NEON USER GUIDE TO STABLE ISOTOPES IN SURFACE WATERS (DP1.20206.001) AND STABLE ISOTOPES IN GROUNDWATER (DP1.20276.001)

PREPARED BY	ORGANIZATION
Zachary Nickerson	AOS
Tanya Chesney	AOS
Caren Scott	AOS
Samantha Weintraub	TOS

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	02/05/2018	Initial Release
B	11/01/2020	Included general statement about usage of neonUtilities R package and statement about possible location changes, updated associated documents, updated generic site maps and isotope sampling locations, added section on sampling design changes, updated internal quality flagging information, updated language on seepage and flow-through lake sampling locations, updated data relationship information to include proper sample numbers expected and update the name of the POM external lab summary table
B.1	06/01/2021	Updated sampling design changes, added quality flagging choices for external laboratory data
C	02/08/2022	Updated section 4.3 Data Revision with latest information regarding data release
D	10/06/2022	Added information about the decarbonation via acid digestion SOP developed by the external lab for the particulate organic matter (POM) isotope samples, updated information pertaining to laboratory quality assurance and uncertainty, added sampling design changes
E	02/18/2025	Updated the url for spatial data in section 3.7. Added information about the new neonUtilities Python package.

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, 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 isotopes in surface water (DP1.20206.001) and Stable isotopes in groundwater (DP1.20276.001). 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 in surface waters (DP1.20206.001) (AD[08]) and NEON Data Variables for Stable isotope in groundwater (DP1.20276.001) (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 (DP0.20090.001) (AD[04]), NEON Raw Data Validation for Stable isotope in surface waters and groundwater field data (DP0.20206.001) (AD[05]), NEON Raw Data Validation for Plant and algae external lab chemistry (DP0.20065.001) (AD[06]), and NEON Raw Data Validation for H20 isotopes external lab data (DP0.20205.001) (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 Data Products Catalog
AD[04]	Available with data download	NEON Raw Data Validation for Water chemistry, isotopes, dissolved gas, and microbes sampling, Level 0 (DP0.20090.001)
AD[05]	Available with data download	NEON Raw Data Validation for Stable isotope in surface waters and groundwater field data (DP0.20206.001)
AD[06]	Available with data download	NEON Raw Data Validation for Plant and algae external lab chemistry (DP0.20065.001)
AD[07]	Available with data download	NEON Raw Data Validation for H20 isotopes external lab data (DP0.20205.001)
AD[08]	Available with data download	NEON Data Variables for Stable isotope in surface waters (DP1.20206.001)
AD[09]	Available with data download	NEON Data Variables for Stable isotope in groundwater (DP1.20276.001)
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]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[14]	Available on NEON data portal	NEON Ingest Conversion Language Function Library
AD[15]	Available on NEON data portal	NEON Ingest Conversion Language
AD[16]	Available with data download	Categorical Codes csv

2.2 Acronyms

Acronym	Definition
H2O	Water
C	Carbon
N	Nitrogen
O	Oxygen
POM	Particulate organic matter
S2	Aquatic sensor set 2
°C	Degrees Celcius
1N	1 Normal
HCl	Hydrochloric acid
min	Minute
ft	Feet
m	Meter
EA	Elemental analyzer
IRMS	Isotope ratio mass spectrometry
R	Stable isotope ratio
H	Heavy, rare isotope
L	Light, more common isotope
SOP	Standard Operating Procedure
API	Application Programming Interface
GPS	Global Positioning System
UI	User Interface
Nicl	NEON's Ingest Conversion Language

3 DATA PRODUCT DESCRIPTION

The Stable isotopes in surface waters data product (DP1.20206.001) and Stable isotopes in groundwater data product (DP1.20276.001) provide stable isotope ratios 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 isotope ratios in water (H2O) and carbon and nitrogen stable isotope ratios in particulate organic matter (POM).

Naturally occurring surface and groundwater stable isotope ratio 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 isotope ratios 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 1a). In rivers, dip sampling in the thalweg at 0.5 m depth is used (Figure 1b). For all stream types, water samples are collected immediately downstream of the most downstream sensor set, S2 or buoy, so that sensor measurements can be validated with water samples.

