From a Grid or Coverage to a Hydrograph: Unlocking Your GIS Data for Hydrologic Applications

by

E. James Nelson

Norman L. Jones

Christopher Smemoe

Introduction

While many aspects of hydrologic modeling are ideally suited for GIS, others are not. In particular most hydrologic modeling studies are constrained to the use of industry standard, time proven, lumped parameter hydrologic models. A lumped parameter model does not take advantage of the spatial distribution of physical properties in a watershed such as soil types, land use, etc. While a GIS is extremely useful in compiling and storing such data, it is not an ideal environment or user interface for converting spatial data into a format which can be used to run a lumped parameter model directly, or as a batch process. For this reason much of the hydrologic data developed using a GIS have remained locked inside the GIS and unusable by engineers wishing to perform hydrologic analyses with traditional lumped parameter models such as HEC-1 and TR-20. A new method to provide a common gateway between GIS data stored as points, lines, and polygons will make it possible to exploit the strengths of both a GIS and an application customized for hydrologic modeling. In this way the GIS can be used to do what it does best: develop and store hydrologic data and at the same time provide a mechanism to use these data as part of a modeling project.

Hydrologic Data Development

Hydrologic data development can be defined as the process of compiling hydrologic parameters from digital or map sources. If developed using a GIS, these data can then be used as a layer in a larger GIS, and/or in combination with analytical models. Hydrologic data of particular importance include:

- Watershed and sub-basin boundaries and areas
- Streams and stream lengths
- Basin and stream slopes
- Land use
- Soil type
- Runoff coefficient

Aspect
Runoff distances (used primarily to compute a basin response time to a rainfall event)

Traditionally these hydrologic data have been derived using paper maps (sometimes requiring several maps to be taped together), and several tedious man hours locating drainage divides and computing basin average values. However, GIS tools, including the new Spatial Analyst in ArcView 3.0, are well suited for automating these processes.

The primary data layer required for hydrologic data development is an elevation grid, but soil type and land use grids/polygons are also important for determining basin infiltration/runoff characteristics. An elevation grid is a two-dimensional grid of elevations with a fixed x and y spacing (resolution). The USGS provides digital elevation data from 1:250,000 scale maps on the internet (EROS Data Center). USGS 7.5 minute DEMs may also be purchased through a USGS data center or other data publishing firms. Figure 1 shows a contoured 7.5 minute digital elevation data model from which hydrologic data may be developed.

Figure 1 Contoured Digital Elevation Model (DEM).

Watershed Delineation from an Elevation Grid

The fundamental aspect of hydrologic data development is the delineation of watershed and sub-basin boundaries. While there are several variations to the original algorithm, the primary method involved in delineating watershed boundaries from DEM data is the eight point pour method developed by Puecker et al. (1975). The basic idea in Puecker's method is to assign a flow direction to each grid cell based on the neighboring cell with the lowest elevation. Figure 2 illustrates how this might be done.

Figure 2 Grid Cell Flow Direction Assignment.

It is important to note that raw DEM data often contain pits, where no neighboring grid cell has a lower elevation. Therefore pre-processing algorithms to fill pits (raise the elevation of the pit until a "pour point" occurs) must be part of the overall process for defining a flow vector grid (Garbrecht, 1995). Figure 3 illustrates how the general flow pattern of a terrain model emerges when displaying the directions of a flow direction grid.

Figure 3 Flow Direction Grid.

Once the flow direction grid has been computed, a flow accumulation grid is created. This is done by counting the number of contributing cells to each cell in the grid (cells whose flow path eventually passes through the cell). Cells which are potentially part of a stream network will have a larger flow accumulation value, whereas cells near watershed boundaries and where overland flow dominates will have a low flow accumulation value. Figure 4 shows a flow accumulation grid overlayed on a contour map. Only the cells whose flow accumulation is greater than 50 cells are displayed, and it can be seen how they clearly define the stream channels present in this portion of the elevation grid.

