Continuous Basin Characterization using EDNA

Kristine L. Verdin
SAIC / EROS Data Center

Susan K. Greenlee
USGS/EROS Data Center

Abstract

The Elevation Derivatives for National Applications (EDNA) database has been developed, in part, to provide modelers with consistent DEM-derived layers for use in basin characterization. As an alternative to traditional basin characterization approaches, an innovative technique, using the EDNA flow direction matrix, has been developed which promises faster characterization and the ability for characterization above any location, not just predefined basin outlets. This technique is described as well as a practical application completed for the Pacific Northwest.

Introduction

The Elevation Derivatives for National Applications (EDNA) project is a multi-agency effort to develop standard topographically-derived layers for use in hydrologic and environmental modeling. The EDNA takes advantage of the seamless and filtered characteristics of the National Elevation Dataset (NED) to create a hydrologically conditioned Digital Elevation Model (DEM) useful for modeling applications. The goals of the project are to create a hydrologically conditioned DEM at 30-meters cellsize for the United States, systematically extract a set of standard derivatives and vertically integrate the data layers with other USGS framework datasets such as the National Hydrography Data (NHD) and the Watershed Boundaries Dataset (WBD).

The EDNA database provides, along with other layers, a nationwide flow direction grid. This grid provides the pixel-to-pixel connectivity for every location in the Conterminous United States (CONUS). Knowing this connectivity forms the basis for extraction of standard hydrologic derivatives, but also provides the means to trace both up and downstream with facility. The standard ARC/INFO implementation of the flow accumulation function (which uses the flow direction grid to derive the contributing area above any pixel) has been used to accumulate both climatological and land cover layers. The procedures developed to facilitate this work are described below. Through these accumulated variables, characterization of a basin above any location is simply a matter of querying the accumulated grid.

Data:

Elevation Derivatives for National Applications (EDNA)

The Elevation Derivatives for National Applications (EDNA) database a multi-layered database containing both vector and raster layers of significance to hydrologic and environmental modelers. Derived from a version of the NED, a DEM of the land surface, EDNA's base layer is essentially a DEM of the drainage surface of the United States. Using this base layer, a suite of standard hydrologic derivatives, both raster and vector, have been generated. The projection used for EDNA is the National Albers Equal Area projection:
1st Standard Parallel 29º30'
2nd Standard Parallel 45º30'
Central Meridian 96ºW
Latitude of Origin 23ºN
The raster layers are developed with a 30-meter cellsize.

Shown in Figure 1 is a schematic of the EDNA Stage one processing steps. The NED begins the EDNA processing, but the derivative layers are extracted from the hydrologically conditioned DEM (the drainage surface). The EDNA layers include: Aspect, Compound Topographic Index (ln(A/Tan(slope)), Nation-wide Contours (from the NED), Filled DEM, Flow Accumulation, Flow Direction, Reach Catchments, Shaded Relief, Sinks (those filled to obtain the drainage surface), Slope and Synthetic Streamlines.

EDNA's raster and vector layers, derived from a DEM, are vertically consistent – for example, the extracted drainage lines ***always*** flow from a higher elevation to a lower elevation, the reach catchments ***always*** follow the drainage divide, etc. This consistency allows for transfer of valuable information from the DEM onto the EDNA-derived drainage lines and watersheds. Examples of the types of attributes on EDNA drainage lines and watersheds are stream gradient, minimum and maximum elevation within the watershed, average slope, average elevation, and many others.

EDNA has been progressing on a three stage development track. The first stage was undertaken as a joint effort between the USGS and the National Weather Service's National Severe Storms Laboratory (NSSL). Using standard ARC/INFO implementations of the filling, flow direction and watershed delineation tools (Jensen and Domingue, 1988), Stage one used semi-automated techniques to produce a “first-cut” at a hydrologically conditioned DEM and its derivatives. Natural sink features, when recognized as such, were maintained in the landscape. This stage in the development process was completed for the Conterminous United States in 2002. More details of the development process are found in Franken et al, 2001.

