Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis

S.K. Jenson and *J. O. Domingue*
TGS Technology, Inc., EROS Data Center, Sioux Falls, SD 57198

ABSTRACT: Software tools have been developed at the U.S. Geological Survey's EROS Data Center to extract topographic structure and to delineate watersheds and overland flow paths from digital elevation models. The tools are special purpose FORTRAN programs interfaced with general-purpose raster and vector spatial analysis and relational data base management packages.

The first phase of analysis is a conditioning phase that generates three data sets: the original DEM with depressions filled, a data set indicating the flow direction for each cell, and a flow accumulation data set in which each cell receives a value equal to the number of cells that drain to it. The original DEM and these three derivative data sets can then be processed in a variety of ways to optionally delineate drainage networks, overland paths, watersheds for user-specified locations, sub-watersheds for the major tributaries of a drainage network, or pour point linkages between watersheds.

The computer-generated drainage lines and watershed polygons and the pour point linkage information can be transferred to vector-based geographic information systems for further analysis. Comparisons between these computer generated features and their manually delineated counterparts generally show close agreement, indicating that these software tools will save analyst time spent in manual interpretation and digitizing.

INTRODUCTION

DIGITAL ELEVATION MODELS (DEMs) can be used to derive a wealth of information about the morphology of a land surface (U.S. Geological Survey, 1987). The algorithms traditionally included in most raster processing systems use neighborhood operations to calculate slope, aspect, and shaded relief (Klingebiel *et al*., 1988) and points of inflection (Peucker and Douglas, 1975). While watersheds and overland flow paths are closely related to slope, aspect, and inflection information, they also present non-neighborhood problems such as determining direction of flow in the interior of large flat area. To overcome these limitations, software has been developed that uses neighborhood techniques as well as iterative spatial techniques that can best be visualized as region-growing procedures. They provide an analyst with the ability to extract from DEMs information on morphologic features and properties, specifically topographic depressions and flow directions, that may be further processed to delineate application-specific watersheds and overland flow paths.

These tools are relatively computer-intensive but require little analyst intervention, thus minimizing analyst time. In addition, the resultant products have the advantage of precise registration with the DEM. While the algorithms are essentially raster based, the products (watershed polygons, drainage line networks, and tabular attribute information defining watershed linkages) can readily be converted to vector form.

BACKGROUND

Douglas (1986) gives an excellent description of techniques that have been developed to define ridges, channels, watersheds, and other hydrologic features from DEMs. These techniques are generally based on neighborhood operations where calculations and decisions are made for a cell based on the values in the eight cells that are spatially adjacent in the raster. For instance, a cell that is equal in elevation to all of its neighbors' elevations, meet the criteria to be classified as a member of a flat area. The approach described here has some similarities to previous approaches, but differs in that depressions and flat areas are fully accommodated.

Previous research has almost universally recognized that depressions, areas surrounded by higher elevation values, in the DEM data are the nemesis of determining hydrologic flow directions because the depressions must fill before the flow can continue. Some depressions are data errors introduced in the surface generation process, while others represent real topographic features such as quarries or natural potholes. A few researchers have attempted to remove depressions by smoothing the DEM data (O'Callaghan and Mark, 1984; Mark, 1983). The smoothing approach removes shallow depressions, but deeper depressions remain. A second approach is to "fill" depressions by increasing the values of cells in each depression to the value of the cell with the lowest value on the depressions boundary. Algorithms for filling depressions with the second approach, presented by Marks et al. (1984) and Jenson and Trautwein (1987), will be discussed later. Collins (1975) presented an algorithm for filling depressions; however, it fails to fully accommodate flat areas as shown by Douglas (1986).

These algorithms follow the second approach, that depressions be filled as the first step in the analysis. They then become flat areas across which water can be routed. It is assumed that these and other flat areas may be quite large and have more than one outflow point.

An important goal in the design and implementation of the algorithms was that there be as few data set size restrictions as possible. An upper limit of 4,000 samples per line in a DEM is the current limitation in the implementation of the algorithms.

CONDTIONING PROCEDURES

An initial conditioning phase produces three data sets that are of general utility for all subsequent steps. The three data sets, in the order that they are produced, are a DEM with depressions filled, a data set indicating the flow direction for each cell, and a flow accumulation data set in which each cell receives a value equal to the total number of cells that drain to it.

FILLING DEPRESSIONS IN A DEM

DEMs almost always contain depressions that hinder flow routing. The objective of the first step in the conditioning phase

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 54, No. 11, November 1988, pp. 1593-1600.

0099-1112/88/5411-1593$02.25/0

is to create an adjusted "depressionless" DEM in which the cells contained in depressions are raised to the lowest elevation value on the rim of the depression. Each cell in the depressionless DEM will then be part of at least one monotonically decreasing path of cells leading to an edge of the data set. A path is composed of cells that are adjacent horizontally, vertically, or diagonally in the raster (eight-way connectedness) and that steadily decrease in value. In the special case where flow routing is of interest within a depression, the original DEM values would be used rather than the depressionless DEM, and the flow paths within the depression would terminate at the bottom of the depression rather than at the data set edge.