Figure 4 Flow Accumulation Grid.

At this point the watershed and sub-basin outlet points for watershed delineation must be identified. Traditionally this has been a difficult process because it requires interaction with a user in order to select the appropriate point(s). Many grid based watershed delineation algorithms simply create sub-basins for each branch in the stream, where the branches may have controlling parameters such as thresholds on contributing area and/or number of grid cells making up the branch. However, as graphical user-interfaces have become more powerful, watershed and sub-basin outlets can be graphically selected by a user. It is important to be able to hand-pick outlet points because often the hydrologic model needs to be developed for a location corresponding to a road crossing or where a detention structure is to be built. Such locations rarely coincide with outlets determined from traditional, automated methods.

Once the outlet locations are specified, watershed and sub-basin delineation can be performed. The process is similar to the definition of flow accumulations in that the flow direction from each grid cell is traced until either an outlet cell or the edge of the grid is encountered. If an outlet cell is found then the grid cell is assigned the id of the basin for that outlet point. If the edge of the flow direction grid is found then the cell is assigned a "no data" value, meaning that it does not contribute flow to any of the defined outlets. The process of assigning basin ids to grid cells can be optimized by first assigning the basin ids to all stream cells upstream from the outlet points. Then, whenever a traced grid cell flow path encounters a stream cell it can be assigned the same basin id as the stream cell encountered (Figure 5). Figure 6 shows a delineated watershed which has one interior sub-basin outlet point. Separate sub-basins were determined for each upstream branch of the interior outlet point.

Figure 5 Watershed ID Assignment for Grid Cells.

Figure 6 Delineated Watershed and Sub-basins.

With watershed and sub-basins determined for each cell in the grid, area, slope, aspect, stream length, and several other geometric parameters can be determined. Separate layers for land use and soil may also be overlayed with the watershed grid to determine infiltration/runoff parameters for each basin.

Hydrologic Data Storage

While a grid is convenient for the development of hydrologic data, it is inefficient for data storage. This is because with a grid data structure there is a many to one relationship for most basin parameters. For example the watershed and sub-basin grid would likely contain a basin id value for each grid cell. Since several hundred or several thousand grid cells may belong to the same basin there are "many" grid cells with "one" value or in other words a many to one relationship.

A more convenient and efficient way to store both stream and basin data is through vectors and polygons. Each stream segment is represented as a single line (arc) in the GIS and each watershed or sub-basin is represented as a polygon. Additional information about the streams or basins is stored as attributes of the arc or polygon. This approach is particularly useful when using lumped parameter models as discussed in the next section. Developed hydrologic data can then be stored as hydrography data layers in a larger municipal or agency wide GIS, or accessed by hydrologists for use in analytical hydrologic models.

For hydrologic modeling there are three primary data layers required:

- A point layer representing watershed or sub-basin outlet points

Figure 7 Point Features (shown by triangles) Used to Define Basin Outlets.

- A line (arc) layer defining the stream and routing network of the watershed

Figure 8 Line Features Used to Define Streams.

- A polygon layer representing the watershed and/or sub-basins of the watershed

Figure 9 Polygon Features Used to Define Basins.

There are several other layers such as lakes, spillways, highways, etc. which may be a part of the hydrography data, but for the purposes of the link being made between GIS data and hydrologic models we will focus on these three layers. When stream networks and basin boundaries are determined using a grid approach described previously it can be "vectorized" into a line network and polygon once the hydrologic data development is complete. It is important to note that basin boundaries or stream networks may already exist in vector format and may be used directly as one of these three primary data layers.

Lumped Parameter Models

There are two basic classes of analytical hydrologic models: distributed and lumped. While distributed models are the focus of much current research, lumped parameter models continue to be the mainstay for hydrologic modeling projects. This is because a hydrologic analysis typically has to be checked by some regulatory agency such as a state highway agency or the Federal Emergency Management Agency (FEMA). These agencies have not accepted distributed models because of the effort involved in calibration and verification. Further, they want to be able to compare new models or results with previous studies performed using one of the accepted lumped parameter techniques.

