

# Automated Grid Element Ordering for GIS-Based Overland Flow Modeling

Michael B. Smith and Mitja Brilly

Civil Engineering Department, Hydraulics and Hydrology Section, University of Ljubljana, 28 Hajdrihova Ulica, 61000 Ljubljana, Slovenia, Yugoslavia

**ABSTRACT:** A group of previously developed processing steps for digital elevation models has been expanded to include the automatic numbering of raster cell elements within a hydrologic cascade. The tools are a series of FORTRAN computer programs which produce a hydrologically ordered cell-to-cell flow sequence which is subsequently used to dictate the order of computations within a distributed parameter model for overland flow computations. Coupling of the numbering algorithm with a rainfall-runoff model to generate runoff hydrographs for a small agricultural watershed and a complex urban catchment demonstrates the utility of the method. The algorithm serves an important link between digital elevation models and GIS-based distributed parameter hydrologic models.

## INTRODUCTION

WITH THE ADVENT OF POINT, grid, and triangulated irregular network (TIN) digital elevation models (DEM) has come a profusion of topographic analysis procedures and hydrologic models based on GIS data structures. The algorithm described here fits into this rapidly developing arena of DEM processing and GIS based hydrologic models. Commercially available GIS packages routinely offer options for many types of DEM processing. Analysis options include steps for watershed boundary delineation, slope angle and aspect computations, elevation interpolation from points or contours, cut and fill estimates, and filtering operations.

Other researchers have developed their own processing methodologies and assembled them into "toolboxes" of analysis utilities. These include procedures to analyze grid cell DEMs and TIN models. TIN models have received significant attention in recent years, while grid cell elevation models continue to be popular because they can be easily coupled to remotely-sensed data structures.

Distributed parameter hydrologic models are those which attempt to quantify the spatial variability of catchment characteristics. Typically, the catchment is discretized into unit elements, with the hydrological response of each element computed within a proper time and spatial sequence. Thus, these models are well suited for investigating the effects of land-use changes on runoff hydrographs, tracing the paths of pollutant movement through a watershed (Young *et al.*, 1989), and modeling sediment transfer (Gupta and Solomon, 1971). However, accounting for spatial variations of hydrologically significant parameters requires the use of large databases and, thus, computers with high memory capacity and fast processing speeds. Recent advances in micro-computer technology have facilitated higher processing speeds and operating systems which are not bounded by previous computer memory limitations (Anderson and Burt, 1990). Thus, distributed parameter hydrologic models are now being used in the microcomputer environment in connection with geographic information systems (Johnson, 1989).

One of the necessary input parameters for certain types of distributed parameter modeling is the cell-to-cell connectivity sequence. When utilizing a kinematic or a cascade of reservoirs approach to compute overland flow, the computations must proceed in the correct hydrological sequence. This means that, before any cell is examined and a rainfall-runoff transformation is performed, all cells upstream of the current cell must be analyzed. In addition, distributed parameter modeling requires that this sequence be followed within each time step. When wa-

tersheds are small or when cell sizes are large, the connectivity matrix can be derived by hand. However, when large watersheds are examined or larger numbers of cells are involved, an automatic procedure is needed. The purpose of this paper is to describe an automatic method for numbering grid cell elements for distributed parameter hydrologic modeling. The method was developed as part of ongoing research to develop a GIS-based hydrologic model for urban areas. In order to accurately model street flows and storm sewer overflows, it is anticipated that cell sizes will be on the order of 10 to 20 metres.

## BACKGROUND

Jenson and Domingue (1988) present a detailed description of conditioning procedures and analyses to be applied to digital elevation models. Depression detection and filling, flow direction computations, and assigning flow accumulation values are conditioning procedures used to create data sets for further analysis. These conditioned data sets include a depressionless digital elevation model in which each cell has assigned to it an outflow direction. Using these conditioned data sets, operations such as watershed delineation, automatic subwatershed delineation, and determination of watershed linkages are performed. The algorithms were tested on several large watersheds, and good agreement was achieved between computed and actual topographic structure. Mark (1984) presents similar procedures for basin detection and flow network delineation. Quinn *et al.* (1991) developed a procedure to assign multiple flow paths exiting from a cell rather than a single value as done by others. Sensitivity of flow path generating algorithms to grid resolution was also studied. Collins (1981) describes a family of algorithms that operate on a dense digital elevation model to locate features such as road intersections.

