NEON USER GUIDE TO RIPARIAN COMPOSITION AND STRUCTURE (DP1.20275.001)

PREPARED BY	ORGANIZATION
Dylan Monahan	AQU
Caren Scott	AQU
Sarah Elmendorf	DPS
Kaelin M. Cawley	AQU
Stephanie Parker	AQU

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	12/11/2017	Initial Release
B	12/23/2020	Included general statement about usage of neonUtilities R package and statement about possible location changes. Updated taxonomy information. Minor updates were made to Spatial Sampling Design and Spatial Resolution and Extent sections. Riparian plot size was reduced in 2020 from 20X50 meters to 10x10 meters at wadeable stream sites, and 10X15 meters at river and lake sites

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 dry weights of litter functional groups 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 Riparian Composition and Structure - 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 Variables for Riparian composition and structure (DP1.20275.001) (AD[04]), 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: Riparian Habitat Assessment (AD[06]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Validation for riparian composition and cover (DP0.20275.001) (AD[03]), 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	Variables csv
AD[05]	NEON.DOC.001152	Aquatic Sample Strategy
AD[06]	NEON.DOC.003826	AOS Protocol and Procedure: Riparian Habitat Assessment
AD[07]	NEON.DOC.000008	NEON Acronym List
AD[08]	NEON.DOC.000243	NEON Glossary of Terms
AD[09]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[10]	Available on NEON data portal	NEON Ingest Conversion Language Function Library
AD[11]	NEON.DOC.004842	AOS Commissioning Test Report: Riparian Habitat Assessment Process Quality
AD[12]	NEON.DOC.004848	AOS Commissioning Test Report: Riparian Habitat Assessment Data Quality
AD[13]	Available on NEON data portal	NEON Ingest Conversion Language
AD[14]	Available with data download	Categorical Codes csv

2.2 Acronyms

Acronym	Definition
EMAP	Environmental Monitoring and Assessment Program
EPA	U.S. Environmental Protection Agency
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index

3   DATA PRODUCT DESCRIPTION

Riparian composition and structure products provide an estimate of the riparian vegetation, human impacts, dominant plant species, and bank characteristics, which buffer the banks of NEON Aquatic lakes, rivers, and streams. The protocol is derived from methods developed by the U.S. Environmental Protection Agency (2008), U.S. Fish and Wildlife Service (1979, 1997, 2009) and U.S. Department of Agriculture (1959).

3.1   Spatial Sampling Design

The Riparian habitat assessment protocol is executed at NEON aquatic sites. At wadeable streams, 10 riparian transects are evenly distributed throughout the 1 km biological sampling reach, on each bank of the transect, bisected by the transect in the riparian area of both banks is a 10x10 meter riparian plots (Figure 1). At lakes, 10 evenly spaced 10X15 meter plots are established around the lake perimeter (Figure 2). Large river sites include 5 right bank and 5 left bank 10X15 meter plots evenly spaced throughout the 1 km reach (Figure 3). The riparian area surveyed at each transect extends up to 10 meters at wadeable stream and 15 meters at rivers/lakes from the shoreline/bank towards the terrestrial systems (less if the plot is visibly obstructed) and 10 meters (5 meters each side of central point) along the shoreline.

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/
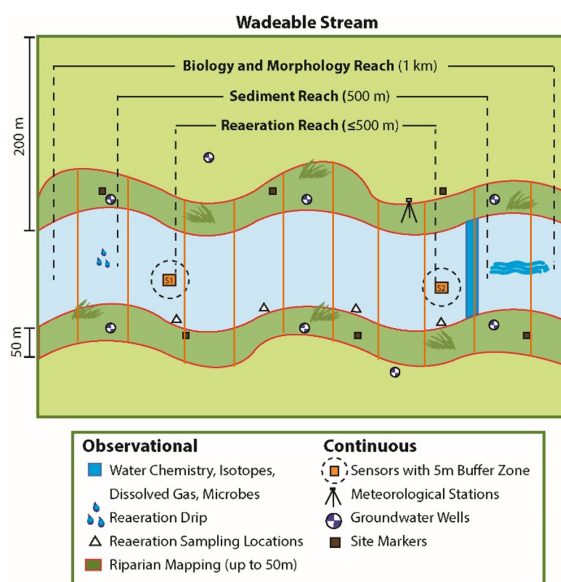

