# NEON USER GUIDE TO BATHYMETRIC AND MORPHOLOGICAL MAPS (DP4.00132.001)

| **PREPARED BY** | **ORGANIZATION** |
| --- | --- |
| Brandon Jensen | AQU |
| Caren Scott | AQU |
| Stephanie Parker | AQU |
| Kaelin M. Cawley | AQU |
| Zachary L. Nickerson | AQU |

# CHANGE RECORD

| REVISION | DATE | DESCRIPTION OF CHANGE |
| --- | --- | --- |
| A | 07/17/2018 | Addition of Data Variables Information |
| B | 05/25/2020 | Included general statement about usage of neonUtilities R package and statement about possible location changes. Updated taxonomy information. |
| B.1 | 04/20/2021 | Updated information on the data tables published in this data product that are not involved with the interannual geomorphology surveys (AIS site survey data). Removed information on the data tables that are no longer published in this data product, but are now published in the Stage-discharge rating curves (DP4.00133.001) and Continuous discharge (DP4.00130.001) data products. |
| C | 03/02/2022 | Updated information on data package contents for suveys produced with the Biosonics MX echosounder rather than the Humminbird models. Updated section 4.3 Data Revision with latest information regarding data release. |
| E | 01/19/2026 | Updated information in the Special Considerations section on interpreting dates and additional details about GPS processing in the Uncertainty section. |

# 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 4 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 unprocessed sonar timeseries data (Level 0). Raw data that have been quality checked via the steps detailed herein, simple metrics that emerge from the raw data, and procesed maps and shapefiles are considered Level 4 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 L4 data creation.

## 1.2 Scope

This document briefly describes the steps needed to generate the L4 data product Bathymetric and Morphological Maps - and associated metadata from input data, for a more detailed description refer to NEON Standard Operating Procedure: Post-Processing of Bathymetric and Side Scan Sonar Data from NEON Lakes and Non-Wadeable Streams, NEON.DOC.004856 (AD[09]). 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 Bathymetric and Morphological Maps (DP4.00132.001) (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: Bathymetry and Morphology of Lakes and Non-Wadeable Streams (AD[07]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Validation for Bathymetric and Morphological Maps (DP0.00132.001) (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 as the corresponding L4 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 | Validation csv |
| AD[05] | Available with data download | Variables csv |
| AD[06] | NEON.DOC.001152 | NEON Aquatic Sampling Strategy |
| AD[07] | NEON.DOC.001197 | AOS Protocol and Procedure: Bathymetry and Morphology of Lakes and Non-Wadeable Streams |
| AD[08] | NEON.DOC.004825 | NEON Algorithm Theoretical Basis Document: OS Generic Transitions |
| AD[09] | NEON.DOC.004856 | NEON Standard Operating Procedure: Post-Processing of Bathymetric and Side Scan Sonar Data from NEON Lakes and Non-Wadeable Streams |
| AD[10] | NEON.DOC.000008 | NEON Acronym List |
| AD[11] | NEON.DOC.000243 | NEON Glossary of Terms |
| AD[12] | Available on NEON data portal | NEON Ingest Conversion Language Function Library |
| AD[13] | Available on NEON data portal | NEON Ingest Conversion Language |
| AD[14] | Available with data download | Categorical Codes csv |
| AD[15] | NEON.DOC.005424 | Algorithm Theoretical Basis Document: OS Data Quality Control |

## 2.2 Acronyms

| Acronym | Definition |
| --- | --- |
| DGPS | Differential Global Positioning System |
| EBK | Empirical Bayesian Kriging |
| GIS | Geographic Information System |
| GPS | Global Positioning System |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| NDVI | Normalized Difference Vegetation Index |
| NICL | NEON's Ingest Conversion Language |
| m | Meter |
| km | Kilometer |
| WAAS | Wide Area Augmentation System |
| WebUI | Web user Interface |

# 3 DATA PRODUCT DESCRIPTION

The Bathymetric and Morphological Map data product (DP4.00132.001) provides depths of lakes and non-wadeable streams (rivers), sediment and substrate characteristics, as well as the presence or absence, approximate abundance, and distribution of underwater flora. All data following processing are reported at the spatial resolution of a single latitude (x), longitude (y), elevation (depth; z), and date/time of a single collection event.

Maps produced from surveys collected prior to 2020: High accuracy depth (bathymetric) maps are obtained using a suite of hydroacoustic instrumentation interfaced with a differential global positioning system (DGPS) mounted on a vessel. The Wide Area Augmentation System (WAAS) is a form of DGPS that provides enhanced position accuracy (<3 m). Acoustic images of biota, substrate, and benthic structures are obtained with sonar which are georeferenced and mosaic in post-processing. The resulting master mosaic imagery coupled with ground truth observations are used to classify the substrate type and benthic habitat features. These may include sand, silt, clay, boulder, bedrock, and large woody debris. Summary data include volume and area calculations at specific bathymetric intervals or whole body of water estimates as well as areal calculations of habitat features.

