

APPLICATIONS OF HYDROLOGIC INFORMATION AUTOMATICALLY EXTRACTED FROM DIGITAL ELEVATION MODELS

SUSAN K. JENSON

U. S. Geological Survey, EROS Data Center, Sioux Falls, South Dakota 57198, U.S.A.

ABSTRACT

Digital elevation models (DEMs) can be used to derive a wealth of information about the morphology of a land surface. Traditional raster analysis methods can be used to derive slope, aspect, and shaded relief information; recently-developed computer programs can be used to delineate depressions, overland flow paths, and watershed boundaries. These methods were used to delineate watershed boundaries for a geochemical stream sediment survey, to compare the results of extracting slope and flow paths from DEMs of varying resolutions, and to examine the geomorphology of a Martian DEM.

KEY WORDS Digital elevation models Geographic information systems Drainage basin Geomorphology Topography

BACKGROUND

DEMs (U.S. Geological Survey, 1987a) are raster representations of land surfaces that have historically been of general utility to provide slope, aspect, and shaded relief information for spatial analysis (Klingebiel *et al.*, 1988). In addition to these analysis capabilities that are available in general-purpose raster analysis systems, software has been developed at the U.S. Geological Survey's EROS Data Center to delineate depressions, overland flow paths, and watershed boundaries from DEMs. The software consists of twelve programs that can be combined to produce a variety of application-specific derivative datasets. The recommended usage of the software is to first process the DEM through a conditioning phase and then follow a processing flow tailored to the application. The software was designed to process very large DEMs, to automate the process as much as possible, and to accommodate DEMs with artifacts and large flat areas. Jenson and Domingue, 1988, provides numeric and visual examples and a more rigorous description of the algorithms than is given here. A comparison to other procedures that extract information from DEMs (Band, 1986; Collins, 1975; Mark, 1983; Marks *et al.*, 1984; O'Callaghan, 1984), is given in Jenson and Domingue, 1988, and also in Douglas, 1986.

The conditioning phase consists of three steps. First, depressions in the DEM are 'filled' by raising the values of cells in depressions to the value of the depression's spill point. Depressions are usually artifacts of the DEM generation process, but may also be bonafide depressions, such as potholes. The original DEM can be subtracted from the depressionless DEM to provide a mask of the depressions and the depth of each cell in each depression. The second preprocessing step is the computation of the flow direction for each cell in the depressionless DEM. The direction water will flow out of each cell is encoded to correspond to the orientation of one of the eight cells that surround the cell. In flat areas, the flow directions are iteratively calculated so that the flow path traverses the flat and continues downhill at one of the flat's spill points. Flow directions are thus quantized into eight possibilities, and cannot represent large facets oriented at angles other than these eight; however, the process has repeatedly performed well on land surfaces characterized by well-established drainage. The third conditioning step is the computation of the flow accumulation value for

each cell. This is simply the count for each cell of how many upstream cells would contribute to it based on their flow directions. Because all cells in a depressionless DEM have a flow path to the dataset edge, the pattern formed by highlighting cells with a flow accumulation value greater than some threshold value delineates a fully-connected drainage network. As the threshold value is decreased, more cells become highlighted and the density of the drainage network increases.

After the conditioning phase, the datasets can be further processed to delineate watersheds. The outflow points, or 'seeds', for which watersheds are desired can be selected by the analyst. Alternatively, an automatic procedure can initiate 'seeds' at all confluences of a flow accumulation drainage network defined by a user-specified density. All seeded watersheds are delineated in one execution of the watershed program. A watershed linkage table containing the label of each watershed's downstream watershed and the spill point location and elevation can be automatically generated.

The software recognizes cells with a value of zero as the study area 'mask'. If drainage networks and watersheds are desired that are interior to a depression such as a large crater, the analyst can indicate a cell or cells in the bottom of the depression that are to be considered 'mask' by changing their elevation value to zero. With this situation, flow directions and flow accumulation drainage lines will run down the sides of the depression and end at the 'mask' cells.

The watersheds, drainage networks, and linkage tables that are generated by this software are readily convertible to vector and tabular form for geographic information system analysis. Even though some manual editing may be required, the automatic generation of this hydrologic information from DEMs saves much time in manual delineation and digitizing. An additional advantage of the derived information over manually-delineated information is that it is precisely registered to the DEM, allowing more confidence in the locational accuracy of elevation data used to calculate slopes for stream segments.

WATERSHED DELINEATION FOR SAMPLE SURVEYS

Watershed boundaries are desirable in geochemical stream sediment surveys so that the chemical concentrations at the sample locations can be related to the lithological composition of the contributing watersheds (Bonham-Carter *et al.*, 1987). A sixty-by-forty kilometre subarea of a geochemical stream sediment study near Tonopah, Nevada was selected to illustrate how automated DEM techniques can be used to delineate these watersheds. The first step of the analysis was to perform the conditioning steps producing a depressionless DEM, a flow directions dataset, and a flow accumulation dataset.

Figure 1a is shaded relief of the DEM with stream sediment sample locations marked by white dots. The cell at each sample location has a unique identification number. The data source is a 1:250 000 scale U.S. Geological Survey DEM that has been interpolated to a 200 metre (ground equivalent) cell size. Figure 1b shows the shaded relief with automatically-generated flow accumulation values that are greater than 100 cells superimposed in white. Figure 1c shows the locations of cells with flow accumulation values greater than 100 cells in gray and the sample locations in white. The drainage network and the sample locations do not coincide precisely, primarily due to imprecision in recording and encoding the sample locations. A program was developed as part of the automated DEM analysis software to relocate stream sediment samples so that they could become 'seeds' for the watershed delineation program. In Figure 1d, the flow-accumulation drainage network and most of the sample locations have been forced to coincide by moving non-coincident sample locations to the nearest cell with a flow accumulation value greater than 100 cells. The program that relocates the samples requires two parameters; the flow-accumulation-value threshold and the maximum distance a sample is allowed to move, 100 cells and ten cells, respectively, in this example. These parameters are intended to be determined by visual examination with general purpose raster display software. The program reports how far each sample was moved, a list of samples that could not be moved within the maximum distance tolerance, and instances where two sample locations are occupying the same cell. In this example, 139 of the 194 samples were moved, and 16 sample locations became coincident.