For digital elevation models represented by triangulated irregular networks, Gandoy-Bernasconi and Palacios-Velez (1990) present methods for the automatic numbering of unit elements while Jones *et al.* (1991) present procedures for TIN-based watershed delineation. Djokic and Maidment (1991), Hong and Eli (1985), and Silber *et al.* (1987) developed hydrologic models that use parameters derived from a TIN DEM structure. Moore *et al.* (1991) present an excellent description of analyses and applications using digital elevation models.

Van Deursen and Kwadijk (1990) also developed a series of programs for deriving hydrologically significant information from grid-type digital elevation models and assembled them within a hydrological toolbox. Procedures were developed for interactive pit and depression removal and for defining drainage paths. Also included in the toolbox are hydrological building

blocks or sub-models that the user can select. Thus, the user selects not one hydrological model but instead composes a model to suit individual needs and project objectives.

APPROACH

PRELIMINARY PROCESSING

The methods presented here are generally based on the conditioning procedures described by Jenson and Domingue (1988) to develop a depressionless digital elevation model. For the purposes of this paper, it is assumed that a digital elevation model has been processed to remove pits and depressions. Before the actual application of the numbering algorithm, several analyses must be performed. Figure 1 presents a schematic of the analysis procedures outlined here.

Flow directions are assigned on the basis of steepest descent away from the current cell *x* and are specified using the numbering convention of Greenlee (1987) shown in Figure 2. The numerical codes are based on powers of two so that the position of each adjacent cell remains uniquely defined throughout analysis. Quinn *et al.* (1991) noted that a grid scale of 50 m or coarser may lead to erroneous results when using a single value direction approach. However, given the anticipated cell resolution of 10 to 20 metres in urban areas, a single direction value

should be appropriate. As an example, Table 1a presents the DEM subset used by Jensen and Domingue (1988), while Table 1b shows the subsequent direction data set. Those cells forming the border of the data set are assigned direction values pointing away from the interior. Figure 3 presents pictorially the direction values.

To determine the watershed draining to a user specified outlet cell, a catchment growing procedure is used. First, the algorithm examines all cells neighboring the selected outlet cell. Those cells flowing into the outlet cell become the preliminary watershed drainage area. In a recursive procedure, the neighbors of this starting group of watershed elements are examined to see if they flow into previously detected watershed cells. Those cells contributing to the watershed area are designated as watershed cells. When no neighbors of existing watershed cells can be found that flow into the watershed, the drainage area is considered to be fully defined. Using the watershed grow procedure, the drainage area for outlet cell (11,9) is outlined in Table 1c. While the watershed grow procedure is similar to the method of Jenson (1988), any other means to define the watershed boundary in a raster format may be used. Also required for the numbering algorithm is a neighbor data set, in which each cell is assigned an integer value denoting the number of inflowing neighbor cells. Table 1d shows the derived neighbor data set. Using the direction data set, the neighbor algorithm determines the number of inflowing neighbors for each cell and also locates start cells, or those cells into which no neighboring cells flow. Most often, these start cells form the border of a watershed. However, local maxima within a watershed can also form start cells. Because no neighboring cell flows into them, these start cells form the beginning of overland flow paths and are thus named start cells. Start cells appear in Table 1d as those cells having a value of zero and in Figure 3 as shaded cells. Cells having a neighbor value of 2 or more are called junction cells.

