Data Set Derivation for GIS-Based Urban Hydrological Modeling

Michael B. Smith and Andrej Vidmar

Abstract

Procedures are described which combine commercially available programs with newly-developed algorithms to derive databases for urban hydrological modeling. While methodologies have been previously developed to automatically derive hydrologic parameters for modeling natural watersheds, less research has been devoted to urban area data definition. Common digitizing/rasterizing programs are used to develop raw data sets which are then input to algorithms to derive input parameters for a comprehensive geographic information system (GIS) based distributed parameter hydrologic model. An emphasis is placed on developing and processing an urban digital elevation model (DEM). In tests on an urban watershed, the automatically derived databases were input to a distributed parameter hydrologic model to predict the watershed response to three rainfall events. Excellent agreement was found between measured outlet hydrographs and those derived using manually developed data and data derived using the automated procedures. Automated procedures potentially offer a savings in time spent on data derivation and facilitate a more effective application of hydrologic models in urban areas.

Introduction

Engineers are routinely involved in urban hydrologic studies. Quite often, the focus of such studies is to determine the increased storm water runoff caused by human activities such as urban and suburban development. When the engineer is concerned only with information at the outlet of the watershed, then simplified hydrologic models can be used to provide the required data. One group of such models are termed *lumped parameter models*, because they use input data derived by spatially averaging hydrologic parameters over the area of interest. However, in urban areas, the engineer is quite often interested in phenomena occurring in the watershed interior. For example, the street network is of interest because frequently it carries a large volume of storm runoff. This concept is of vital importance when designing local flood control systems. One approach to defining interior processes, such as street flow, is to use *distributed parameter hydrologic models*. In contrast to lumped parameter models, these models attempt to partition the watershed into unit elements of homogeneous hydrologic parameters. Thus, a typical application of such a model involves dividing the catchment into many elements and computing their individual hydrologic response. Depending on the model, several parameters must be defined for each cell, resulting in a large quantity of required input data.

Yet, it is this very characteristic of distributed parameter hydrologic models that often renders them inefficient for everyday operational hydrology. As Sircar (1986) points out:

> Regardless of the potential of spatially distributed models in present engineering, the practical use of these approaches has been limited. Indeed, even the research directed toward the development of practical spatially distributed modeling has not attracted the attention that it currently deserves. The primary reason for this limited interest has been concern over the enormous time and cost to acquire the necessary input descriptions of the terrain surface and other variables.

In addition to large data requirements, the application of these models to urban areas is further complicated by the complexities of developed areas. Land cover is very heterogeneous, and surface flow does not always follow paths of steepest terrain descent. Subtle differences in surface relief can be responsible for significant changes in drainage patterns. Moreover, storm sewer systems may link sub-watersheds that may not normally be related by surface topography. Previously, these complexities could only be converted to meaningful input data for hydrologic models using tedious manual methods.

Consequently, the problem facing urban hydrology is not the sophistication of the hydrologic models, but rather the often daunting process of deriving the necessary input data in a timely and effective manner. In order for distributed parameter hydrologic models to be applied practically in an operational manner, automated or semi-automated data collection and processing techniques must be developed.

In response to these considerations, the purpose of the present study is to develop and evaluate a basic strategy for deriving land-use and topographic data for GIS-based urban hydrologic modeling. In this strategy, commercial digitizing/rasterizing and digital terrain modeling (DTM) software are combined with algorithms developed by the authors to derive data sets for a distributed parameter hydrologic model. An emphasis is placed upon deriving and processing an urban digital elevation model. As manual digitizing is likely to remain the prevalent data entry technique in the early 1990s (Carstensen and Cambell, 1991), the proposed strategy hinges on the use of a commercial digitizing package for initial data definition.

Background

Existing Hydrologic Models Linked to GIS and CAD

Huber *et al.* (1991) developed a linkage between a well known urban hydrologic model and both the ARC/INFO GIS and the AutoCAD drafting package. Similarly, the city of Long

Hydraulics Department, Faculty of Architecture, Civil Engineering, and Survey, University of Ljubljana, Hajdrihova Ulica 28, 61000 Ljubljana, Slovenia.
M.B. Smith's present address is 6838 Ducketts Lane, Elkridge, MD 21227.

Photogrammetric Engineering & Remote Sensing, Vol. 60, No. 1, January 1994, pp. 67–76.

0099-1112/94/6001–67$03.00/0
©1994 American Society for Photogrammetry and Remote Sensing

Beach, California has recently installed a GIS-based stormwater drainage planning system using in-house developed hydrologic models (Urenda, 1992). Delaplace and Price (1991) discuss current trends in linking existing urban hydrologic models to computer-aided drafting (CAD) systems and GIS in the United Kingdom.

Terrain Information

Almost all hydrologic models require an estimate of the surface slope. Digital terrain modeling can provide the engineer with tools to derive terrain parameters. Collins (1981) developed a series of algorithms to analyze a dense DEM to locate hydrologically important features such as street intersections, street-stream crossings, and other combinations of linear features. Thorpe (1988) developed a strategy for computing contours in urban areas considering the presence of break lines. Djokic (1991) and also Djokic and Maidment (1991) outlined an approach for storm sewer system evaluation using a system based on Triangular Irregular Network (TIN). Huber *et al.* (1991) were able to automatically compute certain terrain parameters using ARC/INFO and a TIN structure.

Land-Cover Information

For urban hydrological modeling, impervious and pervious areas must be identified because each has an entirely different response to rainfall. Meier and Lakatos (1987) developed an innovative digitizing procedure for deriving data coverages for urban stormwater modeling. Satellite imagery can be used to update street network files when used in combination with computer aided design and drafting (CADD) packages (Baur, 1991). Terstriep and Lee (1989) discuss an ongoing effort using automated scanner technology and low altitude infrared photographs to determine house size and density, and impervious area location for urban hydrologic analysis.

Rationale and Choice of Data Structure

An important consideration in coupling hydrological models with GIS is the choice of data structure. Currently, there are three basic DTM-hydrologic model combinations. Some examples of each are provided in the following list:

- Triangulated Irregular Network (TIN) (Goodrich *et al.*, 1991)
- Stream Path/Contour (Moore and Grayson, 1991)
- Grid/Raster (Johnson, 1989)

Moore *et al.* (1991) and Goodrich *et al.* (1991) provide further discussion on the advantages and applications of each type of structure to natural areas. However, it is not readily clear which type would be most appropriate for urban applications. Djokic (1991) stated:

> It could be argued that, with enough resolution, any terrain modeling approach can be used to represent urban terrain, but problems arise when implementing any of the digital terrain models (raster/grid, TIN, or contours) used for rural watersheds because the level of detail required for automated urban terrain analysis is enormous and makes a fully automated approach very difficult.

