NEON USER GUIDE TO SOIL INORGANIC NITROGEN POOLS AND TRANSFORMATIONS (NEON.DP1.10080)

PREPARED BY	ORGANIZATION	DATE
Samantha Weintraub	FSU	12/04/2017

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	12/04/2017	Initial Release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

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

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

1.2 Scope

This document describes the steps needed to generate the L1 data product Soil inorganic nitrogen pools and transformations (NEON.DP1.10080), which encompasses the procedures involved in measuring soil extractable inorganic nitrogen (N) concentrations as well as net N mineralization and nitrification rates in incubated samples. This document also provides details relevant to the publication of the data products via the NEON data portal, with additional detail available in the file NEON Data Variables for Soil inorganic nitrogen pools and transformations (NEON.DP1.10080) (AD[05]), 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 Biogeochemical and Microbial Sampling (AD[07]), or TOS Standard Operating Procedure: Wetland Soil Sampling (AD[08]) if the site is a wetland. The raw data that are processed in this document are detailed in the file NEON Raw Data Validation for Soil inorganic nitrogen pools and transformations, Level 0 (NEON.DP0.10080) (AD[04]), provided in the download package for this data product. Please note that raw data products (denoted by 'DP0') may not always have the same numbers (e.g., '10033') as the corresponding L1 data product.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Associated Documents

AD[01]	NEON.DOC.000001	NEON Observatory Design (NOD) Requirements
AD[02]	NEON.DOC.000913	TOS Science Design for Spatial Sampling
AD[03]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products Catalog
AD[04]	NEON.DP0.10080.001 _dataValidation.csv	NEON Raw Data Validation for Soil inorganic nitrogen pools and transformations, Level 0 (NEON.DP0.10080)
AD[05]	NEON.DP1.10080.001 _variables.csv	NEON Data Variables for Soil inorganic nitrogen pools and transformations (NEON.DP1.10080)
AD[06]	NEON.DOC.000906	TOS Science Design for Terrestrial Biogeochemistry
AD[07]	NEON.DOC.014048	TOS Protocol and Procedure: Soil Biogeochemical and Microbial Sampling
AD[08]	NEON.DOC.004130	TOS Standard Operating Procedure: Wetland Soil Sampling
AD[09]	NEON.DOC.000008	NEON Acronym List
AD[10]	NEON.DOC.000243	NEON Glossary of Terms
AD[11]	OS_Generic_Trans itions.pdf	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[12]		NEON's Ingest Conversion Language (NICL) specifications

2.2 Acronyms

Acronym	Definition
KCl	potassium chloride
N	Nitrogen
NH_4^+	ammonium
NO_3^-	nitrate
NO_2^-	nitrite

3 DATA PRODUCT DESCRIPTION

The Soil inorganic nitrogen pools and transformations (NEON.DP1.10080) data product provides extractable inorganic N data for soil cores collected using TOS Protocol and Procedure: Soil Biogeochemical and Microbial Sampling (AD[07]), or TOS Standard Operating Procedure: Wetland Soil Sampling (AD[08]) if the site is a wetland. Inorganic N measurements implement the guidelines and requirements described in the TOS Science Design for Terrestrial Biogeochemistry (AD[06]). All inorganic N data are reported at the spatial resolution of a single soil sample, which is collected from a unique x,y coordinate (+/- 0.5 meters) and horizon (mineral or organinc) within a NEON plot. The temporal resolution is that of a single collection date.

Net N transformation rates are estimated using in-situ incubations of either covered cores (the majority of plots) or buried bags (wetlands). In either case, 'initial' soil cores are collected and extracted with 2M potassium chloride (KCl) along with a set of procedural blanks. At the same time, a covered core (2" diameter) or buried bag (4 mm thick) is deployed at each x,y location (+/- 0.25 m). After a two-to-four week incuabtion period, depending on site climate and time of year, 'final' samples are recovered, then extracted with 2M KCl along with their own procedural blanks. Blank-corrected differences in final versus initial inorganic N concentrations can be used to calculate net N mineralization, while blank-corrected difference in final versus intial NO_3^- is net nitrification.

