STORM RUNOFF ESTIMATION USING A GEOGRAPHIC INFORMATION SYSTEM

Mark R. Dickman and Oktay Güven

Auburn University
Highway Research Center
238 Harbert Engineering Center
Auburn, AL 36849-5337

June 1997

ABSTRACT

The purpose of this project was to determine the feasibility of using the Arc/Info geographic information system by the Environmental Systems Research Institute for performing typical storm runoff analyses necessary for storm drainage design. The Soil Conservation Service graphical peak flow method in Technical Release 55 (SCS, 1986) was used to determine the peak flow rates of the Pepperell Creek basin in Opelika, Alabama.

To apply the Soil Conservation Service peak flow method, the watershed area and the average areal curve number had to be estimated. Arc/Info calculated the watershed area after manipulating the digital elevation model of the location. Also, Arc/Info computed the average areal curve number of the watershed by combining the digital land use and land cover data with the digital hydrologic soil types.

The results were then compared to those estimated using traditional methods by the Alabama Department of Transportation. The Alabama Department of Transportation used regression equations determined for urban streams in Alabama.

ACKNOWLEDGEMENTS

This project was supported by the Auburn University's Department of Civil Engineering and the Highway Research Center. The material presented in this report is based on the Master of Science thesis of Mark R. Dickman submitted to Auburn University. Overall supervision was provided by Oktay Güven, with Professors Joel G. Melville and Brian L. Bowman, both of the Department of Civil Engineering at Auburn University, serving as members of the thesis committee.

The authors thank Charles Powell and Tom Flournoy of the Alabama Department of Transportation who graciously supplied their information about the test area. Also, the assistance of Dr. Lorraine Wolf and Jonathan Collier of the Auburn University Geology Department, Jeff Graves of the Auburn University College of Engineering Computing Services, and Kate Ozalas of the Alabama Agricultural Experimental Station is greatly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Geographic information systems (GIS) have been used for various analyses of spatially distributed data. The power of a GIS lies in its ability to reference features to a geographic location just as a conventional map does. However, a GIS has the ability to easily manipulate and combine different data sets in many ways that a map cannot. To employ a GIS for an analysis, the user must determine the objectives which will then define the data necessary for the project. Each data set has to be georeferenced so that each model has a common map projection and scale to ensure that one coordinate represents the same location in each model. The Arc/Info program by the Environmental Systems Research Institute (ESRI), the GIS used for this project, has menu driven subroutines to conduct many types of analyses. This particular GIS has the ability to convert different data types into a format that can be used in various geographic related data manipulations. Arc/Info can then provide several types of output such as color maps for presentations.

The primary purpose of the present study was to investigate the feasibility of using the Arc/Info GIS for performing typical storm runoff analyses required for storm drainage design applications. A hydrologic analysis was performed using Arc/Info to determine the peak flow of the Pepperell Creek basin in Opelika, Alabama, according to the Soil

Conservation Service (SCS) graphical peak flow method described in Technical Release 55 (SCS, 1986).

The first step of the analysis was to delineate the Pepperell Creek watershed from the digital elevation model (DEM) of this location with Arc/Info. Next, the soil type coverage was combined with the land use coverage to obtain the average areal runoff curve number (CN) for the delineated watershed. The watershed area and the curve number were used to estimate the peak flow rates for the 25 year, 50 year and 100 year design storms. These peak flows were compared to the results of an analysis using traditional methods of the same watershed conducted by the Alabama Department of Transportation (ALDOT) as a part of the plan to reroute US Highway 280.

Most of the digital data was available free of charge from public ftp sites. The digital elevation model (DEM) and the land use and land cover data were acquired from the U.S. Geological Survey (USGS). The Natural Resources Conservation Service (NRCS) provided the soil type data. The U.S. Census Bureau had the road, stream, and Lee County political boundaries.

II. BACKGROUND

Literature Review

Several recent articles described applications of geographic information systems (GIS) to conduct hydrologic analyses. The advantage of using a GIS for hydrologic modeling was the ability of the GIS to move away from the lumped parameter approach to acknowledging the spatial variability within a watershed (Greene and Cruise, 1995). Ross and Tara (1993) diagrammed the steps for a hydrologic analysis with a GIS. Using a hypothetical watershed, Warwick and Haness (1994) applied a GIS to determine the parameters necessary for the U.S. Army Corps of Engineers HEC-1 hydrologic model and to compute runoff hydrographs. Bryant and Schmitzer (1994) analyzed the differences between planimetric and true surface areas of GIS generated watersheds created from DEMs of different scales. Krug (1994) integrated thematic databases into a single GIS database.

