

NEON USER GUIDE TO STABLE ISOTOPE CONCENTRATIONS IN SURFACE WATERS (NEON.DP1.20206) AND STABLE ISOTOPE CONCENTRATIONS IN GROUNDWATER (NEON.DP1.20276)

PREPARED BY	ORGANIZATION	DATE
Tanya Chesney	AOS	02/05/2018
Caren Scott	AOS	02/05/2018
Samantha Weintraub	TOS	02/05/2018

<i>Title:</i> NEON User Guide to Stable isotope concentrations in surface waters (NEON.DP1.20206) and Stable isotope concentrations in groundwater (NEON.DP1.20276)	<i>Date:</i> 02/05/2018
<i>Author:</i> Tanya Chesney	<i>Revision:</i> A

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	02/05/2018	Initial Release

TABLE OF CONTENTS

1	DESCRIPTION	1
1.1	Purpose	1
1.2	Scope	1
2	RELATED DOCUMENTS AND ACRONYMS	2
2.1	Associated Documents	2
2.2	Acronyms	2
3	DATA PRODUCT DESCRIPTION	3
3.1	Spatial Sampling Design	3
3.2	Temporal Sampling Design	4
3.3	Theory of Laboratory Measurements	5
3.4	Variables Reported	5
3.5	Spatial Resolution and Extent	6
3.6	Temporal Resolution and Extent	6
3.7	Associated Data Streams	6
3.8	Product Instances	7
3.9	Data Relationships	7
3.10	Special Considerations	8
4	DATA QUALITY	8
4.1	Data Entry Constraint and Validation	8
4.2	Automated Data Processing Steps	9
4.3	Data Revision	9
4.4	Quality Flagging	10
4.5	Analytical Facility Data Quality	10

LIST OF TABLES AND FIGURES

Figure 1	Generic layout of NEON surface and groundwater sampling locations in wadeable streams, non-wadeable streams, and lakes	4
Figure 2	Schematic of the applications used by field technicians to enter aquatic stable isotope field data	9

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, for example the d18OWater, are considered the lowest level (Level 0). Raw data that have been quality checked via the steps detailed herein, as well as simple metrics that emerge from the raw data, are considered Level 1 data products.

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

1.2 Scope

This document describes the steps needed to generate the L1 data products Stable isotope concentrations in surface water and Stable isotope concentrations in groundwater. 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 Stable isotope concentrations in surface waters (NEON.DP1.20206) (AD[08]) and NEON Data Variables for Stable isotope concentrations in groundwater (NEON.DP1.20276) (AD[09]), 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 and processed according to AOS Protocol and Procedure: Stable Isotope Sampling in Surface Waters and Groundwater (AD[10]). The raw data that are processed in this document are detailed in the files NEON Raw Data Validation for Water chemistry, isotopes, dissolved gas, and microbes sampling, Level 0 (NEON.DP0.20090) (AD[04]), NEON Raw Data Validation for Stable isotope concentrations in surface waters and groundwater field data (NEON.DP0.20206) (AD[05]), NEON Raw Data Validation for Plant and algae external lab chemistry (NEON.DP0.20065) (AD[06]), and NEON Raw Data Validation for H2O isotopes external lab data (NEON.DP0.20205) (AD[07]), 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., '20093') 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.001152	NEON Aquatic Sampling Strategy
AD[03]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products Catalog
AD[04]	NEON.DP0.20090.001_dataValidation.csv	NEON Raw Data Validation for Water chemistry, isotopes, dissolved gas, and microbes sampling, Level 0 (NEON.DP0.20090)
AD[05]	NEON.DP0.20206.001_dataValidation.csv	NEON Raw Data Validation for Stable isotope concentrations in surface waters and groundwater field data (NEON.DP0.20206)
AD[06]	NEON.DP0.20065.001_dataValidation.csv	NEON Raw Data Validation for Plant and algae external lab chemistry (NEON.DP0.20065)
AD[07]	NEON.DP0.20205.001_dataValidation.csv	NEON Raw Data Validation for H2O isotopes external lab data (NEON.DP0.20205)
AD[08]	NEON.DP1.20206.001_variables.csv	NEON Data Variables for Stable isotope concentrations in surface waters (NEON.DP1.20206)
AD[09]	NEON.DP1.20206.001_variables.csv	NEON Data Variables for Stable isotope concentrations in groundwater (NEON.DP1.20276)
AD[10]	NEON.DOC.002905	AOS Protocol and Procedure: Stable Isotope Sampling in Surface Waters and Groundwater
AD[11]	NEON.DOC.000008	NEON Acronym List
AD[12]	NEON.DOC.000243	NEON Glossary of Terms
AD[13]	OS_Generic_Transitions.pdf	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[14]		NEON's Ingest Conversion Language (NICL) specifications

