AUTOMATED DERIVATION OF HYDROLOGIC BASIN CHARACTERISTICS FROM DIGITAL ELEVATION MODEL DATA

Susan K. Jenson*
Technicolor Government Services, Inc.
Sioux Falls, South Dakota 57198

BIOGRAPHICAL SKETCH

Susan Jenson is a Senior Applications Scientist at the U.S. Geological Survey's EROS Data Center in Sioux Falls, South Dakota. She received her B.S. in mathematics from the University of South Dakota in 1973.

ABSTRACT

Digital elevation model (DEM) data in a raster format can be used to automatically derive the drainage characteristics of an area. A procedure has been designed that is capable of operating on matrices of elevation data having no algorithmically imposed size limit, while performing within the resolution and accuracy tolerances of the DEM data.

Each cell is processed as the center of a 3- by 3-cell spatial window in the raster elevation data. If a cell is a local minimum in comparison with two of its non-adjacent neighbors, it is labeled as a drainage cell. The linkages of the drainage cells within user-specified distance and elevation thresholds are established in a separate process. The products of these processing steps are digital masks of the drainage cells and the watershed basins, both in raster format.

A drainage cell mask derived using this procedure is useful in computing slope values for a raster data base. Slope has traditionally been calculated for each cell by fitting a plane through the eight nearest cells. However, if the terrain represented by these cells is V-shaped, such as a gully, a plane does not fit well; in fact, the desired slope value is the slope along the bottom of the gully, regardless of the steepness of the gully sides. The automated drainage process will label such a cell as a drainage cell, and its slope can then be computed from the elevation values of neighboring drainage cells.

INTRODUCTION

The U.S. Geological Survey (USGS) has developed an operational procedure to digitally delineate watershed

---

Publication authorized by the Director, U.S. Geological Survey on November 15, 1984.

*Work performed under U.S. Geological Survey contract no. 14-08-0001-20129.

boundaries from raster elevation data and gaging station locations. The project had the following conditions:

1. The procedure must be able to process matrices of elevation data having no algorithmically imposed size limit.

2. The procedure must be adaptable to the accuracy and noise level of the data.

3. The procedure must be able to recognize noncontributing areas of a watershed. For example, a quarry or sinkhole and the area that drains to it must be labeled differently than the surrounding area.

4. The computer implementation must be transportable and memory-efficient.

ALGORITHMS AND IMPLEMENTATION

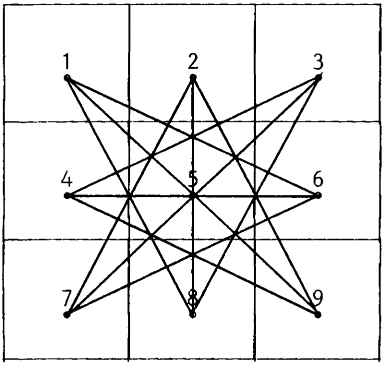
The algorithms described here have three major steps: identifying drainage cells, grouping drainage cells, and linking groups of drainage cells. Two raster products, a mask of drainage cells for an area and a mask of the watershed basins for an area, are produced. The two products are closely related in that the drainage mask is necessary for delineation of basins and the basin-making process establishes the drainage cell linkages. The computer program makes approximately 10 passes through the data, reading lines of elevation data and reading and writing lines of drainage labels.

The data source for figures 2 through 7 is USGS DEM data collected with the Gestalt Photo Mapper II system and within a vertical accuracy of 7 meters weighted root mean square error (Elassal and Caruso, 1983).

In order to define basins based on gaging stations, the elevation value at each gaging station location is made negative. For instance, if a gaging station location has an elevation of 1,100 meters, that elevation value is changed to -1,100 meters. This has the effect of introducing a hole, or pit, in the surface. Since the algorithms function within elevation tolerances, they recognize these pits and their watersheds as being disconnected from the surrounding watersheds by the same logic that isolates natural non-contributing areas.