The second stage in EDNA’s development uses the Stage one data to identify and flag potential errors in the EDNA DEM. Because the goal of the EDNA DEM is to represent the drainage surface, some of the “errors” found in Stage two are due to discrepancies between the land surface and the drainage surface. A highway overpass is a good example of such an “error”. In the case of a road grade, the elevation desired in NED is, indeed, the elevation of the road. The elevation desired in EDNA, however, is the elevation of the culvert which passes flow under the road. However, other Stage 2 errors do result from errors in the original NED. Artifacts and limited vertical resolution in the original DEM along with flat topography can cause erroneous hydrologic derivatives. These are flagged by the Stage two cooperators. In some areas of the country, the Stage 2 process has also been used to develop preliminary watershed and subwatershed delineations, useful for the WBD effort (Kost et al, 2002). The Stage 2 effort relies on local expertise to adjudicate discrepancies between the EDNA derivatives and other data layers, such as Digital Raster Graphics (DRGs), NHD, or local data layers.

Information gathered in the Stage 2 effort is used in Stage 3 to correct the EDNA, where necessary, and generate a correct set of hydrologic derivatives. Tools have been developed which help facilitate the editing of the DEM (Kost et al, 2002). This final extraction of accurate hydrologic derivatives allows for vertical integration with the layers used in the Stage 2 quality assessment. The Stage 2 and Stage 3 efforts are underway, with work being completed on a project-by-project basis.

Methodology:

Flow Accumulation for the Nation

While use of the hydrologic functions in ARC/INFO for extraction of drainage structure and other hydrologic derivatives from a DEM has been applied to many modeling scenarios, the development of the EDNA posed many unique challenges. Foremost among them was the size of the dataset. The DEM which provided the input to EDNA, the NED, alone is 60GB. Obviously, a form of segmentation of the problem into smaller pieces was necessary.

In order to partition the United States into manageable pieces, the Cataloging Unit (Seaber, 1978) was chosen as the fundamental EDNA processing unit. Following processing through the Stage 1 procedures, the EDNA layers were reassembled to create a seamless, nation-wide database.

For most of the hydrologic derivatives, the standard Stage one processing suffices. However, processing each Cataloging Unit without regard for its upstream neighbors wreaks havoc on an areal-dependent derivative such as the flow accumulation. To address this problem, the unique structure of the EDNA Stage one data was used along with the flow direction grid and the synthetic stream lines to, in effect, route the upstream area through the downstream Cataloging Unit.

Shown in Figure 2 is a map of the 2,109 Cataloging Units (8-digit sub-basins) into which the CONUS is divided (Seaber, 1978). During the EDNA Stage 1 processing, some attempt was made to maintain the structure of the original Cataloging Units, particularly the manner in which they are hydrologically connected. The EDNA Stage 1 drainage units are, in many ways, an improvement over the original Cataloging Units. In many areas of the country, the drainage divides extracted from the 1:24,000 scale EDNA are an improvement over the existing 1:250,000 cartographically-derived subbasins. Of particular interest to this work is the manner in which the EDNA-derived drainage units provide the unit-to-unit connectivity necessary to do up- and down-stream tracing.

Figure 3 shown the EDNA-derived drainage units for Hydrologic Region 17. Shown on this figure, as well, is the unit-to-unit connectivity. It is this connectivity that allows for the downstream routing of the flow accumulation grids.

A subregion comprised of five Cataloging Units (Figure 4) will be used to illustrate the downstream routing. The processing of the flow accumulation grid is done on a Cataloging Unit basis and progresses from upstream to downstream according to the EDNA-derived drainage unit connectivity. As can be seen from the hydrologic connectivity vectors, three of these Cataloging Units (17010301, 17010302 and 17010304) are headwaters units. For these units, no adjustment needs to be made to the flow accumulation, since there are no unaccounted-for contributions from upstream units. The values derived during the Stage one process are valid.

Progressing down through the hydrologic connectivity shows that Cataloging Unit 17010303 receives contribution from all of the headwater units. Cataloging Unit 17010305 receives contribution, then, from its upstream neighbor 17010303. Therefore, the flow accumulation grids calculated during Stage one need adjustment in order to properly account for the upstream contribution.

Shown on Figure 5 are the EDNA-derived drainage units, drainage pour points, and synthetic streamlines. The drainage pour points provide the linkage necessary to translate values from upstream Cataloging Units to their receiving Units.