Grid DEMs correspond to the storage structure of computers, leading to simpler handling as a two-dimensional array of elevations. Consequently, grid-based terrain modeling algorithms tend to be relatively straightforward (Weibel and Heller, 1991; Goodrich *et al.*, 1991). Recognizing the relative simplicity of developing and analyzing a grid DEM and gridded data structures, it was decided to use a square cell structure for the data layers. Concurrent development of a grid-based hydrologic model also necessitated the choice of a grid data structure. Thus, urban terrain, flow directions, and land cover will be represented as a matrix of grid cells.

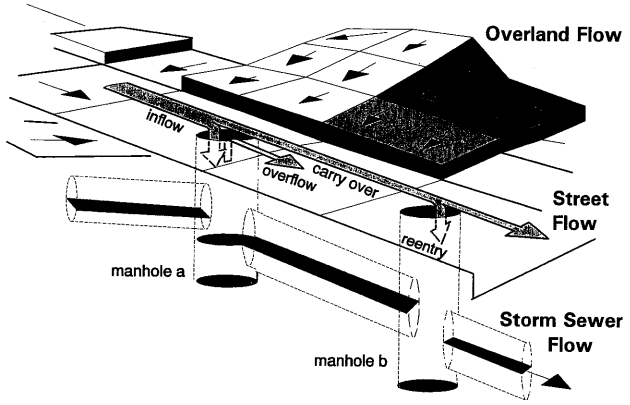


Figure 1. Simplified surface and storm sewer system used by the hydrologic model.

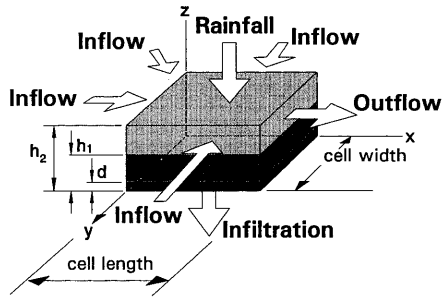


Figure 2. Computational element dimensions and hydrologic processes.

Hydrologic Model Overview

In order to evaluate the proposed procedures, a distributed parameter hydrologic model was linked to the derived data layers to perform urban runoff analysis. Details of the distributed parameter hydrologic model can be found in Smith (1992, 1993). In general, the model operates on an urban watershed that has been partitioned into unit elements as shown in Figure 1. The model computes three major types of flow in an urban area: overland flow, street flow, and storm sewer flow. Given the significant storm water carrying capacity of streets, an emphasis is placed on delineating the street network. In the rasterized data set, street cells are tagged with an attribute value of one, while all other cells are assigned a zero value. Outflow from street cells is computed using triangular gutter flow equations. Flow from the street centerline to the gutter is not computed. Sheet flow is assumed for overland flow cells. Cells containing inlets to the storm sewer system are allowed to lose or gain storm water depending on whether the sewer system is flowing full. Storm sewer overflows are allowed to travel down hill in response to local topography. As seen in Figure 1, any type of terrain analysis for urban drainage is complicated by the

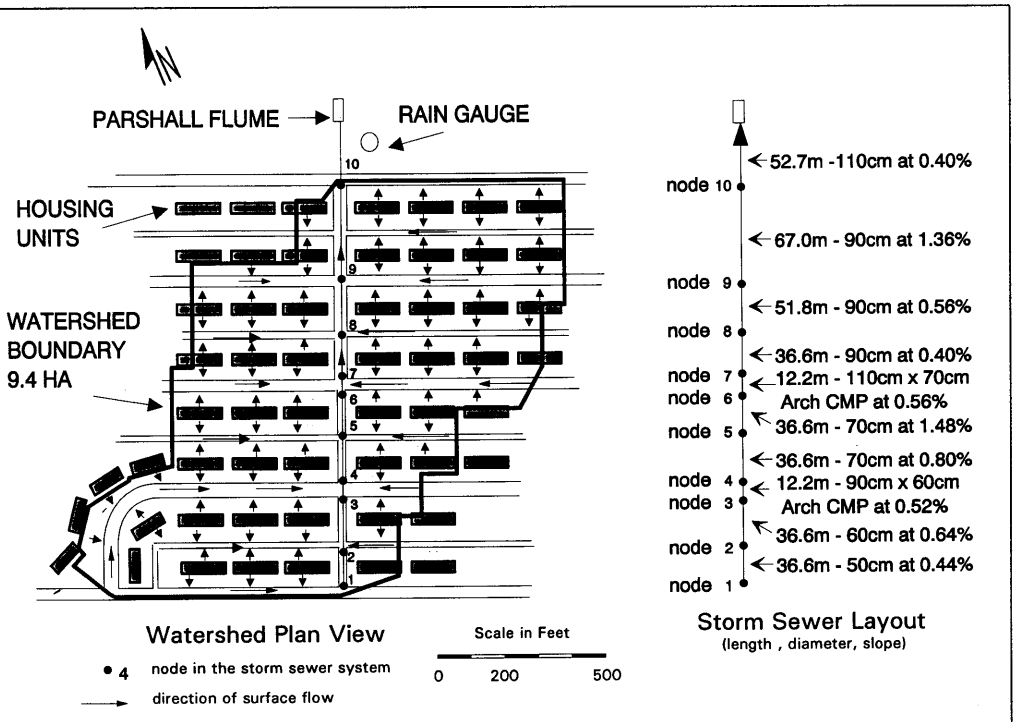


Figure 3. Layout of the Gray Haven watershed near Baltimore, Maryland.

storm sewer network, which conducts runoff to other parts of the basin.

The processes of rainfall addition, infiltration, inflow from neighboring cells, and cell outflow are computed for each computational element as shown in Figure 2. In a typical application of the distributed parameter model, a rainfall event is divided into time steps, and the hydrologic response of each element within the entire watershed is computed before advancing to the next time step. At the beginning of each time step, a unique flow depth h_1 exists in each cell, while h_2 represents the flow depth at the end of each time step considering the various inflows and outflow. The parameter d represents a surface dependent storage depth. Different values of d are assigned to cells with impervious land cover and cells with infiltrating land cover. Cell outflow may occur in one of eight possible directions away from the cell.

In the current stage of development, the hydrologic model does not consider the impact of buildings on the flow regime. Rather, rain falling cells coded as buildings or roofs are simply transferred to adjacent cells according to the general topography of the immediate area. Outflow from cells coded as roofs is computed as overland flow.

Thus, the model requires raster representations of impervious and pervious land cover, streets, and topography. Infiltration is soil dependent, so a rasterized soils data layer is required as well when the soil type varies. While this type of modeling is very data intensive, it allows the analyst to monitor distinct hydrologic processes occurring in the watershed interior.

