NEON USER GUIDE TO WADEABLE STREAM MORPHOLOGY (DP4.00131.001)

PREPARED BY	ORGANIZATION
Nick Harrison	AQU
Kaelin M. Cawley	AQU
Zachary Nickerson	AQU

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	12/07/2017	Initial Release
B	07/17/2020	Included general statement about usage of neonUtilities R package and statement about possible location changes.
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 (rapid habitat assessment data and 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	03/08/2022	Updated section 4.5 Data Revision with latest information regarding data release
D	10/06/2022	Added information pertaining to rapid habitat assessment data
E	03/20/2025	Updated data product structure, including .CSV files migrated from the cloud-stored data packages to be downloadable in data packages directly from the NEON Data Portal; included descriptions of new tables and table relationships; updated information on post-processing methods in this Level 4 data product. 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 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, total station survey data derived from a single geomorphology survey, 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 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 describes the steps needed to generate the L4 data product, Wadeable Stream Morphology - the characterization of stream channel geometry and demarcation of biological habitats and stream channel features within NEON aquatic stream reaches - 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 Wadeable Stream Morphology (DP4.00131.001) (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: Wadeable Stream Morphology (AD[06]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Validation for Wadeable Stream Morphology (DP0.00131.001) (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., '00131') 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.002652	NEON Data Products Catalog
AD[03]	Available with data download	NEON Raw Data Validation for Wadeable Stream Morphology (DP0.00131.001)
AD[04]	Available with data download	NEON Data Variables for Wadeable Stream Morphology (DP4.00131.001)
AD[05]	NEON.DOC.001152	NEON Aquatic Sample Strategy Document
AD[06]	NEON.DOC.003162	AOS Protocol and Procedure: Wadeable Stream Morphology
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]	Available on NEON data portal	NEON Ingest Conversion Language
AD[12]	Available with data download	Categorical Codes csv
AD[13]	NEON.DOC.005399	AIS Standard Operating Procedure: Conducting AIS Site Surveys at NEON Aquatic Sites
AD[14]	NEON.DOC.005402	NEON Algorithm Theoretical Basis Document (ATBD): Post-Processing of Stream Morphology Data from NEON Wadeable Streams

2.2 Acronyms

Acronym	Definition
AIS	Aquatic Instrumented Subsystem
DP	Data Product
GPS	Global Positioning System
L0	Level 0
L4	Level 4
UTM	Universal Transverse Mercator

3 DATA PRODUCT DESCRIPTION

The wadeable stream morphology data product provides raw survey data, maps, shapefiles, and metric tables that quantify stream channel geomorphology and bed composition and delineate biological habitats within the aquatic reach boundaries (approximately 1,000 meters in stream length) of wadeable streams at NEON aquatic sites. Raw survey data are collected with high-resolution total station survey equipment at each NEON wadeable stream site. Survey maps and channel metrics are produced and calculated using raw survey data (Level 0) that is geo-referenced to a global coordinate system (Level 4).

3.1 Spatial Sampling Design

Geomorphology surveys encapsulate the entirety of the aquatic reach, which at most NEON sites is equivalent to approximately 1,000 meters in stream length.

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/

3.2 Temporal Sampling Design

All survey data are reported at the temporal resolution of a single survey, which, depending on the complexity of the aquatic reach and/or environmental conditions during the survey, can last 3-10 days.

3.3 Variables Reported

All variables reported from the field (L0 data) are listed in the file, NEON Raw Data Validation for Wadeable Stream Morphology (DP0.00131.001) (AD[03]). All variables reported in the published data (L4 data) are also provided separately in the file, NEON Data Variables for Wadeable Stream Morphology (DP4.00131.001) (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 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

A raw survey data file includes each of the individually mapped points collected by the total station during the geomorphology survey. Each point contains a Northing, Easting, and elevation coordinate relative to

fixed benchmarks installed at the downstream extent of the reach (where the surveys typically begin). Mapped points are distributed at a high resolution (typically less than 1m) throughout the extent of the aquatic reach. Points are mapped along the main channel to capture thalweg (or the deepest part of the stream) elevation, along the edge of water to capture wetted width, and along select transects that run perpendicular to the channel in order to capture cross-sectional area. Additional points are collected at stream features that locally influence fluvial processes (i.e. large woody debris jams, mid-channel bars, etc.).