Maps produced from surveys collected beginning in 2020: High accuracy depth (bathymetric) maps are obtained using a single-beam, down looking sonar unit and on-board GPS unit. The sonar data was processed alongside ground truth observations to create bathymetric maps and classify the substrate type and benthic habitat features. These may include sand, silt, clay, boulder, bedrock, and large woody debris. Summary data include volume and area calculations at specific bathymetric intervals or whole body of water estimates as well as areal calculations of habitat features.

## 3.1 Spatial Sampling Design

Bathymetric and morphological surveys are executed at NEON lakes and non-wadeable streams. A vessel-mounted acoustic sonar system and GPS is driven along the shoreline as well as across lake sites in a gridded pattern (Figure 1). Bathymetric surveys cover the entire lake with transects spaced up to 35 m apart. At non-wadeable stream sites, the vessel is driven in a series of tracks parallel to shore and in zig-zag patterns (Figure 1) also up to 35 m apart, and covers the entire permitted reach of approximately 1 km.

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/

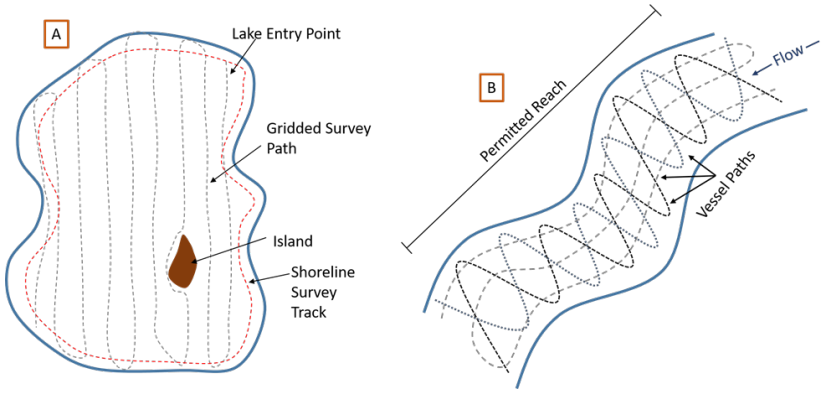


Figure 1: Schematic of survey pattern undertaken in the vessel for (a) lake grid survey, with the shoreline (in red), and island as well as (b) non-wadeable stream survey pattern.

## 3.2 Temporal Sampling Design

Bathymetric surveys at lakes and non-wadeable streams are completed at a minimum of every 5 years during biological sampling bout 2 and during the period of peak greenness, defined as the range of dates where Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) is within 90% of the local seasonal maximum (AD[02]). If an extreme event results in substantial physical change to the morphology or submerged habitat, an out-of-cycle survey may be requested. It is recommended that subsequent sampling events are scheduled to occur +/- 2 weeks of first sampling event when practicable. The survey duration will vary for each site based on water body area (size) and seasonal conditions (wind, lightning, etc.). It is advised that surveys are started as early in the day as possible to complete the survey. Surveys that take longer than one day to complete will resume the following day. Typically, non-wadeable stream sites are surveyed in one day; however, lake sites can take up to five days. If field conditions deteriorate, then survey dates may not be consecutive in order to allow for more ideal conditions. Random ground-truth points are collected at the time of the initial sonar survey each year and targetted ground-truth points are collected within a month of the initial survey to improve the analysis and processing of the sonar data to create metrics and maps.

## 3.3 Variables Reported

All variables reported from the field (L0 data) are listed in the file, NEON Raw Data Validation for Bathymetric and Morphological Maps (DP0.00132.001) (AD[04]). All variables reported in the published data (L4 data) are also provided separately in the file, NEON Data Variables for Bathymetric and Morphological Maps (DP4.00132.001) (AD[05]).

Field names have been standardized with Darwin Core terms (http://rs.tdwg.org/dwc/; accessed 12 December 2017), the Global Biodiversity Information Facility vocabularies (http://rs.gbif.org/vocabulary/gbif/; accessed 12 December 2017), the VegCore data dictionary (https://projects.nceas.ucsb.edu/nceas/projects/bien/wiki/VegCore; accessed 12 December 2017), 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.

### 3.4 Spatial Resolution and Extent

Maps produced from surveys collected prior to 2020: Raw side-scan and down-imaging sonar data are recorded using a dual-beam echosounder at a frequency of 200/83 kilohertz at beam angles of 6° and 19&deg. It is combined with WAAS GPS technology which increases the horizontal accuracy.

