NEON USER GUIDE TO AQUATIC PLANT BRYOPHYTE CHEMICAL PROPERTIES (DP1.20063.001)

PREPARED BY	ORGANIZATION
Tanya Chesney	DPS
Stephanie Parker	AOS

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	07/dd/2017	Initial Release
B	05/25/2020	Included general statement about usage of neonUtilities R package and statement about possible location changes. Updated taxonomy information. Included temporal optimization and opportunistic data return for stable isotopes. Updated spatial figure to indicate changes to lake littoral and inflow/outflow sensor locations.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

This document provides an overview of the data included in this NEON Level 1 data product, the quality controlled product generated from raw Level 0 data, and associated metadata. In the NEON data products framework, the raw data collected in the field, for example, the fresh mass of the sample, 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 Aquatic plant bryophyte chemical properties- carbon and nitrogen concentrations of aquatic plant and bryophytes from benthic collections in lakes, rivers, and wadeable streams and associated metadata from input data. 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 Publication Workbook for AOS Aquatic Plants and Bryophytes: QA/QC of Chemistry Data (DP1.20063.001) (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 AOS Protocol and Procedure: Aquatic Plant, Bryophyte, Lichen, and Macroalgae Sampling (AD[07]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Ingest Workbook for AOS Aquatic Plant, Bryophyte, and Macroalgae Clip Harvest, Level 0 (DP0.20066.001) (AD[03]) and NEON Raw Data Ingest Workbook for AOS Aquatic Plant and Algae External Lab Chemistry Data (DP0.20065.001) (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.002652	NEON Data Products Catalog
AD[03]	Available with data download	Validation csv
AD[04]	Available with data download	Validation csv
AD[05]	Available with data download	Variables csv
AD[06]	NEON.DOC.001152	NEON Aquatic Sampling Strategy
AD[07]	NEON.DOC.003039	AOS Protocol and Procedure: Aquatic Plant, Bryophyte, Lichen, and Macroalgae Sampling
AD[08]	NEON.DOC.000008	NEON Acronym List
AD[09]	NEON.DOC.000243	NEON Glossary of Terms
AD[10]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[11]	Nicl Language.pdf	NEON Ingest Conversion Language Function Library
AD[12]	Available on NEON data portal	NEON Ingest Conversion Language
AD[13]	Available with data download	Categorical Codes csv

2.2 Acronyms

Acronym	Definition
SCUBA	Self-contained underwater breathing apparatus
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey

3 DATA PRODUCT DESCRIPTION

Aquatic plant bryophyte chemical properties data include total carbon and nitrogen of dried plant samples collected as part of Aquatic Plant, Bryophyte, and Macroalgae Clip Harvest (DP1.20066.001). Aquatic plants and bryophytes are sampled via clip harvest once per year at each NEON aquatic site (AD[06]). Sampling dates are based on a combination of variables, including hydrology in streams or ice on/ice off dates in lakes, accumulated degree days (temperature), and riparian greenness (phenology). For additional information see sampling design NEON Aquatic Sampling Strategy (AD[06]) and protocol AOS Protocol and Procedure: Aquatic Plant, Bryophyte, Lichen, and Macroalgae Sampling (AD[07]).

3.1 Spatial Sampling Design

Clip harvest collections are made at 10 locations per site. In lakes and non-wadeable streams (rivers), these locations are selected by randomized GPS points (Figure 1) within the zone of plant colonization at the site (i.e., within the depth ranges that plants are known to colonize based on light attenuation in the lake or river).

At wadeable stream sites, each clip harvest (Figure 2) is collected within 4 m of the nearest point transect (DP1.20072.001) at a total of 10 locations: 5 in the dominant habitat type, 5 in the second-most dominant habitat type. Transects are established in alternating habitat types along the reach (e.g., transect 1 in a pool, transect 2 in a riffle, transect 3 in a pool). If it is not possible to separate each transect by a different habitat type, transects may be spaced at least 10m apart within the same habitat type/unit. Transects are established during the first sampling bout by domain technicians and coordinates are collected using a high-precision GPS unit (e.g., Trimble with a tornado antenna). Plot markers are established on one bank, with the transect extending perpendicular to stream flow across the wetted channel. Transects remain in place until significant morphological changes occur in the stream (e.g., bed-moving spate) that causes the transect to be unusable. After such geomorphological changes, a transect may be re-established in a new location. See AOS Protocol and Procedure: Aquatic Plant, Bryophyte, Lichen, and Macroalgae Sampling (AD[07]) for additional details on strategy and SOPs.