In the illustration of "the benefits derived by the use of a GIS" by Ross and Tara (1993), an outline was presented of the hydrologic modeling processes necessary to implement a GIS. The first step was to gather the digital data required to do the analysis such as the land use and cover, soil types, topography, and the hydrography. Next, the GIS operations were conducted on the digital topography to define the basin. To estimate the runoff, the GIS was used to perform the map overlays and spatial analysis to develop the input data for the hydrologic models while the investigator entered the various rainfalls into the hydrologic model.

Warwick and Haness (1994) digitized the essential, hypothetical coverages in an attempt to determine the efficacy of using a GIS for hydrologic modeling. The coverages were a watershed boundary, topography, stream hydrography, rainfall, hydrologic soil types and land use. The GIS computed the area of the watershed, and the hydrologic soil types and the land use were combined by the GIS to estimate the areal average Soil Conservation Service runoff curve number for the basin.

Bryant and Schmitzer (1994) compared the true and planimetric surface areas computed by a GIS due to different scales of U. S. Geological Survey (USGS) digital elevation models (DEMs). DEMs of different scales mean that the number and spacing of elevation points changed for each surface model. From a DEM, the authors created a grid of the watershed to calculate the planimetric surface area. For the true surface area computation, a Triangular Irregular Network (TIN) was constructed of the same watershed. The watershed grid and TIN procedure created a file listing the surface areas of their respective models.

Description of the Study

The watershed for this hydrologic study was the drainage basin of the new U.S. Highway 280 bridge over Pepperell Creek in Opelika, Alabama. Figure 1 shows the location from a reproduction of a USGS 7.5 Minute Topographic Map. This location was