Approach

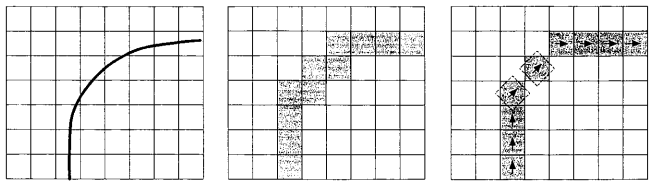
Introduction

Testing of the developed strategies was performed by using data from the Gray Haven urban drainage basin (9.4 ha) located near Baltimore, Maryland and shown in Figures 3. Gray Haven was monitored as part of a research effort by the American Society of Civil Engineers Water Resources Research Program (Tucker, 1969). Detailed rainfall-runoff data are available and have been used to evaluate other urban hydrologic models (Wenzel and Voorhees, 1980; Marsalek *et al.*, 1975). As shown in Figure 3, the watershed has a simple layout and contains a single storm sewer line in the center of the basin which outlets to a Parshall flume. Details of the storm sewer pipes required by the hydrological model are shown on the right side of Figure 3. Nodes 1 through 10 represent manholes in the storm sewer network where pipes of two different sizes or slopes are joined. In addition, the nodes represent points where surface flow enters the sewer network through a series of street gutter inlets. Storm sewer plans, street profiles, and aerial photos were available from the Baltimore County Department of Public Works. In addition, two on-site visits were made to note flow directions and other hydrologic characteristics.

Street Networks

A rasterized street network is required by the hydrologic model in order to compute the proper cell outflow shown in Figure 2. In the course of hydrologic model execution, the model recalls each cell in the watershed by referencing its coordinates. If the current cell is a street element, outflow is computed using flow equations for triangular gutters. Otherwise, the model computes outflow as simple overland flow.

Due to the small size of the study area, a 1:600-scale base map was easily prepared which showed the locations of the street and alley network and housing units. Given the basically orthogonal nature of the street system, the coordinate system was rotated to allow the majority of the streets to be aligned with the x and y axis of the digitizing tablet. This step is not crucial to the analysis but resulted in a more eye



(a) (b) (c)

Figure 4. Vector to raster conversion of curved street center line. (a) Digitized street centerline overlaid with grid. (b) Rasterized centerline. (c) Manually thinned centerline. Dashed lines indicated cell orientation assumed by the hydrologic model. Arrows indicate direction of surface flow along street.

pleasing grid cell representation and also facilitated easier checking and registration with other data bases. A commercially available digitizing program (ROOTS, 1991) was used to digitize the centerline of all streets as line segments. All digitized street centerlines were tagged with an attribute value of one, while all non-street areas were assigned a value of zero. In a separate data layer, the outlines of all housing units were digitized as polygons.

An on-screen grid preview option within the digitizing program facilitated the choice of effective grid cell size. Using the ROOTS software, the street centerlines and the housing unit layer were displayed together on the same screen. Given the significant role of streets as storm water conveyances, a grid size was chosen which would provide a detailed street network description. A cell size of 35 feet (10.67 metres) effectively fit the width of the 36-foot wide streets and allowed for a one-cell width linear street map to be produced. In addition, this cell size also seemed to provide for a best fit considering the overall coverage of the housing units, unpaved areas, and the alignment of the street system. For example, the average street width in the Gray Haven development is 36 ft, the average housing unit width is 31 ft, and the distance between the housing units and the street is 37 ft.

Using the on-screen grid preview option of ROOTS, the entire 35 ft grid network was shifted in relation to the street network to provide the best street delineation while simultaneously covering the housing units and impervious areas with a minimum of rasterizing error. Within the ROOTS software, the polygon rasterizing algorithm assigns attribute values using a center point approach. In other words, the land-cover attribute of the polygon existing at the center of each grid cell is assigned to the cell in the rasterized data layer. Panning through the digitized street network overlaid with both the preview grid and housing units layer facilitated the best placement of the grid, considering the center point rasterizing process of the housing units and pervious areas (polygons) and the line segment rasterizing function of the street network. With the user-defined study boundary and the rasterizing function of ROOTS, the road network was rasterized into a matrix of 45 rows and 48 columns. This matrix size was subsequently chosen as the standard grid data layer format.

Vector to raster conversion of linear features typically produces maps in which the gridded feature is more than one element wide. This occurred in the Gray Haven study when the curved portion of one road centerline was rasterized and is similar to the situation in Figure 4, which illustrates the process of line segment rasterization and thinning. In the ROOTS software, the line segment rasterizing algorithm takes each grid cell in contact with the street centerline in Figure 4a and labels it a street cell as shown in Figure 4b. Given the small size of the rasterized data layer, it was a simple matter to import the matrix in ASCII form into a word processor and manually thin the line. Manual thinning was accomplished after visually examining the digitized street centerline overlaid by the housing units layer and the 35 ft grid. Automated thinning procedures similar to those developed by Greenlee (1983) and Peuquet (1981a, 1981b) to produce a line skeleton (a single connected line) should be incorporated if more complex street networks exist in the study area.

Thinning the line segment of Figure 4b to form the line skeleton of Figure 4c does not have serious hydrologic implications due to simplifying assumptions within the hydrologic model. When computing overland or street flow, the hydrologic model always uses the cell width shown in Figure 2, even when the outflow is in one of the four diagonal directions. By using the cell width, the hydrologic model essentially aligns the grid cell with the direction of cell outflow and assumes the cell orientation shown by the dashed cell outlines and flow direction arrows in Figure 4c. Consequently, the hydrologic nature of flow in the street is preserved by using a line skeleton. Outflow from cells removed in the thinning process is then computed using overland flow equations. Experience with the hydrologic model has shown that differences in the runoff response resulting from one or two cells being classified as overland versus street cells would be insignificant.

Surface Permeability

A pervious/impervious area data layer was derived in a manner similar to the rasterized street network. Using 1:600-scale land-cover maps and aerial photography, apartment block outlines and parking areas were delineated on the 1:600-scale street map previously described and treated as impervious areas. All other surface cover was classified as pervious. Subsequent digitization and rasterization produced a 45 row by 48 column matrix of ones and zeros reflecting impervious and pervious areas, respectively.

Storm Sewer System

As input, the hydrologic model developed by Smith (1992, 1993) requires the cell coordinates of the major junctions of the storm sewer system as well as the various characteristics of the major conduits presented in Figure 3. Given the simplicity and small size of the Gray Haven system, it was a simple exercise to manually specify the coordinate pairs and create the required input file.