Northing, Easting, and elevation data contained in raw survey files are relative to a local Cartesian coordinate plane (X,Y, and Z, respectively) defined by the fixed benchmark used to orient the total station at the beginning of the survey. Fixed benchmark locations are globally referenced (WGS 84 reference coordinate system) and locally projected (UTM Zone xNorth) using global positioning instrumentation to an accuracy of 10-30 centimeters of elevation. The degree of additional uncertainty associated with each survey will vary and is dependent on operator error, site-specific conditions, and environmental factors. Uncertainty associated with each survey is included within the data product package. During post-processing, GPS data are utilized to convert Northing and Easting values to latitude and longitude and elevation values to meters above mean sea level. All geo-referenced survey data are considered Level 4.

3.5 Temporal Resolution and Extent

Geomorphology surveys are conducted at each site once every five years. During years when there is not a geomorphology survey, a rapid habitat assessment survey is conducted in the same reach that includes a pebble count survey and geo-referencing and characterization of each habitat unit within the stream reach.

3.6 Associated Data Streams

Cross-section survey data measured at the stream-discharge transect, staff gauge elevation measured during geomorphology surveys, and rating curve identification metadata are collected during geomorphology surveys and published in the Stage-discharge rating curves data product (DP4.00133.001) and Continuous discharge data product (DP4.00130.001). These data identify and assess potential shifts in stage-discharge relationships at given sites and are used to relate streamflow to stage level.

3.7 Product Instances

One geomorphology survey will be conducted at each NEON wadeable stream site every five years, resulting in approximately seven geomorphology surveys per year across the NEON Observatory. One rapid habitat assessment survey will be conducted at each wadeable stream site during each year that there is not a geomorphology survey.

3.8 Data Relationships

This data product consists of three distinct types of total station surveys: geomorphology survey, rapid habitat assessment survey, and AIS site survey. Regardless of the type of survey, every survey bout will contain one record in geo_surveySummary. Survey type is distinguished in the **surveyBoutTypeID** field of

geo_surveySummary. The other tables available through the download of this data product depends on the type of survey conducted. **Table 1** below provides a description of all tables available via download from the NEON Data Portal and from which survey type users should expect available data.

Table 1: Data tables available in downloads of this data product from the NEON Data Portal and the survey types with which they are associated

Table	Description	Associated Survey Type
geo_featureInfo	Data associated with mapped cross-section transects	geomorphology
geo_geomorphicFeatureCount	The type and total count of all geomorphic features mapped during the survey	geomorphology
geo_mappedPointErrors	Data that describe any survey errors that need to be resolved during post-processing	geomorphology
geo_misclosureInfo	Data associated with survey misclosure	geomorphology
geo_missingLine	Data associated with the missing line workflow used to orient the first total station location during the survey	geomorphology
geo_pebbleCount	Sediment particle size data collected during pebble count sampling	geomorphology, rapid habitat
geo_pebbleFieldData	Sediment particle size distribution summary data associated with pebble count sampling during the survey	geomorphology, rapid habitat
geo_processedSurveyData	Survey data generated during post-processing, including point classifications and real world coordinates	geomorphology
geo_rapidHabitatAssessment	Data associated with the rapid habitat assessment portion of the stream morphology protocol	rapid habitat
geo_rawSurveyData	Raw data generated during the survey, including a unique identifier and raw coordinate values for each point mapped	geomorphology
geo_surveyPoints	Data associated with mapped cross-section transects used to delineate hydrologic controls	geomorphology, AIS site survey

geo_surveySummary	Summary data associated with the survey, including summary statistics such as survey completeness, physical channel characteristics, relationships between upstream and downstream sensor sets, and data download URLs	geomorphology, rapid habitat, AIS site survey
geo_thalwegByHabitatID	Data that describe each stream habitat unit mapped during the survey	geomorphology
geo_thalwegByHabitatType	Summary level data for stream habitat units mapped during the survey	geomorphology
geo_thalwegLongProfile	Data associated with points mapped on the streambed along the thalweg	geomorphology
geo_totalStation	Data that describe each total station location setup, including control point and thalweg point identifiers, geomorphic features surveyed from each location, and associated error metrics	geomorphology
geo_transectBankfullWidths	Data that describe channel widths measured at mapped cross-section transects	geomorphology
geo_trimbleData	Field collected spatial data	geomorphology