Figure 1: Riparian habitat assessment locations for wadeable streams
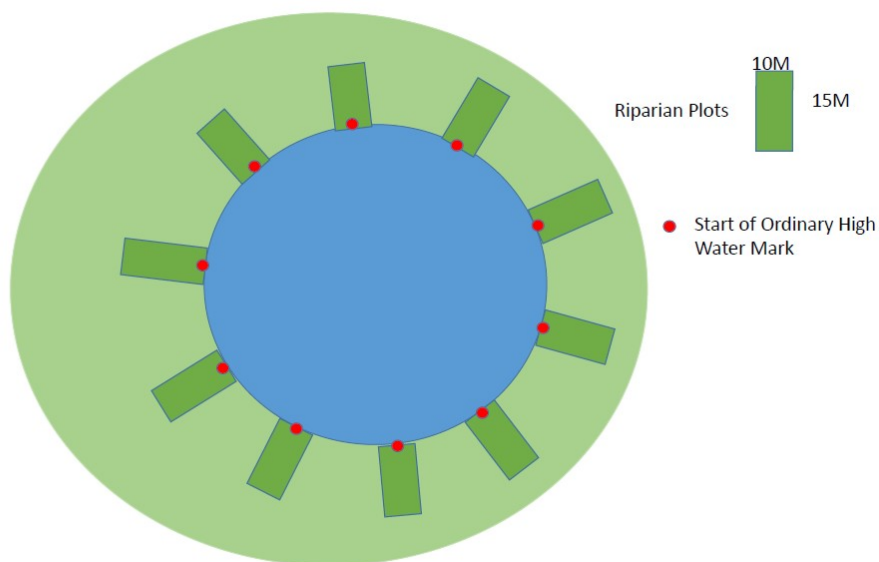
Figure 2: Riparian habitat assessment locations for lakes
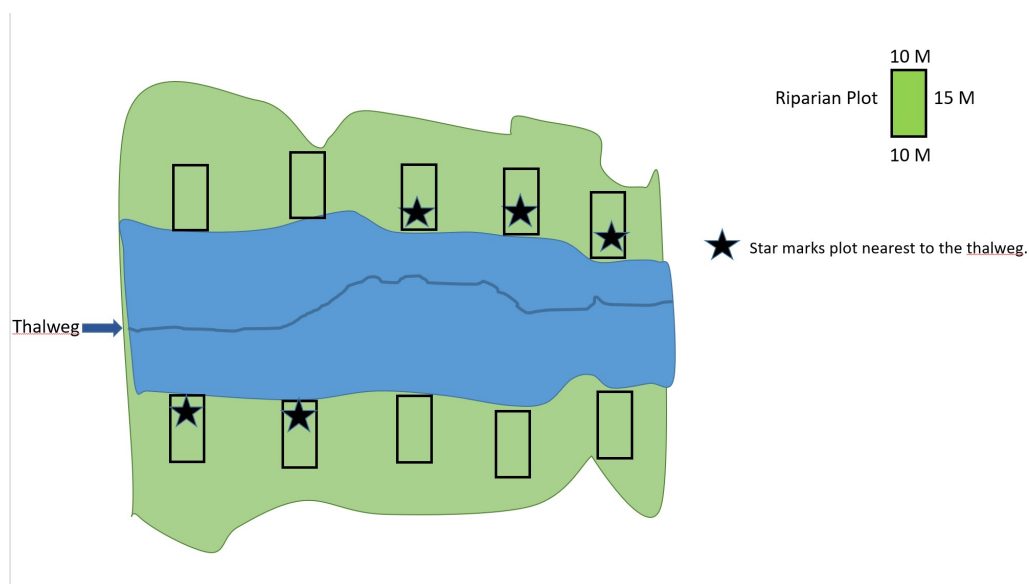

Figure 3: Riparian habitat assessment locations for rivers

3.2 Temporal Sampling Design

Riparian habitat assessment occurs once per year, during periods of peak-greenness as determined by the range of historical dates when MODIS NDVI greenness is within 90% of site maximum.

3.3   Sampling Design Changes

Riparian plot size was reduced in 2020 from 20X50 meters to 10x10 meters at wadeable stream sites, and 10X15 meters at river and lake sites.

3.4   Variables Reported

All variables reported from the field or laboratory technician (L0 data) are listed in the file, NEON Raw Data Validation for riparian composition and cover (DP0.20275.001) (AD[03]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Riparian composition and structure (DP1.20275.001) (AD[04]).

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 basic spatial data included in the data downloaded include the latitude, longitude, and elevation of a fixed riparian location (based on type of aquatic site, see below) where sampling occurred + associated uncertainty due to GPS error. Each observation includes looking at an area approximately 10 meters wide (along the bank) and 10 meters deep at streams and 15 meters deep at lakes and rivers (perpendicular to the bank, into the riparian zone).

Wadeable stream: The 1 kilometer permitted wadeable stream reach is divided into 10 equally spaced riparian transects, with a 10X10 meter riparian plot on each bank, bisected by the transect. Riparian observations occur at each riparian plot, resulting in 20 observations per stream per year. The riparian plot starts at the bank and runs 10 meters back towards the upland.

Rivers: The 1 kilometer permitted river reach is divided into 10 equally spaced 10X15 meter riparian plots, 5 along each bank. Riparian observations occur at each point, resulting in 10 observations per river per year. The basic spatial data refers a point within the plot (which is situated 10 meters from the bank).