Measurements of inorganic nitrogen pools and transformation rates help to reveal drivers of variation in nutrient availability and limitation, soil microbial activity, carbon-nitrogen interactions, and potential for nitrogen losses at the plot, site, and continental scales. They also provide essential data for understanding changes in soil and microbial N dynamics over time.

3.1 Spatial Sampling Design

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

3.2 Temporal Sampling Design

Analyses of inorganic N pools and transformation rates occur on an every-5-years interval at each terrestrial NEON site. During 'on' years, sampling occurs at least once at all sampling sites (8-10 Observatory-wide) during the local period of historic peak greenness. At *nearly* all sites, N transformation measurements occur two additional times, usually bracketing the peak greeness window and aimed at capturing seasonal transitions in plant and microbial activity. Sampling frequency is reduced in arctic and boreal sites, which are only sampled during peak greenness.

Within 24 hours of collection, initial and final soil samples are extracted with 2M KCl (along with a set of procedural blanks), filtered, and frozen. Once the initial/final set have both been processed, samples and blanks are shipped frozen to an external laboratory for analysis.

3.3 Theory of Measurements

Ammonium is measured using the salicylate-nitroferricyanide method, and nitrate + nitrite is measured using cadmium reduction and sulfanilamide. For the latter, NEON samples are not run with and without the reduction step, thus all measurements are technically nitrate + nitrite, though nitrite is often a minor to absent constitutent of natural soils. Standard operating procedures for the external laboratory performing inorganic N analyses can be found on the NEON Data Portal (http://data.neonscience.org/home), in the Resources > Data Documentation > External Lab Protocols section.

Both ammonium and nitrate + nitrite concentrations are reported in units of milligrams nitrogen per millileter KCl. In order to blank-correct the data, then calculate micrograms N per gram soil (per day for net rates), users will need to combine data from four NEON tables spanning two data products, then conduct a series of simple calculations. These calculations are described in the Appendix, and an R package to perform the calculations using the appropriate data table inputs is available here: https://github.com/NEONScience/NEON-Nitrogen-Transformations.

3.4 Variables Reported

All variables reported from the NEON or external laboratory technician (L0 data) are listed in the file NEON Raw Data Validation for Soil inorganic nitrogen pools and transformations, Level 0 (NEON.DP0.10080) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the file NEON Data Variables for Soil inorganic nitrogen pools and transformations (NEON.DP1.10080) (AD[05]).

Field names have been standardized with Darwin Core terms (http://rs.tdwg.org/dwc/; accessed 16 February 2014), the Global Biodiversity Information Facility vocabularies (http://rs.gbif.org/vocabulary/gbif/; accessed 16 February 2014), the VegCore data dictionary (https://projects.nceas.ucsb.edu/nceas/projects/bien/wiki/VegCore; accessed 16 February 2014), where applicable. NEON TOS spatial data employs the World Geodetic System 1984 (WGS84) for its fundamental reference datum and Earth Gravitational Model 96 (EGM96) for its reference gravitational ellipsoid. Latitudes and longitudes are denoted in decimal notation to six decimal places, with longitudes indicated as negative west of the Greenwich meridian.

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

3.5 Spatial Resolution and Extent

The finest resolution at which spatial data are reported is a unique x,y sampling location and horizon within a NEON plot.

sampleID (unique ID given to the individual soil sampling location and horizon, based on x,y coordinates relative to the southwest corner of the plot) → **plotID** (ID of plot within site) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain.

The basic spatial data included in the data downloaded include domain, site, and plot. However, by joining inorganic N data tables to the field metadata present in the Soil physical properties (Distributed, periodic)

(NEON.DP1.10086) data tables and following the instructions contained in the User Guide for that product (http://data.neonscience.org/api/v0/documents/NEON_soils_userGuide_vA), sample easting and northing for each unique soil sample location can be determined. Shapefiles of all NEON Terrestrial Observation System sampling locations can be found in the Document Library: http://data.neonscience.org/documents.

3.6 Temporal Resolution and Extent

The finest resolution at which temporal data are reported is the **collectDate**. The total number of sampling events per year per site will be 1-3. Sampling will be conducted once every 5 years per site.

