horizonName**, calculate **cTot** as follows:
        - Check whether **ctot** is less than zero. Iif so, change **ctot** to zero, if not, leave **ctot** unchanged. Then:

$$Z_{L1} = 10Z_{L0} \tag{1}$$

where

Z_{L0} is Level 0 **cTot**
Z_{L1} is Level 1 **cTot**

- If **biogeoHorizonProportion** < 1 for the same **pitID** and **horizonName**, calculate **cTot** as follows:
    - Check whether **ctot** is less than zero. If so, change **ctot** to zero, if not, leave **ctot** unchanged. Then:

$$Z_{L1} = 10Z_{L0_i}P_i + 10Z_{L0_j}P_j \tag{2}$$

where

Z_{L0} is Level 0 **cTot**
P is **biogeoHorizonProportion**
Z_{L1} is Level 1 **cTot**
i corresponds to the first soil sample within the same **pitID** and **horizonName**
j corresponds to the second soil sample within the same **pitID** and **horizonName**.

2. Calculate **nTot**, **sTot**, **eoc**, **gypsum**, and **caco3**
    - Use the **cTot** algorithm described above, replacing **cTot** with the relevant parameter, to calculate each of the following Level 1 data products: **nTot**, **sTot**, **eoc**, **gypsum**, and **caco3**
3. Calculate **alMjelm**
    - Do not calculate **alMjelm** if **bulkDensSampleType** = Audit or **biogeoSampleType** = Audit for the same **pitID** and **horizonName**.
    - If **biogeoHorizonProportion** = 1 for the same **pitID** and **horizonName**, calculate **alMjelm** as follows:
        - Check whether **alMjelm** is less than zero. If so, change **alMjelm** to zero, if not, leave **alMjelm** unchanged. Then:

$$Z_{L1} \equiv Z_{L0} \tag{3}$$

where

Z_{L0} is Level 0 **alMjelm**
Z_{L1} is Level 1 **alMjelm**

- If **biogeoHorizonProportion** < 1 for the same **pitID** and **horizonName**, calculate **alMjelm** as follows:
    - Check whether **alMjelm** is less than zero. If so, change **alMjelm** to zero, if not, leave **alMjelm** unchanged. Then:

$$Z_{L1} = Z_{L0_i}P_i + Z_{L0_j}P_j \tag{4}$$

where

Z_{L0} is Level 0 **alMjelm**
P is **biogeoHorizonProportion**
Z_{L1} is Level 1 **alMjelm**

i corresponds to the first soil sample within the same **pitID** and **horizonName**

j corresponds to the second soil sample within the same **pitID** and **horizonName**.

4. Calculate **caMjelm, feMjelm, kMjelm, mgMjelm, mnMjelm, naMjelm, pMjelm, siMjelm, srMjelm, tiMjelm, zrMjelm, phCacl2, phH2o, ec12pre, caNh4d, kNh4d, mgNh4d, naNh4d, cecdNh4, alSatCecd33, baseSumCecd10, bsesatCecd10, ececCecd33, alKcl, feKcl, mnKcl, bSatx, brSatx, caSatx, clSatx, co3Satx, ecSatp, flSatx, h2oSatx, hco3Sx, kSatx, mgSatx, naSatx, no2Satx, no3Satx, pSatx, phSp, resist,** and **so4Satx**
    - Use the **alMjelm** algorithm described above, replacing **alMjelm** with the relevant parameter, to calculate each of the following Level 1 data products: **caMjelm, feMjelm, kMjelm, mgMjelm, mnMjelm, naMjelm, pMjelm, siMjelm, srMjelm, tiMjelm, zrMjelm, phCacl2, phH2o, ec12pre, caNh4d, kNh4d, mgNh4d, naNh4d, cecdNh4, alSatCecd33, baseSumCecd10, bsesatCecd10, ececCecd33, alKcl, feKcl, mnKcl, bSatx, brSatx, caSatx, clSatx, co3Satx, ecSatp, flSatx, h2oSatx, hco3Sx, kSatx, mgSatx, naSatx, no2Satx, no3Satx, pSatx, phSp, resist,** and **so4Satx**

5.4 Run the following processing steps for NEON Raw Data Ingest Workbook for Soil Physical Properties (Megapit) (mgc_perarchivesample)

None

5.5 Special Considerations

None

6 UNCERTAINTY

The uncertainty section of this ATBD is not being implemented during the initial release of data derived using this ATBD.