Digital Terrain Model

The most difficult and time consuming segment of the present study was the derivation and processing of a grid DEM of the Gray Haven watershed. Two-foot contour information reflecting the post construction terrain conditions was available for only half the watershed. For the remaining portion, only pre-construction contours were available. The original street profiles were obtained from the Baltimore County Department of Public Works. Using these data, point elevations were scaled off every 25 feet on all road centerlines. Alleys were assumed to have the same slopes as the adjacent streets. The same elevations derived for the road and alley centerlines were assigned to points at the road and alley edges perpendicular to the centerline. Ground elevations for points other than streets and alleys were estimated from photographs and the on-site examinations of the watershed. In general, the watershed gently slopes to roads and alley seg-

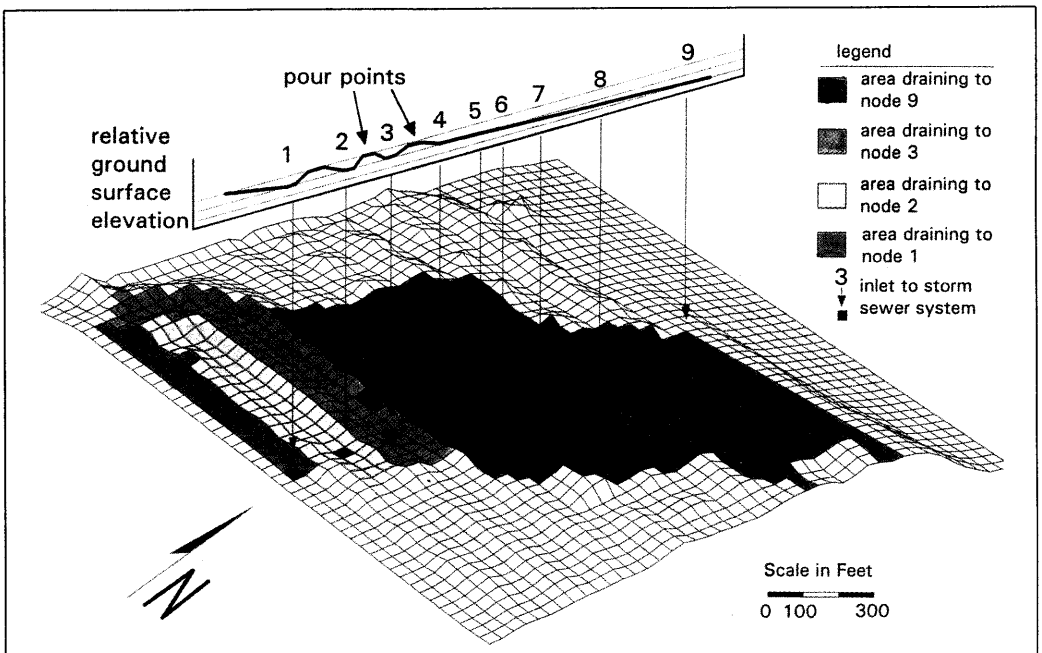


Plate 1. Grid DEM for the Gray Haven study area showing the location of sub-watersheds and storm sewer nodes.

ments. Sharp slope breaks do occur on a few portions of the terrain as one moves from a housing unit to the street.

All elevation points were subsequently digitized and an *x, y, z* ASCII coordinate file was created consisting of 1474 points. Using these data, a commercially available program (QuickSurf, 1991) was utilized to derive the initial 35-foot grid DEM shown in Plate 1. As it is useful with sharp breaking surfaces such as man-made objects, the zero derivative option for the standard grid method within QuickSurf was used. As output from the program, each corner of each quadrilateral in Plate 1 is assigned an elevation. Software written by the authors was used to derive an average elevation for the center of each cell, and resulted in a somewhat smoother surface. Overlaying the grid cell center elevations with the original digitized ground elevations and road centerlines in the ROOTS program revealed that at two alley locations pits were inadvertently created. Most likely, this was the result of sparse point elevation data in the vicinity. Manual correction of the ground elevation at these two locations ensured that all streets and alleys sloped toward the sewer system as shown in Figure 3.

Plate 1 also shows that the total area of the watershed is comprised of four distinct sub-watersheds. Most urban hydrologic models require that only major inlets to the storm sewer system be specified for each sub-watershed. As a result, the storm sewer network in Figure 3 was simplified by combining nodes 3 and 4, resulting in a total number of nine storm sewer nodes as shown in Plate 1. The graph of relative ground surface elevation along the storm sewer line indicates that nodes 2 and 3 are located at sub-watershed low points. Storm runoff not entering the sewer system at nodes 2 and 3 would remain in the vicinity of the nodes. Runoff not entering nodes 4 through 8 in Plate 1 would tend to flow down the center street to node 9 as shown by the graph of relative ground surface elevation. Consequently, nodes 4 through 9 were retained as they would affect the flow in the center street. Runoff not captured by the storm sewer inlet at node 1 would continue down the street in an easterly direction and out of the sub-watershed.

Flow Direction Determination

For the hydrologic model, an estimate is required of the surface slope of each cell. As with other grid-based DEM processing algorithms, a steepest-descent analysis of the eight surrounding neighbor cells was used to determine a flow direction for each cell. The direction coding procedure developed by Greenlee (1987) is used to numerically describe the eight flow directions. However, a modified steepest-descent algorithm was developed which constrains the direction search in the vicinity of street cells. One of the problems in urban areas is that surface flow does not always follow the line of steepest terrain decent as predicted by algorithms for natural watershed analysis. Small elevation differences created by street gutters can significantly alter the direction of surface flow and conduct runoff to other parts of the watershed.

To handle such situations, the algorithm developed by the authors performs a constrained steepest-descent search. As input, the direction algorithm requires both the rasterized street network and the grid DEM shown in Plate 1. During analysis, if the current cell is a street cell, then the algorithm will only consider neighboring street cells as candidates for the next downstream cell. Thus, a constrained direction data set with 45 rows and 48 columns was derived. Surface slopes corresponding to the cell flow directions were written to a file for input into the hydrologic model.

Cell Inflow Counting

An intermediate data layer required for further DEM processing consists of a neighbor inflow count for each cell. This algorithm is described in Smith and Brilly (1992). Briefly stated, a three- by three-cell window is applied to the direction data set. For the current cell in the middle of the window, a count is made of all cells flowing into the current

TABLE 1. COMPUTED VERSUS PUBLISHED SUB-WATERSHED DRAINAGE AREAS FOR GRAY HAVEN.

Sub-Basin Draining to Node	Automatically Computed Drainage Area (Acres)	Published Drainage Area from Tucker (1969) (Acres)
1	1.12	0.99
2	2.22	2.97
3	2.95	2.76
9	16.70	16.57
total	22.99	23.29

cell. In this way, an inflow neighbor data layer is constructed in which each cell has an integer value denoting the number of inflowing adjacent elements. This procedure also locates start cells, or elements that have no inflowing neighbors. These cells are significant in that they form the beginning of overland flow paths. Cells having an inflow neighbor value of two or more are labeled as junction cells. Output from this step consists of a global set of coordinate pairs of start cells and junction cells.

Watershed Delineation

In a much referenced work, Jenson and Domingue (1988) discuss methodologies for watershed delineation and pour point identification. One watershed links to another at a pour point, or the lowest elevation on the common boundary between two watersheds. For example, surface flow ponding at node 2 in Plate 1 would eventually reach the elevation equal to the pour point elevation between nodes 2 and 3 and flow into the third basin. Currently, natural area DEM processing techniques fill these depressions when deriving down slope flow paths so that each cell has a link to the watershed outlet.

