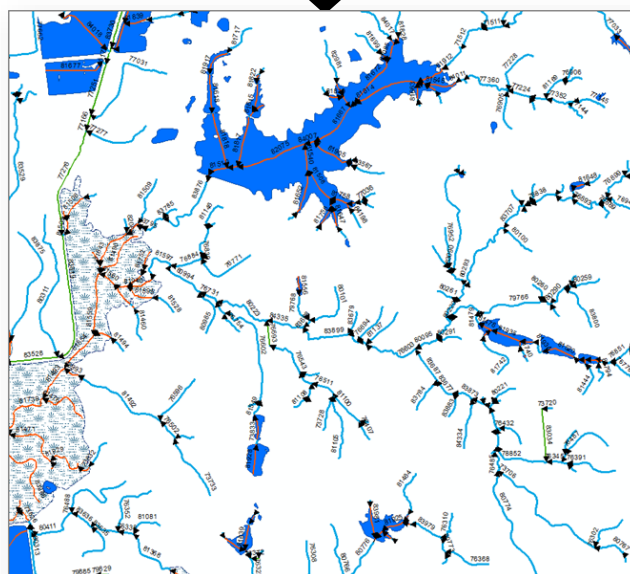
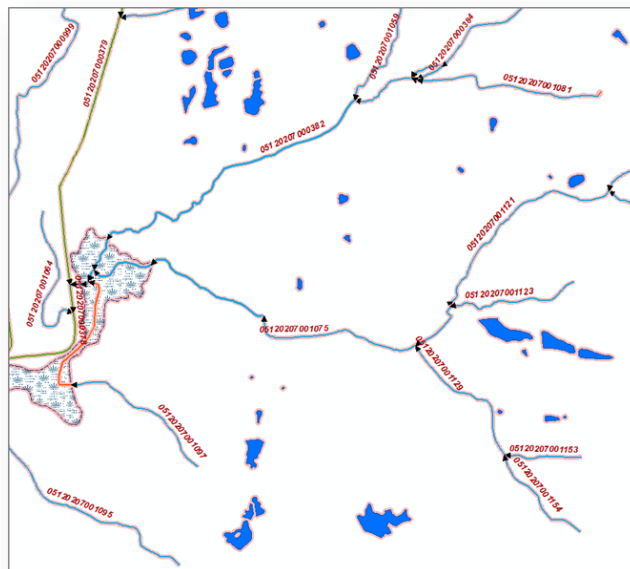


# Preparing Flowline Data for the NHD GeoConflation Tool





## Table of Contents

Overview of the National Hydrography Dataset.....	3
Hydrologic Modeling with Digital Elevation Models.....	4
NHD Stewardship Model.....	5
Bulk NHD Updates.....	7
Appendix: Workflow Diagrams .....	11



# Overview of the National Hydrography Dataset

The National Hydrography Dataset (NHD) is the authoritative national database for spatial information about surface water within the United States and is a component of U.S. Geological Survey's (USGS) The National Map. The NHD consists of vector-based geospatial data that represent the nation's drainage networks and related features, including rivers, streams, canals, lakes, ponds, glaciers, coastlines, dams, and stream-gages at a map scale of 1:24,000 or better. The majority of NHD data is derived from aerial or satellite image interpretation or GPS field survey. An increasing amount of the NHD, however, is comprised of lidar and IfSAR-derived streamlines that are more spatially and geometrically accurate than the traditional 1:24,000 scale data (Poppenga and others 2013, Poppenga and others 2009).

The primary product within the NHD is the High Resolution dataset, which contains all data mapped at a scale of 1:24,000 or better (<https://on.doi.gov/2Nlfl6O>). Formerly, there were publicly available Medium and Local resolution data, but the Medium data is now a product used and controlled within the USGS alone, and the Local resolution products have been merged with the High Resolution dataset. The NHD is also used as one of 3 primary inputs for the NHDPlus High Resolution Database (<https://on.doi.gov/2rmpcoB>).