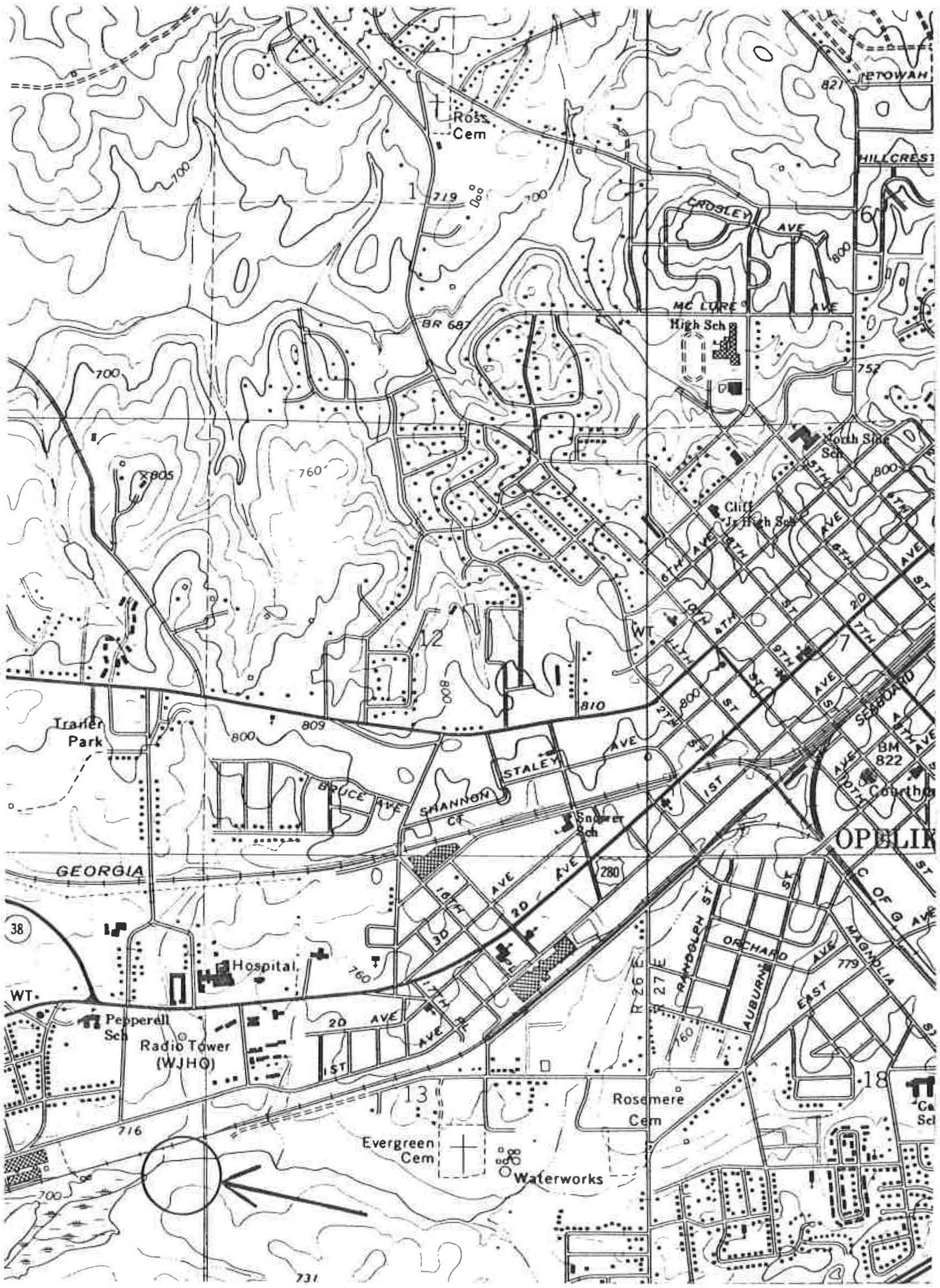

Figure 1. The Pepperell Creek Watershed Area on a USGS 7.5 Minute Topographic Map. The circled area is where the pour point is located.

chosen to compare the basin area and estimated peak discharge derived using the Arc/Info Geographic Information System (GIS) with that estimated using a traditional analysis conducted by the Alabama Department of Transportation (ALDOT).

The GIS was applied to the digital elevation model to determine the watershed area and the curve number necessary to estimate the peak flow rate for the 25, 50, 100 year design storms. The peak flow rates were compared to those obtained by the regression equation analysis by the ALDOT for the 25, 50, 100 year intervals.

III. METHODOLOGY

Overview

The first step for the hydrologic analysis was to determine the parameters necessary to estimate the peak flow rate. For the Soil Conservation Service (SCS) peak flow method (SCS, 1986), the watershed area, the time of concentration and an areal average curve number of the watershed had to be determined. The curve number reflected the average hydrologic soil and land use conditions of the watershed. The 24 hour rainfalls for the 25, 50 and 100 year storm return periods and the rainfall distribution were also needed.

The software used for the project was Arc/Info version 7.0.1 of the Environmental Systems Research Institute (ESRI) that was licensed to Auburn University. The equipment needed to run the software was a Sun Ultra 1 that operates on the Unix platform and was connected to the Auburn University Engineering network.

To estimate the area and time of concentration of the watershed, the western half of the digital elevation model (DEM) of Phenix City, Alabama, was manipulated by the Arc/Info GIS of ESRI. The scale of the DEM was 1 by 1 degree and was acquired from the public USGS ftp site. The DEM is described in more detail under the section describing the Arc/Info data input.

To estimate the watershed curve number, Arc/Info was used to combine the digitized soil type and land use data producing subareas each described by a single hydrologic soil type and a single land use type. Then, the areal averaged curve number for the watershed was calculated by totaling all of the curve numbers of the subareas weighted by the area that they represent.

Arc/Info Digital Input

Appendix II lists the information gathered from each set of digital data. To aid in the explanation, words in all capital letters and italics are Arc/Info commands or items in a coverage file. Words in small letters and italics are coverage names used by the author. For the exact procedures, see Appendix III.

Digital Elevation Model

A Digital Elevation Model (DEM) is a digitized topographic map. A DEM is a matrix of elevation values at regularly spaced intervals along the lines of latitude and longitude. Downloaded from the USGS ftp site at edcftp.cr.usgs.gov, the DEM file phenix_city-w.gz was transferred in the binary mode from the directory pub/data/DEM/250/P. The DEM was free of charge from this public access site.

The DEM used for this study was a 1 degree by 1 degree scale which corresponded to the west half of the 1 by 2 degree USGS topographic map of the Phenix City, Alabama, area. For the 1 degree DEM, the grid positions were in spherical coordinates, and the elevations were in meters, relative to mean sea level, according to the

National Geodetic Vertical Datum of 1929 (NGVD 29). This file contained an array of 1201 by 1201 elevation points in the file. Spaced every three arc seconds along the lines of latitude and longitude, the elevation points are about 75 meters apart parallel to the lines of latitude for the latitude of Phenix City and 90 meters apart along the lines of longitude (USGS, 1993).

The *DEMLATTICE* command in Arc/Info converted the DEM file to an Arc/Info coverage called a lattice. In a lattice representation of a surface, elevation values between sample points were approximated by interpolating between adjacent sample points. Figure 2 shows a representation of an elevation lattice and demonstrates the regular spacing of the lattice points (ESRI, 1995).