2.2 Acronyms

Acronym	Definition
H2O	Water
C	Carbon
N	Nitrogen
O	Oxygen
POM	Particulate organic matter

3 DATA PRODUCT DESCRIPTION

The Stable isotope concentrations in surface waters (NEON.DP1.20206) and Stable isotope concentrations in groundwater (NEON.DP1.20276) data products provide stable isotope values for surface and groundwater samples collected using AOS Protocol and Procedure: Stable Isotope Sampling in Surface Waters and Groundwater (AD[10]). These procedures implement the guidelines and requirements described in the NEON Aquatic Sampling Strategy (AD[02]). All data are reported at the spatial resolution of a single water sample, collected from a unique stationID within a sampled water body. The temporal resolution is that of a single collection date.

Field sampling strategies are specific to the type of waterbody and are described further below. Stable isotope data are produced by external laboratories and include measurements of hydrogen and oxygen stable isotopes in water and carbon and nitrogen stable isotopes in particulate organic matter (POM).

Naturally occurring surface and groundwater stable isotope concentration data allow researchers to assess element cycling, food web dynamics, nutrient transfer, and hydrological modeling within aquatic ecosystems and watersheds. Measuring long-term trends in surface and groundwater stable isotopes is part of the overall NEON biogeochemistry goal to understand how major nutrient and carbon fluxes within and across air, land, and water systems change over 30 years.

3.1 Spatial Sampling Design

In wadeable streams, dip sampling in the thalweg is used to obtain stable isotope samples, assuming the stream channel is completely mixed (Figure 1, top left). In non-wadeable streams, dip sampling in the thalweg at 0.5 m depth is used (Figure 1, top right). For all stream types, samples are collected as close as possible and downstream of sensor sets.

In lakes, up to three locations are designated for sampling: one in the deepest part of the lake, one near the most prominent inlet, and one near the outlet. Samples are taken from variable depths dependent on the degree of lake stratification and are collocated near lake buoy, inlet, and outlet sensor infrastructure (Figure 1, bottom left).

Groundwater stable isotope samples are collected from up to eight 2" diameter, shallow (<100 ft depth) groundwater observation wells on the perimeter of sampled waterbodies (Figure 1). A subset of wells are selected for sampling on a site-by-site basis during each sampling event. Periodic changes to the selected subset of wells may occur during the life of the Observatory and are guided by various parameters, including changes in hydrologic conditions (dry wells, changes in hydrologic flow paths) and status of infrastructure (damaged wells).