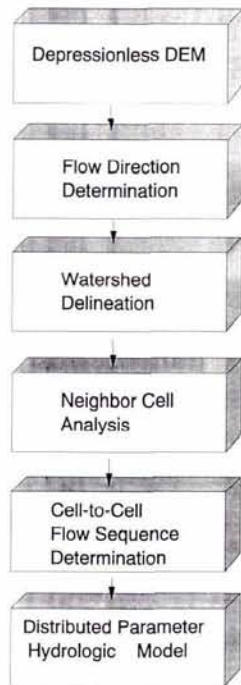
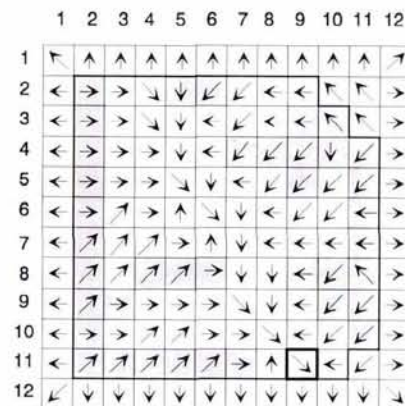


FIG. 1. Processing steps for deriving the computational sequence.

64	128	1
32		2
16	8	4

FIG. 2. Flow direction data codes for cell *x*.



legend

- overland flow direction
- start cell of overland flow path
- watershed outlet
- watershed boundary

FIG. 3. Pictorial representation of the flow directions listed in Table 1b.



## FLOW PATH DELINEATION

In general, the process of element numbering consists of following overland flow paths from each start cell to the watershed outlet. The algorithm makes many passes through the watershed, tracing overland paths from start cells to the outlet. As each path is traced, two different types of path segments are assembled. Start cell paths begin with a start cell and end with the first detected junction cell. Junction paths begin with a junction cell and end at the next downstream junction cell. Values are assigned to each junction cell depending on the number of times the algorithm has traced an overland path through it. These values are used later to reorder the path segments into the proper hydrological order.

For a watershed having  $n$  start cells, the algorithm makes  $n$  passes through the drainage area. Pass number one begins with the algorithm tracing a path from the first start cell using the directions stored in the direction data set. This first path is assigned a path number of one. Upon reaching the first junction cell, the algorithm ends path number one and begins path number two. All cells in the start cell flow path are numbered and cell coordinates are stored in FORTRAN variables as follows:

IVARIABLE(PATH NUMBER, CELL NUMBER) = CELL I COORDINATE

JVARIABLE(PATH NUMBER, CELL NUMBER) = CELL J COORDINATE

For path number two, the algorithm assembles cells into junction cell paths until the next downstream junction cell is reached. Cell coordinates are stored in a manner similar to start cell paths.

Each time the algorithm reaches a junction cell, a new path number must be assigned because the presence of a junction cell indicates upstream cells whose hydrologic influence must be considered before continuing downhill. At each junction cell, the algorithm also updates a counter named *inflow*, indicating the number of passes that have proceeded through the cell. Thus, the junction cell nearest the outlet cell will have the highest value of *inflow* which is also equal to the number of start cells in the watershed. While each start cell has a path number associated with it, each junction cell is assigned both a path number and an *inflow* value.

Processing continues in pass one until the algorithm reaches the outlet cell. At this point, pass two is begun and the algorithm traces the overland path beginning with the second start cell. Cells are assembled into the start path and junction paths as previously described. However, during the second and all subsequent passes, an additional operation is performed on each junction cell. If the algorithm has previously passed through the junction cell, i.e., if the *inflow* value is one or greater, then no junction cell paths are assembled, as that path number and all cell coordinates have previously been established and assembled. However, the *inflow* value for each junction is incremented as the path is traced to the outlet.

Upon completion of the  $n$ th pass through the watershed, the *inflow* values shown in Table 1e result. It should be noted that the values increase as one moves towards the outlet. Because in this example there are 38 start cells, the algorithm has made 38 passes through the drainage area. Thus, the *inflow* value for the outlet is also 38.

## PATH SEGMENT ORDERING

It is important to note at this point that cells are not individually ordered as such, but are located within path segments which themselves are ordered. Within each path segment, however, each cell is ordered according to its proper place in the path.