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/

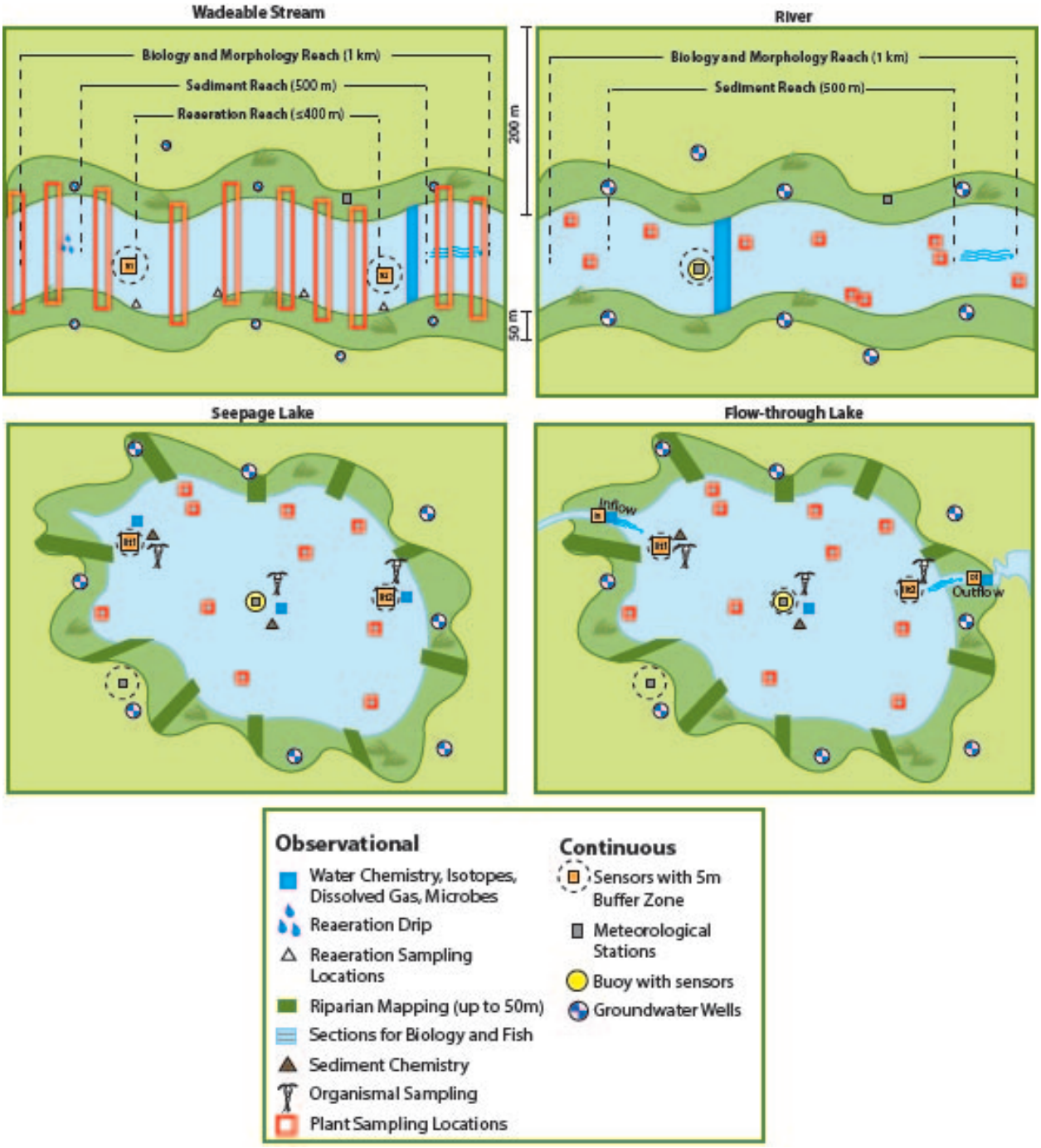

Figure 1: Aquatic plant sampling locations in all types of aquatic sites (wadeable streams, rivers, and lakes).