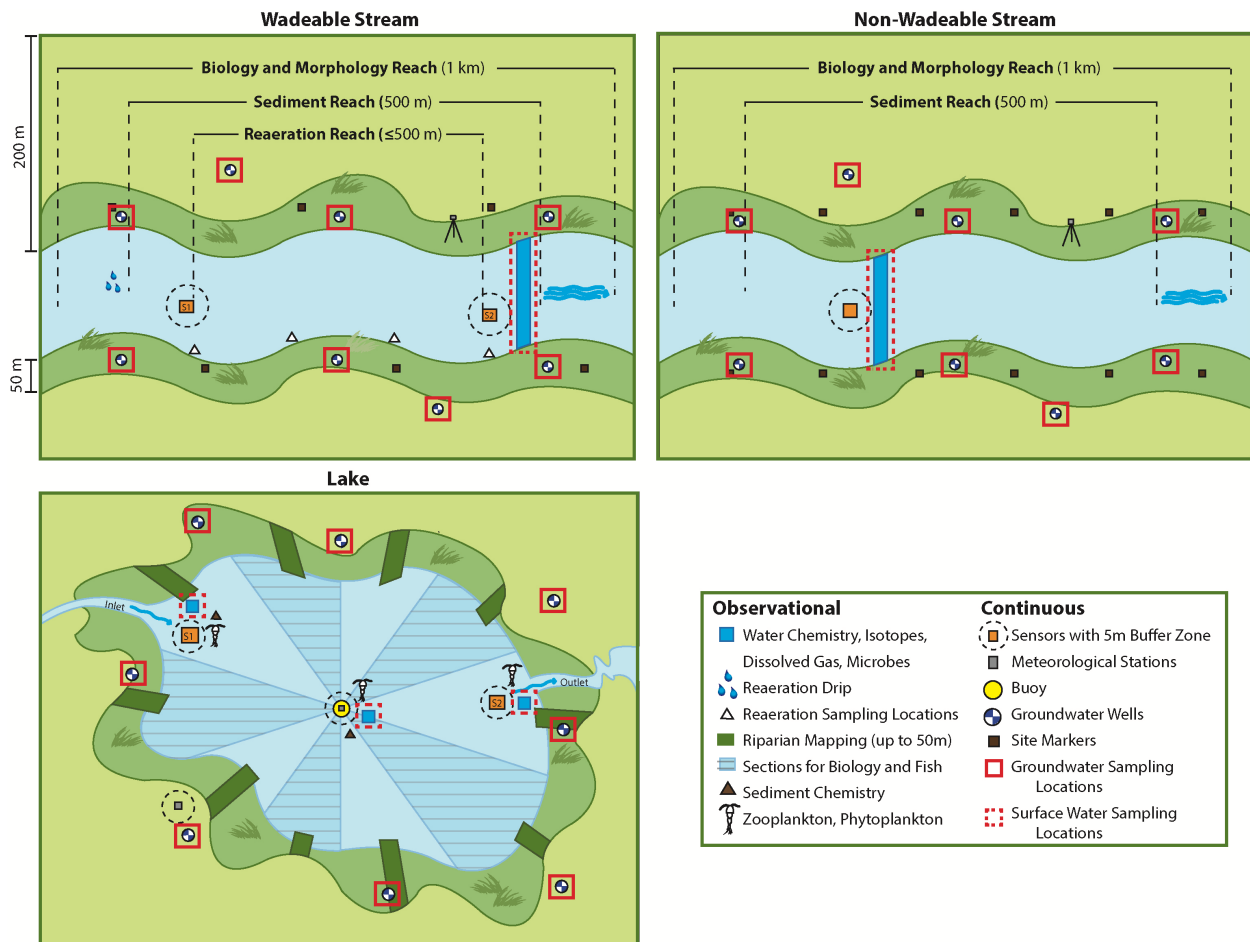


Figure 1: Generic layout of NEON surface and groundwater sampling locations in wadeable streams, non-wadeable streams, and lakes

3.2 Temporal Sampling Design

For streams, sampling occurs up to 26 times per year, approximately every other week but guided by historical stream discharge data for each NEON site. For example, wadeable streams with little or no flow during the summer dry-season or that are completely frozen during the winter are sampled more intensively during wet periods or snowmelt. When applicable, stream samples are collected to coincide with NEON atmospheric wet deposition sampling (Tuesdays) as well as Observatory-wide sampling efforts.

Lake samples will be collected 12 times per year - approximately monthly and during shoulder seasons to capture ice-on/ice-off and lake turnover events. When applicable, lake samples are also preferentially collected to coincide with other aqueous sampling efforts.

Groundwater samples are collected up to twice per year, roughly during early spring and late fall and based on historic cumulative discharge. Samples will be collected within +/- 1 day of surface water sampling events where

possible. Groundwater samples are analyzed for H₂O isotopes only.

All water samples are filtered as soon as possible following collection, preferably within 3-6 hours. Water is held at room temperature in a bottle wrapped with parafilm (to prevent evaporation). For surface water, filters were either frozen (prior to 2018) or oven-dried at 65C (from 2018 onwards). Both sample types are then shipped to analytical facilities for isotopic measurements.

3.3 Theory of Laboratory Measurements

Deuterium and oxygen stable isotopes in water are measured using cavity ringdown spectrometry. Carbon and nitrogen stable isotopes in POM are measured via combustion and elemental analysis (EA) coupled to isotope ratio mass spectrometry (IRMS).

Isotopes are measured as the abundance ratio of a heavy, rare isotope (H) to a light, more common isotope (L), relative to those same ratios in a standard reference material.

$$\delta = [(R_{sample}/R_{standard} - 1)] \times 1000$$

where R = H/L. For all NEON aquatic stable isotope data, deuterium and $\delta^{18}\text{O}$ values are reported on the VSMOW-VSLAP scale, $\delta^{15}\text{N}$ values are reported relative to atmospheric N₂, and $\delta^{13}\text{C}$ values are reported relative to Vienna Pee Dee Belemite.