The procedure by which the depressionless DEM is made is described in Table 1. This procedure is a good example of the philosophy of developing tools that can be combined in a variety of ways because several of the steps, such as delineating watersheds, are independently useful. While the product of the procedure described in Table 1 is a depressionless DEM, the procedure incorporates other procedures to generate necessary intermediate data sets. An earlier version of this procedure (Jenson and Trautwein, 1987) is similar but iterative and, hence, slower. Examples of an original DEM and the depressionless DEM that results from this procedure are in Table 2a and 2b, respectively. Two spatially separated depressions were identified with a maximum depth of 5 feet. The full data set is given in Plate 1. The depression-filling procedure developed by Marks *et al*. (1984) is similar in that, as an intermediate process, it finds watersheds for cells that have no neighbors lower in elevation, identifies cells lower in elevation than the lowest boundary elevation for their watershed, and encodes these cells as being flat for use in the basin delineation process. Their procedure does not, however, include logic for finding looping depressions such as might occur when many depressions are located on a surface that is relatively flat. In their watershed generation process, one watershed is grown at a time, and a flat area is assigned in its entirely to the first watershed that touches it. The watershed generation process presented here differs in that many watersheds are generated in one program execution and flat areas are allowed to be subdivided if they have more than one outflow point.

TABLE 1. FILLING DEPRESSIONS IN A DEM

Step	Procedure
1	Fill single-cell depressions by raising each cell's elevation to the elevation of its lowest elevation neighbor if that neighbor is higher in elevation than the cell. This is a simple case and filling them reduces the number of depressions that must be dealt with.
2	Compute flow directions (Table 3)
3	For every spatially connected group of cells that has undefined flow directions because it would have required an uphill flow, find the group's uniquely labeled watershed from the flow directions.
4	Build a table of pour point elevations between all pairs of watersheds that share a boundary (Table 6).
5	For each watershed, mark the pour point that is lowest in elevation as that watershed's "lowest pour point." If there are duplicate lowest pour points, select one arbitrarily.
6	For each watershed, follow the path of lowest pour points until either the data set edge is reached (go to step 7) or the path loops back on itself (go to step 6a). 6a. Fix paths that loop back on themselves by aggregating the watersheds which comprised the loop, deleting pour points between group members from the table, recomputing "lowest pour point" for the new aggregated watershed, and resume following the path of lowest pour points.
7	In each watershed's path of lowest pour points, find the one that is highest in elevation. This is the threshold value for the watershed. Raise all cells in the watershed that are less than the threshold value to the threshold value.

FLOW DIRECTIONS

The second procedure of the conditioning phase builds the flow direction data set (Table 3). The flow direction for a cell is the direction water will flow out of the cell. It is encoded to correspond to the orientation of one of the eight cells that surround the cell (x) as follows:

64	128	1
32	x	2
16	8	4

For example, if cell x flows to the left in the matrix, its flow direction will be encoded as a 32. Flow direction encoding is done in powers of two so that surround conditions correspond to unique values when the powers of two are summed for any unique set of neighbors. There are four possible conditions to consider in determining flow direction (Table 4). Condition 1 occurs when all eight neighboring cells have elevations higher than center cell. The flow direction will be encoded as negative for such a cell, indicating an undefined flow direction. Condition 1 cells are single-cell depressions. They will not be present after the first step of the depression-filling procedure but are included in the flow direction procedure for completeness. Condition 2 is the case where the distance-weighted drop from the center cell is higher for one cell in the neighborhood over all of the other seven and the flow direction is assigned to this cell. Distance-weighted drop is calculated by subtracting the neighbor's value from the center cell's value and dividing by the distance from the center cell, $\sqrt{2}$ for a corner cell and one for a noncorner cell. Most cells are condition 2 cells. For condition 3, when two or more cells are equal in having the greatest distance-weighted drop, the flow direction is assigned logically using a table look-up operation. For example, if three adjacent cells along one edge of the neighborhood have equal drops, the center cell is logically chosen and assigned as the flow direction. If two cells on opposite sides have equal drops, as in Table 4, condition 3, one is arbitrarily chosen. When all cells are equal or greater in elevation compared to the center cell, as in condition 4, determining the flow direction is the most time consuming. In this case, the cell is located in a flat area and the direction to the outflow-point is not known. After cells with the first, second, and third conditions are resolved, the fourth condition cells are resolved in an iterative process. In each iteration, cells are assigned to flow to a neighbor if the neighbor has a defined flow direction that does not point back to the tested cell. In this way, flow direction assignments iteratively grow into the flat area from the flats' outflow points until all cells have flow directions assigned.