Identifying Drainage Cells

The first step is to label each cell in the raster elevation data by its drainage characteristics. The premise for this step is that if an area's elevation profile is V-shaped, it will channel water and should be part of a drainage network. Each cell is isolated as the center of a 3- by 3-cell neighborhood and each neighborhood's elevations are examined in cross-section. As shown in figure 1, there are 12 possible cross sections which intersect at least a portion of the center cell. Cross sections such

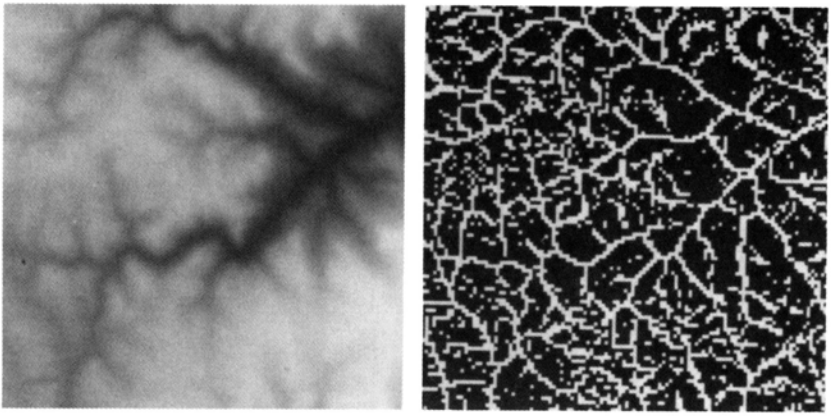


Symmetric Cross Sections	Asymmetric Cross Sections
1-9	1-6
2-8	1-8
3-7	2-7
4-6	2-9
	3-4
	3-8
	4-9
	6-7

Figure 1.--The 12 possible cross sections intersecting at least a portion of the center cell of a 3-cell by 3-cell neighborhood.

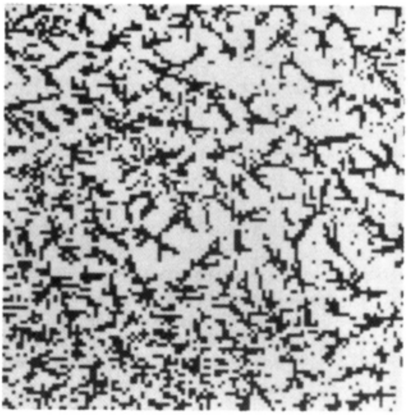
as 1-3 which do not intersect cell five, are not considered because it is cell five's characteristics that are being examined. Cross sections are further categorized as symmetric or asymmetric as shown in figure 1. The symmetric cross sections are more valuable than the asymmetric because they are more strongly controlled by cell five's elevation and hence are more indicative of cell five's characteristics. If both end points of a cross section are higher in elevation than cell five, then the area is V-shaped and cell five is a local minimum. Figure 2 shows a small matrix of elevation data with its symmetric and asymmetric local minima. This matrix is 120 by 120 cells in size with each cell representing 50 by 50 meters on the ground. The 50-meter cell size was produced by bilinearly resampling the original 30-meter USGS DEM cells in order to match the cell size of an existing spatial data base. The symmetric minima visually correlate to paths of drainage and the asymmetric minima thicken the paths. At this point, the mask of drainage cells has been established. This intermediate product may be retained and used on its own merit or processed further to define watersheds.

The drainage cell identification process is implemented as a 3- by 3-cell moving window. The process requires only one pass through the data and minimal processing time.



Elevation Data

Symmetric Minima in White



Symmetric and Asymmetric Minima in White

Figure 2.--A 120- by 120-cell matrix of elevation data with its symmetric and asymmetric local minima.

Grouping Drainage Cells

The second step is to group the drainage cells by drainage channel. The premise here is that if two drainage cells are adjacent, and the cell that is lower in elevation belongs to channel X, then the other cell also belongs to channel X. To begin the grouping process the drainage cell with the lowest elevation is found and given channel label one. This cell's eight adjacent neighbors are then tested. If a neighbor cell is a drainage cell with no channel label and is not lower in elevation than the original center cell, it is given a label of one. The neighbor cells of all cells so marked are tested, as are neighbors of neighbors, and so on, until no more cells qualify to receive the label one. The process is then repeated, with the unlabeled drainage cell with the lowest elevation as the new starting point, until all drainage cells have channel labels. Figure 3 shows elevation data and the drainage cells with their channel labels. The circled cells are defined to be roots for the channels. A root is the lowest elevation cell for its channel and is a starting point in one of the iterations of the channel labeling process.

It is necessary to limit the cells grouped in this step to drainage cells only. If all cells were considered in the process, a channel beginning on one side of a hill could grow through a saddle area and erroneously continue to grow upwards on the other side of the hill.