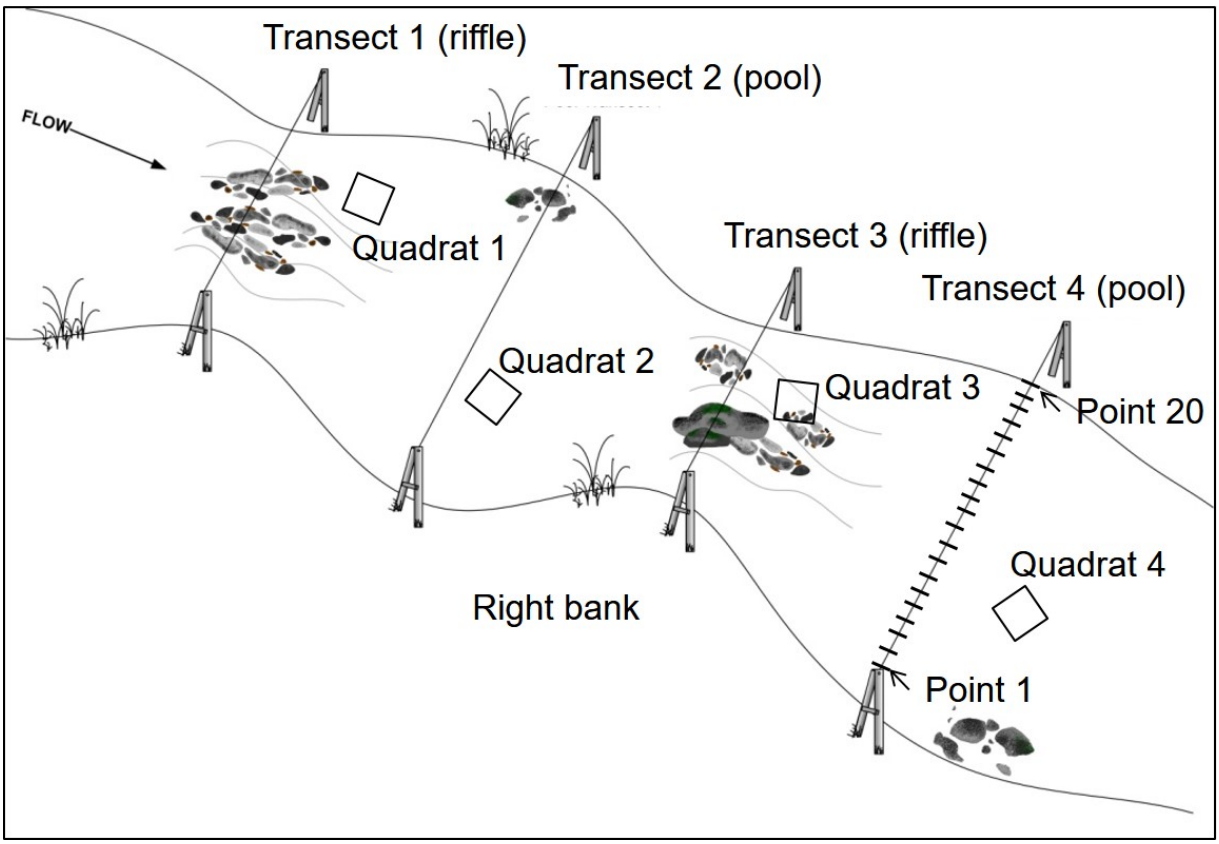

Figure 2: Aquatic plant sampling locations in wadeable streams, showing quadrat locations in proximity to established transects

3.2 Temporal Sampling Design

Aquatic plant clip harvest samples and chemistry subsamples are collected during biological and sediment chemistry sampling, in bout 2 (mid-summer). Data collection for this data product occurs occurs within one day per bout at a given site. See NEON Aquatic Sampling Strategy (AD[06]), AOS Protocol and Procedure: Aquatic Plant, Bryophyte, Lichen, and Macroalgae Sampling (AD[07]) for additional details.

3.3 Sampling Design Changes

Prior to January 1, 2017, data were collected 3 times per year during all biological and sediment chemistry sampling bouts (bouts 1, 2, and 3). Sampling was stopped during bouts 1 and 3 to avoid overharvesting of aquatic vegetation.