After the samples were moved, each sample's watershed was delineated automatically from the flow direction dataset. Raster to vector format conversions were then performed for the watersheds, flow accumulation values, and relocated sample points. A plot of these three datasets, represented by solid lines,

dotted lines, and large dots, respectively, is shown in Figure 2a. At this stage of analysis, the basic categories of information have been generated to facilitate general-purpose spatial modelling. For instance, a vector representation of lithology, as shown in Figure 2b, can be intersected with the watersheds using a general-purpose geographic information system. The watershed composition information computed by intersecting watersheds with lithology, combined with pour-point and watershed connectivity information that were automatically generated, compose a tabular database. At this point the database contains sufficient information to relate each sample to the lithologic composition of all upstream watersheds and test and establish predictive models for observed chemical concentrations as a function of lithologic composition.

While this type of analysis is certainly possible by manually delineating watersheds, digitizing them, and hand-encoding their linkage attributes, the manual approach is quite labour intensive. The automated approach offers a practical way to analyse the large numbers of sample points that are required to provide statistically significant results. The automated results can, of course, be edited and augmented manually.

LARGE AREA ANALYSIS

Need and data availability

Topographic information is vital for a diverse range of land science applications in hydrology, ecology, geophysics, geology, geomorphology, and polar studies; however the availability of fine-scale global and continental digital topographic information is severely limited. The Topographic Science Working Group Report (Topographic Science Working Group, 1988) identified many of these applications and surveyed existing topographic data. Global DEM data with a five arc-minute cell size, termed ETOPO5, is available from the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center (NGDC). For the conterminous United States, a DEM with a 30 arc-second resolution has been assembled by sampling every tenth sample of every tenth line of the U.S. Geological Survey's three arc-second resolution DEMs. This dataset is also available from NOAA's NGDC.

Analysis

Automated hydrologic feature extraction and slope techniques were applied to ETOPO5, to the U.S. 30 arc-second DEM, and to a three arc-second resolution DEM to investigate how well terrain information is preserved at these resolutions.

ETOPO5

The ETOPO5 dataset contains bathymetric information as well as land surface elevations; however, only elevations above sea level were processed in this study. ETOPO5 was processed by the automatic DEM conditioning programs and the resultant drainage lines are shown superimposed on ETOPO5 shaded relief in Figure 3a. The drainage lines were produced by thresholding the flow accumulation dataset at 1500 cells. There is a coarse correspondence with actual major river systems, however many errors are present due to lack of detail in the DEM. One of the reasons for lack of detail in ETOPO5 is that most of the land area elevations were produced by interpolation from ten arc-minute resolution elevations. This lack of detail is illustrated in Figure 3b where flat areas are superimposed on ETOPO5 shaded relief. Areas highlighted in white are blocks of cells that are identical in elevation value.

U.S. 30 arc-second DEM and three arc-second DEMs

The U.S. 30 arc-second DEM was investigated by processing a three and 1/3 degree square area centered on the Delaware River and including portions of the states of Pennsylvania, New Jersey, and New York. Figure 4a shows the 30 arc-second shaded relief of this area with flow accumulation values greater than 200 superposed in white. Figure 4b shows the 1:2 000 000 scale hydrography and state boundaries for this area (U.S. Geological Survey, 1987b). Most rivers, including the Delaware River, indicated by 'B' on Figure 4b were successfully delineated automatically. The Chemung River, indicated by 'A', was erroneously calculated