Traditional use of the NHD ranges from resource management activities (e.g., fishery and forestry management plans); informing environmental regulation (e.g., mineral extraction activities); aiding in disaster response planning (e.g., flooding or chemical spills); to every day mapping requirements for recreation or travel. Federal agencies regularly using the NHD include the USDA, USDOl, USEPA, USDOT, and the USDOE. Each of the 50 states has one or more agencies that are responsible for stewarding content within the NHD for their state. The data is widely used within the tertiary education system across the country, as well as by for profit and non-profit agencies with an interest in the nation's surface water (e.g., Google and The Nature Conservancy).

As the U.S. Forest Service (USFS) continues to collect lidar data for National Forests across the country, there is increasing availability of high resolution (1-meter) digital elevation models (DEM) that can be used to model very accurate stream networks. Although this data has the potential to greatly improve the accuracy of the NHD, many within the USFS see the process of updating the NHD as a difficult, unintuitive process. Moreover, many are unaware of the variety of tools and workflows available to facilitate the submission of large amounts (e.g., multiple HUC10s) of lidar/IfSAR-derived flowline data to the NHD.

Accordingly, this User Guide serves multiple purposes: it provides detailed information on the stewardship model and the tools available to submit new hydrography data for integration with the production NHD dataset; it outlines the proper application of those tools based on the needs of the user; and it serves as an introduction to a set of step-by-step exercises that instruct users how to prepare flowline data for the GeoConflation tool (GCT), which is used by NHD Stewards to conflate existing attributes from NHD data with new, more geometrically accurate flowline data. The

preprocessing steps (pre-GCT) outlined in the exercises can make the NHD update process more efficient by implementing critical quality assurance/quality control (QA/QC) and data management procedures prior to the use of the GCT, thus streamlining the use of the GCT and NHD Update tools. If these QAQC procedures are not implemented prior to the use of the GCT, then the GCT process will fail, and the user will have to address the flagged errors and repeat the GCT workflow. It is important to note that this user guide does not provide instructions on how to use the GCT, nor does it provide the required training needed to gain access to the GCT and other production-based NHD tools. That official NHD training is provided by David Anderson ([danderson@usgs.gov](mailto:danderson@usgs.gov)) at the USGS.

## Hydrologic Modeling with Digital Elevation Models

Lidar and radar data are the primary sources of the high resolution DEMs used to model streamflow in the U.S. Recognizing the numerous ways that lidar data can be applied to inform management decisions, the USFS has collected millions of acres of lidar data in recent years (<https://arcg.is/vme5L>; accessible with FS AGOL account). The 1-meter DEMs that are typically delivered by lidar vendors have been shown to be extremely useful for deriving highly accurate flowlines (Randall and others 2017, Vaughan and others 2013).

Interferometric Synthetic Aperture Radar (IfSAR) is the primary source of radar-derived medium to high resolution digital elevation data. NASA's Shuttle Radar Topography Mission (SRTM) provides near global coverage of 30-meter DEM data (Rabus and others 2003), while higher resolution (5-meter) IfSAR data has been collected across large portions of Alaska (Carswell 2013), a state where lidar acquisitions are complicated by the consistent presence of cloud cover.

The USFS Geospatial Technology and Applications Center (GTAC) has training materials that provide detailed instructions on how to prepare DEMs for hydrologic modeling and how to use ArcMap's downloadable ArcHydro extension (<http://downloads.esri.com/archydro/archydro/Setup/>) to derive streams and rivers from the DEM. This training material can be found here: [https://fsapps.nwcg.gov/gtac/CourseDownloads/LidarDerivatives\\_ZIP.zip](https://fsapps.nwcg.gov/gtac/CourseDownloads/LidarDerivatives_ZIP.zip). Before the streams and rivers can be derived, the DEM needs to be hydro-enforced using ancillary data and tools available in ArcMap (see Appendix). Hydro-enforcement refers to the process of converting a topographic DEM into a hydrologic DEM that permits the simulated flow of water under bridges and through culverts (Poppenga and others 2010, Poppenga and others 2013).