Starting in 2018, stable isotope 13C and 15N data were returned by the external facility because stable isotopes are part of the workflow for delivering total C and N data. Stable isotope data are opportunistic in this data product, and are not a requirement.

3.4 Laboratory Quality Assurance and Uncertainty

External laboratory facilities have been chosen for their use of analytical methods widely adopted by the scientific community. Labs report the long-term analytical precision and uncertainty of standard reference materials analyzed as unknowns for each analyte in a summary file. This allows users to interpret and model the aquatic plant bryophyte chemical properties data in the context of its uncertainty range. Contracted external facilities upload a summary file (asi_externalLabSummaryData) when they begin work for NEON, then again once per year or whenever their information changes (for example, a new instrument is acquired 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.

3.5 Variables Reported

All variables reported from the field technician or laboratory (L0 data) are listed in the files, NEON Raw Data Ingest Workbook for AOS Aquatic Plant, Bryophyte, and Macroalgae Clip Harvest, Level 0 (DP0.20066.001) (AD[03]) and NEON Raw Data Ingest Workbook for AOS Aquatic Plant and Algae External Lab Chemistry Data (DP0.20065.001) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Publication Workbook for AOS Aquatic Plants and Bryophytes: QA/QC of Chemistry Data (DP1.20063.001) (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 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.

3.6 Spatial Resolution and Extent

The finest resolution at which spatial data are reported is near a single transect (wadeable streams) or at a randomized point (lakes and rivers) within a site. For example, data may be collected at a specific depth in the water column of a lake.

Lake and river spatial hierarchy = locationID (ID of sampling point where collection occurred) -> siteID (ID of NEON site) -> domainID (ID of a NEON domain)

Wadeable streams spatial hierarchy = locationID (ID of sampling quadrat near a named transect within site) -> siteID (ID of NEON site) -> domainID (ID of a NEON domain)

The basic spatial data included in the data downloaded include the latitude, longitude, and elevation of the named location at the aquatic site (e.g., the permanent transect location in a wadeable stream) or the

latitude and longitude of the randomized point in a lake or river.

3.7 Temporal Resolution and Extent

The finest resolution at which temporal data are reported is at **collectDate**, the date and time of day when the samples were collected in the field.

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

3.8 Associated Data Streams

This data product is dependent on the field data collected in Aquatic Plant Bryophyte Macroalgae Clip Harvest (DP1.20066.001). Data can be linked to the parent sample through the **fieldID** and **sampleID** fields.

All of the above data products are also loosely related to gauge height data and associated metadata collected on the same sampling day (DP1.20267.001). These data products are linked through the **siteID** and **collectDate**.

Data and maps produced by the Bathymetric and morphological maps data product (DP4.00132.001) may also be relevant to aquatic plant and bryophyte chemistry data.

3.9 Product Instances

At each aquatic site, there will be a variable number of samples collected per year, depending on plant cover at the site. Aquatic plant chemistry samples are collected once per year at each of 10 sampling locations for each taxon collected at those locations. Because data are reported in long format (as opposed to wide), each sample generates records for each analyte measured, 2 analytes per sample (carbon and nitrogen). Starting in 2018, additional chemical analyses of stable isotopes may be included in the download package, but are not a requirement of the data product.

3.10 Data Relationships

A record in apl_domainLabChemistry or apl_algaeExternalLabDataPerSample must have a corresponding record in apl_clipHarvest describing field collection conditions, location, and metadata during sample collection. If **fieldID** is empty in apl_clipHarvest, there will be no additional records in the apl_domainLabChemistry or apl_algaeExternalLabDataPerSample tables. 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.

apl_clipHarvest.csv - > One record is created for each clip harvest observation. If a sample is collected, a **fieldID** is created in this table and linked in subsequent data tables (biomass.csv).

apl_domainLabChemistry.csv - > One or more records for **chemSubsampleID** is created for each **sampleID** processed at the NEON domain lab. A **sampleID** is created for each taxon per location.

apl_algaeExternalLabDataPerSample.csv - > One record is created for each analytical replicate of each analyte for a sample, resulting in multiple entries per sample. Data can be tracked to the fieldData through the parent **sampleID** or the domainLabChemistry data through **sampleID** plus **analyte** plus **replicate**.

apl_plantExternalLabQA.csv - > The data quality table returns data for the blanks and standards used in the lab analysis. This table is filtered by the lab doing the work, and may contain data for analytes used in other data products in addition to the analytes returned in this data product. The field **batchID** is used to connect this table to apl_algaeExternalLabDataPerSample.

asi_externalLabPOMSummaryData.csv - > Summary information for each analytical method are recorded in this table, with **startDate** and **endDate**. These dates can be used to apply to the data in alg_algaeExternalLabDataPerSample using the fields **laboratoryName**, **analysisDate**, and **analyte**.