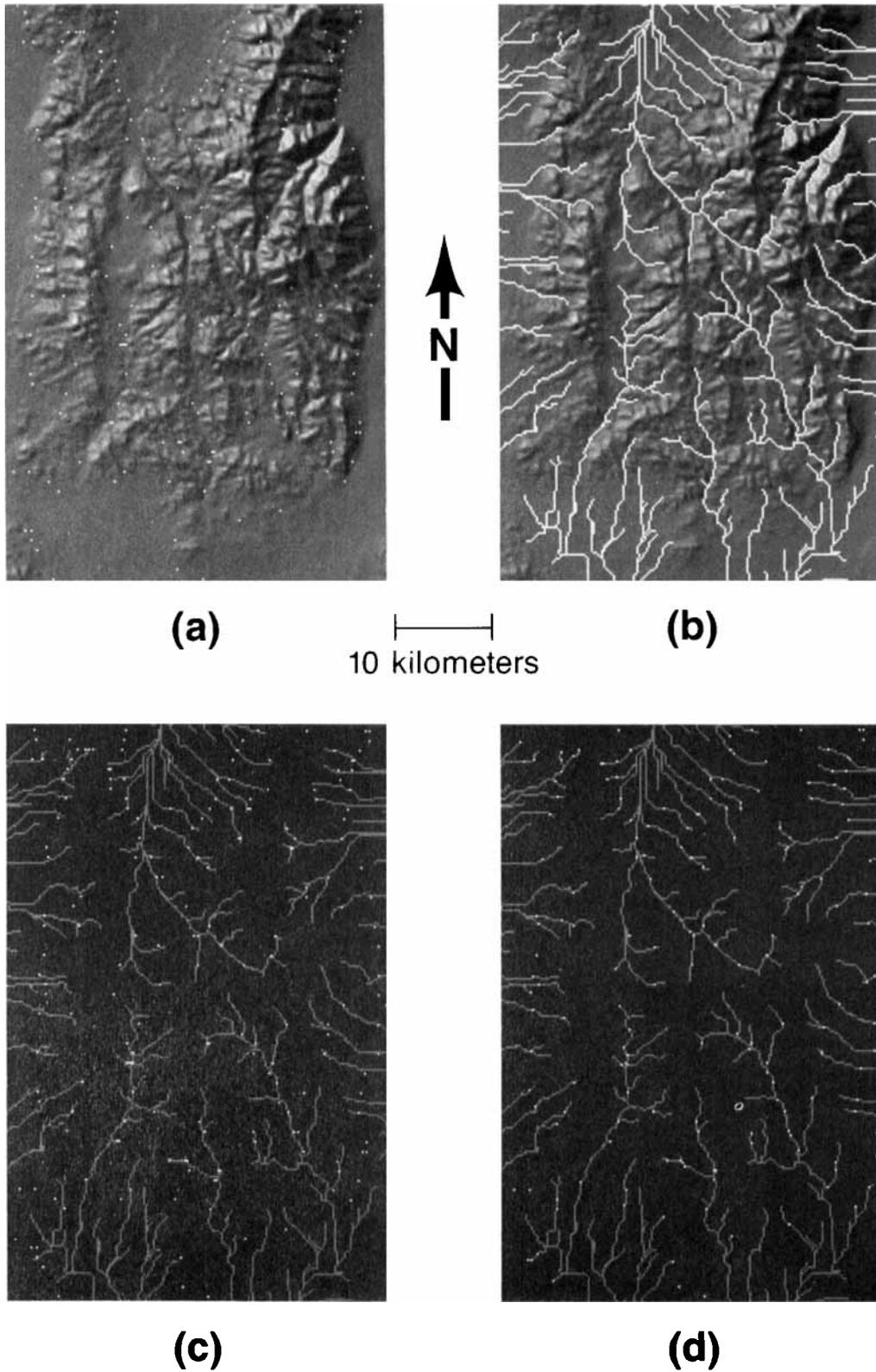


Figure 1. In (a), a shaded-relief representation of a 60 by 40 kilometre DEM near Tonopah, Nevada is shown with geochemical sample site locations shown in white. In (b), the drainage network automatically extracted from the DEM (cells with flow accumulation values greater than 100 cells) has been highlighted in white on the shaded-relief base. In (c), the drainage network and sample locations are shown in gray and white, respectively. In (d), the drainage network is shown in gray and modified sample locations are shown in white, illustrating the results of relocating sample locations to the nearest drainage network cell within a distance of ten cells

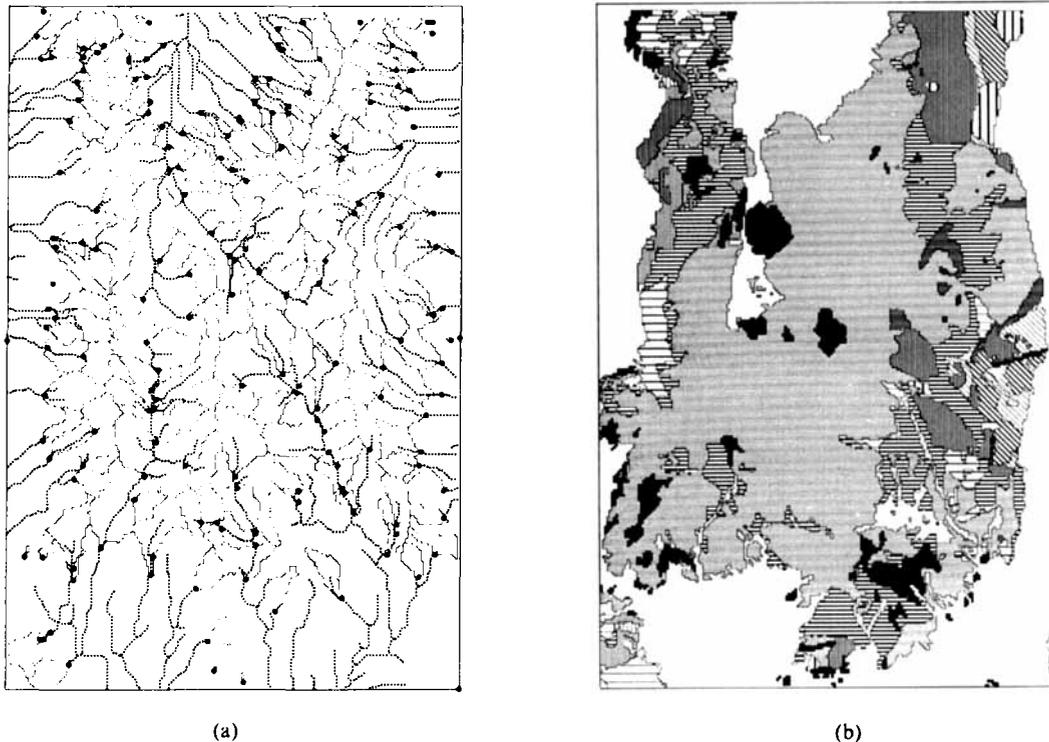
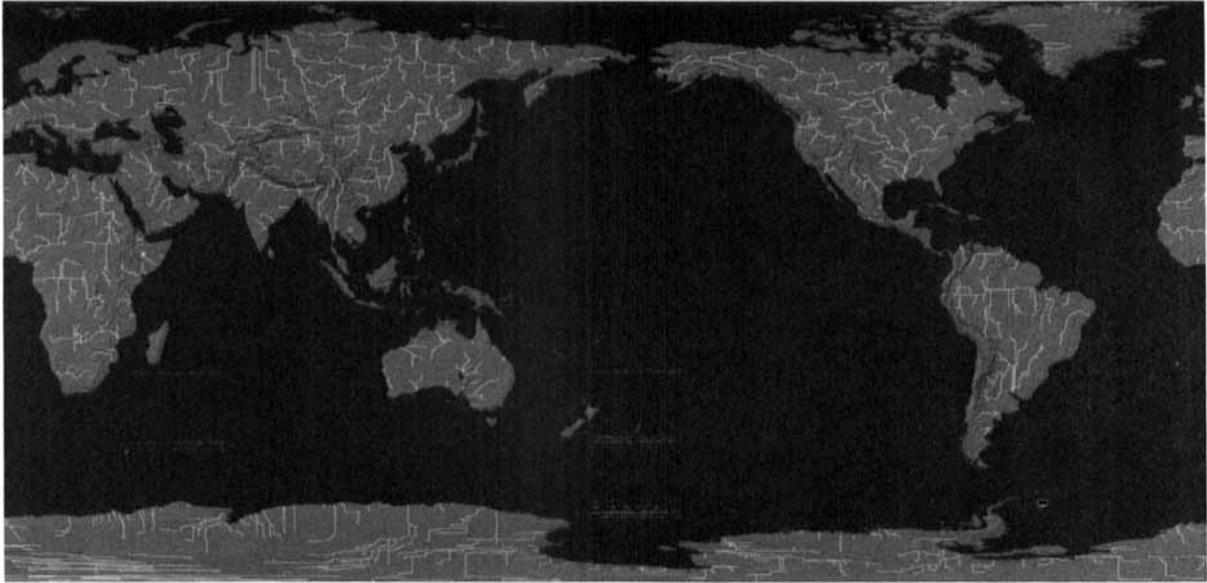