A distributed model takes advantage of the spatial distribution of physical characteristics such as infiltration capacity, land cover, slope, etc. to more accurately model how rainfall is converted to runoff. Distributed models are typically created using grids like those discussed previously for watershed delineation. In this case, a grid for each physical parameter that it is part of the hydrologic model would need to be created. In this way each grid cell could possibly contain a different value for slope, land use, soil type, etc. The distributed model can then compute runoff for each grid cell and route it to the basin outlet point.

A lumped parameter model on the other hand uses a single value or index for each "hydrologic element." Hydrologic elements are comprised of (but not limited to) sub-basin areas, routing reaches, reservoirs, and diversions. Figure 10 shows the relationship between hydrologic elements in a lumped parameter model and the actual watershed they represent.

Figure 10 Lumped Parameter Model Definition of a Watershed.

Data for these elements can be stored in one of the basic GIS layers described above. For example an outlet point for a sub-basin may actually represent a culvert and information such as geometric dimensions, loss coefficients, etc. could all be stored as attributes of the point. Basin parameters such as runoff coefficients, runoff response time, and base flow could all be stored as attributes of the basin polygons. Some lumped parameters, such as average slope or runoff coefficients could be computed using spatial modeling techniques as described in the Hydrologic Data Development section above, while others may simply be entered by the user. Still other parameters such as a rainfall event may not be suited for storage in a GIS at all.

Linking GIS Hydrologic Data with Analytical Models

Because many aspects of hydrologic modeling are well suited to GIS analysis, hydrologic data development/storage tools have been a part of the earliest releases of Arc/Info. While a GIS provides powerful tools for performing distributed hydrologic modeling, it is not structured in a way that is convenient to link with lumped parameter models. Since most regulatory agencies require the use of lumped parameter models, data developed and stored in a GIS have been "locked" to the hydrologist wishing to use pre-compiled spatial data for a hydrologic analysis.

Several attempts including DeBarry et al. (1990) documented cases of hydrologic modeling applications built on top of a GIS. Such interfaces provide the necessary link but have the following restrictions:

- They are typically customized to a single model and/or specific options of a given model
- The interface is built "on top" of the GIS. This means additional licenses (and dollars for the licenses) of the GIS are required in order for the hydrologist to accomplish his work. Further, the hydrologist must become trained on the GIS.
- The interface is often awkward because a GIS is a fairly generic tool for spatial analysis and was not necessarily developed with hydrologic modeling in mind.

Because of these barriers, GIS software for hydrologic modeling has not been widely used, and hydrologic data developed using a GIS have remained locked to the hydrologic modeler.

In an effort to "unlock" data developed with a GIS, and bridge the gap to an environment specifically suited for hydrologic modeling, a set of hydrologic data layers have been defined as a joint project by ESRI and the Engineering Computer Graphics Laboratory(ECGL) of Brigham Young University. These data consist of the three primary layers outlined previously: a point layer for outlet information, a line (arc) layer for stream and routing information, and a polygon layer for basin boundaries. Besides the spatial data defining the positions of the hydrologic features database, other data pertinent to hydrologic models are exported using pre-defined field names.

Using Avenue, ESRI has extended their basic hydrologic modeling tools in the ArcView Spatial Analyst to include a set of customized scripts that do the following:

1. Delineate stream networks and basin boundaries from an elevation grid.
2. Covert the raster representation of the stream network and basin boundaries to line and polygon themes.
3. Compute areas, slopes, aspects, etc. from delineated basins and store as attributes of the vector themes.
4. Using a custom dialog, export the hydrologic themes as three different shape files: one for outlets, one for streams, and one for basins. Any data items defined as attributes of these three themes are also exported.