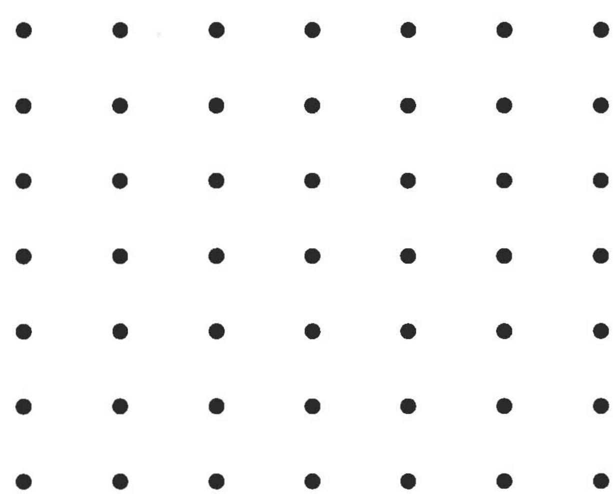

Figure 2. Example of a Lattice of Elevations.

The horizontal and vertical accuracy of the DEM data depend upon the distance between elevation points, the quality of the source data and the digitizing systems.

Within a standard DEM, most terrain feature anomalies, such as peaks and pits, are generalized and smoothed since the elevation points are spaced at regular intervals in the horizontal plane. Since this generalization reduces the ability to resolve specific features smaller than the grid spacing, the actual surface receives a smoothing effect after gridding. For a better representation of these features, the lattice resolution must be increased, or, in other words, the distance between sample points must be decreased (ESRI, 1995). Of course, an increased number of sample points means a larger database and computational time.

Since the horizontal coordinates were in the angular geographic coordinate system and the elevations were in meters, the DEM was not useful for slope, volume, or visibility analysis. With the dem.prj file as listed in Appendix I, the Arc/Info *PROJECT* command converted the spherical coordinates of arc-seconds into the planar coordinates of meters. This procedure also compensated for length variations associated with the spherical projection at the latitude of interest.

Land Use and Land Cover Data

From the USGS, the land use and land cover data describes the vegetation, bodies of water, natural, and man made features of the surface. The file was obtained from the pub/data/LULC/250K/P/phenix_city directory in the edcftp.cr.usgs.gov ftp site at no charge. Initially, the USGS digitized the land use data from NASA high altitude aerial photographs. Early land use maps and field surveys supplemented the land use data as needed (USGS, 1996).

In two dimensional vector format, many lines, or arcs, from a series of ordered (x,y) coordinates created the boundaries of areas called polygons. Each polygon was labeled with an integer to represent a feature code according to the Anderson Land Use and Land Cover classification system created by Anderson and others in 1976. The codes are listed in Table 1. Any man made features, such as building complexes, were shown if the polygon size was larger than 10 acres and had a minimum width of 660 feet. Rural and natural features, such as forests, were mapped when the polygon size was greater than 40 acres and had a minimum width of 1320 feet (USGS, 1996).

Stored according to the Geographic Information Retrieval and Analysis System (GIRAS) format, this data was at the 1:250,000 scale and in a form of the Universal Transverse Mercator (UTM) projection. The different form of the UTM projection means that the coordinates used are not as large as the true UTM coordinates in order to save hard disk space (USGS, 1996).

Table 1. Anderson Land Use Classification Codes.

Land Use Code	Description
1	Urban or built-up land
11	Residential
12	Commercial and services
13	Industrial
14	Transportation, communication, utilities
15	Industrial and commercial complexes
16	Mixed urban or built-up land
17	Other urban or built-up land
2	Agricultural land
21	Cropland and pasture
22	Orchards, groves, vineyards, nurseries, and ornamental horticultural

23	Confined feeding operations
24	Other agricultural lands
3	Rangeland
31	Herbaceous rangeland
32	Shrub and brush land
33	Mixed rangeland
4	Forest land
41	Deciduous forest land
42	Evergreen forest land
43	Mixed forest land
5	Water
51	Streams and canals
52	Lakes
53	Reservoirs
54	Bays and estuaries
6	Wetland
61	Forested wetland
62	Nonforested wetland
7	Barren land
71	Dry salt flats
72	Beaches
73	Sandy areas not beaches
74	Bare exposed rock
75	Strip mines, quarries, gravel pits
76	Transitional areas
8	Tundra
81	Shrub and brush tundra
82	Herbaceous tundra
83	Bare ground
84	Wet tundra
85	Mixed tundra
9	Perennial snow or ice
91	Perennial snowfield
92	Glaciers

Another data type from the USGS Land Use and Land Cover ftp site was the County Political Boundary. This was also available at no charge. This was also at the 1:250,000 scale, a form of the UTM projection and in the GIRAS format. Figure 3 shows the Phenix City, Alabama, 1:250,000 Land Use and Land Cover map with the county boundaries of that quadrangle draped over it. The colors for the Land Use types were assigned by the author based on a generalization of the Anderson Land Use Codes.