Figure 2. In (a), the drainage network and sample locations in Figure 1d are plotted, as dotted lines and large dots respectively, after conversion to a vector format. Automatically-generated watershed boundaries for each sample location were also converted to a vector format and are shown plotted as solid lines. In (b), the lithology types are plotted

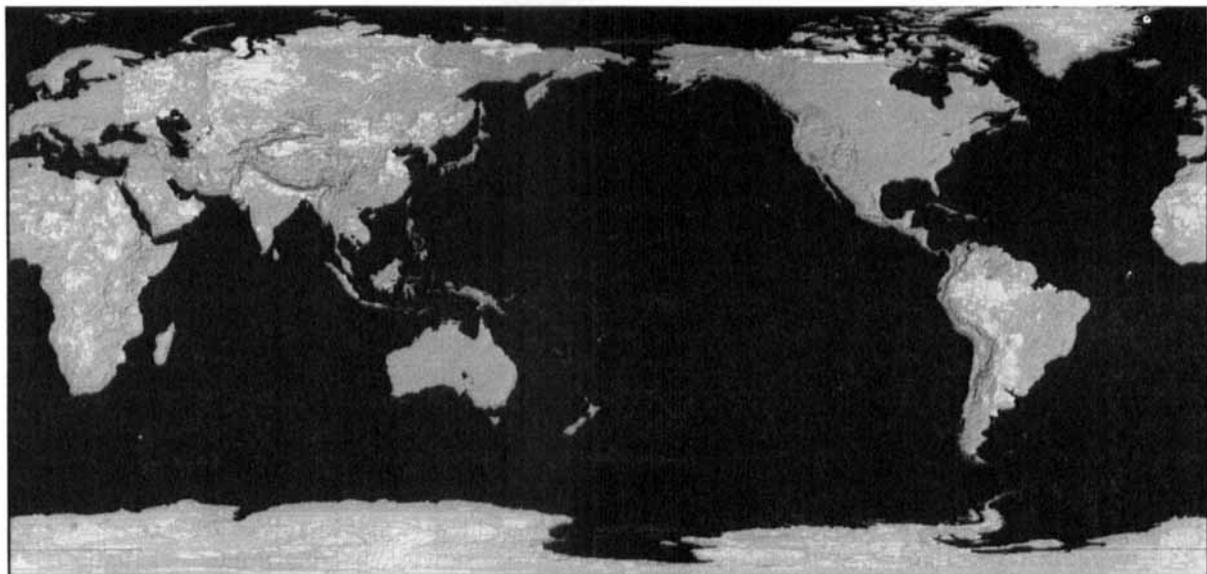
to flow north into the Finger Lakes region of Central New York state because a critical drainage divide is not represented in the DEM. The Chemung area is enlarged in Figure 4c. The USGS three arc-second DEM data for the Elmira 1:250 000 scale map covering the Chemung area was also processed by automated techniques. A shaded relief of the three arc-second DEM for the area coincident with Figure 4c and with flow accumulation values greater than 200 superposed in white is shown in Figure 5(a). The automated calculations for the Chemung River are also erroneous in this dataset. An examination of the DEM shows that the elevation values are all equal to 274 m in the critical watershed divide area. Figure 5b shows the Chemung area shaded relief with all cells equal to 274 m highlighted in white. This figure also illustrates an artifact problem frequently encountered with three arc-second DEM data. Note the three angular indentations into the flat area. They represent areas of erroneously higher elevation values introduced during the DEM interpolation process. These artifacts are associated with areas where map-based topographic contours are widely spaced. These errors and lack of detail cause the automated techniques to incorrectly delineate the Chemung River in this dataset. An examination of the 1:250 000 scale Elmira map contours, from which this DEM was interpolated, shows that this drainage divide is not represented in the map's contour information. The contours of the 1:62 500 scale map do represent the divide, however. If contours from the 1:62 500 scale source were used to define elevation detail in this area at a three arc-second cell size, the Chemung River would be successfully delineated automatically. This experiment indicates that a three arc-second DEM resolution is probably adequate to represent hydrologic features that are significant at a scale of 1:2 000 000, and a 30 arc-second DEM is probably adequate for most features, but that the source contour information used to generate the DEM must be from a scale of 1:62 500 or greater.

Effect of DEM scale on slope calculation

The effect of DEM scale on calculated values of slope was investigated by computing slope statistics by two methods for DEMs of three spatial resolutions. The first method, termed the 'mean' method in this



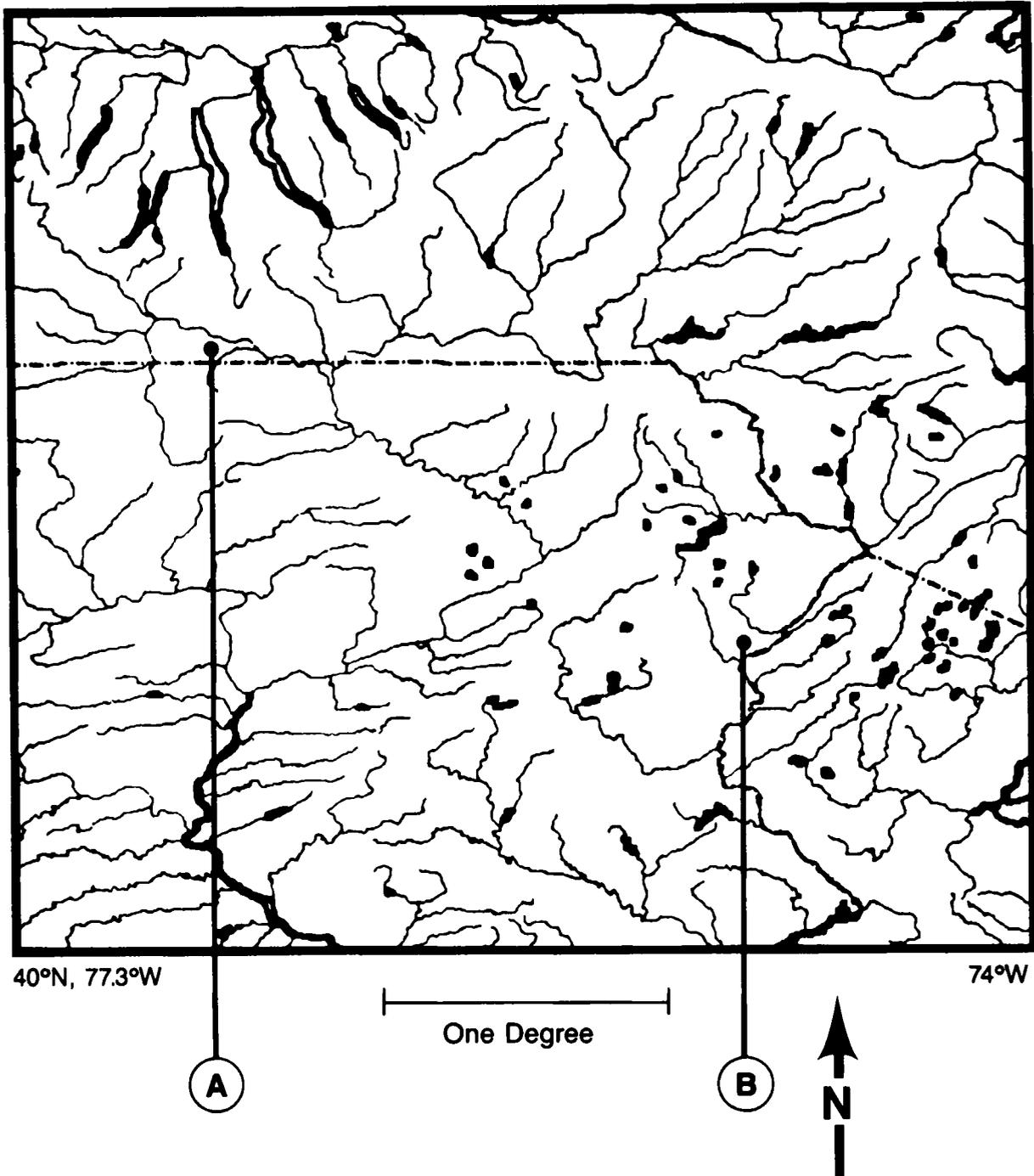
(a)