As seen in Plate 1, the DEM for Gray Haven actually consists of four distinct basins when a system similar to that shown in Figure 1 is assumed. Pour points linking the basins are in a line centered above the storm sewer line. However, in urban hydrologic modeling, it is reasonable that the more significant linkage between sub-watersheds is through the storm sewer system, which is much more efficient as a hydraulic conveyance than overland flow paths. Filling operations would result in sub-watershed 1 being neglected, as its linkage to the other basins is through a sewer conduit, not a pour point. On the other hand, without filling procedures, a watershed growing procedure applied to node 9 in Plate 1 would result in a reduced total watershed, effectively ignoring almost a third of the actual drainage area.

In light of these complexities, a watershed delineation procedure similar to that outlined by Jenson and Dominque (1988) and Smith and Brilly (1992) was developed which considers multiple starting locations. Beginning at a defined node, the algorithm recursively examines the direction and inflow neighbor data sets and locates all cells that have an overland flow path to the node. Subsequent searches are then initiated at other user-defined nodes. For example, the modified algorithm was applied to the Gray Haven area starting at node 9 in Plate 1. Subsequent sub-watershed growing procedures were initiated at nodes 3, 2, and 1, respectively, to define the entire watershed. These cells were used because they represent low points in the sub-watershed and inlets into the storm sewer system. All cells located by the multiple search algorithm are flagged as watershed cells. All other cells are given attributed values of zero to denote them as non-watershed elements. At this point in development, the watershed start cells were manually determined and input into the watershed growing algorithm.

While Huber *et al.* (1992), Urenda (1992), and Djokic and Maidment (1991) relied on digitized drainage area boundaries within a GIS, the proposed method allows the boundaries to be computed automatically.

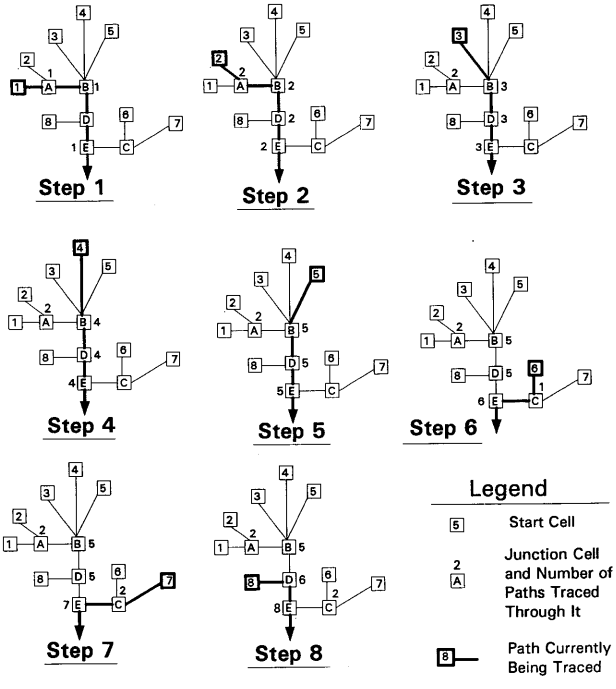
Tucker (1969) manually partitioned the Gray Haven watershed into 17 sub-basins, each leading to a curb inlet to the storm sewer system. In Tucker's very detailed watershed segmentation, the road crown was also used as a sub-watershed boundary. Assuming the idealized surface in Figure 1, this degree of segmentation was not possible, nor is it required for most urban drainage models such as the Storm Water Management Model (Huber and Dickinson, 1988). Such urban drainage models reduce the storm sewer system to a series of lateral conduits and vertical manholes. Combining the component sub-watersheds determined by Tucker (1969) to derive areas similar to those in Plate 1 leads to the comparisons in Table 1. It can be seen that the automated sub-basin growing procedure produces sub-watershed areas that correspond to published data. Tucker's value of 0.99 acres for the first sub-watershed only included half a street width, while the automatically computed area is greater due to the inclusion of the entire street width. The automated algorithm underpredicted the second drainage area while overpredicting the third. This was due to the algorithm wrongly assigning the roof areas to the third sub-watershed rather than to the second. This type of error points to the need to develop more refined procedures to deal with buildings and other obstructions to flow.

Ordering of Computational Elements

A crucial input description in distributed parameter hydrologic modeling is the sequence in which the outflow of each representative element is computed. Before any one cell can be examined, the outflow from all upstream cells eventually leading to the current cell must be computed, requiring a comprehensive cell-to-cell connectivity sequence. Gandoy-Berasconi and Palacios-Velez (1990) developed an approach to order TIN facets within a hydrologic cascade. Smith and Brilly (1992) present a method for grid DEMs. However, their grid-based algorithm is restrictive in that it can only operate on depressionless DEMs, or those watersheds whose depressions have been filled to provide each cell with a flow path to the main watershed outlet. This algorithm would not work using the Gray Haven DEM developed here containing four distinct sub-watersheds. Therefore, as with the watershed delineation algorithm, a significant modification to the grid-based ordering algorithm was performed in order to consider sub-watersheds linked by subsurface drainage elements.

A preliminary step is required in which the nodes representing the sub-watershed low points (which should be storm sewer system inlets) are manually ordered, starting with the most hydrologically distant and ending with the node nearest the main watershed outlet. For Gray Haven, this corresponds to the order of 1, 2, 3, and 9 for the nodes in Plate 1. It should be noted that it is not necessary to specify the remaining nodes in Plate 1 because the local topography permits an overland flow path to either another node or the main sub-basin outlet node 9.

In the modified ordering algorithm, the first step consists of sorting the global lists of start and junction cells (developed in the inflow counting procedure) according to sub-watershed node. Initial sorting is performed by tracing overland flow directions from each start cell until a user-defined sub-watershed node is encountered. Each start cell is assigned to the first node located during a tracing. Thus, start cells are grouped by sub-watershed node. For Gray Haven, four group-



Before Ranking

Node	No. Paths Traced
A	2
B	5
D	6
E	8
C	2

After Ranking

Calling Order in Hydrologic Model	Rank	Path
1	2	A - B
2	2	C - E
3	5	B - D
4	6	D - E
5	8	E - outlet

Figure 5. Illustration of the procedure for ordering computational elements as applied to a sub-watershed of eight start cells and five junction cells.

ings were derived, corresponding to the four major sub-watersheds and nodes 1, 2, 3, and 9 shown in Plate 1. A second series of tracings is subsequently performed, only now according to the start cell groupings derived during the previous tracing. This second tracing is required to sort junction cells by sub-watershed.

Actual ranking of each cell in each distinct sub-watershed is illustrated using Figure 5, which represents a sub-watershed consisting of eight start cells and five junction cells. Beginning with the start cells in each ranked sub-watershed, surface flow paths are traced until the sub-basin node is reached as shown in steps 1 through 8 in Figure 5. Each time a junction cell is reached, a counter is incremented indicating the number of paths that have been traced through it. This process is very similar to the concept of optimal path density as discussed by Berry (1987) and Tomlin (1990). An optimal path is a route between two points that maximizes or minimizes a certain function such as travel time. A map of optimal path density identifies the number of individual optimal paths from a set of dispersed termini to a designated outlet. For example, in Step 2 of Figure 5, the counter for junction cells A, B, D, and E is equal to two, as two paths have been traced to the outlet. In Step 3, the counter for junction cells B, D, and E is incremented to a value of three, corresponding to a third tracing, while the counter for cell A retains a value of 2.