In seepage and flow-through lakes, samples are collected from the deepest part of the lake. Samples are taken from variable depths dependent on the degree of lake stratification and are collocated near the lake's buoy sensor infrastructure (Figure 1c and d). In flow-through lakes, up to 2 additional samples are collected near the inflow and outflow sensor sets (Figure 1d).

Groundwater stable isotope samples are collected from up to 8 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).

As much as possible, sampling occurs in the same locations over the lifetime of the Observatory. However, over time some sampling locations may become impossible to sample, due to disturbance or other local changes. When this occurs, the location and its location ID are retired. A location may also shift to slightly different coordinates. Refer to the locations endpoint of the NEON API for details about locations that have been moved or retired: https://data.neonscience.org/data-api/endpoints/locations/

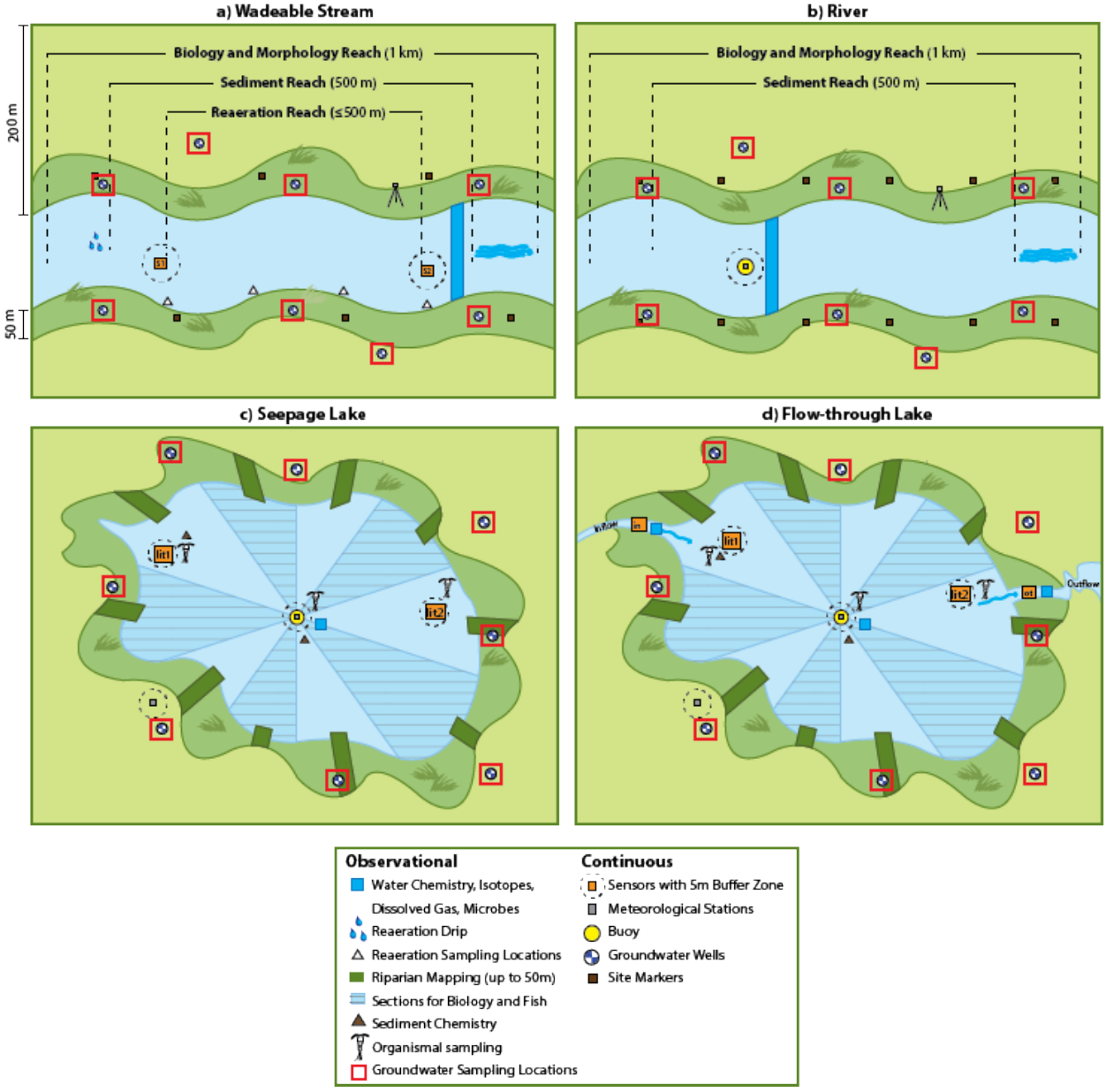

Figure 1: Generic layout of NEON surface and groundwater sampling locations in a) wadeable streams, b) rivers, c) seepage lakes, and d) flow-through lakes. Isotope sampling locations in wadeable streams and rivers are denoted by dark blue quadrangles, and isotope sampling locations in lakes are outlined by red squares. See AD[10] for further details.

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 2 times 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 when 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 are either frozen (prior to 2018) or oven-dried at 65 °C (from 2018 onwards). Both sample types are 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 spectrometery. Particulate carbon and nitrogen stable isotopes in POM are measured via combustion and elemental analysis (EA) coupled to isotope ratio mass spectrometry (IRMS).

Due to high levels of inorganic carbon in the surface water of some NEON sites, samples from certain sites are decarbonated via acid digestion (1N HCl for 30 min) prior to analysis. The following table reports the acid treatment status for each NEON Aquatic site.

Table 1: Acid treatment status for each NEON Aquatic site

NEON Domain ID	NEON Site ID	Acid Treatment
D01	HOPB	N
D02	LEWI	Y
D02	POSE	N
D03	BARC	N
D03	FLNT	N
D03	SUGG	N
D04	CUPE	Y
D04	GUIL	Y
D05	CRAM	N
D05	LIRO	N
D06	KING	Y
D06	MCDI	Y
D07	LECO	N
D07	WALK	Y
D08	BLWA	N
D08	MAYF	N
D08	TOMB	N
D09	PRLA	Y
D09	PRPO	Y
D10	ARIK	N
D11	BLUE	Y
D11	PRIN	Y
D12	BLDE	N
D13	COMO	N