(b)

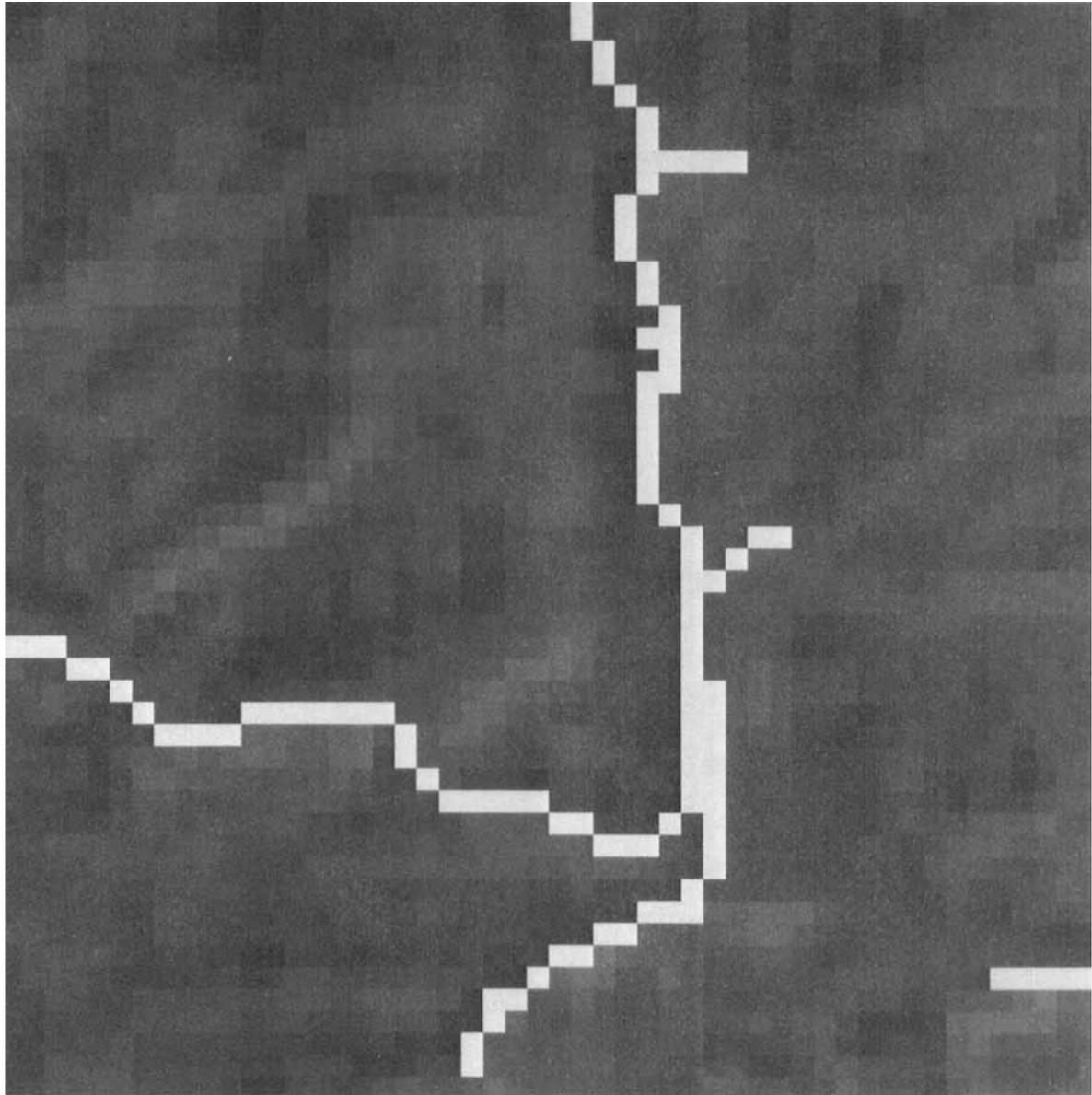
Figure 3. In (a), a shaded-relief representation of the ETOPO5 global DEM is shown with automatically extracted drainage networks highlighted in white. The cells with elevations less than or equal to zero are shown in black. In (b), the ETOPO5 shaded-relief is shown with flat areas highlighted in white

43.3°N



(b)