For data collected in 2017 and beyond, standard operating procedures for laboratories performing surface water and groundwater stable isotope analyses can be found in the NEON Data Portal document library (<http://data.neonscience.org/documents>), in the External Lab Protocols section.

3.4 Variables Reported

All variables reported from the field technician or laboratory (LO data) are listed in the files, NEON Raw Data Validation for Water chemistry, isotopes, dissolved gas, and microbes sampling, Level 0 (NEON.DP0.20090) (AD[04]), NEON Raw Data Validation for Stable isotope concentrations in surface waters and groundwater field data (NEON.DP0.20206) (AD[05]), NEON Raw Data Validation for Plant and algae external lab chemistry (NEON.DP0.20065) (AD[06]), and NEON Raw Data Validation for H₂O isotopes external lab data (NEON.DP0.20205) (AD[07]). All variables reported in the published data (L1 data) are also provided separately in the files NEON Data Variables for Stable isotope concentrations in surface waters (NEON.DP1.20206) (AD[08]) and NEON Data Variables for Stable isotope concentrations in groundwater (NEON.DP1.20276) (AD[09]).

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 AOS 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 spatial resolution at which aquatic stable isotope data are reported is a single sample collected from a unique stationID. Overall, this results in a spatial hierarchy of:

sampleID (unique ID given to the individual water sample) → **stationID** (ID of the sampling location) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain).

StationID in wadeable and non-wadeable (i.e. rivers) streams is indicated as 'ss', and stationIDs for lakes are designated as 'in', 'ot', 'c0', and 'c1','c2', and 'c3', as needed (if center is stratified), with 'c1' being the top layer.

The basic spatial data included in the data downloaded include the latitude, longitude, and elevation of the stationID where sampling occurred, plus associated uncertainty due to GPS error. Shapefiles of all NEON Aquatic Observation System sampling locations can be found in the Document Library: <http://data.neonscience.org/documents>. If for some reason samples cannot be taken at the designated stations, the coordinates for the center of the site will be returned. In this case, the actual coordinates of the sampling location are recorded in the **altLocation**, **altLatitude** and **altLongitude** fields.

For groundwater, users can find the depth of the well as the 'z offset' in the full spatial data, which can be accessed via:

1. The `def.extr.geo.os.R` function from the `geoNEON` package, available here: <https://github.com/NEONScience/NEON-geolocation>
2. The NEON API: <http://data.neonscience.org/data-api>

3.6 Temporal Resolution and Extent

The finest resolution at which aquatic stable isotope temporal data are reported is the **collectDate**, a single date on which stable isotope samples were collected. The total number of sampling events per year is expected to be 26 per wadeable and non-wadeable stream, 12 per lake, and 2 per groundwater well subset per site.

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. Code to stack files across months is available here: <https://github.com/NEONScience/NEON-utilities>

3.7 Associated Data Streams

The super parent sample of the field data from the Stable isotope concentrations in surface water and Stable isotope concentrations in groundwater data products is shared with several other related data products. **siteID**, **stationID** and **collectDate** (or **parentSampleID** of the `xxx_fieldSuperParent` table in each data product) are the linking variables that can be used to associate stable isotope samples and their metadata to related samples found in the

Dissolved gases in surface water (DP1.20097.001), Chemical properties of surface water (DP1.20093), Chemical properties of groundwater (DP1.20092), and Surface water microbe cell count (DP1.20138.001) data products.

3.8 Product Instances

The NEON Observatory contains 34 aquatic sites, consisting of 24 wadeable streams, 3 non-wadeable streams (rivers), and 7 lakes.

Surface and groundwater stable isotope sampling yields one unique sample per stationID per sampling event. There will be up to 26 sampling events per year in wadeable and non-wadeable streams, up to 12 sampling events in lakes, and up to 2 sampling events for groundwater. Thus in wadable streams, there will be up to 26 unique sample records per site per year, non-wadeable streams will also yield up to 26 records per year, lakes will yield 36-60 records per year, depending on stratification, and groundwater will produce 8 records per year. Observatory-wide, this will yield a total of 1226-1394 unique records with stable isotope data per year.

NOTE: Replicate samples may be taken of a small percentage of our samples. If replicate samples are taken, there will be one unique sample per **replicateNumber** per stationID per sampling event, and the sample ID(s) of the replicate sample(s) will have the **replicateNumber** appended to the end.