The NEON Data Portal provides data in monthly files for query and download efficiency. Queries including any part of a month will return data from the entire month. Code to stack files across months is available here: https://github.com/NEONScience/NEON-utilities

3.7 Associated Data Streams

In order to make full utility of the data contained in Soil inorganic nitrogen pools and transformations, users will need to download the field and laboratory metadata for the parent soil samples, contained in the Soil physical properties (Distributed periodic) (NEON.DP1.10086) data product. Most importantly, the table sls_soilMoisture will be required to convert N concentrations in KCl extracts to micrograms nitrogen per gram dry soil. **sampleID** is the variable name needed to join inorganic N tables to soil physical field and laboratory tables.

Initial soil samples used to measure inorganic N pools are also analyzed to produce several microbial data products, including NEON.DP1.10081, NEON.DP1.10109, NEON.DP1.10108, NEON.DP1.10104. During peak greenness, samples are also analyzed for soil carbon and nitrogen concentrations (NEON.DP1.10078) and stable isotopes (NEON.DP1.10100). A full list of related data products and the variable names needed to join and link tables are provided in the User Guide for Soil physical properties (Distributed periodic) (http://data.neonscience.org/api/v0/documents/NEON_soils_userGuide_vA).

3.8 Product Instances

A maximum of 10 plots will be sampled per site. For each soil horizon present (maximum of 2, organic and mineral), 3 samples will be collected per plot per initial or final collection event, with 2 events per bout. Since sites will have 1-3 sampling bouts per year, this will result in 60-360 unique soil samples extracted for inorganic N per site. Approximately 9 terrestrial NEON sites per year will be sampled for inorganic N pools and net transformation rates. Thus, we expect 540-3,240 data records (excluding procedural blanks) per year.

3.9 Data Relationships

TOS Protocol and Procedure: Soil Biogeochemical and Microbial Sampling dictates that each x,y soil sampling location yields a unique **sampleID** per horizon per **collectDate** (day of year, local time). For paired initial and

final cores, **sampleID**s will differ only by **collectDate**. Each sample is then subsampled and extracted for inorganic N, yielding a corresponding **kclSampleID**. Thus, a record from sls_soilCoreCollection in Soil physical properties (Distributed periodic) may have zero or one child records in the ntr_internalLab table in Soil inorganic nitrogen pools and transformations. The information needed to link procedural blanks with samples is provided in ntr_externalLabBlanks. Each child record may then appear from zero to two times in the ntr_externalLab table. Most will appear once, but some may appear twice if analytical replicates were conducted. 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*.

ntr_internalLab.csv - > One record expected per **sampleID**, generates a single **kclSampleID** used to measure inorganic N concentrations

ntr_internalLabBlanks.csv - > One record expected per **kclReferenceID**, used to track which procedural blanks are associated with a batch of samples

ntr_externalLab.csv - > One record expected per **kclSampleID** x **analyticalRepNumber** combination, associated with ammonium and nitrate + nitrite concentration measurements in KCl extracts

ntr_externalSummary.csv -> One record expected per **laboratoryName** x **analyte** x **sampleType** x **labSpecificStartDate** combination. Can use corresponding variables in the externalLab table to associate sample data with relevant uncertainty values.

sampleIDs and **sampleBarcodes** will be generated for each sample. Physical samples will be discarded following analytical measurements.

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 dropdown options, which reduces the number of processing steps necessary to prepare the raw data for publication. A schematic of the data entry application design for internal lab processing of KCl extracts is depicted in Figure 1. 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 document NEON Raw Data Validation for Soil inorganic nitrogen pools and transformations, Level 0 (NEON.DP0.10080), provided with every download of this data product. Contained within this file is a field named ‘entryValidationRulesForm’, which describes syntactically the validation rules for each field built into the data entry application. There is also a field named ‘entryValidationRulesParser’, which described the validation rules for external labs that submit spreadsheets to the NEON database. Data entry constraints are described in Nicl syntax in the validation file provided with every data download, and the Nicl language is described in NEON’s Ingest Conversion Language (NICL) specifications ([AD[12]).