To convert these data types from the GIRAS format into an Arc/Info coverage, the U.S. Environmental Protection Agency (EPA) provided a conversion program called Girasarc2.aml. Written in the Arc Macro Language, the Arc/Info programming language, this program was available for downloading from the EPA ftp site ftp://ftp.epa.gov/pub/EPAGIRAS/meta. Constructed by the USGS Water Resources Division, the program processed the data with a minimum of intervention. In addition to converting the GIRAS files into an Arc/Info coverage, this program transformed the projection of the coverages into the true UTM projection, generated a synthetic neatline based on the mathematically determined corners of the map, and prepared the data for combining with adjacent maps. After completing the application of the Girasarc2.aml, Girasneat.aml removed the boundaries between polygons with the same attribute (EPA, 1996). Girasneat.aml is available from the same EPA ftp site.

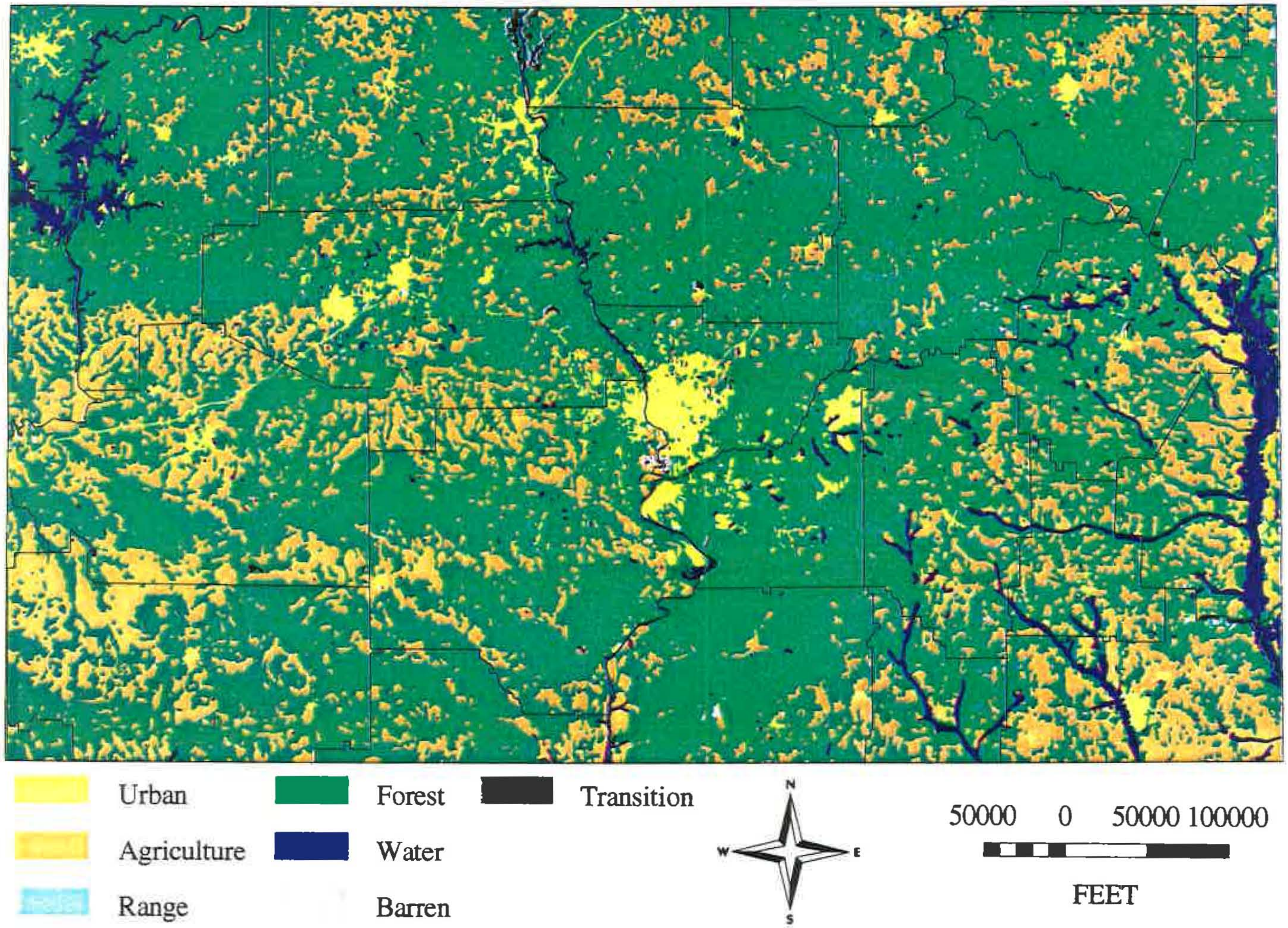

Figure 3. Phenix City, Alabama, 1:250,000 Land Use and Land Cover Map with the County Boundaries Overlaid.