By definition, start cells have no uphill hydrologic contribution. Thus, within the group of start cell paths, each path has the same importance as all other start paths. Therefore, there is no need to place start paths into any specific order. Table 2

TABLE 2. LISTING OF START CELL COORDINATES.

Path No.	Cell No.	Cell Coordinates		Path No.	Cell No.	Cell Coordinates		
		Row	Column			Row	Column	
1	1	2	2	36	1	7	11	
	2	2	3		2	7	10	
	3	2	4		38	1	8	2
	4	3	5			2	7	3
11	1	2	5	40	3	6	4	
	2	3	5		4	6	5	
12	1	2	6	42	1	8	4	
	2	3	5		2	7	5	
13	1	3	2	43	3	7	6	
	2	3	3		1	8	5	
	3	3	4		2	7	6	
	4	4	5		1	8	6	
14	1	3	9	44	2	8	7	
	2	3	8		1	8	9	
	3	3	7		2	8	8	
	4	4	6		3	9	8	
15	5	4	5	46	1	8	10	
	1	3	10		2	9	9	
	2	2	9		3	9	8	
	3	2	8		47	1	8	11
4	2	7	2	7		10		
16	5	3	6	48	1	9	2	
	6	3	5		2	8	3	
	1	4	2		3	7	4	
	2	4	3		49	4	6	5
3	4	4	1	9		3		
17	4	4	5	52	2	9	4	
	1	4	7		3	9	5	
19	2	5	6	53	1	9	10	
	1	4	8		2	10	9	
20	2	5	7	54	3	10	8	
	1	4	9		1	9	11	
	2	5	8		2	10	10	
22	3	6	7	57	3	11	9	
	1	4	10		1	10	2	
25	2	5	10	58	2	10	3	
	1	4	11		1	10	11	
26	2	5	10	59	2	11	10	
	1	5	2		3	11	9	
	2	5	3		1	11	2	
28	3	5	4	60	2	10	3	
	1	5	9		1	11	3	
29	2	6	8	61	2	10	4	
	3	6	7		1	11	4	
32	1	5	11	63	2	10	5	
	2	6	10		3	9	6	
34	1	6	2	64	1	11	5	
	2	6	3		2	10	6	
35	1	6	11	64	3	10	7	
	2	6	10		1	11	6	
	1	7	2		2	10	7	
	2	6	3		1	11	7	
					2	11	8	
					3	10	8	

lists the cells comprising each start path within the example watershed.

Final determination of the flow sequence for junction cell segments is accomplished by compiling an ordered list of junction cells and storing the path number associated with each. Ordering is dictated by the final *inflow* value of each junction cell. Thus, a junction cell having an *inflow* value of three would have its associated path number stored in the third position or level of the ranking variable according to the following assignment statement:

$$\text{rank}(\text{level}, \text{counter}) = \text{path number assigned to junction cell}$$

Because it stores the path number, which is in turn linked to cell coordinates, *rank* serves as a link between the proper hydrological order and cell coordinates. The variable *level* varies from 2 to *n*, the number of start cells. Flow path numbers associated with start cell segments are not stored within the *rank* variable.

Within a watershed, there may exist many junction path segments which exist at the same hydrological level. This concept is similar to stream ordering, in which many stream segments can have the same order. Within *rank*, the variable *counter* stores the number of path segments having the same hydrological level. For the example watershed, Table 3 lists in hydrological order each junction cell path and each cell within each path.

In the context of a distributed parameter hydrological model, the information contained within the *rank* variable serves as a driver to dictate the order of computations for cell paths whose level is two or higher. In essence, all of the path number and cell coordinate information is contained in linked lists which are traced during each time step. Figure 4 presents a generalized set of FORTRAN nested do-loops which recall the coordinates of each cell in hydrological order. The call statements represent calls to subroutines to extract hydrologically important parameters from other databases and then perform a rainfall-runoff transformation.