Lakes: The shoreline of the lake is divided into 10 equally spaced 10x15 meter riparian plots. Riparian observations occur at each plot, resulting in 10 observations per lake per year. The basic spatial data refers to the plot (which is starts at the ordinary high water line, and run 15 meters back into the uplands. The basic spatial data refers to the point within the plot (which is situated 10 meters from the bank), and the uncertainty includes the width of the observation, and the GPS error.

This results in a spatial hierarchy for streams of:

bankLR (left or right bank) →**namedLocation** (unique ID given to the individual transect) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain).

This results in a spatial hierarchy for lakes and rivers of:

namedLocation (unique ID given to the individual transect) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain).

3.6   Temporal Resolution and Extent

The finest resolution at which temporal data are reported is the **startDate**, the date and time of day at which the assessment of a particular transect began.

3.7   Associated Data Streams

namedLocation and **startDate** are linking variables that tie specific observations and associated metadata to the Riparian Vegetation % Cover data product (DP1.20191.001).

3.8   Product Instances

Each of 10 riparian transects per stream site is expected to be surveyed once per year, one observation per bank, yielding 20 data product instances per stream site per calendar year.

Each of 10 riparian points per river or lake site is expected to be surveyed once per year, yielding 10 data product instances per stream site per calendar year.

3.9   Data Relationships

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. 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 riparian composition and cover (DP0.20275.001), 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[10]).

A schematic of the data entry application design is depicted in Figure 4.

5.2 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[09]).

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.
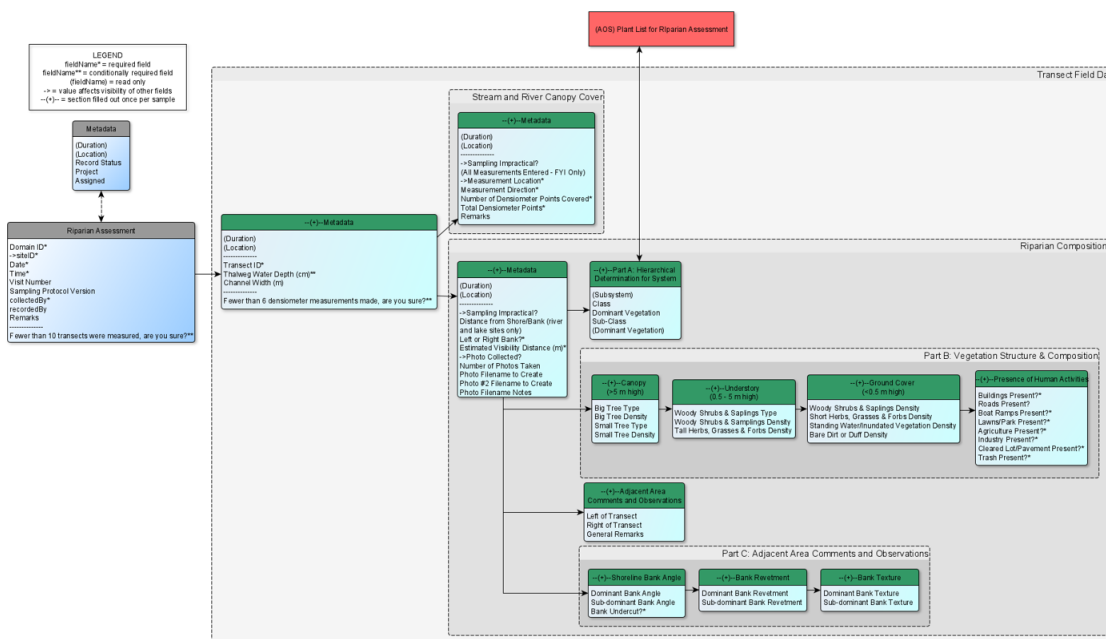
Figure 4: Schematic of the applications used by field technicians to enter riparian field data

5.4   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 this time, there are no known issues applying to the records in this data product.

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)

6   REFERENCES

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

U.S. Department of Agriculture. Strickler, G.S. 1959. Use of the densiometer to estimate density of forest canopy on permanent sample plots, Research Note Number 180, .

U.S. Environmental Protection Agency. 2008. Development for a Statewide Wetland and Riparian Mapping, Assessment and Monitoring Program 2009-2015. 29 pp.

U.S. Fish and Wildlife Service. 1997. A system for mapping riparian areas in the western United States. U. S. Fish and Wildlife Service. Washington, DC. 15 pp.

U.S. Fish and Wildlife Service. 2009. A system for mapping riparian areas in the western United States. U. S. Fish and Wildlife Service. Washington, DC. 7 pp.

U.S. Fish and Wildlife Service. Cowardin, L.M., V. Carter, F.C. Golet, and E.T. LaRoe. 1979. Classification of Wetlands and Deepwater Habitats of the United States. Washington, DC.FWS/OBS 79/31. 103 pp.