Maps produced from surveys collected beginning in 2020: Raw sonar data are recorded using a single-beam, downlooking echosounder at a frequency of 200 kilohertz with a 9&deg beam cone and onboard DGPS.

All maps: Water depths are determined by the first echosounder return and is based on the speed of sound in water and compensated for temperature. The second echo return provides an indication of bottom hardness. More intense returns are suggestive of rocky and compacted sediment where as low intensity returns may indicate soft sediment, vegetation, or fish. The shoreline at the time of sampling is mapped with a handheld GNSS receiver with decimeter accuracy in continuous mode.

The finest scale at which spatial bathymetric data are reported are at a single point (latitidue and longitude) and associated depth for NEON lakes and 1,000 m reaches of NEON non-wadeable streams. In post-processing, the collected point locations and depth contours are constructed using the Empirical Bayesian Kriging (EBK) model parameters for interpolating depth data points in a Geographic Information System (GIS).

### 3.5 Temporal Resolution and Extent

The finest temporal resolution at which bathymetric and morphological data are reported is the **endDate**, a single date on which survey data were collected. Bathymetric and morphological surveys are conducted at each lake once every five years and non-wadeable stream site every 5 years following 3 consecutive years to ensure river bathymetry stability. Should an extreme event result in significant changes in water level, morphology, or habitat, an additional survey may be conducted following a request, once safe conditions allow.

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

### 3.6 Associated Data Streams

At river sites and flow-through lake inflow and outflow locations, cross-section survey data measured at the stream-discharge transect, staff gauge elevation measured during AIS surveys, and rating curve identification metadata can be associated with the Stage-discharge rating curves (DP4.00133.001) data product

and Continuous discharge (DP4.00130.001) data product in order to identify, assess, and/or validate potential shifts in stage-discharge relationships at given sites and to relate streamflow to stage level, respectively.

### 3.7 Product Instances

A minimum of one bathymetric survey will be conducted at each NEON lake every 5 years and at each non-wadeable stream site every three years following three consecutive years to determine river morphology stability. Additional surveys will occur if an extreme event results in substantial physical change to the morphology or submerged habitat. Each sampling event will yield several post-processed data files for bathymetric and morphological data described in (Table 3).

### 3.8 Data Relationships

The basic spatial data that are part of the bathymetric and morphological data product are described in Table 3. Please see Section 3.8.1 Downloading Bathymetric and Morphology Data for information on accessing these data. Variables associated with the collected and published data are described in (Table 4 & Table 6). All NEON Aquatic Observation System sampling locations can be found in the respective Domain Aquatic Site Sampling Design (e.g, Aquatic Site Sampling Design - NEON Domain ##) in the NEON Document Library, Science Designs folder (http://data.neonscience.org/documents).

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.

Table 1: Suite of L4 Data Products Associated with the Bathymetric (BATH) and Morphological Habitat (HAB) Map, RELEASE_2026 and later. Additional files will also be in the folder that are required to open the shapefiles.

| File Type | Description |
| --- | --- |
| AnalysisResults folder | A folder containing all the processed input files |
| vhmcfg extension file | Processed visual aquatic software file |
| DXX_SITE_BATH_YYYY_Map.pdf | Bathymetric map in pdf format |
| DXX_SITE_HABITAT_YYYY_Map.pdf | Habitat map in pdf format |
| DXX_SITE_BATH_YYYY_GE.kmz | Bathymetric map in kmz google earth format |
| DXX_SITE_HABITAT_YYYY_GE.kmz | Habitat map in kmz google earth format |
| postprocessingNotes_YYYY.docx | Freeform data processing notes |
| DXX_SITE_BATH_YYYY_Contour.shp | Bathymetric map shapefile |
| DXX_SITE_HABITAT_YYYY.shp | Habitat map shapefile |
| DXX_SITE_PLANT_YYYY.shp | Plant coverage map shapefile |
| DXX_SITE_BATH_YYYY_shoreline.shp | Shoreline shapefile |
| DXX_SITE_BATH_YYYY_shorePolygon.shp | Shore polygon shapefile |
| DXX_SITE_BATH_YYYY.tif | Raster of interpolated depths |
| DXX_SITE_BATH_processedSonarTracks_YYYY.csv | Processed sonar survey track outputs |

Table 2: Suite of L4 Data Products Associated with the Bathymetric (BATH) and Morphological Habitat (HAB) Map, RELEASE_2025. Additional files will also be in the folder that are required to open the shapefiles.