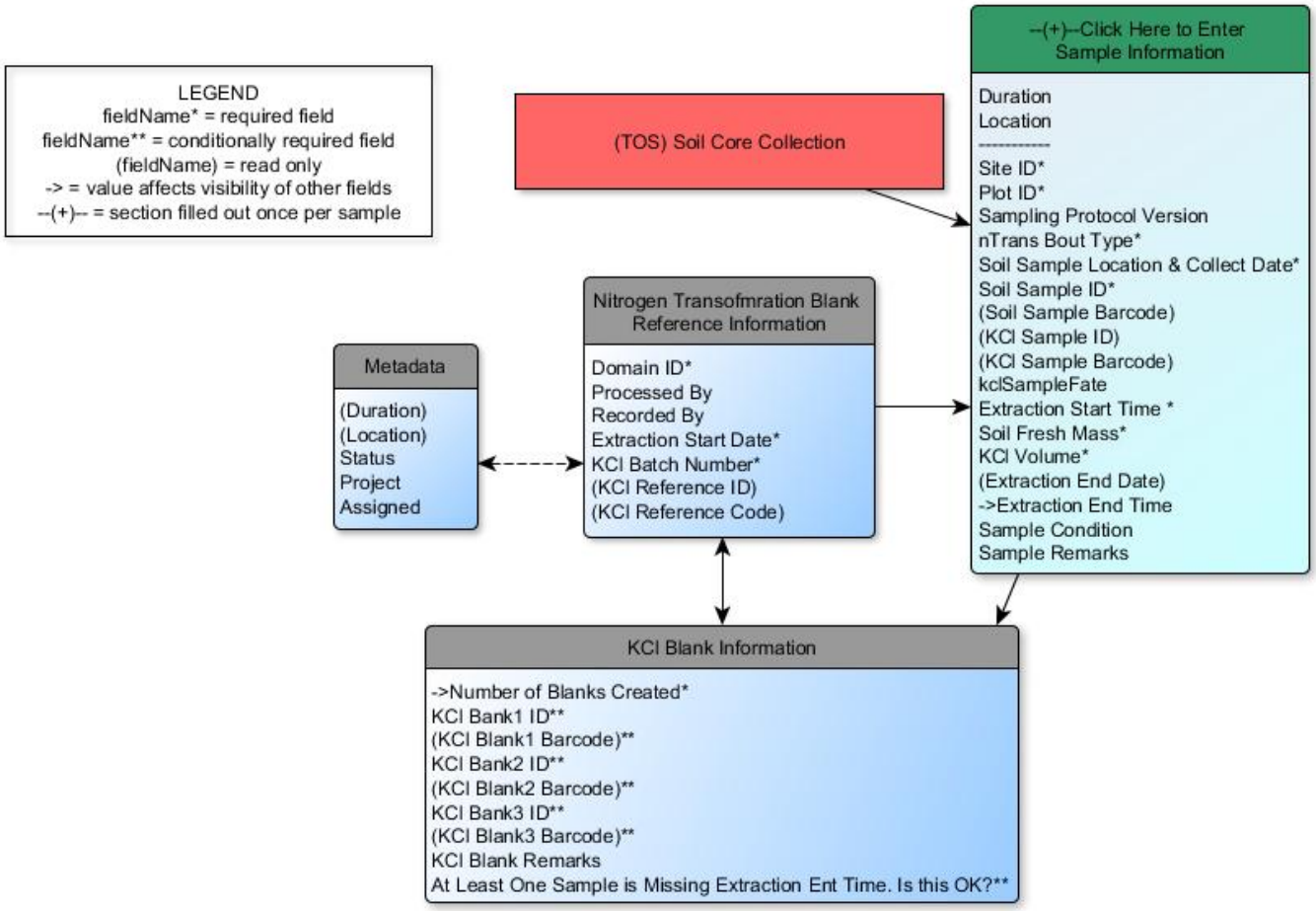

Figure 1: Schematic of the application used by field technicians to enter KCl extraction data

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[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 data record is a quality flag for known errors applying to the record, added by NEON Science upon data review.

In 2017, Type II (1 Mohm) water was used to create 2M KCl solutions and conduct extractions. However, this purity is not sufficient and Type I water (18.2 Mohm) will be used thereafter. 2017 data was thus collected using a deprecated method and has been flagged as described below.