Elevation Data

102	101	101	100	100	102
101	101	99	98	99	103
100	101	97	97	100	101
94	95	101	98	99	100
100	102	103	97	98	99
102	102	100	99	99	99

Channel Labels

0	0	0	1	1	0
0	0	1	1	1	0
0	0	1	1	0	0
①	1	0	1	0	0
0	0	0	②	2	0
0	0	2	2	0	0

Figure 3.--A 6-cell by 6-cell matrix of elevation data is shown with its corresponding channel labels. The zeroes in the matrix of channel labels correspond to cells which are not minima. The two circled cells are channel roots. The computer implementation of this step is based on a last-in/first-out (LIFO) stack. The stack is reinitialized for each channel label. When a cell is labeled, its location is pushed up on the stack. When the stack is completely processed, a new root is sought.

If the elevation values rose smoothly from the root of a drainage up to the tops of all the branches, the channels labeled in the second step would be spatially extensive and only one label would be needed to cover a watershed. In practice, however, the elevation values undulate along a drainage, especially in areas of low gradient. Each undulation causes the labeling process to begin another channel label, as illustrated in figure 3. In figure 3, drainage line two is actually a tributary of drainage line one, but the root cell of line two is artifically low at the joining location. This grouping step uses approximately 4,000 labels to process a 512- by 512-cell data set. Finally, all non-drainage cells are given the label of the channel to which they drain, shown in figure 4. This is accomplished by iteratively labeling cells uphill from the labeled cells until all cells are labeled. The labels now correspond to watersheds for channels which are artificially small due to noise and inaccuracies in the elevation data. A third processing step is now required to establish the linkages among these mini-basins.

Elevation Data

102	101	101	100	100	102
101	101	99	98	99	103
100	101	97	97	100	101
94	95	101	98	99	100
100	102	103	97	98	99
102	102	100	99	99	99

Extended Channel Labels

1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	2
①	1	1	1	2	2
1	1	2	②	2	2
1	1	2	2	2	2

Figure 4.--The 6- by 6-cell matrix of elevation values from figure 3 is shown with its corresponding extended channel labels. The circled cells are channel roots.

Linking Groups of Drainage Cells

The basic premise for basin linking is that if the root cell of basin X is sufficiently close in distance and in elevation to a cell with a different label, for example Y, and there is a path of sufficiently low elevations connecting the cells, then basin X is actually a continuation of basin Y. In the linking step, these distance and elevation tolerances are empirically determined for a given topographic data set and basin generalization requirement.

To determine basin linkages, each root's spatial neighborhood is examined for cells with a different label, with a path of acceptable elevations, and within the distance and elevation tolerances. In figure 4, channel two will link to channel one with an elevation tolerance of zero, a path tolerance of one, and a distance tolerance of two cells. The linkage could also be made with an elevation tolerance of one, a path tolerance of zero, and a distance tolerance of one cell. The user typically begins with small tolerances and gradually increases them in each iteration until the tolerances approach the accuracy of the data. When channel two links to channel one, all cells with a label of two are relabeled one. Channel two's root is then a member of channel one and has lost its distinction as a root.

When the process is completed, the entire matrix has been subdivided into watersheds. Some will have gaging stations for their roots. Others will have the lowest elevation in a non-contributing area for their roots, or they will be draining off the edge of the data set so the root will be on the data set edge. Still others may be erroneously segmented into sub-watersheds because errors in the DEM data exceeded the ability of the algorithms to establish linkages. In these latter cases, the user must relabel the watersheds using a mapping or renumbering program.

Figure 5 shows the three watersheds that were found for the data in figure 2, with the symmetric local minima superimposed in black.

APPLICATION TO SLOPE CALCULATION

A drainage cell mask derived using this procedure is useful in computing slope values in a raster data base. Slope has traditionally been calculated for each cell by fitting a plane through the eight nearest cells. However, when a cell is part of a drainage path, the desired slope value is the gradient along the drainage path regardless of the steepness of the side slopes. Since drainage cells can be identified in the automated process, the slope calculation can be modified to recognize drainage cells and use only other neighboring drainage cells to calculate their slope. Figure 6 shows a 12-cell by 8-cell matrix of elevation data with slopes calculated traditionally, as well as by special treatment of symmetric local minima drainage cells. Revised percent slopes are shown for symmetric local minima only. The dashes in the table of revised percent slopes represent the cells which are not symmetric local minima. On average, the slope values for drainage cells dropped by 7 percent. Individually, however, there was a great deal of fluctuation. For

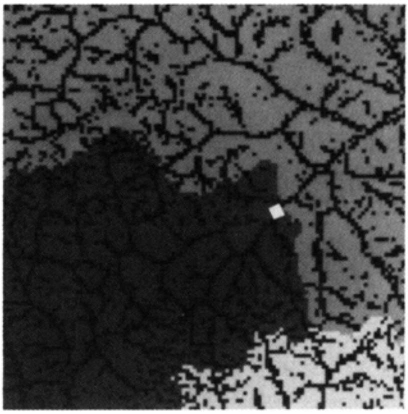


Figure 5.--The three watersheds found for the data in figure 2 are shown here as three shades of gray. The symmetric local minima are superimposed in black. The white cursor (□) shows a gaging station location.

