USGS National Hydrography Dataset Newsletter
Vol. 3, No. 4, February 2004
by Jeff Simley, USGS

The Geodatabase Conversion

The creation of the NHDinGeo is currently processing on schedule and should be available April 1st. Preliminary reviews of the data being converted show that the data is very easy to load and use in ArcMap using standard ArcGIS tools. The next NHD Newsletter will provide the information you will need to obtain and use this new model.

Hydrography Enforced Elevation Data

Hydrologic modeling can make use of digital elevation models to describe the form of the landscape and derive channels, channel cross-sections, channel gradients, and catchment areas. Modelers are also very interested in using the NHD as the basis for the stream network with the elevation model. This raises the issue of the ability of the two datasets to integrate. We can better understand this issue by taking a closer look at USGS elevation data. Basically, there are five types of elevation models produced by the USGS: (1) Digital Elevation Models (DEMs) with a 30-meter matrix derived from direct photogrammetric profiling, (2) DEMs with a 30-meter matrix cartographically derived from contours, (3) DEMs with a 10-meter matrix cartographically derived from contours, (4) National Elevation Data (NED), a seamless national coverage made from the best available (either 30- or 10-meter) cartographically derived DEM's and resampled to a 1 arc-second grid, and (5) NED made solely from 10-meter DEMs and resampled to a 1/3 arc-second grid. The 1 arc-second NED uses source 10-meter DEM's where available, replacing the 30-meter DEMS. The first type of DEM, known as a Level 1, is being phased out with only 100 or so remaining. It can be identified by a processes_code=1 or 2 in the header. The second type, known as a Level 2, has a process_code=3, 4, 5, 6, or 7 and is also typically one megabyte in size. The file size is related to the resolution. This type of DEM is created from the vector contours found on the 7.5-minute series topographic map. The contours on these maps logically "fit" the hydrography. In fact, the contours were normally drafted using the hydrography as a control guide. Since the Level 2 DEM and the NHD are often made from the same source, the two should integrate well, but not always perfectly. To improve the situation, the third type of DEM is being made, also a Level 2, and is typically about 10 megabytes in size. This improves the fit simply because it has more detail. But more importantly, the majority of this type have had hydrography data "burned" into the DEM in a process called hydrography enforcement. This can be identified by in the header with the dem_free_form_text field containing the words "DRAINENF" or "HYDRO ENFORCED". This activity began in the 1997-1999 timeframe. In this process, 1:24,000-scale vector hydrography from the 7.5-minute series topographic map was integrated with the vector contours used to create the elevation matrix. In creating the matrix from the vector data, the pixels immediately adjacent to the vector stream, typically 1-3 cells-wide, are adjusted so that the cell's elevation value is based on a continuous gradient from the larger to smaller contour values intersected by the stream. This overrides the normal cell elevation assignment algorithm and has the effect of "forcing" or "burning" the stream channel into the model. This means that any surface runoff on the DEM will more logically accumulate into streams identified on the topographic map or subsequent revisions. Vector polygons, such as lakes, also force a more consistent elevation assignment to the cells to create the flat lakes we all prefer. As a result, the high-resolution NHD should integrate rather well with this type of DEM since the source used to make the NHD is usually the same source that "enforces" the DEM. Many of the DEMs created since 1999 have improved hydro enforcement where the hydrography data was incorporated during the initial DEM creation as well as "burned" in during DEM editing. Accurate integration can be affected if two different hydrography sources are used. If you are using the medium-resolution NHD, these vectors have been generalized and are not designed to integrate with DEMs or the NED. The first three types of elevation data are produced in 7.5-minute tiles. When combining two adjacent tiles, the edges between the two may not match or may have gaps due to the native-DEM-format UTM grid tilt. To overcome this problem, the NED was created to produce a seamless nationwide coverage by matching the edges and re-sampling to a more useful grid. The NED has become the elevation source of choice for scientists. It is possible to determine if the data you are using from the 1 arc-second NED has been hydro enforced. You can do this by examining the attributes of the spatially registered metadata (which comes with the NED DEM data) for the "FREETEXT" field, which will reflect the source DEMs freeform text field containing the words "DRAINENF" or "HYDRO ENFORCED". You can learn more about the NED at http://gisdata.usgs.gov/NED/default.asp. There are also various post-production processes for hydrography enforcement, which will be covered in a future newsletter.