Uncertainty of *measurement* is inevitable (JCGM 2008, 2012; Taylor 1997). It is crucial that measurement uncertainties are identified and quantified to determine statistical interpretations about mean quantity and variance structure; both are important when constructing higher level data products (e.g., L1 DP) and modeled processes. This portion of the document serves to identify, evaluate, and quantify sources of uncertainty relating to L1 soil property DPs. It is a reflection of the information described in AD[13], and is explicitly described for the various soil properties detailed in the following sections.

6.1 Uncertainty of Soil Sampling and laboratory analyses

Because of the subjective nature of humans and thus, the identification, characterization, and sampling of soil horizons, many sources of uncertainties can be associated with soil properties. These uncertainties introduced by humans can be easily identified but properly quantifying them can be problematic. The aim of the following sections is to identify all known uncertainties inherent in the previously mentioned processes, and if possible, quantify each.

6.1.1 Reported Uncertainty

It should be noted that reported measurement uncertainties outlined in this document comprise *only quantifiable measurement uncertainties*. Quantifiable uncertainties will be computed annually by FIU via the equations in the Appendix (Section 9.2). Once computed, these values will pass from FIU to CI and accompany the measurements discussed in Sections 6.1.4 through 6.1.6. It is urged that the end-user(s) acknowledge that many uncertainties, such as those mentioned in Sections 6.1.2 through 6.1.3 are unquantifiable at this time, and reported uncertainties are most likely underestimated as a result.

6.1.2 Horizon Identification and characterization

Soil horizons at all NEON soil pits are identified and characterized by USDA soil scientists (for more information, please refer to Schoeneberger *et al.* (2012A)). As a result of the subjective nature of this process, identifiable, but currently unquantifiable uncertainties are introduced. For instance, soil horizons and associated characteristics (e.g., depth, thickness, etc.) may be quantified differently among a group of soil scientists; an identifiable uncertainty is introduced. One way to mitigate these issues and *quantify* this uncertainty would be to gather a group ($n \geq 30$) of soil scientists, and have each characterize the horizons of the same soil pit independently of one another. Uncertainty (in the form of *reproducibility*) could then be obtained through this approach. However, such an approach is time and labor intensive and is therefore not completed.

Another identifiable source of uncertainty is spatial representativeness. Given that *only one* soil pit is dug at each NEON site, the spatial variance of horizon depth and thickness cannot be quantified at multiple points throughout the site.

Both examples presented in this section represent identification and characterization uncertainties (of soil horizons) that can be identified but cannot be quantified at this time.

Some aspects of this uncertainty may be gleaned from the soil horizon description, which includes a section describing the boundary between that horizon and the horizon below it (Soil Survey Division Staff 1993). The boundary description consists of two components: distinctness and topography. Distinctness is the depth over which the boundary occurs and is separated into four categories:

- Abrupt: Less than 2 cm thick
- Clear: 2 to 5 cm thick
- Gradual: 5 to 15 cm thick

- Diffuse: More than 15 cm thick

Topography is defined as "the irregularities of the surface that divides the horizons" (Soil Survey Division Staff 1993), and is also separated into four categories:

- Smooth: The boundary is a plane with few or no irregularities.
- Wavy: The boundary has undulations in which depressions are wider then they are deep.
- Irregular: The boundary has pockets that are deeper than they are wide.
- Broken: One or both of the horizons or layers separated by the boundary are discontinuous and the boundary is interrupted.

6.1.3 Collection

Upon completion of the horizon identification and characterization process, each horizon is sampled (collected) independently following the regulations set forth by the National Soil Survey Center (NSSC). Specifically, a sample is collected across a horizon's full depth and breadth, a method known as *Horizon Sampling* (Schoeneberger *et al.* 2012B). Although collecting a sample in such a manner fosters quantification of variability *within* each horizon, uncertainties from the identification and characterization processes propagate into the collection process and can cause horizons to be improperly sampled. For example, if horizon depths and thicknesses are incorrectly quantified (see Section 6.1.1) by the soil scientist, the NEON scientist or technician gathering samples may unknowingly combine soils from neighboring horizons into a sample that represents a *single* horizon. Resulting laboratory analyses (e.g., bulk density and biogeochemistry) may therefore be unrepresentative of the specified horizon of interest. As with the identification and characterization process, this type of uncertainty can *only be identified* at current date.

6.1.4 Biogeochemical Analyses

This uncertainty budget is the most complete of the three (i.e., biogeochemical, bulk density, and particle size distribution) analyses. Known, traceable, reference standards (*truth*) are used to quantify the *trueness* of each chemical and a *reproducibility* analysis is conducted to quantify each chemical's precision.