Once the hydro-enforcement process is complete, flow direction, flow accumulation and stream definition ArcHydro tools are used sequentially to derive streams. This common ArcMap workflow for deriving flowlines from a high-resolution DEM is outlined by Vaughan and others (2013) and Randall and others (2017). Miller and others (2015) outline a similar workflow using an IfSAR-derived DEM along with additional open source GIS software. To submit these DEM-derived flowlines to the NHD, however, requires many additional steps.

# NHD Stewardship Model

The USGS administers and distributes the NHD, and although the USGS does undertake general maintenance activities on the data (i.e., the image integration and network improvement programs of work), states are generally responsible for conducting the bulk of NHD stewardship activities (i.e., feature updates). Stewardship of the NHD can be separated into 2 functional components, federal support and state stewardship. The federal component deals with the provision of infrastructure and technical support, while the state component deals with funding and the provision of resources to actually update the NHD. This division of labor within the stewardship model has come about for a number of reasons, but primarily is the result of database complexity and access to local expertise. Because the NHD is such a large database (i.e., covering the entire extent of the US, as well as borderlands within Canada and Mexico) and involves modeled features from an array of ecosystems, it was decided that a distributed stewardship model was the only way to leverage local ownership and subject matter expertise. As a result, the NHD is administered and regulated at the federal level by the USGS, while stewardship of the data within the NHD is typically the responsibility of agencies at the individual state level. This enables NHD Stewards within each state to utilize a shared common operations platform at the federal level while bringing to bear local knowledge about how surface water typically functions or appears on the landscape within their area of interest.

There are a number of key personnel within the NHD Stewardship framework. USGS Regional points of contact (POCs) serve as technical support for State Stewards and often coordinate training for new stewards; USGS Liaisons help states to coordinate mapping priorities, as well as funding and agreements to update the NHD; State Stewards are typically state or federal employees that oversee partner contributed updates to the NHD within their area of interest – stewards may do the actual update work, as well as oversee a group of sub-stewards who each update the NHD; USGS Technical POCs serve as subject matter experts for specialized stewardship activities and they provide support for all of the actors above, as well as serve as product owners for their respective NHD Toolset (NHD Update Tool, NHD HEM Tool, NHD GCT, etc.). Both USGS Regional POCs and USGS Technical POCs provide training for users and regulate user access to tools developed for the update process. To learn more about the key USGS personnel responsible for a given part of the US, refer to the stewardship website:

[https://www.usgs.gov/core-science-systems/ngp/national-hydrography/stewardship-and-community?qt-science\\_support\\_page\\_related\\_con=0#qt-science\\_support\\_page\\_related\\_con](https://www.usgs.gov/core-science-systems/ngp/national-hydrography/stewardship-and-community?qt-science_support_page_related_con=0#qt-science_support_page_related_con).

## USGS Stewardship Tools

Although there are a variety of tools available to qualified users interested in updating the NHD, it is not always clear what tools are appropriate for a given situation. This section is intended to elaborate on the role of each tool in the update process, as well as clarify what circumstances call for specific tools. The USGS website provides further detail on the tools discussed below and some tools that are not relevant to this user guide, but may be of interest: <https://www.usgs.gov/core-science-systems/ngp/national-hydrography/tools>.

## NHD Update Tool

This is the tool that every NHD Steward in the country uses to submit final updates to the NHD. It contains all of the core functionality required for NHD updates, and it can even be used to do bulk updates (via a stepwise delete and replace workflow) by advanced users. Typically, this tool is used for making manual changes to existing NHD data or replacing up to a few hundred features at once. It's a mandatory tool for anyone who wishes to submit data to the NHD, including Stewards, sub-Stewards, Technical POCs, Regional POCs, and any USGS editors affiliated with the USGS National Geospatial Technical Operations Center.