| File Type | Description |
| --- | --- |
| AnalysisResults folder | A folder containing all the processed input files |
| vhmcfg extension file | Processed visual aquatic software file |
| DXX_SITE_BATH_YYYYMMDD_Map.pdf | Bathymetric map in pdf format |
| DXX_SITE_HABITAT_YYYYMMDD_Map.pdf | Habitat map in pdf format |
| DXX_SITE_BATH_YYYYMMDD_GE.kmz | Bathymetric map in kmz google earth format |
| DXX_SITE_HABITAT_YYYYMMDD_GE.kmz | Habitat map in kmz google earth format |
| postprocessingNotes_YYYYMM.docx | Freeform data processing notes |
| DXX_SITE_BATH_YYYYMMDD_Contour.shp | Bathymetric map shapefile |
| DXX_SITE_HABITAT_YYYYMMDD.shp | Habitat map shapefile |
| DXX_SITE_PLANT_YYYYMMDD.shp | Plant coverage map shapefile |
| DXX_SITE_BATH_YYMMDD_shoreline.shp | Shoreline shapefile |
| DXX_SITE_BATH_YYMMDD_shorePolygon.shp | Shore polygon shapefile |
| DXX_SITE_BATH_YYYMMDD.tif | Raster of interpolated depths |
| DXX_SITE_BATH_processedSonarTracks_YYYY.csv | Processed sonar survey track outputs |

Table 3: Suite of L4 Data Products Associated with the Bathymetric (BATH) and Morphological Habitat (HAB) Map, prior to RELEASE_2025.

| File Type | Description |
| --- | --- |
| BATH Contour.shp | Contour depth lines at a specific interval main shapefile |
| BATH Contour.shx | Contour depth lines index file (required for .shp to open) |
| BATH Contour.dbf | Contour depth lines dBASE attribute file (required for .shp to open) |
| BATH Contour.prj | Contour depth lines projection definition file (required for .shp to open) |
| BATH Contour.xml | Contour depth lines Extensible Markup Language metadata (optional for .shp to open) |
| BATH Contour.sbn | Contour depth lines spatial index file for optimizing spatial queries (optional for .shp to open) |
| BATH Contour.sbx | Contour depth lines spatial index file for optimizing loading time (optional for .shp to open) |
| BATH Contour.cpg | Contour depth lines encoded plain text file describes .shp text character set (optional for .shp to open) |
| BATH Shoreline.shp | Shoreline boundary of water body main shapefile |
| BATH Shoreline.shx | Shoreline boundary index file (required for .shp to open) |
| BATH Shoreline.dbf | Shoreline boundary dBASE attribute file (required for .shp to open) |
| BATH Shoreline.prj | Shoreline boundary projection definition file (required for .shp to open) |
| BATH Shoreline.xml | Shoreline boundary Extensible Markup Language metadata (optional for .shp to open) |
| BATH Shoreline.sbn | Shoreline boundary spatial index file for optimizing spatial queries (optional for .shp to open) |
| BATH Shoreline.sbx | Shoreline boundary spatial index file for optimizing loading time (optional for .shp to open) |
| BATH Shoreline.cpg | Shoreline boundary encoded plain text file describes .shp text character set (optional for .shp to open) |
| BATH EBK.xml | Empirical Bayesian Kriging model parameters for interpolation of depth data points |
| BATH Tracks.csv | Survey track |
| BATH Volume.csv | Volume calculations |
| BATH GE.kmz | Useable bathymetric map viewd in Google Earth |
| BATH Map.pdf | Useable bathymetric map |
| BATH.tif | Raster of interpolated depths |
| HAB.shp | Habitat features classified polygons |
| HAB Area.csv | Habitat area calculations |
| HAB GE.kmz | Useable habitat map viewed in Google Earth |
| Hab map.pdf | Useable habitat map |

Table 4: Maps produced from surveys collected prior to 2020: List of Variables Provided with the Bathymetric (BATH), Morphological Habitat (HABITAT) Map, Survey Tracks (BATH_Tracks), and Shoreline (BATH_shorePolyDigitized) File Types, prior to RELEASE_2025. 'DxxSITE' represents the NEON Domain number for each File Type, for example: D03BARC.