Not all data tables associated with geomorphology surveys in **Table 1** will be available for every survey. Such opportunistic, but not required, tables for geomorphology surveys are as follows:

- **geo_mappedPointErrors:** Table will populate if errors occur during that survey that require edits during post-processing.
- **geo_missingLine:** Table will populate if the Missing Line workflow is conducted during the survey (see details in AD[06])
- **geo_trimbleData:** Table will populate if GPS field data are collected during the survey.

For ease of table joining, each table published in this data product contains an identical **eventID** field formatted as the site ID followed by the year and month that constitutes the **surveyEndDate** (SITE.YYYYMM).

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.

In addition to the data tables available in this data product via direct downloads from the NEON Data Portal, L4 and L0 data packages are available from cloud-stored locations using URL paths contained in direct download files. See **Section 3.9 Special Considerations** for instructions on how to access cloud-stored data packages for this data product.

Table 2: Files available in cloud-stored L0 rapid habitat assessment data packages

File Name	File Format	File Description
RAW data file	.SSF	All raw rapid habitat assessment data collected on a handheld GPS receiver

Table 3: Files available in cloud-stored L4 and L0 geomorphology survey data packages

File Name	File Format	File Description
geomorphicFeaturePoints	.SHP	Shapefile that displays all geomorphic features delineated using single points (i.e. upstream/downstream waterfall boundaries and tributary locations)
geomorphicFeaturePolygons	.SHP	Shapefile that displays all geomorphic features delineated using polygons (multiple points) (i.e. mid-channel bars, islands, large woody debris jams, beaver dams, etc.)
habitatUnits	.SHP	Shapefile that displays habitat unit boundaries throughout the reach
sensors	.SHP	Shapefile that displays sensor set locations within the reach
thalweg	.SHP	Shapefile that displays thalweg line throughout the reach
transects	.SHP	Shapefile that displays transect locations throughout the reach
siteMap	.PDF	PDF map of the aquatic reach that features the reach boundary, transect, aquatic sensor, and staff gauge locations as well as biological habitat delineations along the thalweg line
siteMap	.KMZ	KMZ map of the aquatic reach that can be viewed using Google Earth software. This map features habitat unit delineations, sensor set and trasnect locations, the thalweg profile, and all geomorphic feature points and polygons mapped during the survey
surveyNotes	.DOCX	Word file that documents site specifications, survey observations, and any deviations that occurred during the field survey
postProcessingNotes	.DOCX	Word file that documents site specifications and any deviations that occurred during post-processing
RAW data file	.SSF	Benchmark data collected on a handheld GPS receiver for sites or years without permanent benchmarks

3.9 Special Considerations

3.9.1 Downloading Stream Morphology Data

To download L4 and L0 stream morphology and rapid habitat assessment data packages:

1. Access the NEON data portal and select the desired data product (Stream Morphology Map in Eco-Hydrology and Land Use, Land Cover, and Land Processes Data Products), date range, and location (state and site).
2. Download the L4 dataset and open the **geo_surveySummary.csv**. The full formatted name of these zipped files will appear as "NEON.DOM.SITE.DP4.00131.001.geo_surveySummary.YYYY-MM.basicstransitionDate.csv". Note that "DOM" is the NEON Domain number and "YYYY-MM"" is the year and month of the date range specified.
3. To access cloud-stored L4 data products, links to the location are provided in the **dataFilePath**. To access cloud-based L0 data products, links to the cloud-storage location are provided in the **rawDataFilePath** field.
4. Copy the URL provided in the **dataFilePath** field into a web browser to initiate the download of the L4 survey data package.
5. Copy the URL provided in the **rawDataFilePath** field into a web browser to initiate the download of the L0 survey data package.

4 DATA QUALITY

4.1 Data Entry Constraint and Validation

Data entry constraint and validation apply solely to the metadata associated with this data product. Many quality control measures are implemented at the point of data entry within a mobile data entry application. For example, data formats are constrained and data values controlled through the provision of drop-down options, which reduces the number of post-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 Wadeable Stream Morphology (DP0.00131.001), provided with every download of this data product. Contained within this file is a field named 'entryValidationRulesParser', 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 Ingest Conversion Language (NICL) specifications ([AD[10]).