Additionally, where NEON Science notes problematic sample sets with sample concentrations lower than the blanks (suggesting impure or contaminated water sources), concentration values are set to NA and the samples are flagged as described below.

fieldName	value	definition
dataQF	deprecatedMethod	Data generated using deprecated procedures; Type II water
dataQF	blanksExceedSamples	Blank values higher than samples, concentration data set to NA and corrective actions undertaken

4.5 Analytical Facility Data Quality

All analytical labs that generate inorganic N data include standards run as unknowns alongside NEON samples in order to gauge run acceptability. Long-term analytical precision and accuracy of these standard analyses are reported for each lab to allow users to interpret and analyze inorganic N concentrations and net rate calculations in the context of their uncertainty range. The data table ntr_externalSummary (available in the expanded package) contains the long-term precision and accuracy of lab analyses.

In addition, labs communicate record-level issues with samples or measurements using the quality flags described below. In general, an entry of 0 in a quality flag field means there is no issue to report.

For further information about individual laboratory QA procedures, refer to lab SOPs, found on the NEON Data Portal (http://data.neonscience.org/home) in the Resources > Data Documentation > External Lab Protocols section.

fieldName	value	definition
ammoniumNQF	0	No issues to report
ammoniumNQF	1	Sample ammonium N concentration below the minimum detection limit
ammoniumNQF	2	Other, described in remarks
nitrateNitriteNQF	0	No issues to report
nitrateNitriteNQF	1	Sample nitrate + nitrite N concentration below the minimum detection limit
nitrateNitriteNQF	0	Other, described in remarks

5 APPENDIX

In order to estimate blank-corrected N concentrations in KCl extracts, then convert to micrograms N per gram soil (per day for net rates), users will need to combine data from four NEON tables spanning two data products, then conduct a series of simple calculations. The steps entailed are detailed below. For users comfortable with R, there is a simple R package that will conduct these calculations for you, available for download here: https://github.com/NEONScience/NEON-Nitrogen-Transformations. The def.calc.ntrans function creates several new terms not contained in the file NEON Data Variables for Soil inorganic nitrogen pools and transformations (NEON.DP1.10080) (AD[05]), thus they are defined below:

fieldName	definition	units
blank1ID	identifier of the first blank associated with a set of samples	NA
blank1NH4	ammonium - N concentration of the first blank associated with a set of samples	milligrams per liter
blank1NO3	(nitrate + nitrite) - N concentration of the first blank associated with a set of samples	milligrams per liter
blank2ID	identifier of the second blank associated with a set of samples	NA
blank2NH4	ammonium - N concentration of the second blank associated with a set of samples	milligrams per liter
blank2NO3	(nitrate + nitrite) - N concentration of the second blank associated with a set of samples	milligrams per liter
blank3ID	identifier of the third blank associated with a set of samples	NA
blank3NH4	ammonium - N concentration of the third blank associated with a set of samples	milligrams per liter
blank3NO3	(nitrate + nitrite) - N concentration of the third blank associated with a set of samples	milligrams per liter
blankNH4mean	mean ammonium - N concentration of the blanks associated with a set of samples	milligrams per liter
blankNO3mean	mean (nitrate + nitrite) - N concentration of the blanks associated with a set of samples	milligrams per liter
kclAmmoniumN BlankCor	blank-corrected ammonium - N concentration concentration of a sample, set to zero if negative	milligrams per liter
kclNitrateNitriteN BlankCor	blank-corrected (nitrate + nitrite) - N concentration concentration of a sample, set to zero if negative	milligrams per liter
soilDryMass	dry mass of soil used for KCl extractions	grams
soilAmmoniumNug PerGram	extractable ammonium - N content in sample	micrograms N per gram soil
soilNitrateNitriteNug PerGram	extractable (nitrate + nitrite) - N content in sample	micrograms N per gram soil
soilInorganicNug PerGram	extractable inorganinc N content in sample	micrograms N per gram soil
netNminugPerGram PerDay	net N mineralized over the course of the incubation	micrograms N per gram soil per day
netNitugPerGram PerDay	net nitrification over the course of the incubation	micrograms N per gram soil per day