In order to establish the proper hydrological computation sequence, junction cells are then ranked according to path density, (i.e., the number of traced paths), as shown in the bottom of Figure 5. Those junction cells having higher values of the counter are ranked lower in the computational sequence. Linked to each junction cell is a sequence of cells. Thus, cell paths are ranked, not individual elements. By definition, start cells have no upstream runoff contribution and thus require no ranking within a sub-watershed.

The hydrologic model calling order in Figure 5 shows that no surface flow path is analyzed until after the runoff from all of its uphill constituent flow paths has been computed. First, the runoff from start cells one through eight is computed. The calling order then specifies that the runoff from cells in path A-B be computed, followed by the cells in path C-E, path B-D, and ending with path D-E. In the course of model execution, cell coordinates are recalled in computational order and hydrologic parameters are withdrawn from the rasterized street, slope, and impervious area data layers.

Results and Discussion

In order to evaluate the utility of the procedures, the hydrologic model was applied to the Gray Haven watershed using the automatically computed watershed boundaries, cell slopes, and computational sequences. As the soils were predominantly a sandy loam, uniform infiltration parameters were assigned to the entire watershed. Using these data, outflow hydrographs (plots of runoff discharge versus time) were computed at the outlet of the storm sewer system at the Parshall flume. These hydrographs represent the total watershed response to a rainfall event considering overland flow, street flow, and storm sewer flow. These computed hydrographs were plotted against actual discharges measured at the site. Rainfall-runoff data for storms occurring on 10 June 1963, 20 June 1963, and 14 August 1963 were taken from Tucker (1969) and used for comparisons. Figures 6, 7, and 8 present the computed and actual discharge hydrographs.

In addition, results from earlier studies (Smith, 1992; Smith, 1993) were available in which an earlier version of the hydrologic model was applied using Gray Haven data sets that were derived using traditional manual techniques. In these studies, the flow directions and slopes were assigned and computed by hand using mylar overlays and contour maps. In addition, the discretization of the Gray Haven watershed in these studies was somewhat different, using a slightly larger cell size (40 ft versus 35 ft). The larger cell

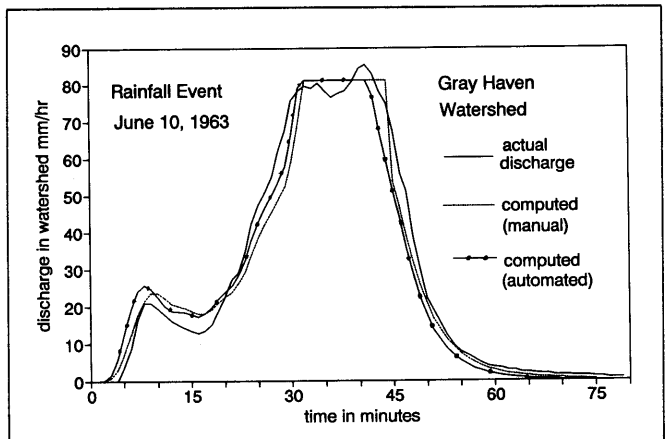


Figure 6. Computed versus actual discharge hydrographs for rainfall event of 10 June 1963.

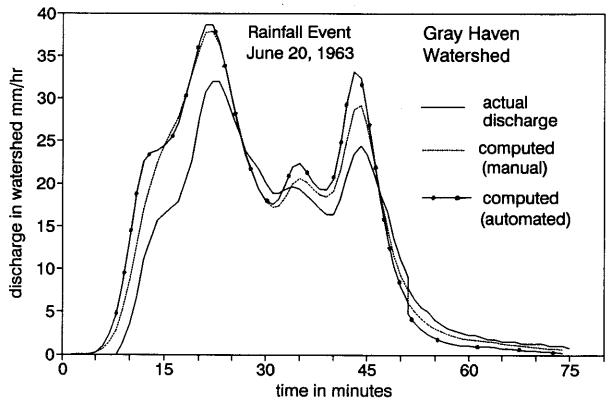


Figure 7. Computed versus actual discharge hydrographs for rainfall event of 20 June 1963.

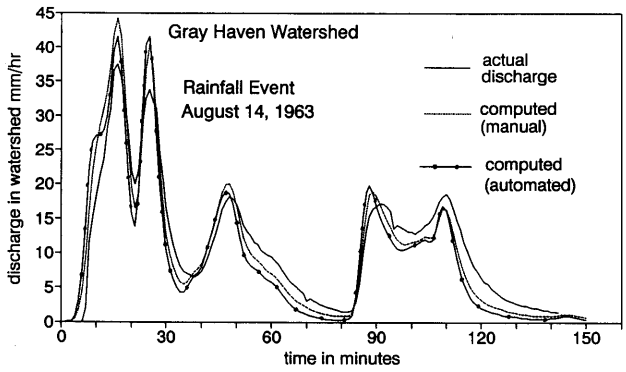


Figure 8. Computed versus actual discharge hydrographs for rainfall event of 14 August 1963.

TABLE 2. ACTUAL VERSUS COMPUTED TIMES FOR MAJOR HYDROGRAPH PEAKS

		Peak Time in Minutes		
			Computed Hydrograph	
Event	Major Peak	Actual Hydrograph	Manual Method	Automated Method
10 Jun 63	2	41	–	–
20 Jun 63	1	21	22	22
	2	44	44	43
14 Aug 63	1	16	16	16
	2	24	25	25
	3	47	49	48
	4	90	90	88
	5	109	109	109

size was chosen on the basis of less accurate maps available at the time of the earlier studies. An earlier version of the hydrologic model was also used in these studies in which a slightly different approach to computing infiltration was used. While the discretization of the watershed and hydrologic model is slightly different, these earlier studies nonetheless represent the application of a distributed parameter hydrologic model using data derived by traditional manual methods. Combined with the results generated in the present study, these earlier studies facilitate a general comparison between two approaches to data set derivation. Figures 6, 7, and 8 also present the outlet hydrographs computed using manually derived watershed parameters.

Green and Stephenson (1980) state that visual comparison of simulated and measured hydrographs provides a quick and often comprehensive means of assessing the accuracy of hydrologic model output. As seen in Figures 6, 7, and 8, both the manually and automatically derived data can be used to accurately reproduce the shape of the actual discharge hydrographs. In all three cases, the slopes of the manual and automated hydrographs correspond to the slopes of the actual hydrographs. The flat-top hydrographs in Figure 6 result from a limitation in the hydrologic model. A significant similarity is that both data versions have a matching tendency to overpredict and underpredict the actual measurements. For example, both approaches overpredict the first four peaks in Figure 8 and then underpredict the fifth peak. This indicates that any differences between the hydrographs are more likely due to the slightly different versions of the hydrologic model rather than to the database derivation methods.