APPLICATION

Testing of the numbering algorithm involved coupling the method with a rainfall-runoff model to generate runoff hydrographs for a small agricultural watershed and a larger urban catchment. A cascade of non-linear reservoirs approach was employed to transform rainfall amounts into runoff rates. During each time step, the rainfall-runoff transformation is applied to each cell in the order dictated by the derived sequence. Runoff generated from one cell becomes inflow to the next cell in the sequence. Details of the hydrological model are not presented here but are well described in Huber *et al.* (1988) and Green (1984).

For the agricultural watershed, topographic information and rainfall-runoff data were taken from a publication of the U.S. Department of Agriculture. (USDA, 1966). Figure 5 shows the watershed and the topographic layout. Due to the simplicity of the watershed, flow directions for each cell were derived by hand and appear in Figure 6. The watershed was modeled as a matrix of 15-metre by 15-metre cells. Combining the automatic numbering algorithm with the hydrologic model produced the

hydrographs shown in Figure 7. It can be seen that the model accurately reproduces the measured hydrograph.

A complex urban watershed was also used to evaluate the utility of the numbering algorithm. Figure 8 shows the 9.4 hectare Gray Haven watershed near Baltimore, Maryland. In order to model street flows and storm sewer overflows, a cell size of 12.1 metres was used, which roughly corresponds to the width of the residential streets. Due to this small cell size, a grid representation of the drainage area resulted in 641 cells. Manual derivation of the cell-to-cell flow scheme would be tedious if not nearly impossible. As with the agricultural watershed, cell flow directions were assigned by hand using two-foot contour

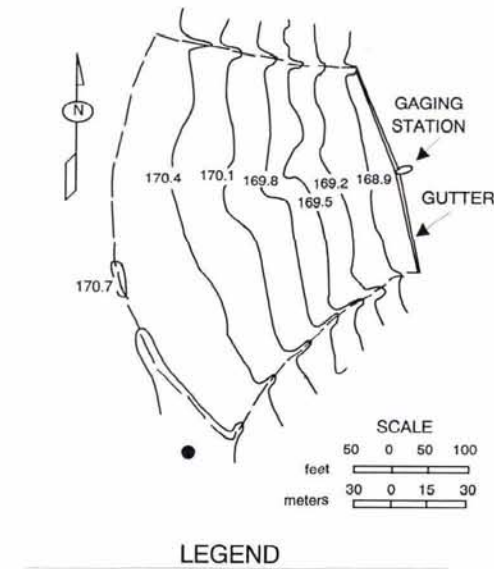
TABLE 3. LISTING OF JUNCTION CELL COORDINATES IN HYDROLOGICAL ORDER.

Level	Counter	Path No.	Cell No.	Cell Coordinates	
				Row	Column
2	1	18	1	5	6
			2	6	6
2	2	23	1	5	10
			2	6	9
			3	7	8
2	3	33	1	6	3
			2	5	4
2	4	39	1	6	5
			2	5	5
2	5	21	1	6	7
			2	7	7
2	6	30	1	6	10
			2	7	9
2	7	41	1	7	6
			2	6	6
2	8	37	1	7	10
			2	7	9
2	9	45	1	9	8
			2	10	8
2	10	55	1	10	3
			2	10	4
2	11	62	1	10	7
			2	10	8
3	1	27	1	5	4
			2	5	5
3	2	56	1	10	4
			2	9	5
4	1	2	1	3	5
			2	4	5
4	2	31	1	7	9
			2	7	8
4	3	50	1	9	5
			2	9	6
5	1	51	1	9	6
			2	9	7
6	1	24	1	7	8
			2	7	7
7	1	3	1	4	5
			2	5	5
12	1	4	1	5	5
			2	6	6
16	1	5	1	6	6
			2	7	7
24	1	6	1	7	7
			2	8	7
25	1	7	1	8	7
			2	9	7
30	1	8	1	9	7
			2	10	8
36	1	9	1	10	8
			2	11	9
38	1	10	1	11	9

```

do 1000 t = 1 to max time by time step
do 900 i = 2 to number of start cells
do 800 j = 1 to number of paths in each level
pathnumber = rank(i,j)
number of cells in path = numcell(pathnumber)
do 700 k = 1 to number of cells in path
recall cell coordinates
c cell i coordinate = ivariable(pathnumber,k)
cell j coordinate = jvariable(pathnumber,k)
c Hydrologic Model
call hydrologic parameters(cell i, cell j, parameters)
call runoff (parameters)
700 continue
800 continue
900 continue
1000 continue
    
```