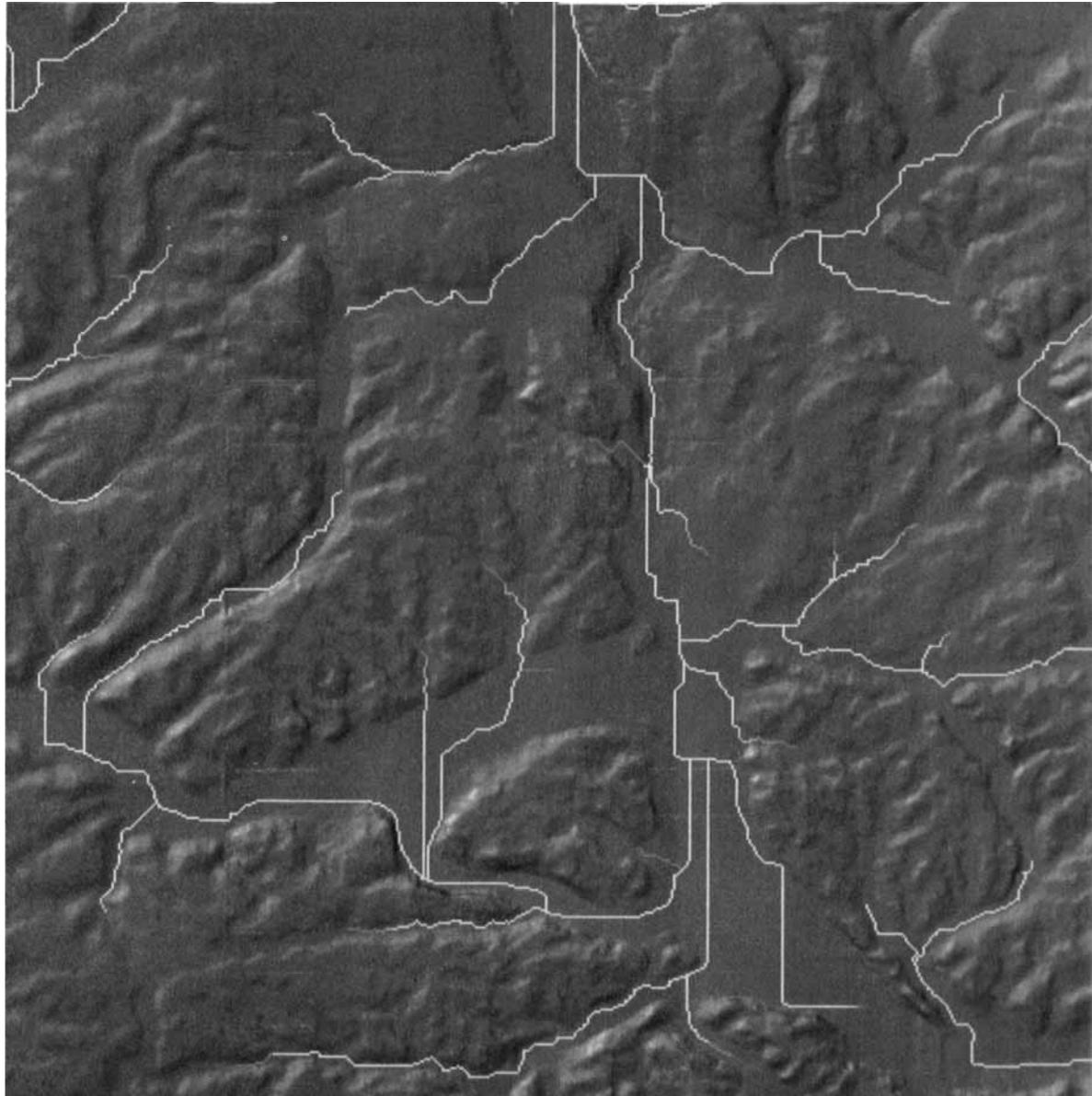
Figure 4 continued. In (b) the 1:2 000 000 scale hydrography and state boundaries are shown for this area in Figure 4 (a). 'A' indicates the Chemung River and 'B' indicates the Delaware River.



5 arc-minutes

(c)

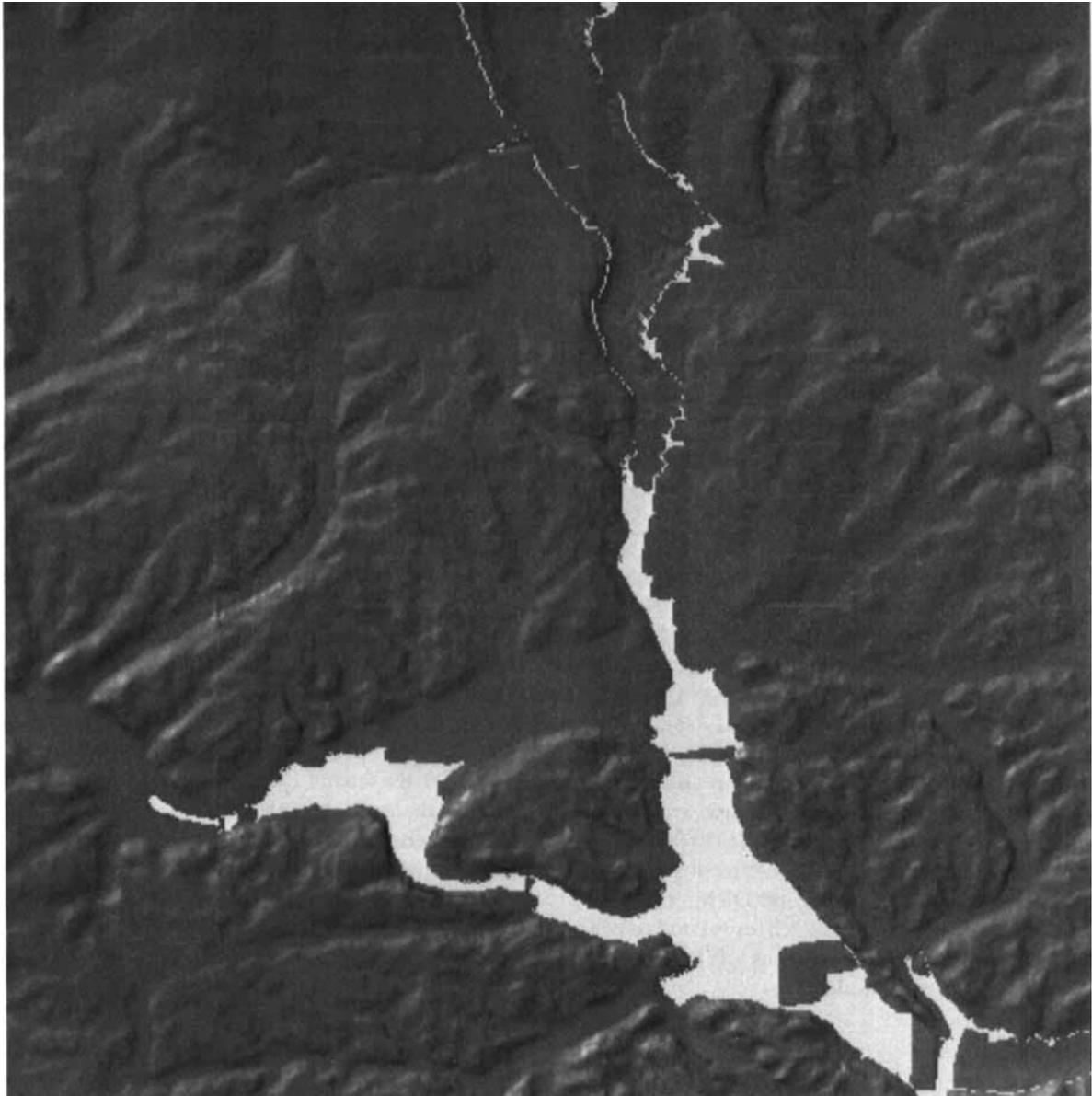
Figure 4 continued. In (c), the Chemung River area indicated by 'A' in Figure 4(b) is enlarged to illustrate the erroneous automatic delineation of the Chemung River in the 30 arc-second DEM.