## NHD GeoConflation Tool (GCT)

The NHD GCT is used to migrate/conflate reachcodes, permanent identifiers and other attribute information from existing NHD content to new, higher resolution NHD content. This tool is typically used to facilitate bulk (i.e., several hundred to tens of thousands) updates to the NHD. However, it does not directly submit final updates to the NHD; instead, the qualified analyst must subsequently use the NHD Utilities and NHD Update Tools to officially submit updates.

In general terms, the GCT ingests two datasets, the new *target* dataset and the original *source* dataset, then transfers attributes from the existing source dataset to the target dataset, thus maintaining existing attribute data while improving the geometry of flowlines (see Appendix). Anderson (2017) provides nuanced information about the GCT workflow.

## NHD Utilities

The NHD Utilities is a software suite of advanced tools for performing specialized tasks with the NHD data. This suite of tools includes: XML Extract, Export to Geodatabase (XML2GDB), Network Builder, Flow Table Builder, M-Value Utility, Geodatabase to Shapefile (PGDB2SHP), Shapefile to Geodatabase (SHP2GDB), and a Merge Database utility.

## Subset by Polygon Tool

The Subset by Polygon tool is an Esri Add-In that allows a user to create a subset of NHD data from any size NHD geodatabase. The tool uses a custom area polygon to define the area of interest (AOI) and will extract all features of the Hydrography feature dataset that intersect the AOI or have the same reachcodes as the intersecting features, allowing the user to see the full extent of the reachcode they are interested in. This tool would be used by Stewards wishing to update a subsection/portion of an NHD checkout job. NHD checkouts occur on the basis of WBD HU8 and HU10 boundaries – this tool lets users interact with a subset of data in those boundaries, without affecting the entire checkout area.

## NHD Flowcheck Tool

The NHD Flowcheck tool is an Esri Add-In that performs a series of functions that validate the flow in the target data (i.e., the data which will be put through the GeoConflation workflow) based on the digitized direction and assigned flow direction values. This tool would be used by any Stewards wishing to undertake the GeoConflation workflow.

# Bulk NHD Updates

Although there are skilled GIS users who can model flowlines from a variety of sources, a fraction of those individuals are trained in the preparation of flowlines for GeoConflation with the NHD or standardized NHD update workflows. In order to foster a greater understanding of this process, this user guide introduces an informal preprocessing workflow implemented prior to using the GCT. This workflow, which was documented by the NHD Principal Steward for Alaska, Mike Plivelich, will make the GeoConflation process more efficient by correcting errors or incongruities within the flowline data that would otherwise invalidate the data for use within the GeoConflation workflow. If implemented correctly, these preprocessing steps will eliminate the need to run the GCT multiple times until all errors are identified and corrected.

To understand where this preprocessing workflow fits in to the larger NHD update workflow, figure 1 maps the general steps from acquiring lidar or IfSAR derived elevation data from a vendor to the eventual submission of DEM derived flowlines to the NHD. GTAC's preprocessing workflow is displayed in red.