The flow direction concept has been employed by both Marks *et al*. (1984) and O' Callaghan and Mark (1984). However, neither included logic for condition 3 cells or extended the technique beyond the neighborhood operation to solve for condition 4 cells. When the flow direction procedure is applied to a depressionless DEM, all cells will have a definable flow direction value because, by filling depressions, the DEM is conditioned so that every cell has a flow path to the data set edge. The flow direction is illustrated numerically in Table 2c, and visually in Plate 1b.

FLOW ACCUMULATION DATA SET

The third procedure of the conditioning phase makes use of the flow direction data set to create the flow accumulation data set, where each cell is assigned a value equal to the number of cells that flow to it

TABLE 2. THE ALGORITHMS FOR EXTRACTING HYDROLOGIC INFORMATION FROM DEM DATA ARE ILLUSTRATED HERRE WITH A NUMERIC EXAMPLE. (A) SHOWS THE ORIGINAL DEM ELEVATION VALUES IN FEET. THESE VALUES WERE EXTRACTED FROM THE DEM SHOWN IN PLATE 1 AT THE LOCATION OUTLINED BY THE WHITE RECTANGLE IN PLATE 1A. (B) SHOWS THE DEPRESSIONLESS DEM WITH THE CELLS THAT ROSE IN VALUE OUTLINED. IN (C), A FLOW DIRECTION CODE IS SHOWN FOR EACH CELL. IN (D), THE FLOW ACCUMULATION VALUES ARE SHOWN. IN (E), THE DELTA VALUES SHOW THE INCREASES IN FLOW ACCUMULATION VALUES AS EACH CELL DRAINS TO ITS DOWNSTREAM NEIGHBOR. THE DATA SET THAT RESULTS FROM THE AUTOMATIC SUB-WATERSHED STARTING PROCEDURE USING A THRESHOLD OF 10 IS SHOWN IN (F). THE FULL SUB-WATERSHEDS ARE SHOWN IN (G).

(a) original DEM

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	778	765	750	740	747	759	765	766	769	776	786	795
2	770	758	745	737	741	751	753	761	777	789	802	814
3	777	763	747	736	735	743	750	767	787	806	820	832
4	786	767	750	737	729	739	752	769	785	797	808	822
5	794	773	756	741	730	732	744	759	772	779	789	806
6	799	782	763	750	737	728	732	745	757	767	782	801
7	802	788	771	761	751	736	729	738	751	764	779	798
8	799	790	780	772	762	746	733	737	754	770	784	794
9	811	799	787	771	757	741	728	730	745	765	779	783
10	823	807	790	774	762	748	733	724	733	750	764	763
11	830	814	801	787	776	761	743	728	725	737	748	751
12	822	818	811	801	791	776	757	739	726	725	735	751

(b) depressionless DEM

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	778	765	750	740	747	759	765	766	769	776	786	795
2	770	758	745	737	741	751	753	761	777	789	802	814
3	777	763	747	736	735	743	750	767	787	806	820	832
4	786	767	750	737	733	739	752	769	785	797	808	822
5	794	773	756	741	733	733	744	759	772	779	789	806
6	799	782	763	750	737	733	733	745	757	767	782	801
7	802	788	771	761	751	736	733	738	751	764	779	798
8	799	790	780	772	762	746	733	737	754	770	784	794
9	811	799	787	771	757	741	728	730	745	765	779	783
10	823	807	790	774	762	748	733	725	733	750	764	763
11	830	814	801	787	776	761	743	728	725	737	748	751
12	822	818	811	801	791	776	757	739	726	725	735	751

(c) flow directions

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	32	128	128	128	128	128	128	128	128	128	128	128
2	32	2	2	4	8	16	16	32	32	64	64	2
3	32	2	2	4	8	32	16	32	32	64	64	2
4	32	2	2	2	8	32	16	16	16	8	16	2
5	32	2	2	2	4	8	32	16	16	16	16	2
6	32	2	1	2	128	4	8	32	16	16	32	2
7	32	1	1	1	2	128	8	32	32	32	32	2
8	32	1	1	1	1	2	8	8	32	16	64	2
9	32	1	2	2	2	2	4	8	32	16	16	2
10	32	2	2	1	1	2	2	4	32	16	16	2
11	32	1	1	1	1	1	2	128	4	32	16	2
12	32	8	8	8	8	8	8	8	8	8	8	2

(d) flow accumulation values

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	2	1	0	0
2	0	0	1	2	0	0	3	2	1	1	0	0
3	0	0	1	2	10	4	2	1	0	0	0	0
4	0	0	1	2	21	3	0	0	0	0	0	0
5	0	0	1	5	35	3	1	1	0	2	0	0
6	0	0	2	2	6	44	4	1	3	2	0	0
7	0	0	1	2	1	3	62	11	6	2	0	0
8	0	0	1	0	0	0	64	1	0	0	0	0
9	0	0	0	1	7	10	76	4	1	0	0	0
10	0	0	2	4	1	1	3	90	1	1	0	0
11	0	0	0	0	0	0	0	1	95	1	0	0
12	0	0	0	0	0	0	0	0	97	0	0	0