5 arc-minutes

(a)

Figure 5. In (a), the shaded-relief representation of the three arc-second DEM of the Chemung River area corresponding to Figure 4c is shown with its automatically extracted drainage network highlighted in white.



5 arc-minutes

(b)

Figure 5 continued. In (b), the cells with an elevation value of 274 m are highlighted in white to illustrate lack of detail and presence of artifacts.

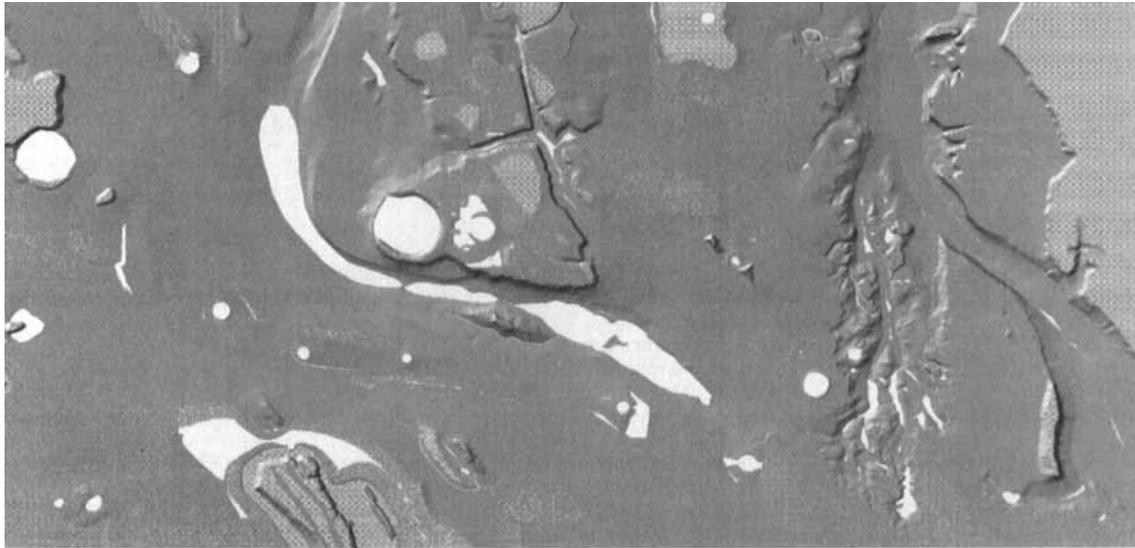
Table I. Comparison of slopes for DEMs of varying resolutions

% Slope	Conterminous United States				Elmira 1/250 000 map area			
	ETOPO5		30" DEM		30" DEM		3" DEM	
	% Mean	% Steepest	% Mean	% Steepest	% Mean	% Steepest	% Mean	% Steepest
0	59.0	58.2	31.0	30.6	5.6	16.4	18.3	21.6
1	21.3	20.2	30.0	29.4	15.8	11.6	12.5	10.1
2	8.4	7.2	11.6	9.4	20.9	10.9	9.4	12.7
3	3.7	4.0	6.0	7.5	18.1	14.0	7.6	6.6
4	1.4	2.4	3.7	3.0	14.5	7.6	6.7	5.4
5	0.5	1.2	2.7	3.4	9.8	9.7	5.6	7.0
6		0.7	2.1	1.6	6.3	5.5	5.3	4.3
7		0.4	1.7	2.4	3.9	8.2	4.8	3.8
8			1.5	1.1	2.3	3.3	4.3	3.4
9			1.2	1.1	1.4	3.3	3.8	5.1
10			1.1	1.1	0.8	3.0	3.3	2.4
11			0.9	0.9	0.4	1.7	2.8	2.1
12			0.8	0.8	0.3	1.4	2.4	1.9
13							2.0	2.9

experiment, determines the slope of each cell in a DEM by fitting a plane to the elevations of its eight surrounding cells (Sharpnack and Akin, 1969). The second method, termed the 'steepest' method in this experiment, determines the slope of each cell as the slope of a line connecting it with the cell of lowest elevation of its eight neighbouring cells. In Table I, the calculated slopes for the U.S. 30 arc-second DEM versus the portion of ETOPO5 that is coincident with the U.S. 30 arc-second DEM are shown, as well as the slope calculations for the Elmira three arc-second DEM versus the portion of the U.S. 30 arc-second DEM that is coincident with the Elmira 1:250 000 scale map. It is apparent from this table that increases in cell size produce lower slope values. These results are consistent with comparisons made by Wolock and Price, 1990. Their analysis of three arc-second DEMs versus approximately one arc-second DEMs in the Delaware River basin showed that slope differences related to cell size were of a great enough magnitude to introduce model prediction errors in applications of TOPMODEL, a topography-based hydrologic model (Wolock and Price, personal communication).