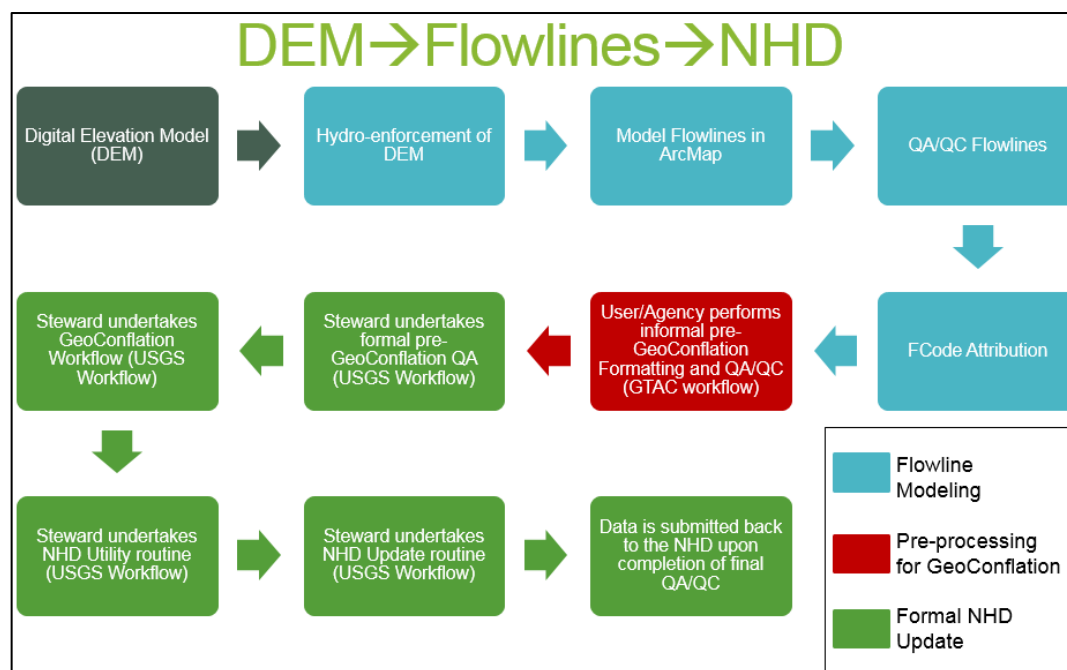


Figure 1: General Workflow Steps

## Prerequisites for Pre-GeoConflation Data Processing

Prior to preparing the data for GeoConflation and NHD Update, there are a couple of prerequisites that must be met, including the QA/QC of derived flowlines and the attribution of FCodes to those flowlines. The QA/QC process is typically performed by a hydrologist who is familiar with local hydrology, and it entails field visits or, more commonly, using high-resolution imagery or DEM derived hillshade rasters to assess the validity of flowlines. One of the primary things they assess is whether modeled flowlines are

likely to contain water. According to experts familiar with this QA/QC process, this largely manual process is the most labor intensive and costly part of the whole NHD update workflow.

Another prerequisite is the interpretation of flowlines and subsequent categorization based on their physical form and function. These categories are expressed as five-digit integer values, known as FCodes. FCodes are hierarchical codes that specify a general feature type (e.g., canal, coastline, pipeline, stream/river, or underground conduit) and then a more specific definition (e.g., perennial stream, intermittent stream, or ephemeral stream). An example of an FCode is 46003, where '460' specifies the 'stream/river' feature type and '03' specifies an 'intermittent' stream/river. A complete list of FCodes used for the NHD features can be found here: <https://bit.ly/2slzeXW>.

To execute the full pre-GeoConflation workflow, you will need an ArcMap 10.5 license and an installation of the Production Mapping extension. We recommend that FS employees contact the Customer Help Desk for assistance with installing ArcMap and the Production Mapping extension.