Data downloaded from the NEON Data Portal are provided in separate data files for each site and month requested. The neonUtilities R package contains functions to merge these files across sites and months into a single file for each table described above. The neonUtilities package is available from the Comprehensive R Archive Network (CRAN; https://cran.r-project.org/web/packages/neonUtilities/index.html) and can be installed using the install.packages() function in R. For instructions on using neonUtilities to merge NEON data files, see the Download and Explore NEON Data tutorial on the NEON website: https://www.neonscience.org/download-explore-neon-data

4 TAXONOMY

NEON manages taxonomic entries by maintaining a master taxonomy list based on the community standard, if one exists. Through the master taxonomy list, synonyms submitted in the data are converted to the appropriate name in use by the standard. The master taxonomy for plants is the USDA PLANTS Database (USDA, NRCS. 2014. https://plants.usda.gov). Taxon ID codes used to identify taxonomic concepts in the NEON master taxonomy list are alpha-numeric codes, 4-6 characters in length based on the accepted scientific name. Each code is composed of the first two letters of the genus, followed by the first two letters of the species and first letter of the terminal infraspecific name (if applicable) then, if needed, a tiebreaking number to address duplicate codes. Genus and family symbols are the first five (genus) or six (family) letters of the name, plus tiebreaking number (if needed). Symbols were first used in the Soil Conservation Service's National List of Scientific Plant Names (NLSPN) and have been perpetuated in the PLANTS system. The portions of the PLANTS Database included in the NEON plant master taxonomy list includes native and naturalized plants present in NEON observatory sampling area including the Lower 48 U.S. States, Alaska, Hawaii, and Puerto Rico. NEON plans to keep the taxonomy updated in accordance with USDA PLANTS Database starting in 2020 and annually thereafter.

The master taxonomy list includes geographic range and nativity as described by the USDA PLANTS Database. A list for each NEON domain includes those species with ranges that overlap the domain as well as nativity designations - introduced or native - in that part of the range. Errors are generated if a species is reported at a location outside of its known range. If the record proves to be a reliable report, the master taxonomy table is updated to reflect the distribution change.

The full master taxonomy lists are available on the NEON Data Portal for browsing and download: http://data.neonscience.org/static/taxon.html.

5 DATA QUALITY

5.1 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. The data entry workflow for collecting aquatic plant clip harvest field data is diagrammed in Figure 3, and the lab workflow is diagrammed in Figure 4.

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 Ingest Workbook for AOS Aquatic Plant, Bryophyte, and Macroalgae Clip Harvest, Level 0 (DP0.20066.001), 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. 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[11]).

5.2 Automated Data Processing Steps

Following data entry into a mobile application of 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[10]).

5.3 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.

5.4 Quality Flagging

The **dataQF** field in each data record is a quality flag for known errors applying to the record. Please see the table below for an explanation of **dataQF** codes specific to this product.