For users who wish to conduct the N transformation calculations on their own, the following table lists the key variables needed as inputs as well as their associated data products and tables of origin:

DPID	tableName	fieldName	units
NEON.DP1.10080	ntr_externalLab	kclAmmoniumNConc	milligrams per liter
NEON.DP1.10080	ntr_externalLab	kclNitrateNitriteNConc	milligrams per liter
NEON.DP1.10080	ntr_internalLabBlanks	kclReferenceID	NA
NEON.DP1.10080	ntr_internalLab	incubationLength	days
NEON.DP1.10080	ntr_internalLab	kclVolume	milliliters
NEON.DP1.10080	ntr_internalLab	soilFreshMass	grams
NEON.DP1.10086	sls_soilMoisture	dryMassFraction	unitless

Before conducting the calculations, users may wish to review the ntr_internalLab and ntr_externalLab tables and decided whether they'd like to remove/ignore any values with anomolous **sampleConditions** or **receivedConditions**. Additionally, they may want to ignore concentration values with certain **ammoniumNQF** or **nitrateNitriteNQF** values.

Step-by-step calculations:

1. Calculate mean blank ammounium - N concentration for each sample set using the 1-3 blanks associated with each **kclReferenceID**.

$$\overline{(NH_4^+ - N)_B} = \frac{(NH_4^+ - N)_{B1} + \cdots + (NH_4^+ - N)_{Bn}}{n}$$

2. Calculate mean blank (nitrate + nitrite) - N concentration for each sample set using the 1-3 blanks associated with each **kclReferenceID**. (nitrate + nitrite) - N is abbreviated as NO_3^--N in the equations below.

$$\overline{(NO_3^- - N)_B} = \frac{(NO_3^- - N)_{B1} + \cdots + (NO_3^- - N)_{Bn}}{n}$$

3. Calculate corrected ammonium - N concentrations for each sample by subtracting the mean value of the blanks for that sample set (use the **kclReferenceID** to associated samples with blanks).

$$(NH_4^+ - N)_{cor} = (NH_4^+ - N)_{Sam} - \overline{(NH_4^+ - N)_B}$$

4. Calculate corrected (nitrate + nitrite) - N concentrations for each sample by subtracting the mean value of the blanks for that sample set (use the **kclReferenceID** to associated samples with blanks).

$$(NO_3^- - N)_{cor} = (NO_3^- - N)_{Sam} - \overline{(NO_3^- - N)_B}$$

5. Convert the **soilFreshMass** used in KCl extractions to soil dry mass by multiplying by the **dryMassFraction**.

$$soilDryMass = soilFreshMass * dryMassFraction$$

6. Convert to micrograms ammonium - N per gram soil by multiplying the corrected ammonium - N concentration by the **kclVolume** used in extractions and dividing by the soil dry mass calculated in the step above.

$$(NH_4^+ - N)_{soil} = \frac{(NH_4^+ - N)_{cor} * kclVolume}{soilDryMass}$$

7. Convert to micrograms (nitrate + nitrite) - N per gram soil by multiplying the corrected (nitrate + nitrite) - N concentration by the **kclVolume** used in extractions and dividing by the soil dry mass calculated above.

$$(NO_3^- - N)_{soil} = \frac{(NO_3^- - N)_{cor} * kclVolume}{soilDryMass}$$

8. Use the **incubationPairID** and **nTransBoutType** variables to arrange cores in initial and final pairs.

9. Calculate net N mineralization using a core pair as follows: total inorganic N in final core minus total inorganic N in initial core, divided by **incubationLength** of the final core.

$$netNmin = \frac{((NH_4^+ - N)_{cor} + (NO_3^- - N)_{cor})_{final} - ((NH_4^+ - N)_{cor} + (NO_3^- - N)_{cor})_{initial}}{incubationLength}$$

10. Calculate net nitrification using a core pair as follows: (nitrate + nitrate) - N in final core minus (nitrate + nitrate) - N in initial core, divided by **incubationLength** of the final core.

$$netNit = \frac{((NO_3^- - N)_{cor})_{final} - ((NO_3^- - N)_{cor})_{initial}}{incubationLength}$$