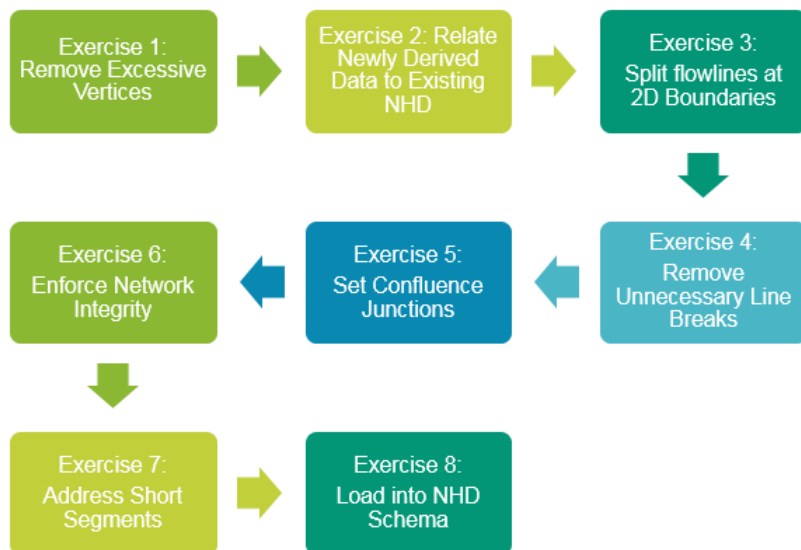
## Pre-GeoConflation Data Processing

There are several topologic criteria that must be met prior to importing the data into the NHD data model. The conditions include flowlines being contiguous features between confluence points; flowlines breaking at relevant polygon boundaries (i.e., lakes, 2D stream features, glaciers, etc.); the removal of duplicate features; and removing short segments and/or cutbacks. In addition, there is a limit on the number of vertices that are allowed per features. With traditional data, this was mostly never a problem, but lidar derived flowlines can be overly complex, and therefore need to be thinned to an acceptable extent to be ingested into the NHD (Anderson 2017, p.22-27).

Once the data have met topologic criteria, they can be loaded into an empty NHD data model template (File Geodatabase (FGDB) format) and coded appropriately. Generally, this involves loading the data into the FGDB by feature type and then using the field calculator to code mandatory attributes in the NHD data template (Anderson 2017, p.27). These steps are detailed in the exercises developed by GTAC, which are outlined in figure 2.



## Pre-GeoConflation Data Processing



**Figure 2: Workflow Diagram for GTAC Workflow**

Upon completion of the pre-conflation data processing, the data are then ready for the USGS standardized pre-conflation QA workflow. This workflow, and that of the subsequent GeoConflation workflow, are available in the form of a user guide document that can be obtained from any of the actors noted in the above section discussing the NHD stewardship model. To request formal GeoConflation training, contact David Anderson: [danderson@usgs.gov](mailto:danderson@usgs.gov).

# References

Anderson, D. 2017. NHD GeoConflation Tool (GCT) User Guide. U.S. Geological Survey, Partner Support Section, Production Branch, National Geospatial Technical Operations Center.

Carswell Jr., W. 2013. The 3D Elevation Program—Summary for Alaska. U.S. Geological Survey Fact sheet 2013-3083. 2p.

Miller, D.; Benda, L.; DePasquale, J. and Albert, D. 2015. Creation of a digital flowline network from IfSAR 5-m DEMs for the Matanuska-Susitna Basins: a resource for NHD updates in Alaska. Report to The Nature Conservancy. Terrain Works Inc.

Poppenga, S.; Worstell, B.; Stoker, J.; and Greenlee, S. 2009. Comparison of surface flow features from lidar-derived digital elevation models with historical elevation and hydrography data for Minnehaha County, South Dakota: U.S. Geological Survey Scientific Investigations Report 2009–5065, 24 p.

Poppenga, S.K.; Worstell, B.B.; Stoker, J.M.; Greenlee, S.K. 2010. Using selective drainage methods to extract continuous surface flow from 1-meter lidar-derived digital elevation data: U.S. Geological Survey Scientific Investigations Report 2010–5059. 1-12. <http://pubs.er.usgs.gov/usgspubs/sir/sir20105059>.

Poppenga, S.K.; Gesch, D.B.; Worstell, B.B. 2013. Hydrography change detection—the usefulness of surface channels derived from LiDAR DEMs for updating mapped hydrography. Journal of the American Water Resources Association. 49(2): 371–389. <http://dx.doi.org/10.1111/jawr.12027>.

Rabus, B.; Eineder, M.; Roth, A. and Bamler, R. 2003. The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. ISPRS journal of photogrammetry and remote sensing. 57(4): 241-262.