(*Table continued on next page)*

TABLE 2. (CONTINUED)

(e) delta values

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	1	8	10	10	1	1	1	1	1	0
3	0	1	1	19	11	6	1	1	1	1	1	0
4	0	1	1	19	14	18	3	1	1	2	2	0
5	0	1	4	30	9	41	2	3	1	1	2	0
6	0	2	3	4	29	18	58	3	8	4	2	0
7	0	2	1	4	2	41	2	51	5	4	2	0
8	0	1	1	1	3	64	12	3	1	1	2	0
9	0	1	1	6	3	66	14	86	3	1	1	0
10	0	2	2	3	9	2	87	5	89	94	1	0
11	0	2	4	1	1	3	1	89	2	94	97	0
12	0	0	0	0	0	0	0	0	0	0	0	0

(f) automatic sub-watershed start data set

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
5	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
6	-1	-1	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1
7	-1	-1	-1	-1	-1	-1	-1	-3	-1	-1	-1	-1
8	-1	-1	-1	-1	-1	-1	-4	-1	-1	-1	-1	-1
9	-1	-1	-1	-1	-1	-1	-5	-1	-1	-1	-1	-1
10	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
11	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

(g) sub-watershed data set

Sample Line	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	1	1	1	1	1	1	1	0	0	0
3	0	1	1	1	1	1	1	1	1	1	0	0
4	0	1	1	1	1	1	2	2	4	3	3	0
5	0	2	2	2	2	2	2	4	4	3	3	0
6	0	2	2	2	2	2	4	4	3	3	3	0
7	0	2	2	2	2	2	4	3	3	3	3	0
8	0	2	2	2	2	4	4	0	0	0	3	0
9	0	2	5	5	5	5	5	0	0	0	0	0
10	0	5	5	5	5	0	0	0	0	0	0	0
11	0	5	5	5	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0

(O'Callaghan and Mark, 1984). Cells having a flow accumulation value of zero (to which no other cells flow) generally correspond to the pattern of ridges. Because all cells in a depressionless DEM have a path to the data set edge, the pattern formed by highlighting cells with values higher than some threshold delineates a fully connected drainage network. As the threshold value is increased, the density of the drainage network decreases. The flow accumulation data set that was calculated for the numeric example is shown in Table 2d, and the visual example is shown in Plate 1c.

APPLICATIONS

Upon completion of the conditioning procedures, the three derived data sets (i.e., depresionless DEM, flow direction, and flow accumulation) may be further processed for specific applications. Five examples will be discussed.

SPECIFIC WATERSHED DELINEATION

Delineation of watersheds requires both a flow direction data set and another "starter" data set. The starter data set consists of background values of –1 inch which "start" cells or groups of cells have been inserted at the outflow points of the desired watersheds, with each start cell or group of cells having its own unique positive values. In creating the starter data set, it is useful to have a raster image processing system to display color-coded flow direction and flow accumulation data sets. A cursor is used to identify the line and sample coordinates of the outflow points when watersheds are to be delineated with respect to the locations of hydrologic stations or the locations where samples are collected for water or stream sediment chemistry. If a watershed is to be delineated for a broad feature such as a dam, a block of cells should be inserted to represent the feature. If a watershed is to be delineated for a depression such as a pothole, the cells isolated by the depression-filling procedure would be used as a "start" group. The flow direction data set is then used in the