It can also be seen in Figures 6, 7, and 8 that the automated hydrographs initially tend to lead both the manually derived and actual discharge hydrographs. Also, the automated procedure generally produces lower minimum values between the peaks compared to the manual methodology. These trends are most likely the result of using a smaller coefficient describing surface roughness in the overland flow equations and the slightly different infiltration approach described earlier.

Excellent correspondence exists among the peak times in all events. Actual and computed peak times are listed in Table 2 and show that there is no significant difference between the hydrographs derived using manual and automated data derivations methods. On occasion, the automated methodology leads to slightly earlier peaks, as on peaks 3 and 4 of the 14 August event. However, these differences are not hydrologically significant and are probably the result of the lower value of the surface roughness coefficient.

Marselek *et al.* (1975) suggested using the integral square error (ISE) as a further measure of goodness of fit between actual and computed discharge hydrographs. The ISE is computed using Equation 1: that is,

$$\text{ISE} = \frac{\left[\sum_{i=1}^{N} (O_i - C_i)^2\right]^{1/2}}{\sum_{i=1}^{N} (O_i)} \times 100 \qquad (1)$$

where O_i is the actual or observed hydrograph value at time i, C_i is the computed hydrograph value at time i, and N is the number of values. Smaller values of ISE suggest better agreement between the actual and computed values of a variable. The authors also suggest the ratings in Table 3 to aid in the evaluation of hydrologic models or methodologies.

For each of the three rainfall events, two ISE indices were computed. The first represents the agreement between the actual hydrograph and that computed using manually derived data. The second ISE index represents the agreement between the actual hydrograph and that computed using the automated procedures. Table 4 presents the ISE values for each storm and methodology.

Using the ISE as an index of goodness-of-fit, and the ratings in Table 3, it can be seen that both data derivation methods generate either excellent or very good results. More-

TABLE 3. RATINGS OF THE INTEGRAL SQUARE ERROR (ISE) (AFTER MARSELEK ET AL., 1975)

Integral Square Error (ISE) in Percent	Rating
0.0 < ISE ≤ 3.0	Excellent
3.0 < ISE ≤ 6.0	Very Good
6.0 < ISE ≤ 10.0	Good
10.0 < ISE ≤ 25.0	Fair
25.0 < ISE	Poor

TABLE 4. VALUES OF THE INTEGRAL SQUARE ERROR FOR HYDROGRAPHS COMPUTED USING THE MANUALLY AND AUTOMATICALLY DERIVED DATA SETS.

	Integral Square Error (ISE)	
Rainfall Event	Manual Vs. Actual	Automatic Vs. Actual
10 Jun 63	2.10	2.15
20 Jun 63	2.11	2.86
14 Aug 63	3.57	4.70

TABLE 5. ACTUAL VERSUS COMPUTED RUNOFF VOLUMES FOR HYDROGRAPHS COMPUTED USING MANUALLY AND AUTOMATICALLY DERIVED DATA SETS

	Runoff Volume in Cubic Meters			Ratio	Ratio
	Actual	Computed			
Event		Manual	Automated	Manual/ Actual	Automated/ Actual
1	2	3	4	5	6
10 Jun 63	3,545	3,372	3,476	0.95	0.98
20 Jun 63	1,412	1,585	1,604	1.12	1.14
14 Aug 63	2,456	2,413	2,353	0.98	0.96

over, for each rainfall event, the manual and automated methodologies produce ISE indices that fall within the same rating class. Both methods produce excellent results for the events of 10 June 1963 and 20 June 1963, while very good results were achieved for the 14 August 1963 event. However, there is a tendency for the automated methodology to produce slightly higher ISE values compared to the manual method for each rainfall event. This trend is probably due to the higher peaks and lower valleys caused by the more recent version of the hydrologic model as previously discussed.

Table 5 indicates that the automated procedures generate nearly the same results as the manual method when considering total runoff volume (area under the hydrograph). Perhaps the more significant tendency is that both methods exhibit matching trends when predicting volumes. For example, the ratios in columns 5 and 6 of Table 5 show that both methods concurrently overpredict and underpredict the actual runoff volumes.

Summary and Conclusions

Semi-automated procedures operating on digitized data sets facilitate the use of distributed parameter hydrologic models in urban areas. While digitizing is a tedious procedure, it can be viewed as a basic methodology for data input at the operational level until other strategies are refined. Application of the proposed procedures and algorithms revealed that some manual interaction is required, as evidenced by the thinning needed to derive the rasterized street network and the manual sub-watershed ordering. In addition, much manual interaction was required to construct the DEM for Gray Haven due to incomplete coverage of the present condition contour maps. Given an accurate flow direction data set, accurate watershed delineation and cell ordering can be achieved.

The data sets derived using the automated procedures were combined with a hydrologic model to simulate runoff rates for three rainfall events. Predicted results were compared to two other sets of results: actual measured runoff rates and those generated in earlier studies using traditional manual methods to develop databases. While slight differences in the hydrological model and in the watershed discretization used in the earlier studies does not permit an explicit statistical comparison of the manual and automated methods, the results nonetheless can be used to indicate the utility of the automated procedures.

Analysis of the runoff hydrographs computed using the two approaches for data derivation indicates that the automated methods produce results that have excellent agreement with results generated using manually derived data. Excellent agreement was also achieved between actual runoff measurements and hydrographs computed using the automatically derived data sets. Thus, compared to traditional manual methods, automated procedures potentially offer a significant reduction in the time spent on database derivation while providing acceptable results for hydrologic modeling.

Further research needs to address the analysis of buildings and other obstructions to flow in urban areas. Furthermore, it was noticed that during the DEM processing runoff can enter or leave the drainage area depending on local topography and storm sewer inlet capture. Perhaps a wider regional approach to urban hydrology might be developed which simultaneously considers adjacent sewer-drained watersheds.

Acknowledgments

Partial support for this study was provided through the government of the Republic of Slovenia. The authors wish to thank Dr. Mitja Brilly for arranging the required facilitates. In addition, the comments of three anonymous reviewers greatly enhanced the focus and clarity of this article.

References

Baur, M.F., 1991. The View From Above - Combining Satellite Imagery with AutoCAD, *Cadence*, October, pp. 56–59.

Berry, J.K., 1987. Fundamental Operations in Computer Assisted Map Analysis, *Int. Journal of Geographical Information Systems*, 1(2):119–136.

Carstensen, L.W., and J.B. Cambell, 1991. Desktop Scanning for Cartographic Digitization and Spatial Analysis, *Photogrammetric Engineering & Remote Sensing*, 57(11):1437–1446.

Collins, S.H., 1981. Algorithms for Dense Digital Terrain Models. *Photogrammetric Engineering & Remote Sensing*, 47(1):71–76.