| File Type | Header Name | Description | Data Type | Units |
| --- | --- | --- | --- | --- |
| DxxSITE_HABITAT_YYYYMMDD_Area.csv | Frequency | Rate of occurance for each habitat type | int | NA |
| DxxSITE_HABITAT_YYYYMMDD_Area.csv | Habitat | Habitat type classification | String | NA |
| DxxSITE_HABITAT_YYYYMMDD_Area.csv | Area_m2 | Calculated area for each habitat type | Real | meter |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Object ID | Unique ID for each track | String | NA |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Longitude | Geographic longitude | real | decimalDegree |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Latitude | Geographic latitude | real | decimalDegree |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Depth | Depth of water | int | meter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | Depth | Depth of water at set intervals | int | meter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | 2d_area | 2 dimensional surface area | int | squareMeter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | 3d_depth | 2 dimensional surface area plus slope | int | squareMeter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | Volume | Calculated water volume at a set depth interval | int | cubicMeter |
| DxxSITE_BATH_YYYYMMDD_ shorePolyDigitized.shp | Name | Lake and non-wadeable stream perimeter | string | NA |
| DxxSITE_BATH_YYYYMMDD_ shorePolyDigitized.shp | Shape_Leng | Lake and non-wadeable stream perimeter length | real | meter |
| DxxSITE_BATH_YYYYMMDD_ shorePolyDigitized.shp | Shape_Area | Lake and non-wadeable stream reach area | real | squareMeter |

Table 5: Maps produced from surveys collected beginning in 2020: List of Variables Provided with the Bathymetric (BATH), Morphological Habitat (HABITAT) Map, Survey Tracks (BATH_Tracks), and Shoreline (BATH_shorePolyDigitized) File Types, prior to RELEASE_2025. 'DxxSITE' represents the NEON Domain number for each File Type, for example: D03BARC.

| File Type | Header Name | Description | Data Type | Units |
| --- | --- | --- | --- | --- |
| DxxSITE_HABITAT_YYYYMMDD_Area.csv | Habitat | Habitat type classification | String | NA |
| DxxSITE_HABITAT_YYYYMMDD_Area.csv | Area_m2 | Calculated area for each habitat type | Real | meter |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Object ID | Unique ID for each track | String | NA |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Latitude_deg | Geographic latitude | real | decimalDegree |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | Longitude_deg | Geographic longitude | real | decimalDegree |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | BottomElevation_m | Depth of water | int | meter |
| DxxSITE_BATH_YYYYMMDD_Tracks.csv | BottomType | Bottom type classification PCA component number | int | NA |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | Depth | Depth of water at set intervals | int | meter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | 2d_area | 2 dimensional surface area | real | squareMeter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | 3d_area | 2 dimensional surface area plus slope | real | squareMeter |
| DxxSITE_BATH_YYYYMMDD_Volume.csv | Volume | Calculated water volume at a set depth interval | real | cubicMeter |
| DxxSITE_BATH_YYYYMMDD_shorePolyDigitized.shp | Name | Lake and non-wadeable stream perimeter | string | NA |
| DxxSITE_BATH_YYYYMMDD_shorePolyDigitized.shp | Shape_Leng | Lake and non-wadeable stream perimeter length | real | meter |
| DxxSITE_BATH_YYYYMMDD_shorePolyDigitized.shp | Shape_Area | Lake and non-wadeable stream reach area | real | squareMeter |

All released data prior to RELEASE_2025: The protocol dictates that each **locationID** is surveyed at least once every 5 years. One record is expected per **startDate** in bat_fieldData. Each **waypoint** may have multiple or no child records associated with **bathymetryDataType** (shoreline boundary, in-water feature, groundtruth targeted, groundtruth random) within the bat_pointcollection record, depending on whether a waypoint was collected during the survey. Each **recordingNumber** record from bat_sonarRecord is expected to have one to several child records for every survey track that is collected. One **dataFileName** and one **rawDataFileName** record will be created in bat_resultsFile which contain the maps and raw data plus file extensions. 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*.

bat_fieldData - > One record expected per **locationID**. Field data associated with the field level record. This table is linked to the bat_pointcollection and bat_sonarRecord tables through the **locationID**.

bat_pointcollection - > Up to several records expected per **waypoint**. Field data associated with the point level record.

bat_sonarRecord - > Several records expected per **recordingNumber**. A single parent location ID (recordingNumber) generates child recording number files.

bat_resultsFile - > One record per **siteID** and **startDate** combination. Contains basic information about the filenames and URLs used to access cloud-storage data product components that comprise the packaged bathymetric and habitat maps.

bat_AISsiteSurveyResultsFile - > One record per **siteID**x**startDate**. Contains basic information about the Aquatic Instrumental System site survey and URLs to access cloud-storage data products used to delineate hydraulic controls for the Stage-discharge rating curves (DP4.00133.001) and Continuous discharge (DP4.00130.001) data products for river sites and Toolik Lake inflow and outflow locations. These surveys will likely occur outside of a usual bathymetry bout and are not constrained by the same peak greenness window scheduling constraint.

