The National Hydrography Dataset
NHDinARC QuickStart

The National Hydrography Dataset (NHD) is a comprehensive set of digital spatial data that contains information about naturally occurring and constructed bodies of water, natural and artificial paths through which water flows, and related hydrographic entities. Within the NHD, **features** are combined to form **reaches**, which provide the framework for **linking (or geocoding)** water-related data to the NHD surface water **drainage network**. These linkages enable the analysis and display of these water-related data in upstream and downstream order.

This document provides a brief overview of the National Hydrography Dataset as implemented in ARC/INFO.[1] It is part of a series of NHD user documentation.

Want to make a map?

Features describe information found on USGS topographic maps.

Features

Features represent bodies of water, paths through which water flows, and related hydrographic entities. They are classified by **feature type** (item FTYPE).

Features are further described by **characteristics**, using these different methods:
- A unique five-digit **feature code** (item FCODE).
- Text aliases for the feature code recorded within the NHD.FCODE lookup table as a series of individual fields and, also, combined into a single **description field** (item DESCRIP).

Features are represented by points, lines, and areas:
- **Points:** NHDPT.PAT contains all point features (except underpasses)–gaging stations, wells, etc. **Underpasses** occur where features cross at different elevations. NHD.NAT contains all underpasses.

- **Lines:**
    - NHD.RATDRAIN contains features that comprise the surface water drainage network and coastline–stream/river, canal/ditch, artificial path, connector, and pipeline. The lines representing network features, through which water flows in a known direction, are oriented in the direction of flow. The lines representing coastlines are usually oriented so that the sea or ocean is to the right.
    - NHD.RATLM contains all other linear features–collectively described as "landmarks."

- **Areas:**
    - NHD.PATWB contains areal (or polygonal) features that contain water, or through which water flows–lake/pond, swamp/marsh, stream/river, canal/ditch, area of complex channels, estuary, ice mass, playa, reservoir, sea/ocean, and wash.
    - NHD.PATLM contains all other areal features–collectively described as "landmarks."

The **common identifier** (item COM_ID) uniquely identifies the occurrence of each feature.

Note: NHDinARC uses ARC/INFO routes and regions to uniquely identify and

manage groups of lines and polygons, respectively, as single entities.

Need to geocode and map observations? Need to model flow and geocoded observations through a drainage network?

Reaches and reach codes provide the starting point.

Reaches

Reaches are segments of surface water with similar hydrologic characteristics. A reach is composed of one or more features. A reach code uniquely identifies each reach (item RCH_CODE). The date of reach code assignment also is available (item RCH_DATE).

There are three types of reaches:
- NHD.RATRCH contains **transport reaches**, which are composed of the **linear** features that make up the surface water drainage network. If the direction of water flow is known, the lines representing the transport reaches are **oriented** in the direction of flow.

- NHD.RATRCH also contains **coastline reaches**, which are composed of the **linear** features that traverse the coastline. The lines representing coastline reaches are usually **oriented** so that the sea or ocean is to the right.

- NHD.PATRCH contains **waterbody reaches**, which are composed of **areal** features. In the initial release of the NHD, waterbody reaches are only composed of **lake/pond** areal features. Waterbody reaches are not part of the linear drainage network.

The **common identifier** (item COM_ID) uniquely identifies the occurrence of each reach.

Flow relationships establish the connectivity between transport and coastline reaches.

Flow relationships between transport and coastline reaches

Flow relationships link together individual **transport and coastline reaches** to form the surface water **drainage network**. Flow relationships describe:
- the **network connectivity** between reaches,
- the **direction of water flow** between transport reaches,
- the **traversal of the coastline** along coastline reaches, and
- the **order of termini** of individual drainage networks along the coastline.

NHD.RFLOW items COM_ID_1, COM_ID_2, DIRECTION, and DIR_TEXT describe the flow relationships.

In addition, **sequence numbers** convey the order in which transport reaches connect to the side of another transport or coastline reach (item SEQUENCE).

Need to identify the main paths through a drainage network?

With level paths, much of the work is already done.

Identifying level paths through the drainage network

A **level path** is the ordered set of transport reaches that trace the main path for a given flow of water.

Stream level identifies the main path to which a particular transport reach belongs (item LEVEL in NHD.RATRCH).

Delta level, which is computed from stream level, identifies main paths of water flow between flow relationships (item DELTA_LVL in NHD.RFLOW).

Geographic names help identify features and reaches and label maps.

Geographic names

A **geographic name** is the proper name, specific term, or expression by which a particular geographic entity is known (item NAME).

An identifier (item GNIS_ID) maintains the link to the **Geographic Names Information System**.

Reaches carry the geographic name most often; **features** only carry the name if they are not part of a reach.

Common identifiers hold together the National Hydrography Dataset and organizations planning to update the data.

Common identifiers

The **common identifier** is a 10-digit integer value that uniquely identifies each feature or reach (item COM_ID).