FIG. 4. Generalized FORTRAN procedure for recalling junction path cells in hydrological order.



**LEGEND**

- Watershed Boundary
- Recording Rain gauge
- 169.5 — Contours 0.305 meters (1 foot)

FIG. 5. Generalized layout of the experimental watershed located in Waco, Texas.

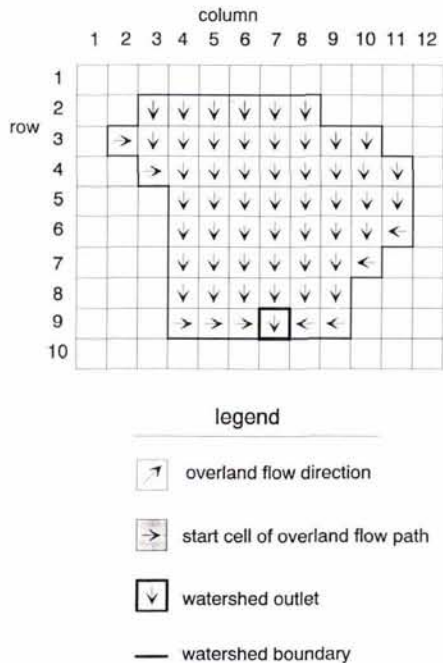


FIG. 6. Grid cell representation of the experimental watershed in Waco, Texas. Cell size is 15m by 15m.

information and personal inspection of the watershed. Rainfall and runoff data were taken from a publication of the American Society of Civil Engineers (Tucker, 1968). Time shift routing was used for flows in the storm sewer system. As seen in Figures 9

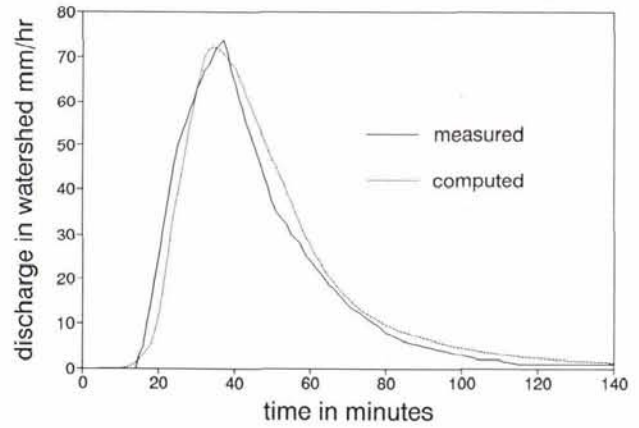


FIG. 7. Computed versus measured runoff hydrographs for the Watershed in Waco, Texas. Storm of 24 April 1957.

and 10, the numbering algorithm facilitates the accurate simulation of the residential watershed.

**CONCLUSIONS**

Proper ordering of elements is required for proper watershed runoff simulation. An algorithm has been developed that computes a hydrologically ordered flow sequence of grid cell elements for input into a distributed parameter hydrological model. As input, the method requires a data set in which each cell is assigned a flow direction that can be traced to the outlet. Coupling the algorithm with a rainfall-runoff model accurately reproduced measured results for two quite different watersheds. The algorithm serves as a useful link between digital elevation models and GIS-based distributed parameter hydrologic models.

**REFERENCES**

Anderson, M. G., and T. P. Burt, 1990. Subsurface Runoff, chapter 11 of *Process Studies in Hillslope Hydrology*, John Wiley and Sons, New York, N.Y., 539 p.