D13	WLOU	Y
D14	SYCA	Y
D15	REDB	Y
D16	MART	N
D16	MCRA	N
D17	BIGC	Y
D17	TECR	Y
D18	OKSR	Y
D18	TOOK	N
D19	CARI	N

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 $\delta^{18}O$ values are reported on the VSMOW-SLAP scale, $\delta^{15}N$ values are reported relative to atmospheric N_2, and $\delta^{13}C$ 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 Sampling Design Changes

2014 - February 2017: In early 2017, NEON stopped shipping POM isotope samples to the Academy of Natural Sciences of Drexel University for analysis due to inaccurate SOPs and began shipping to the SIRFER Lab at the University of Utah. The field **laboratoryName** in the tables asi_POMExternalLabDataPerSample and asi_externalLabPOMSummaryData_pub indicate the name of the external facility at which a record was analyzed. The most recent record to contain *Academy of Natural Sciences of Drexel University* as the **laboratoryName** was the 2017-02-27, but the date of external facility change will vary across the observatory.

2014 - October 2017: At the request of the external facility, the SOP for processing and shipping POM filters from the domain support facilities was updated. The SOP changed from shipping POM filters frozen on ice using overnight to first drying POM filters overnight at 65 °C and shipping at ambient temperature. Laboratory data downloaded in the table asi_POMExternalLabDataPerSample that was collected on or before the 2017-10-18 will have been shipped frozen and on ice.

2014 - 2017: During the first 3 years of sampling, surface water isotope samples were collected at littoral locations in seepage lakes. Beginning in 2018, surface water isotope sampling was discontinued at seepage lake littoral locations

2014 - 2018: During the first 4 years of sampling, vials containing water isotope samples were filled to 90% volume. Beginning in 2019, vials containing water isotope samples were filled to 80% volume to minimize the amount of samples that broke due to freezing during transport.

August 2020: It was discovered that some Particulate Organic Matter (POM) isotope data may have been impacted by the presence of inorganic carbon. Scientists at NEON and the contracted external lab worked together to develop a method to remove inorganic carbon from samples via acidification. Beginning in August 2020, sites that were determined to contain high inorganic carbon are acidified prior to analysis. Users can identify acidified samples using the asi_POMExternalLabDataPerSample:**acidTreatment** field.

