

NEON USER GUIDE TO GROUNDWATER AND ACTIVE LAYER MEASUREMENTS AT PERMAFROST SITES (DP1.20099.001)

| PREPARED BY | ORGANIZATION |
|------------------|--------------|
| Kaelin M. Cawley | AQU |
| Nora Catolico | AQU |

CHANGE RECORD

| REVISION | DATE | DESCRIPTION OF CHANGE |
|----------|------------|--|
| A | 08/08/2018 | Initial Release |
| B | 05/13/2020 | Included general statement about usage of neonUtilities R package and statement about possible location changes. |
| C | 02/08/2022 | Updated section 4.3 Data Revision with latest information regarding data release |
| D | 02/18/2025 | Updated the url for spatial data in section 3.4. Added information about the new neonUtilities Python package. |



TABLE OF CONTENTS

| | | |
|----------|--|----------|
| 1 | DESCRIPTION | 1 |
| 1.1 | Purpose | 1 |
| 1.2 | Scope | 1 |
| 2 | RELATED DOCUMENTS AND ACRONYMS | 2 |
| 2.1 | Associated Documents | 2 |
| 2.2 | Acronyms | 2 |
| 3 | DATA PRODUCT DESCRIPTION | 3 |
| 3.1 | Spatial Sampling Design | 4 |
| 3.2 | Temporal Sampling Design | 4 |
| 3.3 | Variables Reported | 4 |
| 3.4 | Spatial Resolution and Extent | 4 |
| 3.5 | Temporal Resolution and Extent | 5 |
| 3.6 | Associated Data Streams | 5 |
| 3.7 | Product Instances | 5 |
| 3.8 | Data Relationships | 5 |
| 4 | DATA QUALITY | 6 |
| 4.1 | Data Entry Constraint and Validation | 6 |
| 4.2 | Automated Data Processing Steps | 6 |
| 4.3 | Data Revision | 6 |
| 4.4 | Quality Flagging | 6 |

LIST OF TABLES AND FIGURES

| | | |
|----------|--|---|
| Table 1 | Descriptions of the dataQF codes for quality flagging | 7 |
| Figure 1 | Diagram detailing the measurements made for Groundwater and active layer measurements at permafrost sites: (A) is the depth from the top of the groundwater well to the ground surface, (B) is the depth from the top of the groundwater well to the surface of liquid water, and (C) is the depth from the top of the groundwater well to the surface of the active layer at the bottom of the liquid water column. | 3 |

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, such as a single active layer depth measurement, 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 Groundwater and active layer measurements at permafrost sites, the process of manually measuring the distance from the ground surface to liquid water, if present, and distance from the ground surface to the top of the thawed active layer. 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 Groundwater and active layer measurements at permafrost sites (DP1.20099.001) (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 NEON Preventive Maintenance Procedure: AIS Goundwater Wells (AD[07]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Validation for Groundwater and active layer measurements at permafrost sites(DP0.20099.001) (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 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.000913 | TOS Science Design for Spatial Sampling |
| AD[03] | NEON.DOC.002652 | NEON Data Products Catalog |
| AD[04] | Available with data download | NEON Raw Data Validation for Groundwater and active layer measurements at permafrost sites(DP0.20099.001) |
| AD[05] | Available with data download | NEON Data Variables for Groundwater and active layer measurements at permafrost sites (DP1.20099.001) |
| AD[06] | NEON.DOC.001152 | Aquatic Sampling Strategy |
| AD[07] | NEON.DOC.004362 | NEON Preventive Maintenance Procedure: AIS Goundwater Wells |
| 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] | Available on NEON data portal | 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 |
|---------|----------------------------|
| AOS | Aquatic Observation System |



3 DATA PRODUCT DESCRIPTION

Groundwater availability, distribution, and seasonal variability have a substantial influence on ecological processes. The NEON data product Elevation of Groundwater (DP1.20100.001) can be used to enable the study of these influences at most NEON sites; however, it cannot be applied to permafrost ecosystems where frost heaving physically moves the sensors and wells to new positions. In these ecosystems, the active layer, which is the saturated layer above the permafrost that thaws in the summer and freezes again in the fall, is a primary ecohydrologic driver, impacting nutrient cycles, biological processes, and downstream water quality.

The Groundwater and active layer measurements at permafrost sites (DP1.20099.001) data product provides measurements of the active layer at permafrost sites. These include the depth to liquid water from the ground surface, if present, and depth to the active layer thaw collected using NEON Preventive Maintenance Procedure: AIS Goundwater Wells (AD[07]). Similar to groundwater elevation in stationary ecosystems, these data are useful for addressing water table related questions of groundwater availability and seasonal variability.

Field personnel visit wells on a weekly basis during the active layer thaw season and push the screened wells down by hand to the current thaw level. They manually measure the top of well to ground surface distance, top of well to water level, and ground surface to thaw depth (Figure 1). Field ecologists also record 10 additional measurements with a thaw probe within a 2m radius of each well to obtain an average thaw depth. There is also a free-form remarks field where technicians can record any additional notes about the field sampling day.