6.2 Uncertainty Budget

The uncertainty budget is a visual aid detailing i) quantifiable sources of uncertainty, ii) means by which they are derived, and iii) the order of their propagation. Individual uncertainty values denoted in this budget are either provided here (within this document) or will be provided by other NEON teams (e.g., FIU) and stored in the CI data store.

Table 6-1. Uncertainty budget for L1 Soil Property DPs. Shaded rows denote the order of uncertainty propagation (from lightest to darkest).

Source of uncertainty	**Standard uncertainty component $\boldsymbol{u(x_i)}$**	**Value of standard uncertainty**	$\boldsymbol{c_i \equiv \frac{\partial f}{\partial x_i}}$	$\boldsymbol{u_i(Y) \equiv \lvert c_i \rvert u(}$
Biogeochemistry	$u_c(BioGC)$	Eq. (5)	--	--
Reference material	u_{truth}	Eq. (12) or (13)	1	Eq. (12) or (13)
Trueness	u_{true}	Eq. (11)	1	Eq. (11)
Reproducibility	u_{reprod}	Eq. (9) or (10)	1	Eq. (9) or (10)

7 FUTURE MODIFICATIONS AND PLANS

The uncertainty budget will be implemented in the future. The approach for quantifying the uncertainty for the data products in this ATBD may be revised in the future based on the amount and type of data received from the analytical laboratory. Consistency tests and flags may be added to the QA/QC procedures for some or all of the data products in this ATBD. Additional QA/QC procedures may be added.

An Audit Test may be added. This will consist of a routine audit with reference material and/or analysis of lab protocols performed on the external lab providing data by NEON's CVAL team to determine whether data passes or fails this test. If the lab is found to be out of compliance with stated uncertainty values or NEON protocols, incoming data will be flagged until lab proves restored compliance.

8 BIBLIOGRAPHY

Joint Committee for Guides in Metrology (JCGM) (2008) Evaluation of measurement data – Guide to the expression of uncertainty in measurement. pp. 120.

JCGM (2012) International vocabulary of metrology – Basic and general concepts and associated terms (VIM). 3rd Edition. pp. 92

Schoeneberger, P. J., Wysocki D. A., and Chiaretti, J. V., 2012A. Soil Taxonomy. 4.1 – 4.13 *in* National Resources Conservation Service's National Soil Survey Center, editor. Field book for describing and sampling soils. **3**, United States Department of Agriculture, Lincoln, Nebraska, USA

Schoeneberger, P. J., Wysocki D. A., and Chiaretti, J. V., 2012B. Soil Sampling. 8.1 – 8.7 *in* National Resources Conservation Service's National Soil Survey Center, editor. Field book for describing and sampling soils. **3**, United States Department of Agriculture, Lincoln, Nebraska, USA

Soil Survey Staff. 2004. Soil Survey Laboratory Methods Manual - Soil Survey Investigations Report No. 42, Version 4.0. USDA Natural Resource Conservation Service.

Taylor, J. R. (1997) An Introduction to Error Analysis: The Study of Uncertainties in Physical

Measurements. University Science Books, Mill Valley, California. 2nd Ed. pp. 327.

9 APPENDIX

9.1 USDA NRCS method codes

Table 9-1. Method codes for each data product.

Data product	Method code*
cTot	4H2a1
nTot	4H2a2
sTot	4H2a3
alMjelm	4H1b1a1a1
caMjelm	4H1b1a1a2
feMjelm	4H1b1a1a3
kMjelm	4H1b1a1a4
mgMjelm	4H1b1a1a5
mnMjelm	4H1b1a1a6
naMjelm	4H1b1a1a7
pMjelm	4H1b1a1a8
siMjelm	4H1b1a1a9
srMjelm	4H1b1a1a12
tiMjelm	4H1b1a1a11
zrMjelm	4H1b1a1a10
phCacl2	4C1a2a2b1
phH2o	4C1a2a1b1
ec12pre	4F1a1a
gypsum	4E2a1a1
caco3	4E1a1a1
caNh4d	4B1a1b1
kNh4d	4B1a1b3
mgNh4d	4B1a1b2
naNh4d	4B1a1b4
cecdNh4	4B1a1a1a1
alSatCecd33	4B4d1
baseSumCecd10	4B4a1
bsesatCecd10	4B4c1
ececCecd33	4B4b2
alKcl	4B3a1a1
feKcl	Not analyzed
mnKcl	4B3a1a2
bSatx	Not analyzed
brSatx	4F2c1b1a1
caSatx	4F2c1a1
clSatx	4F2c1b1a3

co3Satx	4F2c1c1a1
ecSatp	4F2b1
flSatx	4F2c1b1a4
h2oSatx	4F2a1
hco3Sx	4F2c1c1a2
kSatx	4F2c1a3
mgSatx	4F2c1a2
naSatx	4F2c1a4
no2Satx	4F2c1b1a6
no3Satx	4F2c1b1a5
pSatx	4F2c1b1a7
phSp	4C1a1a2
resist	Not analyzed
so4Satx	4F2c1b1a8
eoc	Not applicable

[*] See USDA NRCS Soil Survey Laboratory Methods Manual (Soil Survey Staff 2004) for method descriptions.