4.2 Using Google Earth to View .KMZ Maps

Geomorphic feature polygon layers may not be immediately visible while using Google Earth software to view .KMZ morphology maps. Controlling the way multiple layers are drawn into Google Earth is a known issue and (to date) a solution has not been identified to stack layers in a customized manner. Therefore geomorphic feature polygons may be hidden by habitat unit polygons (sometimes a black outline of the

geomorphic feature polygon is present). To fully view geomorphic feature polygons, turn off the habitat units polygons by unchecking the box to the left of the feature name.

4.3 Automated Data Processing Steps

Following the entry of metadata 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]).

Following the manual data post-processing steps listed below, shapefile inputs are loaded into R statistical software to calculate metrics, summaries, and final shapefile outputs that are available in the L4 cloud-stored locations.

4.4 Manual Data Post-Processing Steps

Manual data post-processing transforms survey data from raw L0 data to an L4 data product. During this time QAQC procedures are performed, including a careful evaluation of field survey notes (contained in the metadata) that document errors that occurred during the survey. If notes indicate that L0 data need to be altered or deleted during post-processing these changes are addressed and subsequently expressed in the L4 product. Field survey errors include (but are not limited to) mis-labeled survey points, incorrect prism rod heights, and mapped points that require deletion.

In ArcPro or similar software, points from the total station survey are converted to real-world coordinates using a similarity transformation based on the known benchmark locations. Thalweg, habitat polygons, and geomorphic feature layers are created from the total station point data using a combination of ArcPro tools and manual editing based on survey point names. Metric calculations are completed using a custom R code using the shapefile inputs created in ArcPro. See NEON Algorithm Theoretical Basis Document (ATBD): Post-Processing of Stream Morphology Data from NEON Wadeable Streams (AD[14]) for additonal details about the manual and automated processing steps.

4.5 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.

L4 tables include a **processedSurveyVersion** field which links the tabular data to the specific survey data package stored in the cloud.

4.6 Uncertainty

Spatial datasets are not a seamless representation of real-world phenomena. Most sites have benchmark location data which is input to the POC prior to the start of a survey. These serve as reference points for

the subsequently mapped survey points. Permanent benchmark locations and metadata data for transformations can be obtained from the geoNEON R package function 'getLocBySite'.

Download the spatial data for the site of interest and then filter for locationType = "AOS bench-mark named location type".

The **surveyPointID** in the geo_rawSurveyData format is "BM_X" and is associated with the SITE.AOS.benchmark.X in the **namedLocation** column from the getLocBySite function.

For example, **surveyPointID** "BM3" from a HOPB geo_rawSurveyData is associated with the **namedLocation** "HOBP.AOS.benchmark.3".

Spatial data can also be obtained using the NEON locations API. To search for the above example, use: http://data.neonscience.org/api/v0/locations/HOPB.AOS.benchmark.3.

Sites without permanent benchmarks use GPS receivers to acquire positioning information and spatial uncertainty values are reported in the geo_trimbleData table.

A HILTI POS 180 Robotic Total Station and POC data collector are used to spatially map the aquatic reach during geomorphology surveys. The POS 180 contains angle measurement accuracy (DIN 18723) of 3", a distance measurement accuracy pf +- 2 mm + 2 ppm and a magnification of 31x. The measurement range of the instrument is 79" to 9843'. Standard deviation values are reported and documented for each total station setup throughout the survey. The best effort is made to minimize standard deviations to <= 5mm per total station setup, however this is not always possible given certain site conditions (i.e. unstable ground surfaces or extreme weather).

A Trimble GeoXH 6000 or 7000 series GPS unit with Tornado antenna is used to collect survey grade GPS locations throughout the geomorphology survey. 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 Trimble 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 30cm. Due to dense canopy at some sites and/or distance from base stations, not all GPS points surveyed fall within the desired range.

4.7 Quality Flagging

Two document files (surveyNotes.docx and postProcessingNotes.docx) are included in the L4 cloud-stored location if specifications and/or survey deviations specific to the product occurred during field collection or post-processing.

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)