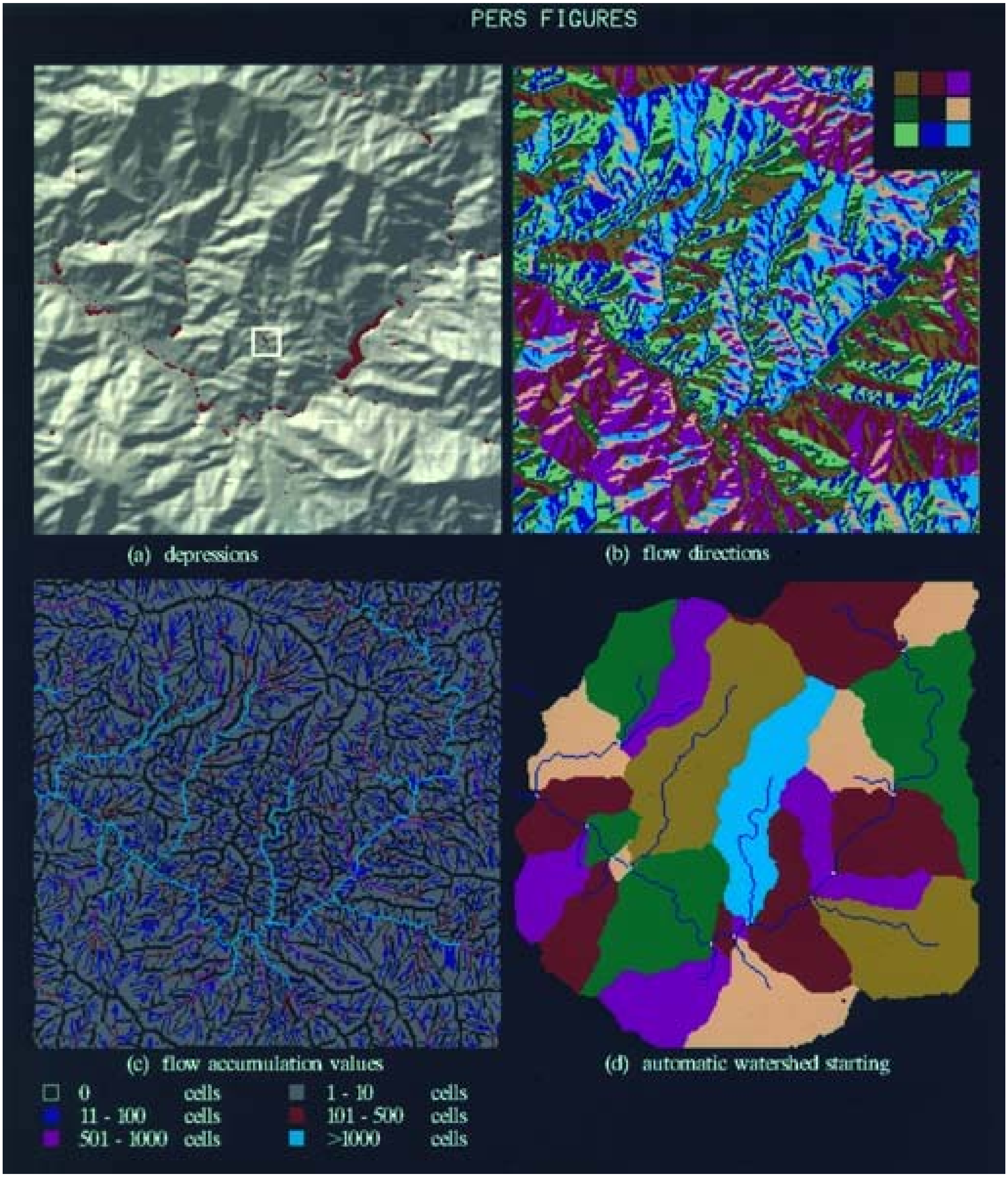

PLATE 1. 248 lines by 248 samples of DEM data are shown here to graphically illustrate extracted hydrologic features. In (a), the areas highlighted in red on a shaded-relief representation of the DEM are cells that were raised in value to create a depressionless DEM. The white rectangle indicates the location of the 12-line by 12-sample subsection used for the numeric example in Table 2. Flow directions for the depressionless DEM are shown in (b). The 3 by 3 color box in the upper right indicates the colors used to represent the eight possible flow directions. (c) shows the color-coded flow accumulation data set, and (d) is the result of automatic sub-watershed starting using a threshold of 1000. The start points are highlighted in white and the cells with flow accumulation values greater than 1000 are superimposed in blue.

TABLE 3. COMPUTING FLOW DIRECTIONS FOR A DEM

Step	Procedure
1	For all cells adjacent to the data set edge or the study area amsk, assign the flow direction to flow to the edge or the mask. This action is taken under the assumption that the study area is interior to the data set.
2	For each cell not assigned a flow direction in step 1, compute the distance-weighted drop in elevation to each of the cell's eight neighbors.
3	Examine the drop value to determine the neighbor(s) with the largest drop and perform one of the following: 3a. If the largest drop is less than zero, assign a negative flow direction to indicate undefined. This situation does not occur for a depressionless DEM 3b. If the largest drop is greater than or equal to zero and occurs at only one neighbor, assign the flow direction to that neighbor. 3c. If the largest drop is equal to zero and occurs at more than one neighbor, assign the flow direction logically according to a table loop-up. 3d. If the largest drop is equal to zero and occurs at more than one neighbor, encode the locations of those neighbors by summing their neighbor location codes. Neighbor location codes are 64 128 1 32 x 2 16 8 4 for any cell x. If all neighbor elevations were equal to the center cell, the center would receive a value of 255. Examples of steps 3a through 3d are given in Table 4, conditions 1 through 4.
4	For each cell not already encoded as negative, 0, 1, 2, 4, 8, 16, 32, 64, or 128, examine the neighbor cells with the largest drop. If a neighbor is encountered that as a flow direction of 1, 2, 4, 8, 16, 32, 64, or 128, and the neighbor does not flow to the center cell, assign the center cell a flow direction which flows to this neighbor.
5	Repeat step 4 until no more cells can be assigned a flow direction.
6	Make the flow direction value negative for cells that are not equal to 1, 2, 4, 8, 16, 32, 64, or 128. This situation will not occur for a depressionless DEM.

watershed generation procedure to iteratively reassign background cells to the value of the "start" cell to which they flow.