September - November 2021: The external laboratory responsible for H2O isotope analysis altered their analytical method and data return to measure and report a full propagation of measurement uncertainty rather than only the precision of replicate injections of the sample. The external laboratory stopped reporting the **d18OsdWater** and **d2HsdWater** fields in gsi_externalLabH2OIsotopes and

asi_externalLabH2OIsotopes and started reporting the **d18OWaterUncert** and **d2HWaterUncert** fields. The date of this change will vary at a given site across the observatory, but any record after 2021-11-11 should only have the full propagation of uncertainty published.

3.5 Laboratory Quality Assurance and Uncertainty

External laboratory facilities have been chosen for their use of stable isotope analytical methods widely adopted by the scientific community. Labs report the long-term analytical precison and accuracy of standard reference materials analyzed as unknowns for each analyte in a summary file. This allows users to interpret and model the stable isotope data in the context of its uncertainty range. Contracted external facilities upload a summary file (asi_externalLabSummaryData, asi_externalLabPOMSummaryData) when they begin work for NEON, then again once per year or whenever their information changes (for example, a new instrument is aquired or a change is detected in analytical precision). Additionally, NEON's Calibration/Validation department has regular procedures for auditing the quality assurance of external laboratories and their reports are available to data users.

In summary files, for every unique combination of **analyte** x **qaReferenceID** x **labSpecificStartDate**, users can view analytical accuracy (**analyteAccuracy**) and precision (**analyteStandardDeviation**). These metrics can be used to interpret and model external laboratory data by matching the **analysisDate** of a laboratory data record to the range between **labSpecificStartDate** and **labSpecificEndDate** in the summary tables by **analyte**.

Additionally, in the gsi_externalLabH2OIsotopes and asi_externalLabH2OIsotopes tables, uncertainties in hydrogen (**d2HWaterUncert**) and oxygen (**d18OWaterUncert**) stable isotope ratio are reported for each record. These values represent a full propagation of uncertainty associated with an analysis, including the precision of replicate injections of the sample, precision of analyses of primary laboratory reference materials in the same analytical batch, and the uncertainty in the 'known' calibrated values of the primary laboratory reference materials.

3.6 Variables Reported

All variables reported from the field technician or laboratory (L0 data) are listed in the files, NEON Raw Data Validation for Water chemistry, isotopes, dissolved gas, and microbes sampling, Level 0 (DP0.20090.001) (AD[04]), NEON Raw Data Validation for Stable isotope in surface waters and groundwater field data (DP0.20206.001) (AD[05]), NEON Raw Data Validation for Plant and algae external lab chemistry (DP0.20065.001) (AD[06]), and NEON Raw Data Validation for H20 isotopes external lab data (DP0.20205.001) (AD[07]). All variables reported in the published data (L1 data) are also provided separately in the files NEON Data Variables for Stable isotope in surface waters (DP1.20206.001) (AD[08]) and NEON Data Variables for Stable isotope in groundwater (DP1.20276.001) (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 Geoid12A

geoid model for its vertical reference surface. 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.

3.7 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:

parentSampleID (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 streams and rivers is indicated in the **namedLocation** field as 'ss', and stationIDs (**namedLocation**) 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 related to the NEON Aquatic Observation System sampling locations can be found on the NEON science webpage at https://www.neonscience.org/data-samples/data/spatial-data-maps. 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

3.8 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 up to 26 per wadeable stream and river, 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

3.9 Associated Data Streams

The super parent sample of the field data from the Stable isotopes in surface water and Stable isotopes in groundwater data products is shared with several other related data products. **siteID**, **namedLocation**

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

3.10 Product Instances

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

Surface and groundwater stable isotope sampling yields one unique sample per **namedLocation** x **collectDate**. There will be up to 26 sampling events per year in wadeable streams and rivers, 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, rivers 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 on a small percentage of surface water isotope samples. If replicate samples are taken, there will be one unique sample per **replicateNumber** x **namedLocation** x **collectDate**, and the sample ID(s) of the replicate sample(s) will have the **replicateNumber** appended to the end.