RELEASE_2025 and after: The protocol dictates that each **locationID** is surveyed at least once every 5 years at lake and river sites. One record is expected per **startDate** in bat_fieldData. Each **waypoint** may have multiple or no child records associated with **bathymetryDataType** (shoreline boundary, in-water feature, groundtruth targeted, groundtruth random) within the bat_pointcollection record, depending on whether a waypoint was collected during the survey. Each **recordingNumber** record from bat_sonarRecord is expected to have one to several child records for every survey track that is collected. One **dataFileName** and one **rawDataFileName** record will be created in bat_resultsFile which contain the maps and raw data plus file extensions. 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*.

D##_SITE_BATH_processedSonarTrack_YYYY_VX.csv -> One record per **collectDateTime**, which corresponds to a single sonar ping.

bat_fieldData - > One record expected per **locationID**. Field data associated with the field level record. This table is linked to the bat_pointcollection and bat_sonarRecord tables through the **locationID**.

bat_groundTruthPoints - > Up to several records expected per **waypoint**. Field data associated with the point level record and in-water features.

bat_processedHabitatResults- > One record expected per **assignedHabitatType**.

bat_processedPlantCoverage - > One record expected per **plantCoverageCategory**.

bat_processedSiteLevelSummary - > One record expected per **eventID**. Contains basic water body metrics, such as **maxDepth**, **waterBodyVolume**, and **waterBodySurfaceArea**, the SOP version used to create the L4 data, and the URLs to access processed and raw cloud-storage files.

bat_processedSonarTrackData - > One record per **downloadFileName**. Contains name and URL information for the processed sonar track file (the first file described in this list).

bat_processedVolumeData - > One record per **referencePlaneDepth**. Contains bathymetric contour area at multiple depths for a water body.

bat_sonarRecord - > One record per **sonarRecordingNumber**. Contains basic information and the names of all sonar files processed to create metrics and maps.

## 3.9 Special Considerations

### 3.9.1 Downloading Bathymetric and Morphology Data

To download L4 bathymetric and morphology data packages:

1. Access the NEON data portal and select the desired data product (Bathymetry and Morphology Map), date range, and location (state and site).
2. Download the dataset and open the .CSV file **bat_processedSiteLevelSummary** contained in the zipped folder. This .CSV file will contain a file named with the format "NEON.D##.SITE.DP1.00131.001.bat_processedSiteLevelSummary.YYYY-MM.basic or expanded.transitionDate.csv". Note that D## is the NEON Domain number and YYYY-MM is the year and month of the date range specified.
3. Links to the processed data storage location are provided in the **dataFileName** field and links to the unprocessed data storage location are provided in the **rawDataFileName** field.
4. For L4 data products, the links to the data storage location are provided in the **dataFileName** field.
5. Copy the URL provided in the **dataFileName** field into a web browser to initiate the download of the L4 survey data packages. Similarly, copy the URL provided in the **rawDataFileName** field into a web browserr for the L0 unprocessed data.
6. Users can also build scripts or use other tools that retrieve data from the NEON API and use the URL in different ways to retrieve the bathymetry data. The neonUtilities package, zipsByURI function, on CRAN is especially useful for these tasks (https://cran.r-project.org/web/packages/neonUtilities/index.html). A python version of the zipsByURI function is in development at the time of this document update.

### 3.9.2 Interpretation of Date Fields

There are several date fields within the Bathymetry data product and not all are specifically meaningful. For instance, the startDate and endDate fields in most of the data tables match and reflect the last

planned day of sonar collection. Because data collection is somewhat dependent on initial analysis of the sonar data, quite a bit of data is loaded to the database prior to the completion of data collection for the survey. The startDate and endDate are chosen to ensure that all of the data is grouped in the database and will be in the same download package without requiring edits to previously ingested data.

For users interested in the dates of sonar data collection, we suggest using the sonarRecordingStartTime and sonarRecordingEndTime information that can be found in the bat_sonarRecord table. For users interested in the dates for the random and targeted ground-truthing data, we suggest using the pointCollectionDate found in the bat_groundTruthPoints table. The eventID field in all tables, which contains only the site and year of the survey, is the best way to join across tables.

The date that is part of the dataFileName used for the zipped data and in the bat_processedSiteLevelSummary table may or may not match the startDate or endDate in field data tables and is not specifically meaningful. The eventID in the corresponding record in the bat_processedSiteLevelSummary table should be used to join with field data tables. Starting with the 2024 surveys the eventID will be used in the filenames in order to make linking processed, zipped data with other tables more straighforward.

# 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 field data entry (Fulcrum) 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 Bathymetric and Morphological Maps (DP0.00132.001) (AD[04]), 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[12]).

## 4.2 Automated Data Processing Steps

Following data entry into a mobile application, 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[08]).

Published data are reviewed for completeness, timeliness, and validity using an internal set of tests and metrics, as detailed in the NEON Algorithm Theoretical Basis Document: OS Data Quality Control

(AD[15]). These quality tests are used to guide process improvements, audits of analytical facilities, and data updates, but do not generate quality flags in published data.