AUTOMATIC DELINEATION OF SUB-WATERSHEDS

For some hydrologic applications, it is necessary to divide a watershed into sub-watersheds defined by major tributaries. Sub-watersheds are watersheds but are referred to as sub-watersheds in the context of also being part of a larger watershed. Band (1986) developed a technique to delineate sub-watersheds by isolating ridgelines in deeply incised terrain and assuming that depressions and flats were to significant features. Definition of a major tributary is data-set and application dependent; however, it may be related to the area of the tributary's sub-watershed. To build a start data set for tributaries with sub watersheds greater in area than in user-specified threshold, the automatic sub-watershed starting procedure given in Table 5 can be used. The procedure inserts sub-watershed "starts" at drainage line intersections that meet a sub-watershed area criteria. When this procedure is followed using the numeric example, step 2 procedures the delta values shown in Table 2e. Delta value is the amount of increase in flow accumulation value in the flow direction. Table 2f shows the results of automatic sub-watershed starting where the cells having both flow accumulation value and delta value greater than an area threshold of ten cells are given unique positive numbers. When this starter data set is subjected to the watershed delineation procedure, the full subwatershed data set shown in Table 2g results.

When the flow accumulation data set in Plate 1c is subjected to the automatic sub-watershed starting procedure using an area threshold of 1,000 cells, and then to the watershed delineation procedure, the sub-watersheds shown in Plate 1d result. In this figure, cells highlighted in blue have flow accumulation values greater than or equal to 1,000; cells highlighted in white have both delta and flow accumulation values greater than 1,000 and thus are the start cells for the sub-watersheds.

TABLE 4. EXAMPLES OF THE FOUR POSSIBLE SURROUND CONDITIONS FOR DETERMINIGN FLOW DIRECTION ARE SHOWN. CONDITION 1 IS A SINGLE-CELL DEPRESSION NOT FOUND IN DEPRESSIONLESS DEMS; CONDITION 2 OCCURS WHEN ONE NEIGHBOR IS THE BEST CANDIDATE; CONDITION 3 OCCURS WHEN MORE THAN ONE NEIGHBOR IS A POSSIBLE CHOICE; AND CONDITION 4 IS WHERE THE CELL IS PART OF A FLAT AREA. WEIGHTED DROPS ARE CALCULATED BY SUBTRACTING THE NEIGHBOR'S VALUE FROM THE CENTER CELL'S VALUE AND DIVIDING BY THE APPROPRIATE DISTANCE, √2 FOR CORNER CELLS OR ONE FOR NONCORNER CELLS.

	Elevation Values			Weighted Drops			Flow Direction
Condition 1	100	102	100	-7.0	-12.0	-7.0	-4
	99	90	92	-9.0		-2.0	
	98	94	92	-5.6	- 4.0	-1.4	
Condition 2	92	91	90	-1.4	- 1.0	0.0	2
	92	90	89	-2.0		1.0	
	94	93	90	-2.8	- 3.0	0.0	
Condition 3	90	91	90	0.0	- 1.0	0.0	2
	89	90	89	1.0		1.0	
	90	93	90	0.0	- 3.0	0.0	
Condition 4	92	91	90	-1.4	- 1.0	0.0	Temporarily encoded as 1 + 2 + 4 = 7, and then resolved iteratively
	93	90	90	-3.0		0.0	
	94	93	90	-2.8	- 3.0	0.0	

TABLE 5. SUB-WATERSHED STARTING BY AREA THRESHOLD

Step	Procedure
1	Define an area threshold to constrain the minimum area of sub-watersheds.
2	Compute a delta value for every cell by subtracting the flow accumulation value of the cell it flows to from its own flow accumulation value.
3	For each cell where both the flow accumulation value and the delta value are greater than the area threshold, assign the cell a unique positive value in the starter data set.
4	Assign all remaining cells a value of –1.
5	Report how many watersheds were started.

TABLE 6. DETERMINING POUR POINTS BETWEEN WATERSHEDS

Step	Procedure
1	Compare each cell in a watershed data set to its eight neighbors. When a cell and its neighbor have different watershed labels, proceed to steps 2 through 5.
2	Compare the elevation values of the cell and its neighbor. The larger of the two elevation values is the elevation of the possible pour point they represent, and the line and sample of the cell with the larger elevation is the pour point location.
3	If this pair of watershed labels is not yet in the table of pour points, make a new table entry by recording the pair of watershed labels and the location and elevation of the pour point.
4	If this pair of watershed labels is already in the pour point table, compare the elevation in the table to the elevation for the possible pour point being examined. If the new elevation is lower, replace the old pour point lines, sample, and elevation with the new ones.
5	Repeat the procedure for all cells.

