USGS National Hydrography Dataset Newsletter
Vol. 3, No. 9, July 2004
by Jeff Simley, USGS

The NHD and ArcGIS 9

The National Hydrography Dataset (NHD) has recently gone through a major change to utilize the Geodatabase provided in ArcGIS 8.3 as the foundation for the NHD model. This change, now available as NHDinGeo, was done to take advantage of advances in GIS technology that could better support the NHD. Now, the company behind ArcGIS, Environmental Systems Research Institute (ESRI), has announced that a version step is being made to ArcGIS 9. What does this mean for the NHD? As far as we know, not much. Certainly there will be performance improvements that will require changes, but much of this will be "behind the scenes" type of changes not particularly evident to the typical user. The biggest concern is that, when NHD does switch to ArcGIS 9, the Geodatabase data produced will not be compatible with those users still using ArcGIS 8.3 platforms. So it is important not to switch too soon. ArcGIS 8.3 era data will be usable in ArcGIS 9. Advancements in ArcGIS include better web capability, the use of scripting languages, improvements to ArcSDE performance, and easier integration of modeling. ArcGIS 9.1 includes the ArcGIS Server to support synchronization of distributed databases. When will NHDinGeo switch? Not anytime soon, but such modernization is inevitable and will help the NHD achieve better performance. Look for updates in future NHD Newsletters.

The NHD and Stream Levels

Stream Level is an attribute used in the NHD to aide in network navigation. Basically, the reach emptying into the ocean is a level 1. The reaches that connect to make the longest route from the ocean are also coded as a level 1. The tributaries off the level 1 route are then assigned a level 2. This process repeats throughout the network of tributaries. The farther removed from the main route, the higher the level value. This is particularly useful in navigation because it is possible to find the main route upstream simply by following the current level. For example, if you are on a level 3 route and you want to find the main path upstream, you simply follow all the level 3's. That will give you the main path. The same thing applies downstream, and will allow you to follow the main path through braided streams. Normally, the main path follows the named path. Level is pre-calculated when you download a dataset. Levels disappeared when the switch was made to Geodatabase, because the old navigation technology was abandoned in favor of ArcGIS Network Analyst. NHDinGeo and NHDGEOinArc users will note there are no levels. However, there is now a growing demand to retain the level attribute. By using level along with the flow table, it is possible to navigate and create routes by simply processing a table outside of the GIS, a technique still useful. So, stream level will return. It is being reinstated in a joint USGS-EPA effort that should be completed in late August. At that time, hydrologic unit-based downloads will include stream level. Look for an update in a future NHD Newsletter on availability.

Preparing Data for NHD Integration

Many of us who use the NHD have forgotten the days when assembling a hydrography GIS meant the painstaking process of piecing together 7.5-minute DLG quadrangles, and the frustration of locating "the missing quad." Those days have never actually disappeared, but have simply been absorbed by a specialized team of about a dozen people within the USGS. The integration of high-resolution hydrography data into the NHD begins with the search for source data, which is still organized by 7.5-minute cells, and as is almost always the case, there is at least one missing quad. Thanks to a strong cooperative program by the USGS and its partners to produce hydrography DLG's in the 1990's, many of the primary areas of interest are covered. Add to that an equally strong program at the U.S. Forest Service to produce hydrography Cartographic Feature Files over Forest Service lands, plus programs by

several States to obtain statewide DLG coverage, and the map of 1:24,000-scale hydrography coverage of the country looks fairly impressive. Unfortunately, the correlation of hydrologic unit boundaries to political boundaries, not to mention quadrangle boundaries, is rare. So when a subbasin is to be produced, there is inevitably not one, but many missing quadrangles. A search for additional data is conducted, and in some cases, this is successful. But more often than not, the missing data must be produced from scratch by scanning 7.5-minute topographic maps. As a general rule, the typical subbasin requires original production of one-third of the quadrangles. This is done by producing Tagged Vector-Hydrography files, a slightly stripped-down version of the DLG. Besides the task of assembling the source data, a lot of effort is put into financing the project. Almost all of the 2,254 subbasins in the U.S. are produced as a cooperative venture with a USGS partner. This means that equitable financial or worksharing arrangements must be established and documented. Then the real fun begins. The source data must be reviewed and "cleaned" to make sure it can be easily conflated with the foundation 1:100,000-scale NHD to produce the 1:24,000-scale high resolution NHD. A typical subbasin will have about 36 quadrangles, but this can expand to well over 100 at times. Within this batch of data there will be varying formats, datums, UTM zones, vintages, standards, and a host of quality variability – all which must be unified into a single seamless NHD subbasin. The process starts with format conversion to a standard used by the pre-processing systems. Then the header must be edited to make sure all fields are correct. Then registration is checked and corrected as needed. Content is quickly checked in correlation with the latest DRG. Edge matching between quadrangles is checked and blunders fixed. Perfecting the edge matches is not attempted due to cost constraints. The largest effort is placed on cleaning topology. First on the list is closing polygons and eliminating overlap. Then linework snapping is checked. After the topology process is complete, a review of the attributes is made. Like many quality assurance functions, this is done largely by automated routines. The quadrangles are then loaded into the NHD production system where additional automated quality control checks are made. Datum shifts are performed and UTM zones are unified. Then polygon centerlines are created for the quadrangle data. These quadrangles are then combined into a single master file for the subbasin. The subbasin boundary is then used to clip the data to its correct hydrologic unit boundary. This involves interactive judgment to make sure features logically connected to the subbasin network are included in the clipped file despite the boundary being used. Those features previously subdivided by quadrangles are collapsed into single features. Finally, the coordinate strings of the linear features are oriented to flow downstream. With this done, that data file is now ready to be conflated with the 1:100,000-scale NHD to assign reaches and eventually evolve into the NHD. The above processes vary to some degree by producer and system. The time it takes to do this also varies considerably by geography, quality of the source data, and the skill of the editor. The work of a knowledgeable editor in preprocessing can play a big role in how smoothly the NHD production process will go. When the NHD is produced by a partnering organization, they will perform the later half of these steps. With production of a nationwide NHD, the steps outlined above will have been accomplished and hydrography coverage of the Nation will have evolved to a new level. Although the NHD is produced subbasin-by-subbasin, the master NHD Geodatabase becomes a completely seamless nationwide coverage. The piecing together of disparate quadrangles into a GIS coverage will soon join the annals of the slide rule and punch card.

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

Thanks to Paul Wiese, Stan Obert, Paul Kimsey, and Sandy Piksa.

Jeff Simley, USGS, assumes full responsibility for the content of this newsletter.