example, note the behavior of the drainage cell at line 8, column 2. This cell's 3-cell by 3-cell neighborhood is examined in figure 7. By the traditional method, the slope was 17 percent. Using only the elevations of other neighboring drainage cells, the slope was 24 percent. Since each cell represents a 50-meter square area, a 17-percent slope is equivalent to a drop per cell of 8.5 meters; a 24-percent slope is equivalent to a drop of 12 meters. The example in figure 7 shows an area which drains to the right. When a plane is fit through the elevation values of drainage cells only (figure 7b), the plane slopes to the right with an average drop of 12 meters. When a plane is fit to all elevation values (figure 7a), the steepness of the plane is mitigated by the values at line 7, columns 2 and 3, and line 9, columns 2 and 3; and the average drop is lessened to 8.5 meters.

Elevation Data, in Meters

COLUMN

LINE	1	2	3	4	5	6	7	8
1	1720	1728	1719	1727	1735	1736	1725	1717
2	1725	1735	1725	1726	1732	1732	1725	1721
3	1730	1735	1729	1723	1726	1724	1720	1718
4	1739	1737	1727	1722	1724	1722	1717	1711
5	1739	1733	1726	1722	1724	1721	1715	1706
6	1735	1732	1722	1712	1716	1713	1712	1698
7	1729	1727	1717	1716	1707	1703	1696	1687
8	1729	1719	1712	1708	1694	1684	1686	1693
9	1739	1729	1718	1700	1691	1692	1706	1723
10	1744	1742	1729	1702	1689	1702	1725	1743
11	1739	1738	1732	1714	1691	1694	1719	1737
12	1732	1730	1726	1720	1699	1692	1711	1732

Percent Slopes by Traditional Method

COLUMN

LINE	1	2	3	4	5	6	7	8
1	23	1	10	15	8	10	18	19
2	18	8	10	7	10	12	12	10
3	14	7	12	4	7	11	12	9
4	10	10	13	3	2	7	12	11
5	11	13	15	9	8	10	16	16
6	11	15	17	11	15	19	24	27
7	11	17	16	16	22	28	24	13
8	15	17	16	22	23	9	14	34
9	26	27	30	28	11	22	46	48
10	16	16	36	38	11	30	40	25
11	12	13	24	39	18	27	43	28
12	17	19	20	27	24	16	41	36

Revised Percent Slopes for Symmetric Local Minima
(Dashes indicate cells which are not minima)

COLUMN

LINE	1	2	3	4	5	6	7	8
1	-	-	10	-	-	-	-	-
2	7	-	8	8	-	-	-	-
3	7	-	-	3	-	-	9	-
4	-	-	-	-	-	-	14	14
5	-	-	-	7	-	-	-	14
6	-	-	-	15	-	-	-	-
7	10	-	-	-	23	-	-	14
8	-	24	15	-	19	27	6	19
9	20	-	-	18	9	11	-	-
10	-	-	-	-	4	-	-	-
11	-	-	-	-	5	7	-	-
12	-	-	-	-	-	1	-	-

Figure 6.--Comparison of slopes calculated traditionally versus along drainages.

Elevation Values Used to Compute Slope

Traditional Slope

	1	2	3
7	1729	1727	1717
8	1729	1719	1712
9	1739	1729	1718

Figure 7(a)

Drainage-Based Slope

	1	2	3
7	1729	--	--
8	--	1719	1712
9	1739	--	--

Figure 7(b)

Percent Slope

Traditional Slope

	1	2	3
7	11	17	16
8	15	17	16
9	26	27	30

Figure 7(c)

Drainage-Based Slope

	1	2	3
7	10	--	--
8	--	24	15
9	20	--	--

Figure 7(d)

Figure 7.--An example of the difference between traditional and drainage-based slope calculation based on the data in figure 6.

CONCLUSION

Digital elevation data have potential for use in creating hydrologic categories in digital spatial data bases. It is now possible to automatically derive spatially referenced drainage cell and watershed masks, and these basic derivative products enable improved calculation of slopes.

REFERENCES

Elassal, A. A. and Caruso, V. M., 1983, "Digital elevation models." U.S. Geological Survey Circular 895-B.