In order to demonstrate how these data can be utilized inside of a hydrologic modeling environment, the ECGL has modified the Watershed Modeling System (WMS) so that it can be used to process the three hydrologic data layers. Specifically, new tools have been added to WMS that make it possible to:

1. Import and read shape files, including both geometric and database attribute information.
2. Construct a topologic model representation of the watershed from the spatial layout of the outlets, streams, and basins as illustrated in Figure 10 above.
3. Convert database attributes to their corresponding data structures so that they can be used inside of WMS for hydrologic modeling.

When the GIS data layers are imported, the dialog shown in Figure 11 is used to set up mapping between the database (dbf) field names and the WMS data structures. WMS automatically maps any variables using the pre-defined names (as would be the case if the data were developed using the ArcView custom hydrologic extensions), and any other names can be manually mapped by the user.

Figure 11 Linking Database Attributes to WMS.

Once the GIS data layers have been imported into the WMS, additional parameters required for a specific hydrologic simulation, such as a rainfall event, time steps, and output options, which would not normally be defined inside of ArcView, can be entered using the customized model interface (Figure 12).

Figure 12 Hydrologic Model Definition in WMS.

Once the hydrologic model has been completely defined it can be executed from a menu command and the results can be viewed inside of WMS. Typical output from a hydrologic model includes a hydrograph from which peak flow and runoff volume for each hydrologic unit can be determined. WMS includes a comprehensive interface to most commonly used models, including HEC-1, TR-20, NFF, and the Rational method. Sample results for a TR-20 analysis are shown in Figure 13. The results from the analysis can be exported back and stored inside the ArcView database.

Figure 13 Results of a Hydrologic Simulation Using TR-20.

The steps described above were discussed in the context of using Arc/Info Grid or the ArcView Spatial Analyst for hydrologic data development and export. However, there are other means of cataloging and developing hydrologic data. For example, historical watershed boundaries and stream networks are often defined manually. The three data layers are defined in a general fashion such that they can be used to link hydrologic data with the WMS environment regardless of how the data were developed.

Conclusions

Geographic Information Systems such as ArcView are excellent tools for developing hydrologic data. Stream networks and watershed and sub-basin boundaries are easily determined using elevation grids. Once boundaries have been delineated, geometric parameters, required as part of input to hydrologic models, can be computed on a cell by cell basis and/or aggregated to provide a single value for each basin, stream, or outlet. While a GIS is an ideal environment for hydrologic data development, it is difficult and cumbersome to link the GIS with traditional hydrologic models that are required for use by regulatory agencies when performing hydrologic analysis. However, both spatial and attribute data from the GIS can be passed to hydrologic modeling applications as three primary data layers: points for outlets, lines for streams, and polygons for basin boundaries. In this way hydrologic modeling parameters which are best computed using GIS can be done in with GIS and then passed as attributes to an application like WMS that is customized for hydrologic modeling, where the remainder of the model parameters can then be defined. In this way the power of a GIS to develop spatial hydrologic data can be "unlocked," saving time and money for both the GIS specialist and the hydrologic modeler. This procedure has been implemented and tested using ArcView and WMS.

References

DeBarry, Paul A. and J.T. Carrington, July 1990, "Computer Watersheds," ASCE Civil Engineering.

Garbrecht J., and L.W. Martz, 1995, "TOPAZ: An Automated Digital Landscape Analysis Tool for Topographic Evaluation, Drainage Identification, Watershed Segmentation, and Subcatchment Parameterization: Overview," U.S. Department of Agriculture, Agricultural Research Service, ARS Publication No. NAWQL 95-1, 17 pp., October 1995.

Puecker, T.K., and D.H. Douglas, 1975, "Detection of surface-specific points by local parallel processing of discrete terrain elevation data," Computer Graphics and Image Processing, Vol 4. pp. 375-387

E. James Nelson
Assistant Professor
Brigham Young University
368 CB
Provo, Utah 84602
Telephone: (801) 378-7632
Email: jimn@byu.edu