Collins, S. H., 1981. Algorithms for dense digital terrain models. *Photogrammetric Engineering & Remote Sensing*, Vol. 47, No. 1, pp. 71-76.

Djokic, D., and D. R. Maidment, 1991. Terrain analysis for urban stormwater modelling. *Hydrological Processes*, Vol. 5, pp. 115-124.

Gandoy-Bernasconi, W., and O. Palacios-Velez, 1990. Automatic cascade numbering of unit elements in distributed hydrological models. *Journal of Hydrology*, Vol. 112, pp. 375-393.

Green, I. D. A., 1984. WITWAT Stormwater Drainage Program - Version II. *Report No. 2/1984*, University of the Witwatersrand, Republic of South Africa, pp. 12-20.

Greenlee, D. D., 1987. Raster and vector processing for scanned line-work. *Photogrammetric Engineering & Remote Sensing*, Vol. 53, No. 10, pp. 1383-1387.

Gupta, S. K., and S. I. Solomon, 1977. Distributed numerical model for estimating runoff and sediment discharge of ungaged rivers -1. The information system. *Water Resources Research*, Vol. 13, No. 3, pp. 613-636.

Hong, H., and R. N. Eli, 1985. Accumulation of streamflow in complex topography. *Computer Applications in Water Resources, Proceedings of the Specialty Conference, A.S.C.E. Water Resources Planning and Management Division*, Buffalo, New York, 10-12 June, pp. 196-205.

Huber, W. C., R. E. Dickenson, and T. O. Barnwell, 1988. *Storm Water Management Model, Version 4 Users Manual*. U. S. Environmental Protection Agency, Athens, Georgia, pp. 448-489.

Jenson, S. K., and J. O. Domingue, 1988. Extracting topographic structure from digital elevation data for geographic information system

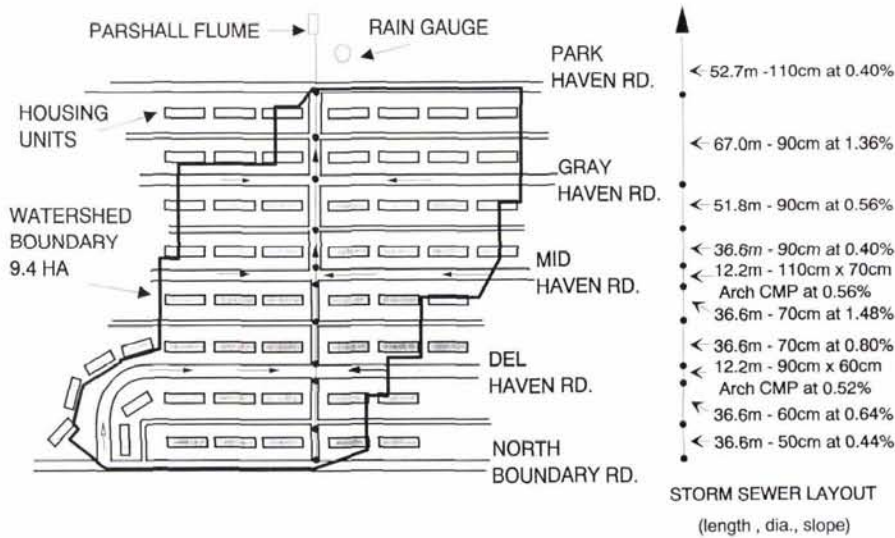


Fig. 8. Gray Haven watershed near Baltimore, Maryland. Arrows indicate flow along streets and in the sewer system.