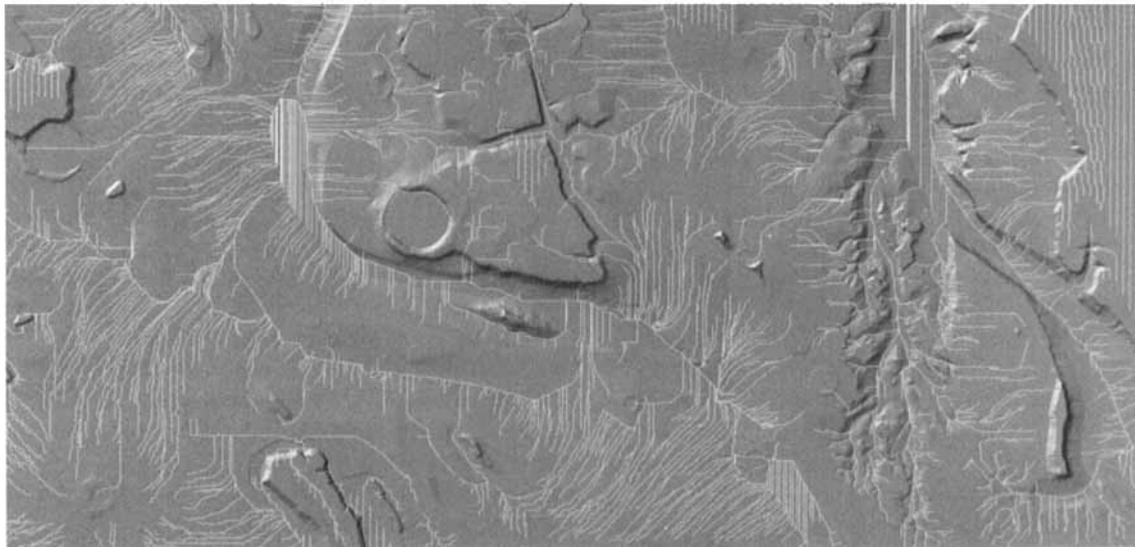
AUTOMATED ANALYSIS OF A MARTIAN DEM

Automated information extraction analysis was applied to a DEM of the Kasei Valles area of Mars to provide objective surface information. The 235 by 470 kilometre DEM, based on stereogrammetry, photogrammetry, and shadow measurements of Viking images, has a cell size of 200 (ground equivalent) metres. Robinson and Tanaka, 1990, present an analysis of the Kasei Valles as an enormous outflow-channel system on Mars that was created by a catastrophic flood. Automatically delineated depressions are shown highlighted in white against a shaded relief representation of the surface in Figure 6a. This depression analysis efficiently delineates impact craters that can be further processed with general-purpose geographic information systems to derive depth, area, and volume information. Jenson and Trautwein, 1987, presents an analogous example in pothole terrain on Earth. The systematic pattern of depressions present in some flat areas indicates subtle artifacts of the DEM generation process. For this DEM, the drainage network based on flow accumulation values greater than 500 are shown superposed on shaded relief in Figure 6b. The automatic formation of a realistic drainage network in the valley region, compared with non-fluvial features elsewhere, supports the hypothesis that the valley was created by hydrologic processes.



100 Kilometers

(a)



100 Kilometers

(b)

Figure 6. In (a), a shaded-relief representation of a 235-by-470 kilometre DEM of the Kasei Valles area of Mars is shown with automatically extracted depressions highlighted in white. In (b), the drainage lines automatically extracted from the DEM are shown highlighted in white

SUMMARY

The automatic delineation of depressions, watersheds, and drainage networks from DEMs has been shown to be a practical labour-saving device for hydrologic applications. The quality of the information that can be automatically derived, as well as the quality of slope information, is a function of both the horizontal and vertical resolution of the DEM.

REFERENCES

- Band, L.E. 1986. 'Topographic partition of watershed with digital elevation models', *Water Resources Research*, **22**,(1), 15-24.
- Bonham-Carter, G. F., Rogers, P. J., and Ellwood D. J., 1987. 'Catchment basin analysis applied to surficial geochemical data, Cobequid Highlands, Nova Scotia', *Journal of Geochemical Exploration*, **29**, 259-278.
- Collins, S. H. 1975. 'Terrain parameters directly from a digital terrain model', *Canadian Surveyor*, **29**,(5), 507-518.
- Douglas, D. H. 1986. 'Experiments to locate ridges and channels to create a new type of digital elevation model', *Cartographica*, **23**(4), 29-61.
- Jenson, S. K. and Domingue, J. O. 1988. 'Extracting topographic structure from digital elevation data for geographic information system analysis', *Photogrammetric Engineering and Remote Sensing*, **54**(11), 1593-1600.
- Jenson, S. K. and Trautwein, C. M. 1987. 'Methods and applications in surface depression analysis', *Proceedings of Auto-Carto 8*, Baltimore, Maryland, 137-144.
- Klingebiel, A. A., Horvath, E. H., Reybold, W. U., Moore, D. G., Fosnight, E. A., and Loveland, T. R. 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 pp.
- Mark, D. M. 1983. 'Automated detection of drainage networks for digital elevation models', *Proceeding of Auto-Carto 6*, **2**, Ottawa, Ontario, Canada, 288-298.
- Marks, D., Dozier, J., and Frew, J. 1984. 'Automated basin delineation from digital elevation data', *Geo-Processing*, **2**, 299-311.
- Robinson, M. S. and Tanaka, K. L. 1990. 'Magnitude of a Catastrophic flood event at Kasei Valles, Mars', *Geology*, **2**, 902-905.
- Sharpnack, D. A. and Akin, G. 1969. 'An algorithm for computing slope and aspect from elevations', *Photogrammetric Engineering and Remote Sensing*, **35**, 247-248.
- Topographic Science Working Group 1988. *Topographic science working group report to the land processes branch, earth science and applications division*, NASA Headquarters, Lunar and Planetary Institute, Houston, 64 pp.
- U.S. Geological Survey, 1987a. 'Digital Elevation Models', *U.S. Geological Survey Data User's Guide*, **5**, 38 pp.
- U.S. Geological Survey 1987b. 'Digital line graphs from 1:2 000 000-scale maps', *U.S. Geological Survey National Mapping Program, Technical Instructions, Data Users Guide*, **3**.