User Experience with the NHD in Alaska

The National Hydrography Dataset is being produced for Alaska from 1:63,360-scale source material, which has undergone a recent photogrammetric revision program. As each of the 133 subbasins is produced, they undergo a quality assurance review by the USGS prior to distribution. A follow-on quality review effort is conducted by the Alaska Geographic Data Committee's Hydrography Subcommittee to provide a user's perspective on the suitability of the data in scientific analysis. The conclusion is that the Alaskan NHD meets the requirement to represent Alaskan hydrography for cartographic purposes, with some limits on network navigation and attribute addressing. Any flow and connectivity errors will be corrected later when the new NHD in Geodatabase coverage is available. Other errors in the data set are being addressed through a local maintenance project. While the Alaska NHD is suitable for most broad scale work, errors in the data can cause problems at the very local level. For users in Alaska, stream naming is very important and any errors must be resolved. These errors typically involve (1) unnamed reaches that create a gap in the consistent naming of a stream, (2) a name completely missing from a stream, (3) a misnamed stream, (4) unnamed reaches in combination with missing reaches, (5) a name routed up a tributary rather than the main path. From three sub-basins currently completed in the maintenance project, a total of 5,282 named reaches were observed and 182 of those were in error. For example, in the 19020501 subbasin, the following errors were found: (1) Horsepasture Creek was not applied, (2) the name Spray Creek was applied to the wrong set of reaches, (3) one of the 347 reaches within the Susitna River is missing the name, (4) questionable associations of Coal Creek/ Fox Creek, (5) some ambiguous endpoints of Tolson Creek, and (6) Willow Creek - a short tributary to Valdez Creek - does not appear in the data. Names in the NHD are based on the Geographic Names Information System (GNIS) and applied in a semi-automated process during production. The cartographer applying the name may make an error due to unfamiliarity with the local geography, which is why a review by the user community is important. Such problems can be expected in production because of the large volume of data and the complexity of stream networks found in hydrography. This is why the NHD was designed around the requirement for data maintenance, using feature identifiers so that individual features could be edited and replaced when found in error. This "transactional maintenance" greatly simplifies the maintenance process. This process is still evolving, so the Alaskan quality review program makes the necessary edits for local use and then records the features affected so that permanent changes can be made later. In a fully mature transactional maintenance scenario, the edit made by a data steward will directly correct all data holdings in periodic uploads. Other problems being edited include: (1) missing reaches creating an interruption of the network, (2) a spatial gap in a feature [com_id], (3) missing features, e.g. ice-ocean or an entire stream-river, (4) connecting reaches that should not have been included (5) errors in assigning the correct hierarchal level to reaches in the network and, (6) reaches migrated to an adjacent sub-basin. NHD leveling, similar to stream order, can be in error. This is an automated process that can be thrown off-track in complex situations such as stream braiding. It is important that local stewards review the results. A single error can have a cascading effect in the network. In one example, an error resulted in 9,997 of 26,639 reaches being misleveled. Often a few simple corrections can clear the remaining problems. Good statistics, and a good understanding of those statistics, will become available as the Alaska NHD maintenance project progresses. This will allow users and producers to determine how actual error rates correspond to product specifications. The experience in Alaska is likely similar to problems found elsewhere. Producing the NHD, like any production process, is going to have a certain degree of error due to normal variation. The program strives to minimize that variation, but some error will always exist until the user community applies their specialized expertise.

Using the NHD at the USGS

Experts in water resources at the USGS recently got together to discuss the use of the NHD in science applications. The perspective of the group largely revolved around hydrologic modeling. Here is a synopsis of the second half of the discussion.

<u>Hydrologic Modeling Considerations</u> – One hydrologic modeling system is the USGS Modular Modeling System (MMS), which is based on the premise that each modeling problem is unique and best served by a customized model created from a library of modularized components chosen by the user. It is freely available over the web and used by the USGS, other bureaus, and universities. The RiverWare system models a network of nodes where each node generates parameters based on a localized model. Here other spatial databases are used such as vegetation density, land use, soils, and radiation to account for the rate of runoff. The user can select any location and an accumulation basin is determined for that point. This basin looks at a series of parameters and the model generates the predicted flow data at the requested point. Often, modeling begins as a drainage network, typically derived from Digital Elevation Models. Then the characteristics of the network channels are determined such as slope, length, roughness, and cross-section. Models determine the characteristics of a basin, which can be an aggregate of the thousands of underlying catchments. Models can be exclusively surface oriented, ground water oriented or include the interaction dynamics of the two. Hydrologic modeling can be based on spatial data or simply an array of mathematical parameters. Even when modeling is performed using spatial data, the sophisticated portrayal of the stream network can end up getting reduced down to a generalized set of parameters that feed mathematical algorithms. We have now come to the intersection of hydrologic modeling, GIS, and the NHD. It is clear that hydrologic modeling and GIS have become integrated, and now choices have to be made about where the spatial data comes from: elevation-derived streams, the NHD, or other sources. Up to now, elevation-derived streams have been used, but the advent of the NHD, particularly the high-resolution NHD, is getting the attention of the hydrologic modelers because it represents how water has actually been observed on the earth, as opposed to its theoretical location. It seems clear that there is a role for elevation data and the NHD. This then highlights the importance of the successful integration of the two.