Figure 6 shows a close-up of the junction of Cataloging Units 17010301, 17010302 and 17030303. The EDNA drainage pour points are located at the pour points of the upstream unit. Through the EDNA processing, these pour points also have knowledge of the downstream stream segment to which they flow. Using this knowledge, we are able to translate downstream through the stream network, and subsequently any raster dataset, the flow accumulation values. In Figure 6, we can see that the Stage one processing has produced a flow accumulation grid for the receiving Cataloging Unit, 17010303, which provides values which are simply an artifact of from the Stage one processing steps. To adjust the flow accumulation values, the erroneous drainage area resulting from the Stage one processing buffer is removed from the flow accumulation grid. The correct contribution from the upstream unit is added back in along the main downstream flow path. The trace downstream through the receiving Cataloging Unit's stream network is facilitated through the use of the Pfafstetter codification scheme (Verdin and Verdin, 1998). Shown in Figure 7 are the downstream flow paths in need of flow accumulation adjustment for Cataloging Unit 17010303. This procedure is done for each upstream Cataloging Unit, with each correction being added to the previous one. Once the algorithm has cycled through all the drainage pour points, the flow accumulation has been adjusted for all upstream contribution. The final flow accumulation (true flow accumulation) is then ready to be translated downstream to its receiving unit; in this case, Cataloging Unit 17010305.

The true flow accumulation grid for the CONUS has been created through use of these procedures. A simple sampling of this grid provides the drainage area above any 30-meter pixel in EDNA.

One complication arose through this processing. The values of the true flow accumulation grid become too large to be stored within a 16-bit integer data structure. Seven Cataloging Units (from the mouth of the Mississippi up river) have had their values stored as floating point. Loss of integer accuracy is the ramification of this work-around. A 32-bit integer data structure would solve this small problem.

Application to Low-Head Power Assessment for CONUS

The above described work, developing the procedures to trace the flow accumulation down through the EDNA grids, led to the natural extension of using the same procedures to trace any areal-dependent variable. Using different data layers as weights, the flow accumulation function could be used on almost any variable. The EDNA-derived drainage units could still be the processing units and the procedures to route the variable downstream applied.

Work funded through the Department of Energy's Idaho National Engineering and Environmental Laboratory (INEEL) provided a nation-wide application of these procedures to climatological variables. The INEEL had been tasked with developing estimates of the undeveloped low-head power potential for the CONUS. In order to perform this work, estimates of the mean annual flow for and the head drop along each stream segment in the CONUS was required. EDNA was perfectly suited to provide this information.

The head drop along each stream segment was a straight forward calculation, given than every EDNA reach is attributed with elevations on its from- and to-nodes. The calculation of the average annual flow for each reach, however, provided an opportunity to apply the weighted flow accumulation procedures. The estimates of mean annual streamflow were obtained through application of regression equations in the literature (Vogel et al., 1999) with separate sets of equations developed for each of the 18 hydrologic regions of the conterminous United States. These regression equations provide estimates for natural streamflow conditions, and require determination of drainage area, mean annual precipitation, and mean annual temperature for the upstream area.

As an example of the form of the regression equations, the mean annual streamflow for the analysis of Region 17 was:

$$\mathbf{Q_m = e^{-10.18} A^{1.00269} P^{1.86412} T^{-1.1579}} \quad \textit{(Equation 1)}$$

Where:

Q_m is the estimated mean annual flow (m^3/sec)

V_m is the estimated variance of the mean annual flow (m^3/sec^2)

A is the drainage area above a point (km^2)

P is the mean annual precipitation for the drainage area (mm/yr)

T is the mean annual temperature for the drainage area (1/10°F).

Every independent variable in Equation 1 was obtained with EDNA through the application of the flow accumulation function using different weights. The drainage area above the location, A, was calculated simply by translating the flow accumulation value (in pixels) into an appropriate area. The mean annual precipitation and temperature values for every location, P and T, were determined through use of the flow accumulation function using the average annual temperature and precipitation grids from the PRISM Spatial Climate Layers for the United States (Daly et al, 1997) as weights. Use of the flow accumulation function with the PRISM layer as weights resulted in a continuous representation of the accumulated precipitation or temperature above every pixel. These accumulated values were then routed down through the Cataloging Units in the same manner as the flow accumulation grids.