3.11 Data Relationships

The protocol dictates that each **siteID** x **namedLocation** combination is sampled at least once per event (one record expected per **parentSampleID** in asi_fieldSuperParent or gsi_fieldSuperParent). A record from asi_fieldSuperParent may have 0 to 3 child records in asi_fieldData, depending on whether a sample is collected and whether duplicate samples are collected. A record from gsi_fieldSuperParent may have 0 or 1 child record gsi_fieldData, depending on whether a water sample is 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 or any other table within the data product. Each record from asi_fieldData is expected to have 1 wide-format (1 record per multiple analytes) child record in asi_externalLabH2OIsotopes (table reporting deuterium and $\delta^{18}O$ isotope ratios in filtered surface water) and up to 4 long-format (1 record per analyte) child records in asi_POMExternalLabDataPerSample (table reporting $\delta^{15}N$ and $\delta^{13}C$ isotope ratios along with C and N masses from surface water POM filters). Each record from gsi_fieldData is expected to have 1 wide-format child record in gsi_externalLabH2OIsotopes (table reporting deuterium and $\delta^{18}O$ isotope ratios in filtered groundwater). However, 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*.

Stable isotope sample IDs and barcodes will be generated for each child sample at every sampling event. After shipment to external labs are complete, any physical sample that remains will be discarded.

Stable isotopes in surface waters:

asi_fieldSuperParent - > 1 record expected per **parentSampleID**

asi_fieldData - > 1-3 record(s) expected per **parentSampleID**. Triplicate samples are expected to be collected 3 times per year per **namedLocation**. Each record generates a single subsample for H2O isotope analyses, **isotopeH2OSampleID**, and up to 2 filters for POM isotope analyses, **isotopePOMSampleID** and **isotopePOMRep2SampleID**.

asi_externalLabH2OIsotopes - > 1 record expected per **isotopeH2OSampleID**, associated with external laboratory H2O isotope analyses.

asi_POMExternalLabDataPerSample - > 4 records expected per **sampleID**, associated with external POM isotope analyses.

asi_externalLabSummaryData - > 1 record expected per **analyte** x **method** x **laboratoryName** x **labSpecificStartDate** combination, used to associate sample data with relevant uncertainty values.

asi_externalLabPOMSummaryData - > 1 record expected per **analyte** x **method** x **laboratoryName** x **labSpecificStartDate** combination, used to associate sample data with relevant uncertainty values.

Stable isotopes in groundwater:

gsi_fieldSuperParent - > 1 record expected per **parentSampleID**

gsi_fieldData - > 1 record expected per **parentSampleID**. Generates a single subsample for water isotope analyses, **isotopeH2OSampleID**

gsi_externalLabH2OIsotopes - > 1 record expected per **isotopeH2OSampleID**, associated with external laboratory water isotope analyses

asi_externalLabSummaryData - > 1 record expected per **analyte** x **method** x **laboratoryName** x **labSpecificStartDate** combination, used to associate sample data with relevant uncertainty values.

Data downloaded from the NEON Data Portal are provided in separate data files for each site and month requested. The neonUtilities package in R and the neonutilities package in Python contain functions to merge these files across sites and months into a single file for each table. The neonUtilities R 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. The neonutilities package in Python is available on the Python Package Index (PyPi; https://pypi.org/project/neonutilities/) and can be installed using pip. For instructions on using the package in either language to merge NEON data files, see the Download and Explore NEON Data tutorial on the NEON website: https://www.neonscience.org/download-explore-neon-data.

3.12 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 (DP0.20090.001) (AD[04]), NEON Raw Data Validation for Stable isotope in surface waters and groundwater field data (DP0.20206.001) (AD[05]), NEON Raw Data Validation for Plant and algae external lab chemistry (DP0.20065.001) (AD[06]), and NEON Raw Data Validation for H20 isotopes external lab data (DP0.20205.001) (AD[07]), 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.