Soil Type Data

The Natural Resource Conservation Service (NRCS) in Auburn, Alabama, provided, at no cost, the digitized soil type data for Lee County which was constructed in Map Information Assembly and Display System (MIADS). This data was stored in raster format which means that the county was divided into evenly spaced cells with one soil type assigned to each cell (Spratling, 1996, Personal Interview).

The Lee county soils data was at the 1:20,000 scale and in the UTM projection (Spratling, 1996, Personal Interview). However, each cell was not georeferenced. This means that the creators of the soil type maps did not ensure that the cell coordinates in their model represented the same location in another model (ESRI, 1994).

Each cell was 15.44 acres, or 250 meters by 250 meters (Spratling, 1996, Personal Interview). The feature values are in Table 2 with the corresponding hydrologic soil codes listed by the Soil Survey of Lee County, Alabama (SCS, 1981). The soil type codes referred to the soil names and were referenced to the hydrologic soil groups in the Soil Survey of Lee County of 1981. Figure 4 shows the soil type distribution for Lee County, Alabama. The Arc/Info command *MIADSARC* converted the soils information file from the MIADS format into an Arc/Info coverage.

Table 2. Soil Codes with the Corresponding Hydrologic Soil Groups.

Hydrologic Group	Soil Codes
A	4,5,40,41,42
B	2,3,7,8,9,10,14,16,17,18,19,20,22,23,24,
	25,26,28,29,30,31,32,33,34,39
C	6,11,12,13,27,36,37,38,
D	15,21