3.9 Data Relationships

The protocol dictates that each siteID x stationID combination is sampled at least once per event (one record expected per parentSampleID in asi_fieldSuperParent or gsi_fieldSuperParent). A record from asi_fieldSuperParent or gsi_fieldSuperParent may have zero or one child records in asi_fieldData or gsi_fieldData, depending on whether a water sample was collected. In the event that a water sample cannot be taken, a record will still be created in fieldSuperParent tables, but the **samplingImpractical** field will be something other than NULL and there will be no corresponding record in fieldData tables. Each record from asi_fieldData is expected to have one child record in asi_externalLabH2OIsotopes and one child record in asi_POMExternalLabDataPerSample. Each record from gsi_fieldData is expected to have one child record in gsi_externalLabH2OIsotopes. However, duplicates and/or missing data may exist where protocol and/or data entry aberrations have occurred; *users should check data carefully for anomalies before joining tables*. **sampleIDs** and **sampleBarcodes** will be generated for each sampling event. After shipment to external labs is complete, any physical sample that remains will be discarded.

Stable isotope concentrations in surface waters:

asi_fieldSuperParent.csv - > One record expected per **parentSampleID**

asi_fieldData.csv - > One record expected per **parentSampleID**. Generates a single subsample for water isotope analyses, **isotopeH2OSampleID**, and up to two filters for POM isotope analyses, **isotopePOMSampleID** and **isotopePOMRep2SampleID**

asi_externalLabH2OIsotopes.csv - > One record expected per **isotopeH2OSampleID**, associated with external laboratory water isotope analyses

asi_POMExternalLabDataPerSample.csv - > One record expected per **sampleID**, associated with external POM isotope analyses

<p>Title: NEON User Guide to Stable isotope concentrations in surface waters (NEON.DP1.20206) and Stable isotope concentrations in groundwater (NEON.DP1.20276)</p>	<p>Date: 02/05/2018</p>
<p>Author: Tanya Chesney</p>	<p>Revision: A</p>

asi_externalLabSummaryData.csv -> One record expected per **analyte x method x laboratoryName x labSpecific-StartDate** combination, used to associate sample data with relevant uncertainty values.

apl_externalLabSummaryData.csv -> One record expected per **analyte x method x laboratoryName x labSpecific-StartDate** combination, used to associate sample data with relevant uncertainty values.

Stable isotope concentrations in groundwater:

gsi_fieldSuperParent.csv -> One record expected per **parentSampleID**

gsi_fieldData.csv -> One record expected per **parentSampleID**. Generates a single subsample for water isotope analyses, **isotopeH2OSampleID**

gsi_externalLabH2OIsotopes.csv -> One record expected per **isotopeH2OSampleID**, associated with external laboratory water isotope analyses

asi_externalLabSummaryData.csv -> One record expected per **analyte x method x laboratoryName x labSpecific-StartDate** combination, used to associate sample data with relevant uncertainty values.

3.10 Special Considerations

None to report.

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 (field data) and web user interface (UI, lab data). 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 (Figure 2). 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 Water chemistry, isotopes, dissolved gas, and microbes sampling, Level 0 (NEON.DP0.20090), NEON Raw Data Validation for Stable isotope concentrations in surface waters and groundwater field data (NEON.DP0.20206), NEON Raw Data Validation for Plant and algae external lab chemistry (NEON.DP0.20065), and NEON Raw Data Validation for H2O isotopes external lab data (NEON.DP0.20205), 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.

Additionally, there is a field named 'entryValidationRulesParser', which describes the validation rules for external labs that submit spreadsheets to the NEON database. Data entry constraints are described in NiCl syntax in the validation file provided with every data download, and the NiCl language is described in NEON's Ingest Conversion Language (NICL) specifications ([AD[14]).