9.2 Measurement Uncertainty Equations

9.2.1 Measurement Uncertainty

The following subsections present the uncertainties associated with *individual pressure observations*. It is important to note that the uncertainties presented in the following subsections are *measurement uncertainties*, that is, they reflect the uncertainty of an *individual* measurement. These uncertainties should not be confused with those presented in Section 6.1.2. We urge the reader to refer to AD[11] for further details concerning the discrepancies between quantification of measurement uncertainties and L1 uncertainties.

NEON calculates measurement uncertainties according to recommendations of the Joint Committee for Guides in Metrology (JCGM) 2008. In essence, if a measurand *y* is a function of *n* input quantities $x_i \ (i = 1, \dots, n),\ \ i.e., y = f(x_1, x_2, \dots, x_n)$, the combined measurement uncertainty of *y*, assuming the inputs are independent, can be calculated as follows:

$$u_c(y) = \left(\sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) \right)^{\frac{1}{2}} \quad (5)$$

where

$\frac{\partial f}{\partial x_i}$ = partial derivative of *y* with respect to x_i

$u(x_i)$ = combined standard uncertainty of x_i.

Thus, the uncertainty of the measurand can be found be summing the input uncertainties in quadrature. For temperature measurements, the sources of uncertainty are depicted in Figure 1. The calculation of these input uncertainties is discussed below.

Repeatability:

$$\bar{X}_i = \frac{1}{n}\sum_{k=1}^{n} X_{i,k} \tag{6}$$

AND:

$$u_{repeat}(measurand\ units) = \left[\frac{1}{(n-1)}\sum_{k=1}^{n}(X_{i,k} - \bar{X}_i)^2\right]^{\frac{1}{2}} \tag{7}$$

OR:

$$u_{repeat}\ (relative) = \frac{u_{repeat}(x_i)}{|\bar{X}_i|} \tag{8}$$

Where the input quantity X_i is estimated from n independent observations $X_{i,k}$ obtained under identical measurement conditions.

Reproducibility:

$$u_{reprod}(measurand\ units) = \left[\frac{1}{(n-1)}\sum_{k=1}^{n}(X_{1,k} - X_{2,k})^2\right]^{\frac{1}{2}} \tag{9}$$

OR:

$$u_{reprod}(relative) = \left[\frac{1}{(n-1)}\sum_{k=1}^{n}\frac{(X_{1,k} - X_{2,k})^2}{\left(\frac{X_{1,k} + X_{2,k}}{2}\right)^2}\right]^{\frac{1}{2}} \tag{10}$$

Where $X_{1,}$ *and* X_2 are replicate measurements with from *n* independent observations for X_k having varying conditions of measurement. Laboratory data will be provided by FIU and stored in the CI data store to calculate both repeatability and reproducibility.

Trueness:

$$u_{true}(measurand\ units) = \left[\frac{1}{(n-1)}\sum_{i=1}^{n}(X_i - X_{i,truth})^2\right]^{\frac{1}{2}} \tag{11}$$

Where X_i is estimated from the n independent observations of the *truth (e.g., reference standard)* $X_{i,truth}$ obtained under identical measurement conditions. A reference standard is a sensor or material used for the calibration of other standards / materials (JCGM 2012).

Truth (uncertainty of):

This metric is taken directly from calibration certificates or manufacturer specifications. It represents the uncertainty of the reference material or standard and may be presented as a combined or expanded uncertainty. If given as a combined uncertainty it can be plugged directly into the final combined uncertainty equation. If this uncertainty is provided at an expanded confidence level, it must be divided by the appropriate expansion factor to convert it to a combined uncertainty prior to propagation:

$$u_{truth}(measurand\ units) = \frac{u_{certified\ value}\ (\%\ CL)}{k_p} \tag{12}$$

OR:

$$u_{truth}(relative) = \frac{u_{truth}}{X_{truth}} \tag{13}$$

Where, k_p is the expansion factor at the designated confidence level and X_{truth} is the value of the traceable reference material.

10 CHANGELOG