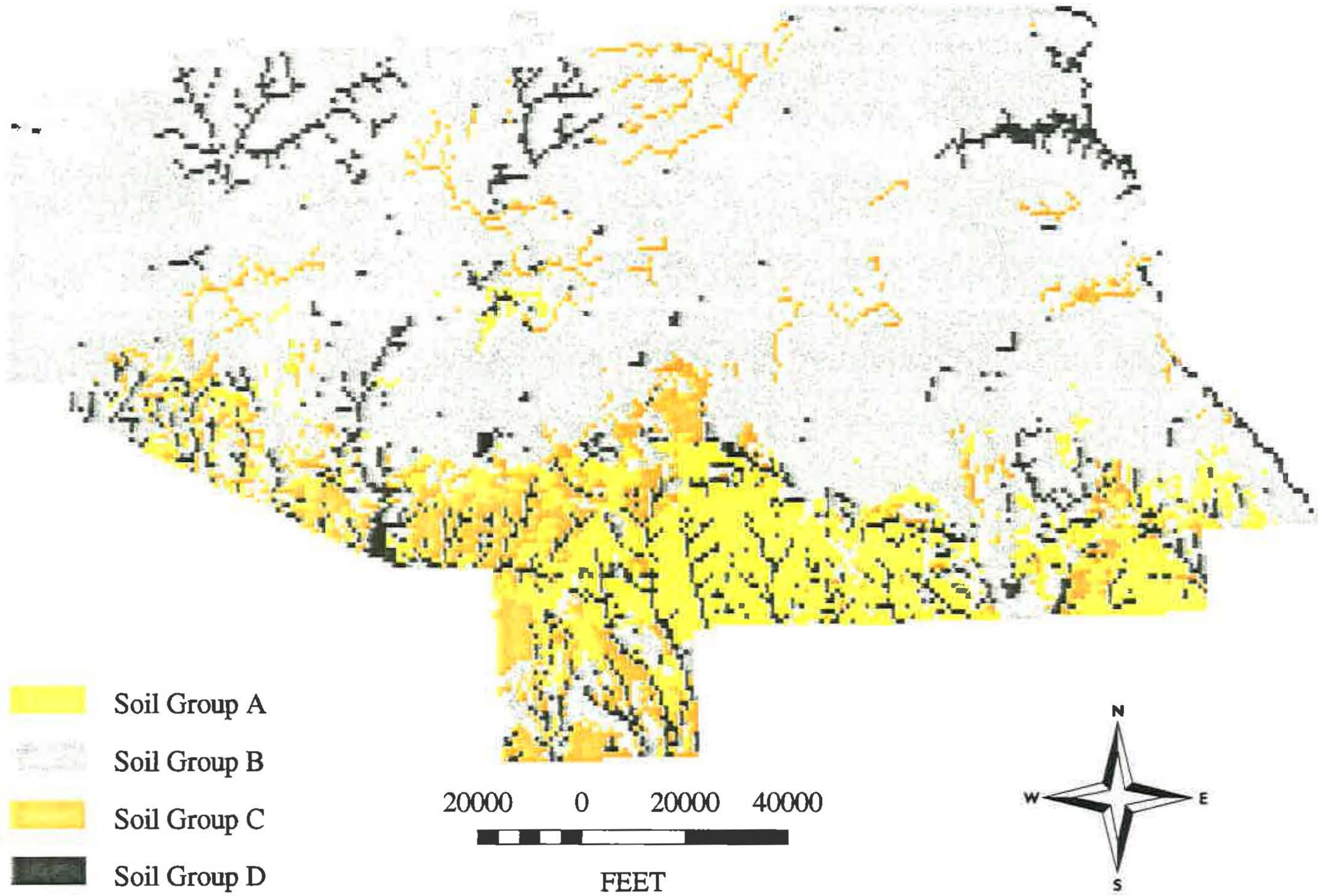

Figure 4. Hydrologic Soil Groups of Lee County, Alabama.

TIGER Files

The streams, roads, and boundary information came from the US Census Bureau and were named TIGER/Line (Topographically Integrated Geographic Encoding and Referencing) files. The TIGER/Line files were at the 1:50,000 scale and in the UTM projection (U.S. Census Bureau, 1992).

The TIGER digital line graph (DLG) files were created from the USGS 1:100,000 digital line graphs (DLG) and the USGS 1:24,000 scale quadrangles. Later, these were updated by aerial photographs (U.S. Census Bureau, 1992).

TIGER/Lines represent features, or spatial objects, found on a map. Spatial objects have geometry, which is a coordinate location and shape, and topology which is the relationship between points, line objects and polygons. The topological structure of the TIGER/Line data base defines the location and relationship of streets, rivers, railroads and other features to each other and to other geographic entities. Each digitized line is made of chains, and each chain has one node, or point, to mark the beginning and another to mark the end of a complete chain. Also, nodes connect chains together (U.S. Census Bureau, 1992).

The TIGER/Line files were available from the Census Bureau for a fee. The data used for this project was obtained from the Auburn University Alabama Land Resource Information Center of the Alabama Agricultural Experiment Station at no charge. Figure 5 is the combination of streams, roads, and boundaries.

To convert the TIGER/Line files to an Arc/Info coverage, the command *DXFINFO* was used to list the layers in the file, the kinds of features in each layer, and