### 4.3 Manual Data Processing Steps

Manual data processing transforms survey data from raw L0 data to an L4 data product. During this time QAQC procedures are performed which include a careful evaluation of field survey notes (contained in the metadata) that document errors that occurred during the survey. If survey notes indicate that L0 data needs to be altered or deleted during post-processing these changes are addressed and subsequently expressed in the L4 product. 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[08]). Additional post-processing procedures are detailed in NEON Standard Operating Procedure: Post-Processing of Bathymetric and Side Scan Sonar Data from NEON Lakes and Non-Wadeable Streams (AD[09]). Additionally, random and targetted ground-truth data points are collected to inform bottom type classifications.

### 4.4 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.

### 4.5 Uncertainty

Spatial datasets are not a seamless representation of real-world phenomena.

Maps produced from surveys collected prior to 2020: A HumminBird 1198c SI Combo sonar with an XHS 9 HDSI 180 T transducer with down imaging (dual eam 200/83 kHz) and side imaging sonar (455/800 kHz) are used to map NEON aquatic lakes and non-wadeable stream reaches during bathymetric and morphology surveys. The internal GPS inside the MX surface unit is a Garmin OEM DGPS which provides <3 meters horizontal accuracy when differential correction is available and <15 meter when it is not. The qualatiative 'surveyPointCoordinateQuality' outputs (standard, differential) have been converted quantiative measurements (15, 3) in the processedSonarTrack .csv file. See the Garmin OEM DGPS spec sheet for more information: https://static.garmin.com/pumac/GPS_18x_Tech_Specs.pdf

Maps produced from surveys collected beginning in 2020: A Biosonics MX aquatic habitat echosounder (200 kHz, single beam, down-looking) is used to map NEON aquatic lakes and non-wadeable stream reaches during bathymetric and morphology surveys.

All surveys: A GNSS receiver is used to collect GPS locations for the shoreline boundary, in-water features, and during ground truthing activities, throughout the bathymetric and morphology survey. Starting in

2021, if a shoreline boundary cannot be collected by walking the permieter of the lake or river, a track is collected as close to the shore as possible and then buffered by three meters to create the shoreline and shorelinePolygon files. Differential correction is applied during post-processing to improve GPS accuracy and reduce atmospheric errors by comparing the time signature at a fixed base station (typically CORS [Continuously Operating Reference Station]) nearby the rover file from the GNSS receiver unit. The resulting file defines a horizontal and vertical accuracy using the root mean square error based on a 68% confidence level. NEON has aimed to utilize post-processed GPS positions that are within a horizontal and vertical precision of 10cm. Due to dense canopy at some sites and/or distance from base stations, not all GPS points surveyed fall within the desired range.

### 4.6 Quality Flagging

The **dataQF** field in each data record is a quality flag for known errors applying to the record. Please see below (Table 6) for an explanation of **dataQF** codes specific to this product.

Table 6: Options for dataQF and their explanations.

| 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)

## 5 Overlaying map layers in R

The zipped L4 folder contains shapefiles along with two kmz and pdf maps. One set for the habitat and plant coverage and the second set for the bathymetry. Custom maps can be created by turning various layers on and off in google earth if both kmzs are open. Custom maps can also be created from shapefiles by combining different layers in GIS software, such as arcMap or arcPro, or programmatically through code, such as R or python. Below is an example of overlaying the plant coverage and bathymetry contour shapefiles in R. The code can be adapted for other combinations of layers. For additional information use the Contact Us form (https://www.neonscience.org/about/contact-us) on the NEON Data Portal and select DP4.00132.001 in the Data Product drop-down list.