Common identifiers are the **basis for relating**:

- underpasses and features (items ABOVE_ID and BELOW_ID in NHD.NAT),
- reaches (items COM_ID_1 and COM_ID_2 in NHD.RFLOW),
- features and reaches (item RCH_COM_ID in NHD.RATDRAIN and NHD.PATWB),
- linear features and areal features that contain them (item WB_COM_ID in NHD.RATDRAIN), and
- features and reaches and their digital update units (items COM_ID and DUU_ID in NHD.DUU2FEA).

They are used to **communicate and share corrections and updates** among organizations.

Never change a common identifier.

What coordinate and measurement systems are used?

Coordinates and related measures

Latitude and **longitude** values are used for horizontal coordinates. The horizontal datum is the **North American Datum of 1983**.

Lengths (item METERS) and **areas** (item SQ_KM) of features and reaches, computed from a projected coordinate system, are supplied for convenience.

Elevations of surfaces (item ELEV) and the associated **stage** where water pools (item STAGE) are supplied for some features.

Curious about the data?

Metadata and digital update units

Metadata and digital update units provide the answers.

Metadata, or data about data, describe the content, quality, condition, or other characteristics of data. Metadata are recorded in **text files**.

In the initial release of the National Hydrography Dataset, metadata are organized into:

- a **general set of metadata** which applies to the entire NHD (file NHD.MET)
- **digital update units** (DUU), which are collections of features and (or) reaches to which a set of metadata applies. Two standard types of digital update units are quadrangle and cataloging units:
  - **quadrangle** digital update units contain metadata for features, and
  - **cataloging unit** digital update units contain metadata for reaches.

NHDDUU.PATDOM items DUU_ID, DUU_NAME, and DUU_DATE identify the unique identifier, file name (without the .MET extension), and creation date, respectively, for an individual digital update unit.

NHD.DUU2FEA associates features and reaches (item COM_ID) with their digital update unit(s) (item DUU_ID).

Data Quality and Lineage

Metadata associated with the **digital update units** are the best source of information for a particular location. **Variations** within the National Hydrography Dataset make it difficult to make general statements about data quality.

Statements of data quality **vary geographically**. Different general statements are made for (1) most of the conterminous United States, Hawaii, Puerto Rico, and the Virgin Islands, (2) the Pacific Northwest, and (3) Alaska.

For the conterminous United States (excluding the Pacific Northwest), Hawaii, Puerto Rico, and the Virgin Islands:
- Lineage:
  - Two main sources of information are USGS digital line graphs and EPA Reach File Version 3 data. The most prominent map scale for sources is 1:100,000; larger scale data are being integrated as part of data update activities.
  - Production steps included creating features, integrating artificial paths, associating features with reaches and geographic names, orienting features and reaches in the direction of water flow, constructing flow relationships, and validating the results of these efforts.
- Attribute accuracy assessments are based on those for the digital line graph data; geographic names and their identifiers are the same as those found in the Geographic Names Information System.
- Points, nodes, lines, and areas follow topological rules.
- The data reflect the completeness of the source materials; these sources vary in completeness.
- The accuracy of horizontal and vertical positions is based on the assessments for the source materials; these sources conformed to the National Map Accuracy Standard.

Similar approaches are being used to develop data for the **Pacific Northwest** and **Alaska**. Specific statements will be provided when these data are released.

How is the NHD distributed?

Data Distribution

"NHDinArc" data are distributed as tarred and compressed ARC/INFO workspaces. Each workspace contains the data for a single hydrologic cataloging unit. Cataloging units are drainage basins averaging 700 square miles (1,813 square kilometers) in area. Within a workspace, there are three ARC/INFO coverages plus several related INFO tables. There is also a folder containing the metadata text files.

After uncompressing and untarring a workspace, you will see the following folders:

   📂   07080204 (the number of the cataloging unit)
        📁   info
        📁   metadata
        📁   nhd
        📁   nhdduu
        📁   nhdpt

To make it easier to append data for adjacent cataloging units and to develop applications, each workspace contains ALL tables that may occur in the National Hydrography Dataset. Because an individual cataloging unit may not contain all items that have been defined for the National Hydrography Dataset, its workspace may have some tables that are empty and are 0 bytes in size.

Please use decompression software that will decompress files that are 0 bytes in size. If you use software that does not decompress files that are 0 bytes in size, such as WINZIP,[1] you will be able to view and use the data, but may have difficulties when appending data from other cataloging units or running others' applications software. A set of shareware utilities (gzip.exe and untar.exe) that meets this requirement is available through the National Hydrography Dataset Web site (http://nhd.usgs.gov). Gzip.exe will uncompress the tar.gz file and untar.exe will expand the uncompressed file into the appropriate directories and files. There are other such utilities available on the Internet.

Additional information

Visit the National Hydrography Dataset Web site at http://nhd.usgs.gov to obtain more comprehensive user documentation, data, tools, training materials, and technical support. Data are available in both NHDinARC and Spatial Data Transfer Standard (SDTS) formats.

_____

[1] Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.