WATERSHED LINKAGES

A table of linkages can be produced for a watershed (or subwatershed) data set (Table 6). A watershed links to another at a pour point, the point of lowest elevation on the common boundary between the two watersheds. To accommodate data sets with internal (or closed) drainage, or where two or more pour points leading to different watersheds are very close or equal in elevation, all possible pour points are computed. The procedure uses the watershed and depressionless DEM data sets to produce a pour point table. For each watershed, all bordering watersheds are examined, including the watersheds with a label of zero indicating that they flow to the data set edge, and a pour point is found on each watershed-to-watershed border. The line and sample coordinates and elevation of each pour point are recorded in the table. Table 7 shows the table generated for the watersheds in Table 2g.

The pour points that are lowest in elevation are indicated by the parenthesized basin numbers. For instance, basin two's lowest pour points are to basin one and four, both at an elevation of 733 feet. Basin five has one pour point lower than all others, thereby linking it to the zero watershed at elevation 728 feet, line 9, sample 7. Allowing multiple pour points is necessary to handle areas of poorly defined drainage. In one application of the software to pothole terrain, multiple pour points were common. In applications to terrains where drainage is well defined and where watersheds are large enough to contain a sloping stream segment, multiple pour points of equal elevation are rarely encountered. The pour point table can be transferred to a relational database management system where the linkages can be used to aggregate watersheds. The line and sample coordinates of the pour points can also be transferred to a vector-based spatial analysis system as point data sets for additional analysis and plotting.

DRAINAGE NETWORKS

As stated previously, a flow accumulation data set can be used to produce a drainage network data set when cells with values greater than a threshold value are selected. The density of the network increases as the threshold value decreases, as shown in Plate 1c. These raster drainage networks will be fully connected topologically if the DEM's depressions are first filled. A raster-to-vector process for lines (Greenlee, 1987) can be used to translate raster line information to a vector format so that it can be transferred to a vector-based geographic information system. Because the lines were generated directly from the DEM, there is no possibility of misregistration between the two data sets. Therefore, elevation values at drainage line intersections can be recorded during the raster-to-vector conversion process and used to accurately calculate stream gradient information.

TABLE 7. THE POUR POINT TABLE COMPUTED FROM THE DEPRESSIONLESS DEM AND WATERSHED DATA SET IN TABLES 2B AND 2G IS SHOWN HERE. PARENTHESIZED WATERSHED NUMBERS INDICATE THE LOWEST, OR TIED FOR THE LOWEST, POUR POINT FOR THE WATERSHED.

	Pour Point		
Watershed Pair	Elevation	Line	Sample
0 - 1	737	2	4
(1) - (2)	733	4	5
1 - 4	785	4	9
1 - 3	797	4	10
(2) - (4)	733	5	6
(3) - 4	738	7	8
0 - (3)	738	7	8
0 - 2	786	4	1
0 - 4	733	8	7
2 - 5	762	8	5
(4) - 5	733	8	7
0 - (5)	728	9	7

OVERLAND PATHS

If an overland path (flow path) is needed for a cell or a set of cells, the flow direction data set can be used to produce the path or paths by following the cell-to-cell linkage until the data set edge is reached. This process is analogous to the watershed process in that, to find watersheds, cells that flow to watershed "start" cells receive the "start" cell's watershed label, and to find overland paths, cells that are flowed to by overland path "start" cells receive the "start" cell labels. One practical application of this process is to track the overland path of pollution from a point source into the drainage network.

If a drainage network is required for an area where channels are characterized as broad, shallow valleys, the flow accumulation data set will produce a drainage network with wide lines in these valleys. An alternative is to use the local minima method of Jenson (1985) or one of several other techniques listed by Douglas (1986) to identify and label cells that are downward inflection points in the depressionless DEM surface. When the overland paths are followed for these cells, a thin drainage network results.

EVALUATION AND DISCUSSION

This software toolbox has been tested in several hydrologic studies. Watersheds have been defined for gaging stations on the Susquehanna River and Genegantslet Creek in New York, for the South Fork of Lower Willow Creek in Montana, and for the dam of Big Tujunga Reservoir in California. The New York studies used 1:250,000 DEMs. The Montana and California studies used 1:24,000 DEMs. In comparisons between the watersheds derived from the digital elevation data using the toolbox algorithms and watersheds manually delineated from topographic maps, agreement was very close.