<u>What Can We Do With Future Capabilities?</u> – The NHD can contribute a number of important characteristics to spatial hydrologic modeling. One is its addressing capability that allows users to give data a locational address on the network, not unlike addresses on a street. Streamgages, water quality sampling sites, and meteorological data can be located on the stream network using the NHD's reach indexing system. With the data addressed to the network, it is then possible to travel throughout the network to and from this data. The NHD also characterizes the stream network. It gives us an idea of which streams tend to be perennial or intermittent. It tells us where streams become wide or braided. It tells us something about density. It tells us the location and perimeter of lakes and ponds, as well as the observed locations of marshes and swamps. It also provides the network of canals and ditches which can modify the natural network. The NHD provides a level of national consistency. Basically the hydrography represented in Florida uses the same rules as the hydrography represented in the State of Washington, thus the similarities and differences are consistent in portrayal. The types of features represented, the attributes, the accuracy, the density, and the topology rules are all generally consistent. There are localized anomalies, but on a broad canvas, the national portrayal of hydrography with the NHD is consistent ocean to ocean, border to border. Another characteristic of the NHD is the strong degree of clean topology and network flow. Common rules for topology of the data are present. Reaches

connect at defined nodes, divergence and convergence is maintained, and the looping of flow is corrected. Also, the names of water features found in the Geographic Names Information System are linked to the NHD to provide valuable reference naming information. (Other major characteristics not mentioned at the meeting include permanent feature identifiers, transactional maintenance mechanisms, and data versioning). In addition to some of these characteristics, many more characteristics can be added to the NHD based on linkage with other datasets, such as watersheds, stream slopes, dams, drinking water intakes, and inundation areas.

Where Do We Go From Here? – Users in hydrology have a significant use for the NHD and can be important contributors to a critique of the NHD for improvement efforts, provide needed design characteristics, review quality, provide links to other users, and create opportunities for the NHD. Here are some issues important to the successful use of GIS in hydrography with an emphasis on the NHD: (1) good integration between the NHD and NED is needed, (2) both flow-enforced and raw NED is needed, (3) different types of flow-enforcement algorithms are needed, (4) some users are enforcing NED without a mechanism to redistribute it, (5) NHD flow must be correct, (6) catchments and characteristics for each reach will be needed, (7) many projects are local and call for higher resolutions of the NHD, (8) the states will push EPA to move towards 1:24,000-scale NHD, (9) there will be some problem with the conflation of medium and high-resolution reaches, (10) regulations will be written based on the NHD hydrography, (11) the NHD is too complicated for some users, (12) a simplified NHD should be available, (13) NHD web services will be important, (14) hydrologists do have a strong interest in the medium-resolution, (15) concern over the lack of maintenance of the medium-resolution, (16) thinning and generalization is needed to create national scales, (17) medium-resolution maintenance would have to be very simple, (18) users need to keep track of data snapshots, (19) high and medium-resolution will not be in sync, (20) EPA will maintain the medium-resolution to assure basic needs, (21) isolated streams need to be linked to the network, (22) perennial-intermittent classifications are important, (23) DRG and NHD integration are important, and (24) the NHD needs to be user-friendly. Two current initiatives will greatly help improve the NHD: (1) the high-resolution program and (2) the move to geodatabase. Subjects requiring more discussion include: used of Elevation Derivatives for National Applications (EDNA), the use of DRG's, and training/education.

High Resolution Status Web Site

As part of the continuing evolution of the NHD, we have a new status site located on the NHD web page http://nhd.usgs.gov. Click on Data and then NHD Status Map, or just click on: http://webhosts.cr.usgs.gov/nhdstatus/viewer.htm. You can also still use http://statgraph.cr.usgs.gov.

Thanks to Pete Schneider, John Kosovich, Mike Kling, Steven Rich, Dennis Tol, George Leavesley, Roland Viger, Dave Wolock, Curtis Price, Keven Roth, and Sandy Piksa.
Jeff Simley, USGS, assumes full responsibility for the content of this newsletter.