Once the accumulated precipitation and temperature were traced through the entire dataset, the average above any pixel was calculated simply. To calculate the average precipitation, for instance, the following formula was applied:

us_precip = accum_precip / (trueflow_acc + 1) **(***Equation 2)*

Where:

us_precip is the average precipitation value (in mm) for the drainage area upstream of the current pixel

accum_precip is the calculated accumulated precipitation pixel value

trueflow_acc is the area upstream of the pixel (augmented by one to account for the current pixel)

The final result of this analysis is two continuous surfaces of average precipitation and temperature for the CONUS. Sampling of any location yields average precipitation or temperature for the entire drainage area above the point. The continuous nature of this dataset facilitates transfer of the average precipitation and temperature data onto the streamlines through a simple cellvalue query.

Discussion

The procedures developed in the EDNA processing have been shown to be of value for the development of continuous parameterization surfaces. These surfaces, such as flow accumulation, average precipitation and average temperature, can expedite the processing time significantly when used in a wide-area investigation.

Previously, if one was interested in obtaining the average precipitation above a location, several processing steps would probably be undertaken:

1. The watershed above the location would need delineation. The work involved in this step has been significantly reduced since the wide-spread application of DEMs and EDNA-like products, but still requires time and effort.
2. This delineation would then be overlaid on the grid of interest -- precipitation or temperature and a function such as zonalmean would be executed.
3. The result would be transferred onto the appropriate streamline.
4. These procedures would need to be repeated for every streamline in the dataset of interest.

In the case of the work done for the INEEL, using the zonalmean-type procedure would have required analysis of a watershed for every reach in the EDNA dataset. The reaches in the EDNA dataset for the CONUS number more than 980,000 For each of these reaches and watersheds, the zonalmean command would have been executed. Procedures such as these can be streamlined through use of pre-calculation and summarization at the time of execution, but they still are very time-consuming when compared with the continuous method of parameterization.

The continuous method of parameterization has the additional advantage of allowing for automatic parameterization above any location, not just predefined watershed outlets. The parameter has been precalculated for every 30-meter pixel in the EDNA dataset. If a sampling location is moved, the new parameter is simply sampled from the grid -- no recalculation of watershed boundaries and zonalmean statistics is required.

Work continues on these continuous parameterization surfaces at the EROS Data Center. We are currently working closely with the Environmental Protection Agency's MidContinent Ecology Laboratory to develop surfaces of the using land cover classes as weights. These will be used in watershed sampling site selection studies.

Kristine L. Verdin
SAIC/EROS Data Center
Sioux Falls, SD 57198

(605) 594-6002
kverdin@usgs.gov

EDNA home page: http://edna.usgs.gov

References

Daly, C., G. Taylor, and W. Gibson, 1997, The PRISM Approach to Mapping Precipitation and Temperature, 10th Conf. on Applied Climatology, Reno, NV, Amer. Meteor. Soc., 10-12.

Franken, S., D. Tyler and K. Verdin, 2001, Development of a National Seamless Database of Topography and Hydrologic Derivatives, Proceedings of the 21st ESRI User's Conference, CD-ROM

Jenson, S.K. and Domingue, J.O., 1988, "Extracting topographic structure from digital elevation data for geographic information system analysis: Photogrammetric Engineering and Remote Sensing, v. 54, p. 1,593-1,600

Kost, J., K. Verdin, B. Worstell and G. Kelly, 2002, Methods and Tools for the Development of Hydrologically Conditioned Elevation Data and Derivatives for National Applications, 2nd Annual Conference on Watershed Modeling, Las Vegas, CD-ROM publication

Seaber, P., F. Kapinos, and G. Knapp, 1987, “Hydrologic Unit Maps”, USGS Water Supply Paper 2294, 63 p.

Verdin, K.L. and J.P. Verdin, 1999, A topological system for delineation and codification of the Earth's river basins, Journal of Hydrology, vol. 218, nos. 1-2, pp. 1-12

Vogel, Richard M., Ian Wilson and Chris Daly, Regional Regression Models of Annual Streamflow for the United States, Journal of Irrigation and Drainage Engineering, May/ June 1999, pp. 148-157