Randall, J.; Schaaf, A.; Clark, J.; Dickinson, J. 2017. Identifying and extracting features from a lidar-derived DEM. RSAC-10113-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 8 p.

Vaughan, R.; Miller, J.; Gubernick, R.; Fisk, H. 2013. Lidar enhancement of NHD flowline dataset in the Eastern Region. RSAC-10024-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 20 p.

## Appendix: Workflow Diagrams

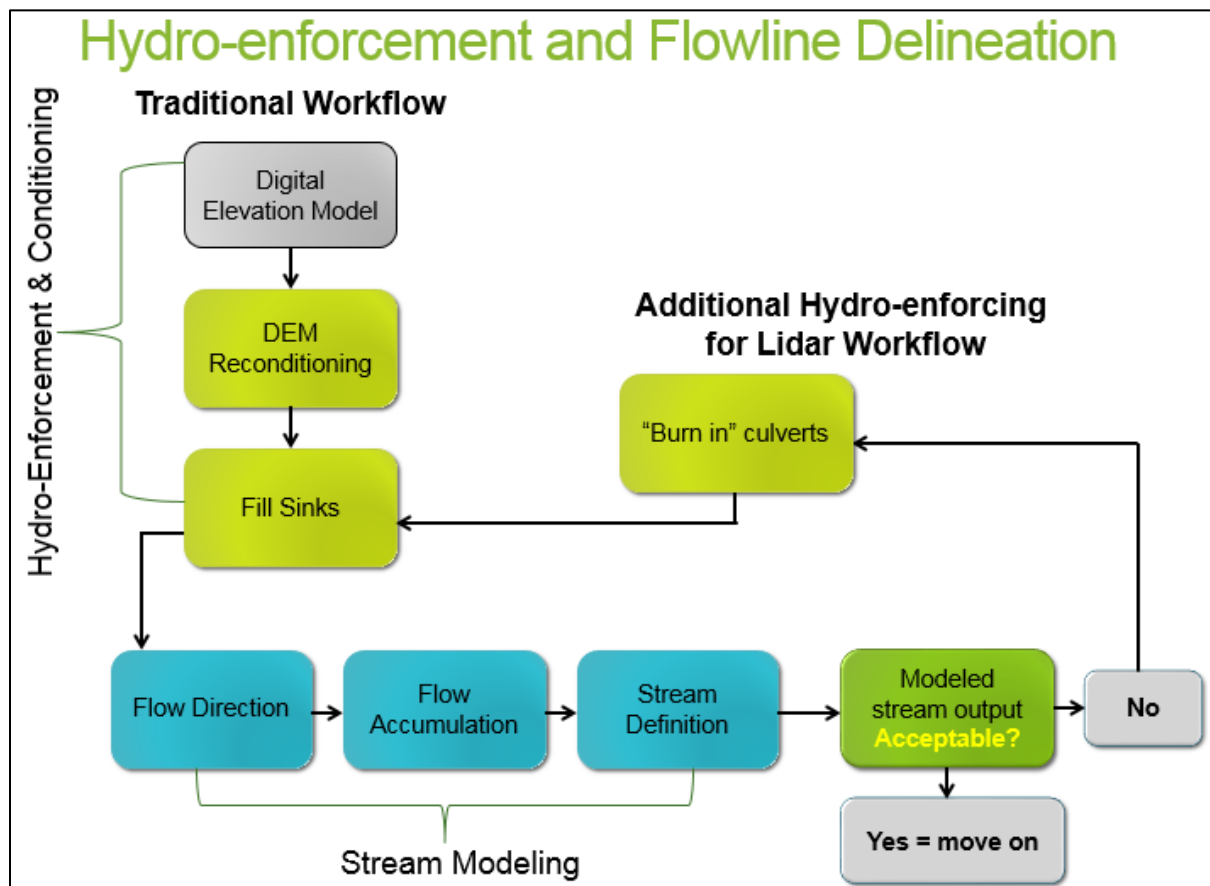


Figure 3: GTAC's Hydro-modeling Workflow

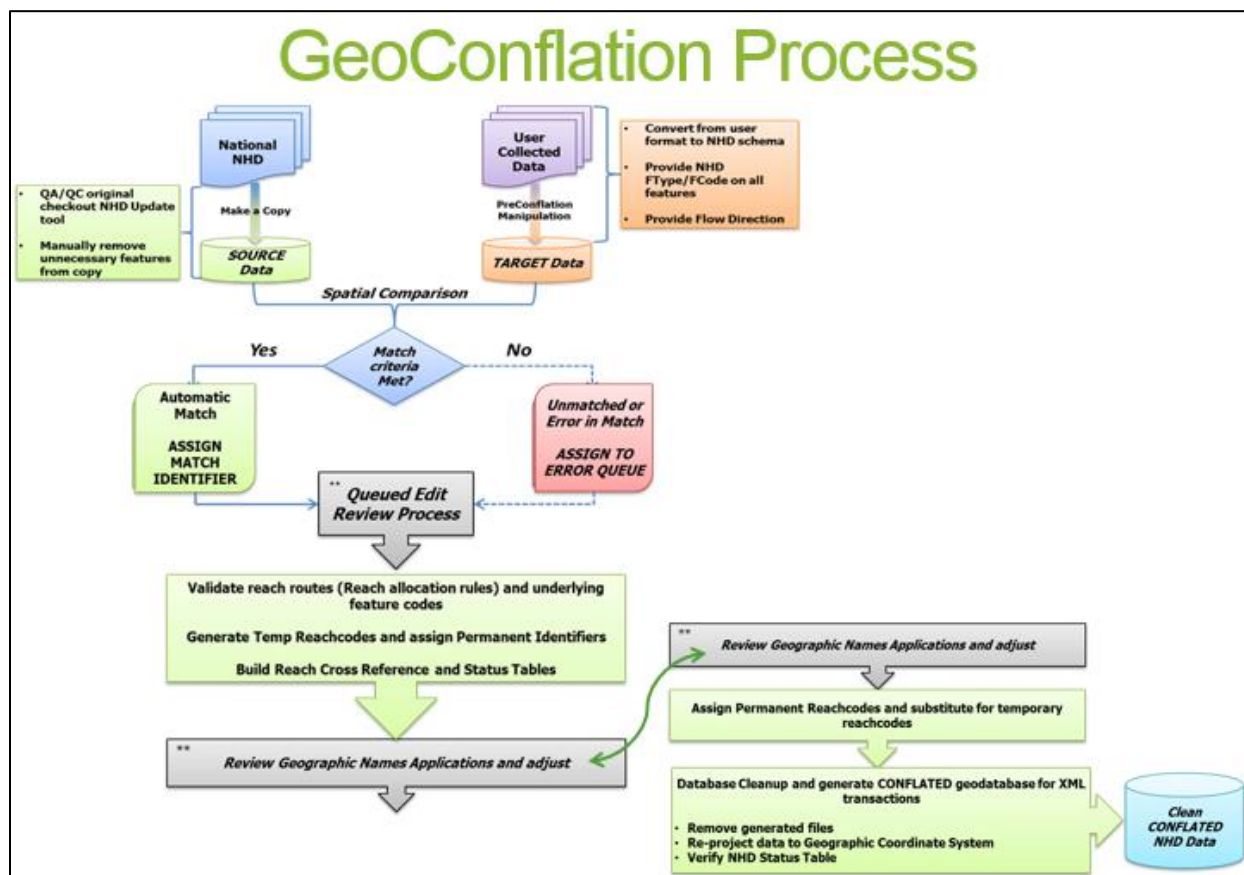


Figure 4: USGS GeoConflation Workflow—courtesy of David Anderson, USGS.