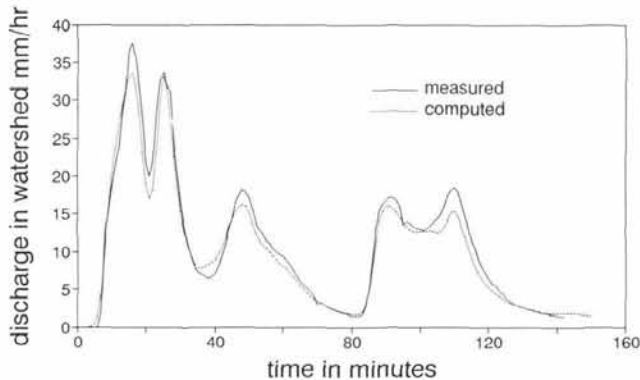


Fig. 9. Computed versus measured runoff hydrographs for Gray Haven watershed. Storm of 14 August 1963.

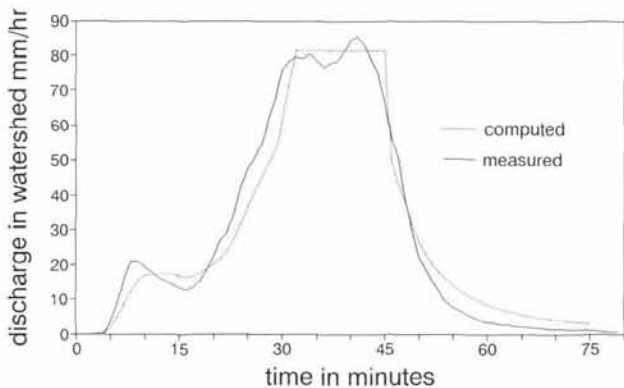


Fig. 10. Computed versus measured runoff hydrographs for Gray Haven watershed. Storm of 10 June 1963.

analysis. *Photogrammetric Engineering & Remote Sensing*, Vol. 54, No. 11, pp. 1593-1600.

Johnson, L. E., 1989. MAPHYD - A digital map-based hydrologic modeling system. *Photogrammetric Engineering & Remote Sensing*, Vol. 55, No. 6, pp. 911-917.

Jones, N. L., S. G. Wright, and D. R. Maidment, 1990. Watershed delineation with triangle-based terrain models. *Journal of Hydraulic Engineering*, ASCE, Vol. 116, No. 10, pp. 1232-1251.

Mark, D. M., 1984. Automated detection of drainage networks from digital elevation models. *Cartographica*, Vol. 21, pp. 168-178.

Moore, I. D., R. B. Grayson, and A. R. Ladson, 1991. Digital terrain models: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, Vol. 5, No. 1, pp. 1-30.

Quinn, P., K. Beven, P. Chevallier, and O. Planchon, 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, Vol. 5, No. 1, pp. 59-79.

Silfer, A. T., G. J. Kinn, and J. M. Hassett, 1987. A geographic information system utilizing the triangulated irregular network as a basis for hydrologic modeling. *Proceedings of Auto-Carto 8, Eighth International Symposium on Computer-Assisted Cartography*, Baltimore, Maryland, 29 March-3 April, pp. 129-136.

Tucker, L. S., 1969. *Availability of Rainfall-Runoff Data for Sewered Drainage Catchments*. Technical Memorandum No. 8., ASCE Water Resources Research Program, New York, N.Y.

USDA, 1963. *Hydrologic Data for Experimental Agricultural Watersheds in the United States, 1956-1959*. Miscellaneous Publication No. 945, U.S.D.A Agricultural Research Service, U.S. Govt. Printing Office, Washington, D.C.

Van Deursen, W. P. A., and K. J. Kwadijk, 1990. Using the watershed tools for modeling the rhine catchment. *Proceedings of the First European Conference on Geographical Informational Systems*, Amsterdam, The Netherlands, pp. 254-262.

Young, R. A., C. A. Onstad, D. D. Bosch, and W. P. Anderson, 1989. AGNPS: a non-point source pollution model for evaluation agricultural watersheds. *Journal of Soil and Water Conservation*, March-April, pp. 168-173.

You can receive a copy of **PE&RS** each month by joining **ASPRS**. Additional membership benefits include discounts on: classified advertisements, publications, exhibit spaces and registration, and mailing lists. To become an **ASPRS** member call Anne Ryan or Sokhan Hing at 301-493-0290 today!