### 5.1 Code

```
require(tidyverse)
require(sf)
require(utils)
require(tidyterra)
library(leaflet)
library(neonUtilities)

# Download bathyemtry data from the 'Bathymetric and morphological maps' data product.
# For this example we will use TOOK data from 2021
bat_TOOK <- neonUtilities::loadByProduct(
  dpID='DP4.00132.001',
  check.size=F,
```

```
  site = "TOOK",
  startdate = "2021-01",
  package='basic',
  token = Sys.getenv('NEON_PAT')) # remove this line if you don't have a token

# Use zips by URL to download zip file contents to directory
workDir <- getwd()
zipDir <- paste0(workDir,"/ECS_zipFiles")
test <- neonUtilities::zipsByURI(filepath = bat_TOOK,
                                 check.size = FALSE)

#make a list of all the shapefile names available
surveyFolderName <- list.files(zipDir, pattern = "L4")
shpFilenamesL4 <- list.files(paste0(zipDir,"/",surveyFolderName),
                             pattern = ".shp$",
                             full.names = T)

# Read in the plant coverage shapefile
plant_shp <- sf::st_read(grep("PLANT",shpFilenamesL4, value= T))

# Read in the contour shapefile
contour_shp <- sf::st_read(grep("Contour",shpFilenamesL4, value= T))

##### Use leaflet package to create custom maps #####

# Create a basemap
basemap <- leaflet() %>% addTiles()

# Transform all sp data to wgs 84 (lat/longs, epsg = 4326) to work with leaflet
plantWGS <- st_transform(plant_shp, crs = 4326)
contourWGS <- st_transform(contour_shp, crs = 4326)

# Plot each shapefile, choosing the color/style that makes sense
# Options include addPolyline, addPolygons, addMarkers
# Users might need to break into multiple maps if many layers

# Create color palettes
# Use RColorBrewer::display.brewer.all() to view some built-in options
# For more color info:
# https://github.com/rstudio/leaflet/blob/main/vignettes/articles/colors.Rmd
pal <- colorNumeric(
  palette = topo.colors(25),
  domain = contourWGS$Contour
)
```

```
pal2 <- colorFactor(
  palette = "Greens",
  domain = plantWGS$Bin
)

# Create a map
basemap %>%
  addPolygons(data = plantWGS,
              fillColor =  ~pal2(Bin),
              stroke = F, #no outline,
              fillOpacity = .35) %>%
  addPolylines(data = contourWGS,
                color = ~pal(Contour),
                weight= 1.5) %>%
  addLegend("bottomright", pal = pal, values = ~Contour,
          title = "Contour (m)",
          data = contourWGS) %>%
  addLegend("topleft", pal = pal2, values = ~Bin,
            title = "Plant Coverage Bins",
            data = plantWGS)
```

## 5.2 Map Output

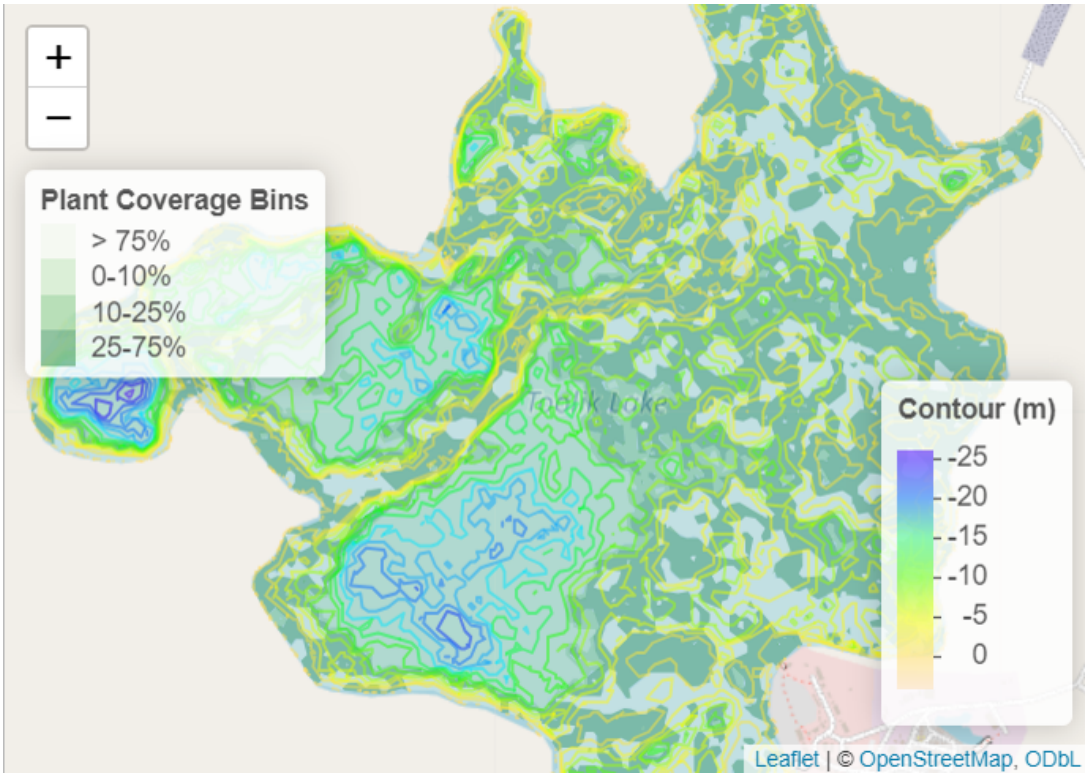


Figure 2: Map created using R leaflet package and Bathymetry zipped shapefiles. An example of overlaying plant coverage and bathymetry topography lines.