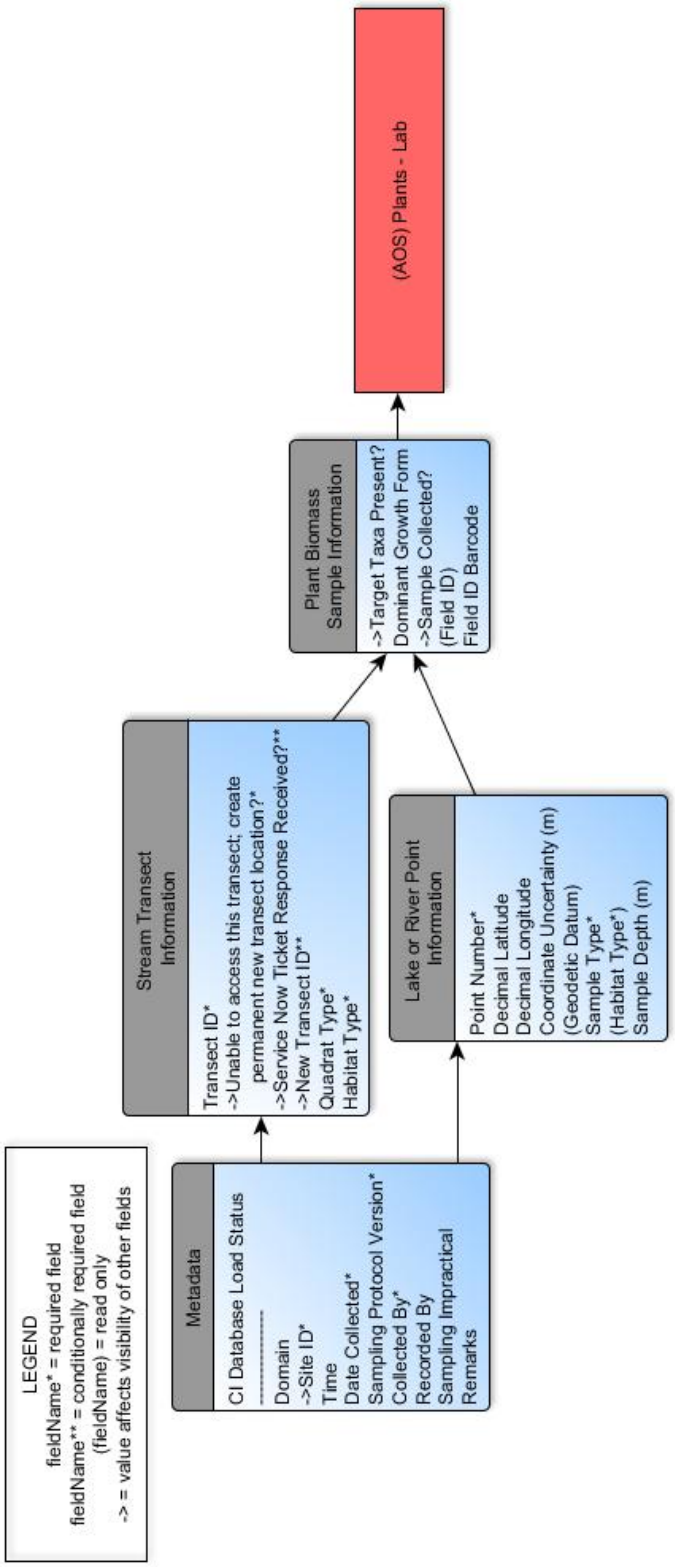

Figure 3: Schematic of the applications used by field technicians to enter clip harvest field data