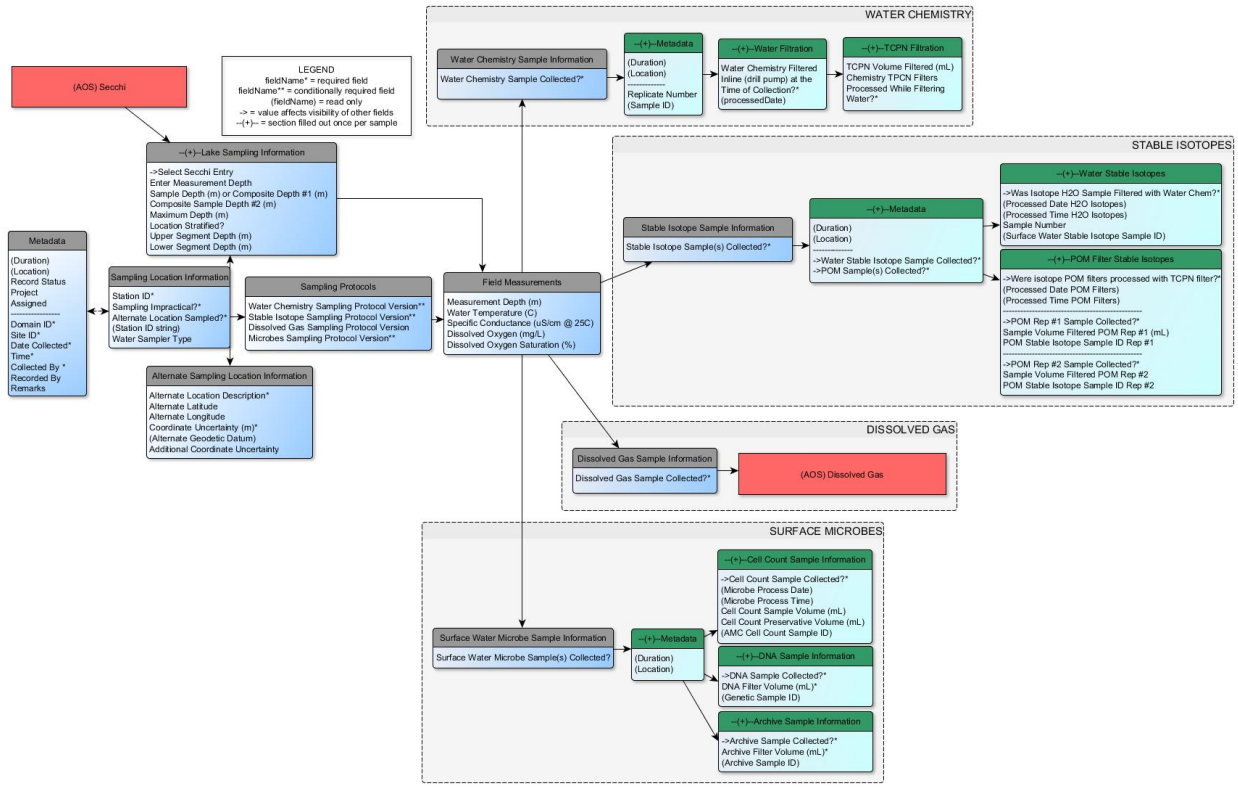

Figure 2: Schematic of the applications used by field technicians to enter aquatic stable isotope field data

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[15]) and function library (AD[14]).

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. Annually, NEON releases a static version of all or almost all data products, annotated with digital object identifiers (DOIs). The first data Release was made in 2021. During the provisional period, QA/QC is an active process, as opposed to a discrete activity performed once, and records are updated on a rolling basis as a result of scheduled tests or feedback from data users. The Issue Log section of the data product landing page contains a history of major known errors and revisions.

4.4 Quality Flagging

The **dataQF** field in each record is a quality flag for known issues applying to the record, added by NEON Science upon data review. Please see the table below for an explanation of **dataQF** code specific to this data product.

Table 2: Descriptions of the dataQF codes for quality flagging in field data

fieldName	value	definition
dataQF	legacyData	Data recorded using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow

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

4.5 Analytical Facility Data Quality

Analytical labs that generate 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 asi_externalLabPOMSummaryData_pub, 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

Table 3: Descriptions of the analytical facility codes for quality flagging H2O samples

fieldName	value	definition
externalLabDataQF	legacyData	Data recorded using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow
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

asi_POMExternalLabDataPerSample

Table 4: Descriptions of the analytical facility codes codes for quality flagging POM samples

fieldName	value	definition
externalLabDataQF	legacyData	Data recorded using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow
externalLabDataQF	Did not meet quality audit requirements for analysis audit	The external lab did not meet the requirements of the NEON external facility audit for the year the data were generated
externalLabDataQF	acidTreatmentSOPNotFollowed	The external lab did not follow the standard operating procedure that indicated samples were decarbonated via acid fumigation