Delaplace, D.M., and R.K. Price, 1991. Urban Drainage Models Linked to CAD and GIS – The UK Experience, *Proceedings of UDT '91 – New Technologies in Urban Drainage*, Dubrovnik, Yugoslavia, 17-21 June, pp. 373–383.

Djokic, D., 1991. *Urban Stormwater Drainage Network Assessment Using an Expert Geographical Information System*, PhD. Dissertation, University of Texas at Austin. (University Microfilms International order no. 9128214)

Djokic, D., and D.R. Maidment, 1991. Terrain Analysis for Urban Stormwater Modeling, *Hydrological Processes*, 5:115–124.

Gandoy-Bernasconi, W., and O. Palacios-Velez, 1990. Automatic Cascade Numbering of Unit Elements in Distributed Hydrological Models, *Journal of Hydrology*, 112:375–393.

Goodrich, D.C., D.A. Woolhiser, and T.O. Keefer, 1991. Kinematic Routing Using Finite Elements on a Triangular Irregular Network, *Water Resources Research*, 27(6):995–1003.

Green, I.R.A., and D. Stephenson, 1986. Criteria for Comparison of

Single Event Models, *Hydrological Sciences Journal,* 31(3):395–411.

Greenlee, D.D., 1987. Raster and Vector Processing for Scanned Linework, *Photogrammetric Engineering & Remote Sensing,* 53(10):1383–1387.

Huber, W.C., and R.E. Dickinson, 1988. *Storm Water Management Model, Version 4: User's Manual,* U.S. Environmental Protection Agency, Athens, Georgia.

Huber, W.C., A.F. Zollo, T.W. Torbox, and J.P. Heaney, 1991. *Integration of the SWMM Runoff Block with ARC/INFO and AutoCAD: A Case Study,* Dept. Environmental Engineering, University of Florida, Gainsville, Florida., Purchase order VN1–320–420000, 121 p.

Johnson, L.E., 1989. MAPHYD - A Digital Map-Based Hydrologic Modeling System, *Photogrammetric Engineering & Remote Sensing,* 55(6):911–917.

Jenson, S.K., and J.O. Domingue, 1988. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis, *Photogrammetric Engineering & Remote Sensing,* 54(11):1593–1600.

Marsalek, J., T.M. Dick, P.E. Wisner, and W.G. Clarke, 1975. Comparative Evaluation of Three Urban Runoff Models, *Water Resources Bulletin,* 11(2):306–328.

Meier, D.W., and D.F. Lakatos, 1987. Modeling Support for Stormwater Management, *Proceedings of the ASCE Conference on Engineering Hydrology* (Arlen D. Feldman, editor), pp. 561–566.

Moore, I.D., and R.B. Grayson, 1991. Terrain-Based Catchment Partitioning and Runoff Prediction Using Vector Elevation Data, *Water Resources Research,* 27(6):1177–1191.

Moore, I.D., R.B. Grayson, and A.R. Ladson, 1991. Digital Terrain Models: A Review of Hydrological, Geomorphological, and Biological Applications. *Hydrological Processes,* 5(1):1–30.

Peuquet, D.J., 1981. An Examination of Techniques for Reformatting Digital Cartographic Data / Part 1: The Raster to Vector Process, *Cartographica,* 18(1):34–48.

———, 1981. An Examination of Techniques for Reformatting Digital Cartographic Data / Part 2: The Vector to Raster Process, *Cartographica,* 18(3):21–33.

QuickSurf, 1991. Company Information. Schrieber Instruments Inc., 4800 Happy Canyon Road, Suite 250, Denver Colorado, 80237.

ROOTS, 1991. Company Information. Decision Images, 196 Tamarack Circle, Skillman, New Jersey 08556.

Sircar, J.K., 1986. *Computer Aided Watershed Segmentation for Spatially Distributed Hydrologic Modeling,* PhD dissertation, Department of Civil Engineering, University of Maryland. (University Microfilms International order no. 8712265)

Smith, M.B., 1992. *A Distributed Parameter Hydrologic Model for Urban Stormwater Protection,* Doctoral Dissertation, Civil Engineering Department, University of Ljubljana, Slovenia. (University Microfilms International order number 9223634)

———, 1993. A GIS-Based Distributed Parameter Hydrologic Model for Urban Areas, *Hydrological Processes,* 7(1):45–61.

Smith, M.B., and M. Brilly, 1992. Automated Grid Element Ordering for GIS-Based Overland Flow Modeling, *Photogrammetric Engineering & Remote Sensing,* 58(5):579–585.

Terstriep, M.L., and M.T. Lee, 1989. Regional Stormwater Modeling, Q-ILLUDAS and ARC/INFO, *Proceedings of the 6th Annual ASCE Conference on Computing in Civil Engineering,* New York, pp. 338–345.

Thorpe, J.A., 1988. Contour Interpolation in Large Scale Urban Mapping Systems, *Proceedings of the 1988 ACSM-ASPRS Annual Convention,* St. Louis, Missouri, 13–18 March, pp. 61–68.

Tomlin, C.D., 1990. *Geographic Information Systems and Cartographic Modeling,* Prentice Hall, New Jersey, 249 p.

Tucker, L.S., 1969. *Availability of Rainfall-Runoff Data for Sewered Drainage Catchments,* Technical Memorandum No. 8. A.S.C.E. Water Resources Research Program.

Urenda, C., 1992. PC-Based GIS Manages Municipal Water Systems', *GIS World,* June, pp. 43–47.

Weibel, R., and M. Heller, 1991. Digital Terrain Modelling, *Principles, Geographic Information Systems,* Vol. 1 (D.J. Maguire, M.F. Goodchild., and D.W. Rhind, editors), John Wiley and Sons, N.Y.

Wenzel, H.G., and M.L. Voorhees, 1980. Adaptation of ILLUDAS for Continuous Modeling, *Journal of the Hydraulics Division,* A.S.C.E., 106(11):1795–1812.

(Received 7 December 1992; accepted 14 January 1993)

COMPUTER PROCESSING OF REMOTELY-SENSED IMAGES: AN INTRODUCTION

This book is an introduction to computer techniques of image processing applied to multispectral images acquired from satellites and aircraft, written at a level suitable both from students and for professionals moving into the field. The book begins with a survey of underlying physical principles, and goes on to consider the nature and sources of remotely-sensed data.

Computer Processing of Remotely-Sensed Images: An Introduction will be of value to all those interested in digital analysis of remotely sensed images in fields such as geography, geology, oceanography, agricultural science, civil engineering and environmental science.

Chapters include:

- Remote Sensing: Basic Principles
- Remote Sensing Platforms and Sensors
- Digital Computers and Image Processing
- Preprocessing of Remotely Sensed Data
- Enhancement Techniques
- Image Transforms
- Filtering Techniques
- Classification

by: Paul M. Mather. 1987. 352 pp. $55 (softcover); ASPRS Members $40. Stock # 4528.

For more information, see the ASPRS Store.