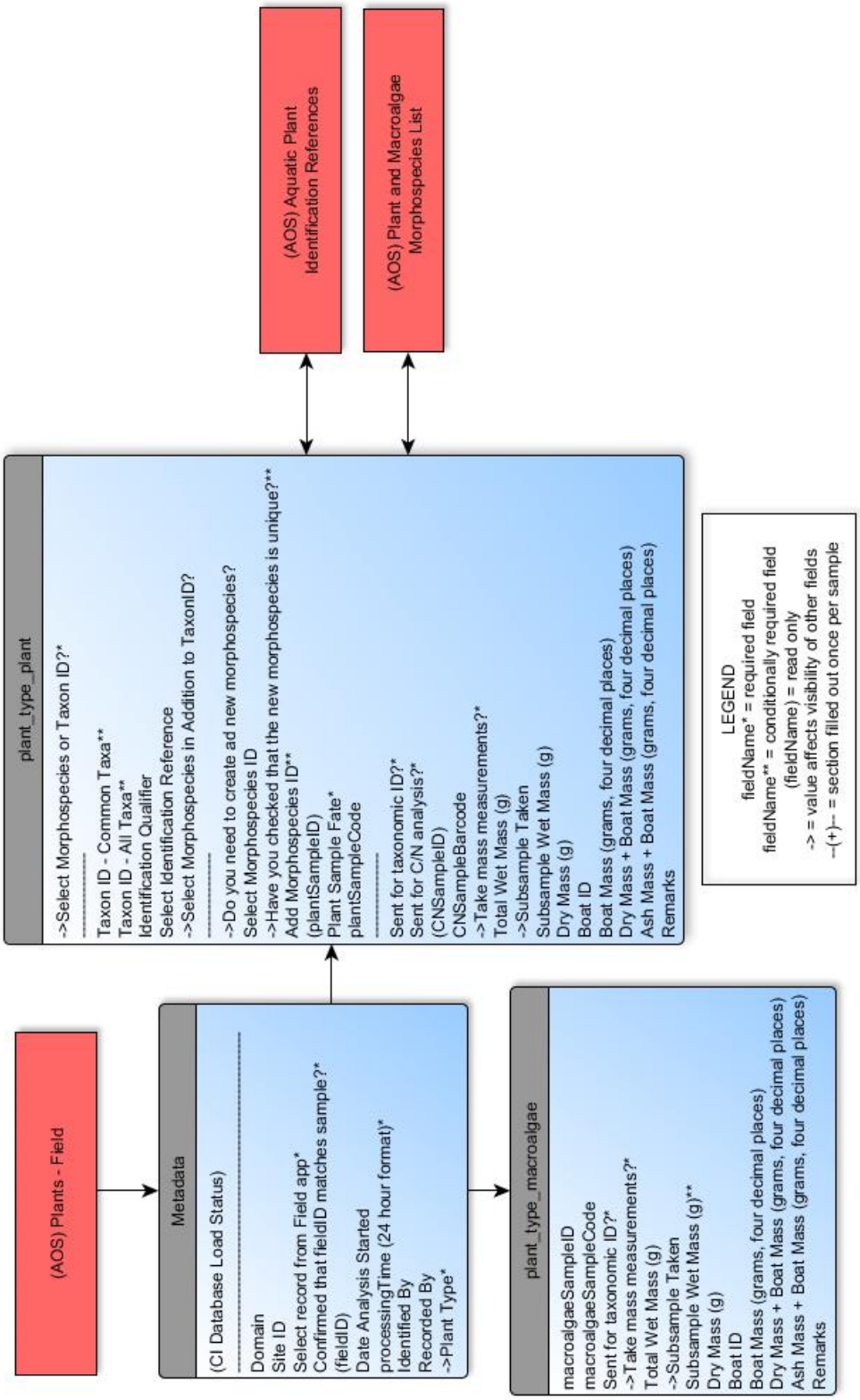

Figure 4: Schematic of the applications used by field technicians to measure aquatic plant biomass data and subsample in the domain lab

Table 1: Descriptions of the dataQF codes for quality flagging

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
dataQF	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

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)

5.5 Analytical Facility Data Quality

Data analyses conducted on plant chemistry data conform to the current data quality standards used by practitioners. Secondary standards or reference material are analyzed in every batch of NEON data. Lab quality data are presented in the table "asi_externalLabPOMSummaryData" that is included in this download package.

6 REFERENCES

Berg, M. S. 2009. Warm water point intercept aquatic macrophyte and fall EWM (Myriophyllum spicatum) bed mapping surveys, Ham Lake (WBIC: 2467700), Burnett County, Wisconsin. St. Croix Falls, Wisconsin.

Bowden, W.B., J.M. Glime, and T. Riis. 2006. Macrophytes and Bryophytes. Pages 381-414 in F.R. Hauer and G.A. Lamberti, editors. Methods in Stream Ecology, Second Edition. Elsevier, Inc., Boston, Massachusetts, USA.

Downing, J. A. and M. R. Anderson. 1985. Estimating the standing biomass of aquatic macrophytes. Canadian Journal of Fisheries and Aquatic Sciences. 42: 1860-1869.

Neuman, J. 2008. Minnesota National Lakes Assessment Project: Aquatic Macrophyte in Minnesota Lakes.

USDA, NRCS. 2014. The PLANTS Database (http://plants.usda.gov, 25 August 2014). National Plant Data Team, Greensboro, NC 27401-4901 USA.

Wisconsin DNR, 2008. 2007 National Lake Survey - Wisconsin Results. http://dnr.wi.gov/lakes/nls/plants. Accessed 27 September 2011.

Yin, Y., J. S. Winkelman, and H. A. Langrehr. 2000. Long term resource monitoring program procedures: Aquatic vegetation monitoring. U.S. Geological Survey, Upper Midwest Environmental Sciences Center, La

Crosse, Wisconsin. LTRMP 95-P002-7. 8 pp. + Appendices A-C.