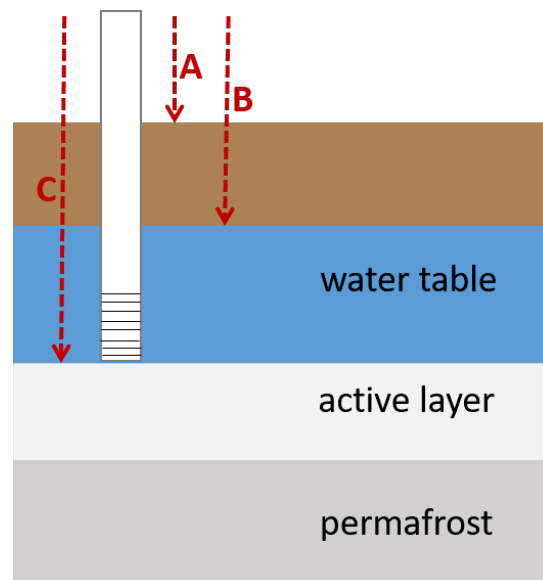


Figure 1: Diagram detailing the measurements made for Groundwater and active layer measurements at permafrost sites: (A) is the depth from the top of the groundwater well to the ground surface, (B) is the depth from the top of the groundwater well to the surface of liquid water, and (C) is the depth from the top of the groundwater well to the surface of the active layer at the bottom of the liquid water column.

3.1 Spatial Sampling Design

Groundwater and active layer measurements at permafrost sites is recorded at all groundwater wells located at the 3 NEON aquatic sites with permafrost (CARI, OKSR, and TOOK). Up to eight shallow groundwater wells per site are installed in inter-tussock areas at each site surrounding the lake and stream reaches to capture groundwater flow paths.

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

Groundwater and active layer measurements at permafrost sites measurements are conducted every week during the spring through fall at aquatic sites with permafrost in order to capture the seasonal change in active layer and liquid water depth.

3.3 Variables Reported

All variables reported from the field or laboratory technician (L0 data) are listed in the file, NEON Raw Data Validation for Groundwater and active layer measurements at permafrost sites (DP0.20099.001) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Groundwater and active layer measurements at permafrost sites (DP1.20099.001) (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/ncceas/projects/bien/wiki/VegCore>; accessed 16 February 2014), where applicable. NEON TOS 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

The finest resolution at which Groundwater and active layer measurements at permafrost sites data are reported is a specific groundwater well. Overall, this results in a spatial hierarchy of:

locationID (ID of the groundwater well) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain).

The groundwater wells are predominantly stationary over time. However, if a well can no longer be sampled, e.g. a well collapses, a new well location may be created to replace an existing well that will no longer be sampled.

Shapefiles related to the NEON Aquatic Observation System sampling locations can be found on the NEON science webpage at <https://www.neonscience.org/data-samples/data/spatial-data-maps>. If users are interested in the geospatial locations of the data relative to a global coordinate system, those can be retrieved using the NEON data API using the **locationID** and the following:

1. The `def.extr.geo.os.R` function from the `geoNEON` package, available here: <https://github.com/NEONScience/NEON-geolocation>
2. The NEON API: <http://data.neonscience.org/api>

3.5 Temporal Resolution and Extent

The finest resolution at which Groundwater and active layer measurements at permafrost sites data are reported is the **collectDate**, a single date on which measurements of active layer and liquid water depth relative to ground height are made. The total number of sampling events per year is expected to be 6 - 7 per aquatic site with permafrost.

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

The data from this L1 data product are substituted for Elevation of groundwater (DP1.20100.001) at permafrost sites due to frost heaving of the wells and sensors.

3.7 Product Instances

The NEON Observatory contains 3 aquatic sites with permafrost (CARI, OKSR, and TOOK), all located in Alaska.

For these three sites, Groundwater and active layer measurements at permafrost sites yields approximately 144 readings per year of the active layer depth and liquid water depth relative to ground surface, assuming that 8 wells are measured at each of the three sites every 2 weeks for a 3 month season.

3.8 Data Relationships

The protocol dictates that wadeable Groundwater and active layer measurements at permafrost sites take place once at each **locationID** per event (one record expected per **locationID** and **collectDate** combination in `agw_groundwaterFieldData`).

`agw_groundwaterFieldData` has one record per groundwater well per **collectDate**

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>.

4 DATA QUALITY

4.1 Data Entry Constraint and Validation

Many quality control measures are implemented at the point of data entry within a mobile data entry application 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 Groundwater and active layer measurements at permafrost sites(DP0.20099.001), provided with every download of this data product. Contained within this file is a field named 'entryValidationRules-Form', 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]]).

4.2 Automated Data Processing Steps

Following data entry into a mobile application or web user interface, the steps used to process the data through to publication on the NEON Data Portal are detailed in the NEON Algorithm Theoretical Basis Document: OS Generic Transitions (AD[14]).

4.3 Data Revision

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

4.4 Quality Flagging

The **dataQF** field in each data record is a quality flag for known errors applying to the record. Please see the *Special Considerations* section of this document for a list of known errors that may be present in the data, and below for an explanation of **dataQF** codes specific to this product.



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 |

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)