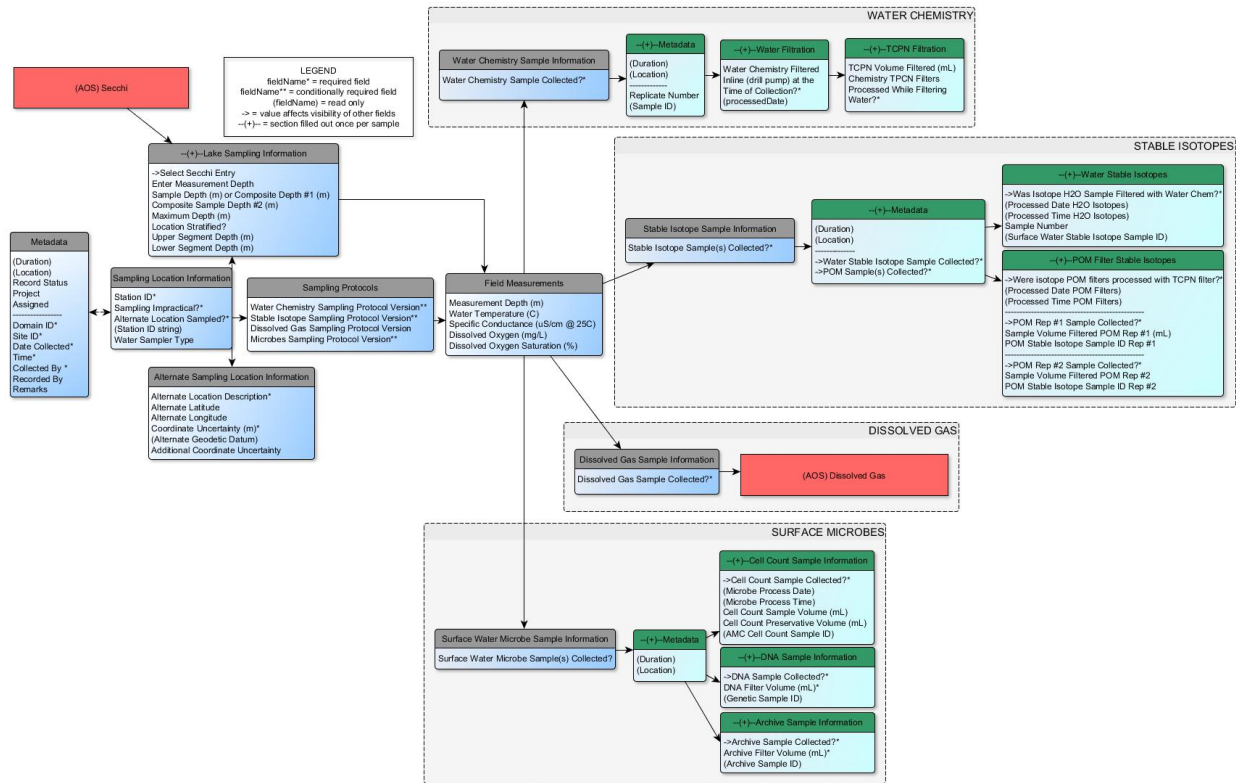


Figure 2: Schematic of the applications used by field technicians to enter aquatic stable isotope field data

Note that field data collected prior to 2017 were processed using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow. Moreover, external laboratory data were also not subject to the same full suite of quality controls.

4.2 Automated Data Processing Steps

Following data entry into a mobile application or web user interface, the steps used to process the data through to publication on the NEON Data Portal are detailed in the NEON Algorithm Theoretical Basis Document: OS Generic Transitions (AD[13]).

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.

<i>Title:</i> NEON User Guide to Stable isotope concentrations in surface waters (NEON.DP1.20206) and Stable isotope concentrations in groundwater (NEON.DP1.20276)	<i>Date:</i> 02/05/2018
<i>Author:</i> Tanya Chesney	<i>Revision:</i> A

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 dataQF entries for the surface water and groundwater stable isotope data products.

4.5 Analytical Facility Data Quality

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

In addition, long-term analytical precision and accuracy of secondary reference material analyses are reported for each lab to allow users to interpret and analyze aquatic stable isotope data in the context of their uncertainty ranges. The data tables `asi_externalLabSummaryData` and `apl_externalLabSummaryData`, which are available in the data product expanded package, contain the long-term precision and accuracy of lab analyses.

For further information about individual laboratory QA procedures, refer to the lab-specific SOPs found in the NEON Data Portal document library (<http://data.neonscience.org/documents>), External Lab Protocols section. NEON’s Calibration/Validation department has regular procedures for auditing the quality assurance of external laboratories and their reports are also available to data users.

`asi_externalLabH2OIsotopes` and `gsi_externalLabH2OIsotopes`

fieldName	value	definition
isotopeH2OExternalLabQF	1	High sample standard deviation (d2H sd >= 0.75 or d18O sd >= 0.2)
isotopeH2OExternalLabQF	0	No issue to report, low sample standard deviations