Line plots produced from the vectorized watersheds in the New York study were planimetered and compared to the manually delineated watersheds. Ninety-seven percent of the Susquehanna River manually delineated watershed coincided spatially with the computer-generated watershed. The automated methods excluded a large portion of the manually delineated Genegantslet Creek watershed. Upon further examination of the topographic maps, hydrologists determined that the manually delineated watershed was in error. After the manually delineated watershed was corrected, the Genegantslet Creek watershed was also in 97 percent spatial agreement (Zariello, personal communication, 1986).

Digital comparisons of the Big Tujunga Reservior watershed and a digitized version of the manually delineated watershed indicated 98 percent agreement (Verdin, personal communication, 1988). Visual comparisons of the Willow Creek watersheds indicated that manual and digital results were also very similar (Stannard, personal communication, 1988). Quantitative comparisons of automatically derived and hand-drawn drainage networks are difficult to make because the automatically derived networks are much more detailed. Visual comparisons of digitized, manually delineated networks and automatically derived networks on a raster display device indicate that the main channels are described almost identically.

The algorithms were also used to derive attributes relating to the water storage capacity of glaciated pothole terrain in the James River Basin in North Dakota. In addition to the raster processing system used in conjunction with the software-tools, this project required the use of a general-purpose vector processing system and tabular database management system. The original DEM data were first processed with the conditioning algorithms to create a depressionless elevation data set, and were then subtracted from the depressionless elevation data. The resulting difference data set was processed through a raster-to-polygon conversion algorithm, which built a depression data set. The depression data set and the difference data set were then used to calculate area and volume for each depression in the tabular data-base management system. Watersheds were delineated using selected depressions from the depression mask as watershed "starts" for each pothole. Finally, all possible pour points were computed for each watershed. Because of the low, rolling topographic relief of the terrain, manual delineation and flow modeling of potholes and their watersheds based on topographic maps would have been extremely time consuming. Automated extraction proved to be the most efficient method.

It has been determined that the accuracy and detail of the hydrologic information that can be automatically extracted from a DEM with these algorithms is directly related to the quality and resolution of the DEM itself. Because the algorithms cannot distinguish between artifacts and real features, a severe artifact may critically disrupt flow paths. If the elevation values of a DEM do not capture the crucial hydrologic features, such as many happen when multiple shallow drainages cross an area of low relief, the resolution of the DEM will not be sufficient for accurate automatic delineation of watershed boundaries between the drainages.

CONCLUSIONS

Software tools that were developed to derive morphologic information from DEMs have proven useful in hydrologic applications. The tools are standard FORTRAN programs in the public domain. In practical usage, however, the tools require general-purpose raster, vector, and tabular display and analysis support capabilities. Research is ongoing to find new applications for the toolbox and to expand the toolbox. Processing of drainage line information in the vector domain holds potential for the automatic extraction of gradient and stream ordering information. The efficient derivation of watersheds for large numbers of stream sediment and hydrologic geochemical samples also presents an extremely useful potential application and a software development challenge.

REFERENCES

Band, L.E., 1986. Topographic partition of watershed with digital elevation models: *Water Resource Research,* Vol. 22, No. 1, pp. 15-24.

Collins, S. H., 1975. Terrain parameters directly from a digital terrain model: *Canadian Surveyor,* Vol. 29, No. 5, pp. 507-518.

Douglas, D. H., 1986. Experiments to locate ridges and channels to create a new type of digital elevation model: *Cartographica,* Vol. 23, No. 4, pp. 29-61.

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

Jenson, S. K., 1985. Automated derivation of hydrologic basin characteristics from digital elevation model data: *Proceedings of Auto-Carto* 7, Washington, D.C., pp. 301-310.

Jenson, S. K., and C. M. Trautwein, 1987. Methods and applications in surface depression analysis: *Proceedings of Auto-Carto 8,* Baltimore, Maryland, pp. 137-144.

Klingebiel, A. A., E. H. Horvath, W. U. Reybold, D. G. Moore, E. A. Fosnight, and T. R. Loveland, 1988. *A Guide for the Use of Digital Elevation Model Data for Making Soil Surveys:* U.S. Geological Survey Open-File Report 88-102, 18 p. [in press].

Mark, D. M., 1983. Automated detection of drainage networks for digital elevation models: *Proceedings of Auto-Carto 6,* Vol. 2, Ottowa, Ontario, Canada, pp. 288-298.

Marks, D., J. Dozier and J. Frew, 1984. Automated basin delineation from digital elevation data: *Geo-Processing,* 2 pp. 299-311.

O'Callaghan, J. F., and Mark, D. M., 1984. The extraction of drainage networks from digital elevation data: *Computer Vision, Graphics and Image Processing,* Vol. 28, pp. 323-344.

Peucker, T. K., and Douglas, D. H., 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.

U.S. Geological Survey, 1987. *Digital Elevation Models:* U.S. Geological Survey Data User's Guide 5, 38 p.