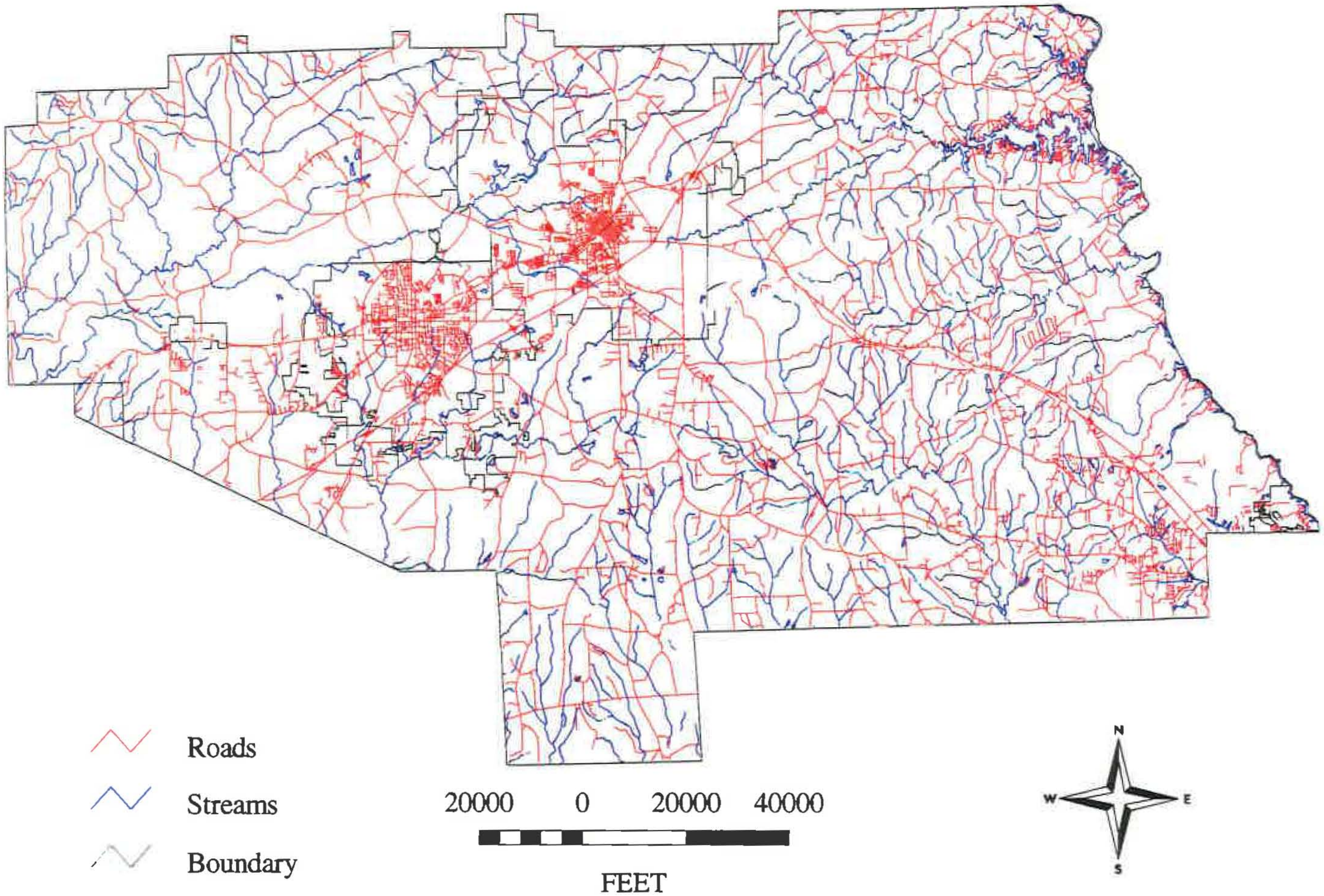

Figure 5. Lee County, Alabama, Streams, Roads and Boundary from the TIGER Files.

the length of the character strings that must be converted. *DXFARC* converted the files into an Arc/Info coverage (ESRI, 1995).

GIS Procedures

This section describes the Arc/Info procedures used to obtain the hydrologic parameters from the digital data. First, the watershed was delineated from the DEM. Then, the slopes within the watershed were determined from the DEM. The curve number of the watershed was calculated from the combination of the hydrologic soil type and the Land Use and Land Cover data. Finally, these parameters were used in the hydrology equations to calculate the peak flow rates for the 25, 50, and 100 year design storms.

Watershed Delineation

Before delineating the watershed, sinks in the *PROJECT*ed grid were *FILL*ed. Sinks, individual cells totally surrounded by much higher cells, would interfere with the flow direction calculations since they would only receive flow and not contribute to downstream flow. Even though some of these sinks could be natural, most were imperfections in the DEM. For coarse DEMs like the type used here, 5% of the cells could be sinks (ESRI, 1994).

To determine the direction of flow from every cell in the grid, the *FLOWDIRECTION* grid was composed. The flow direction was the direct path that runoff would take if rain fell on the DEM. With the *FILL*ed surface as input, the output was a grid containing the direction of flow out of each cell. This was the path of maximum descent from the center of each cell to its steepest downslope neighbor and computed by Arc/Info as

maximum drop = change in elevation / distance.

After the flow direction from each cell was evaluated for the grid, the flow direction grid was entered into the *FLOWACCUMULATION* function to calculate the number of cells that flow into each cell. The flow accumulation grid contained the number of upslope cells that flows into each downslope cell. Cells with a high flow accumulation were areas of concentrated flow identifying stream channels, but cells with zero flow accumulation were topographic highs.

From the *FLOWACCUMULATION* grid, the *STREAMNET* function created stream networks predicted by the DEM. These networks were cells that were identified because they collected flow from a large number of cells. During this operation, each cell in the lattice was assigned a value if the number of cells emptying into it was greater than threshold number of cells. As seen in Figure 6, the derived stream network shows the cells that receive flow from at least 100 cells which matches very well with the Lee County stream coverage.

Overlaying the stream network grid derived from the DEM with the Lee County road and stream coverages, the pour point of the watershed was determined by the *MEASURE WHERE* feature in Arcplot. To determine the watershed grid, the *FLOWDIRECTION* and *PCFILL* output grids and the coordinates of the pour point were entered into the *WATERSHED* function. The *WATERSHED* function used the entered

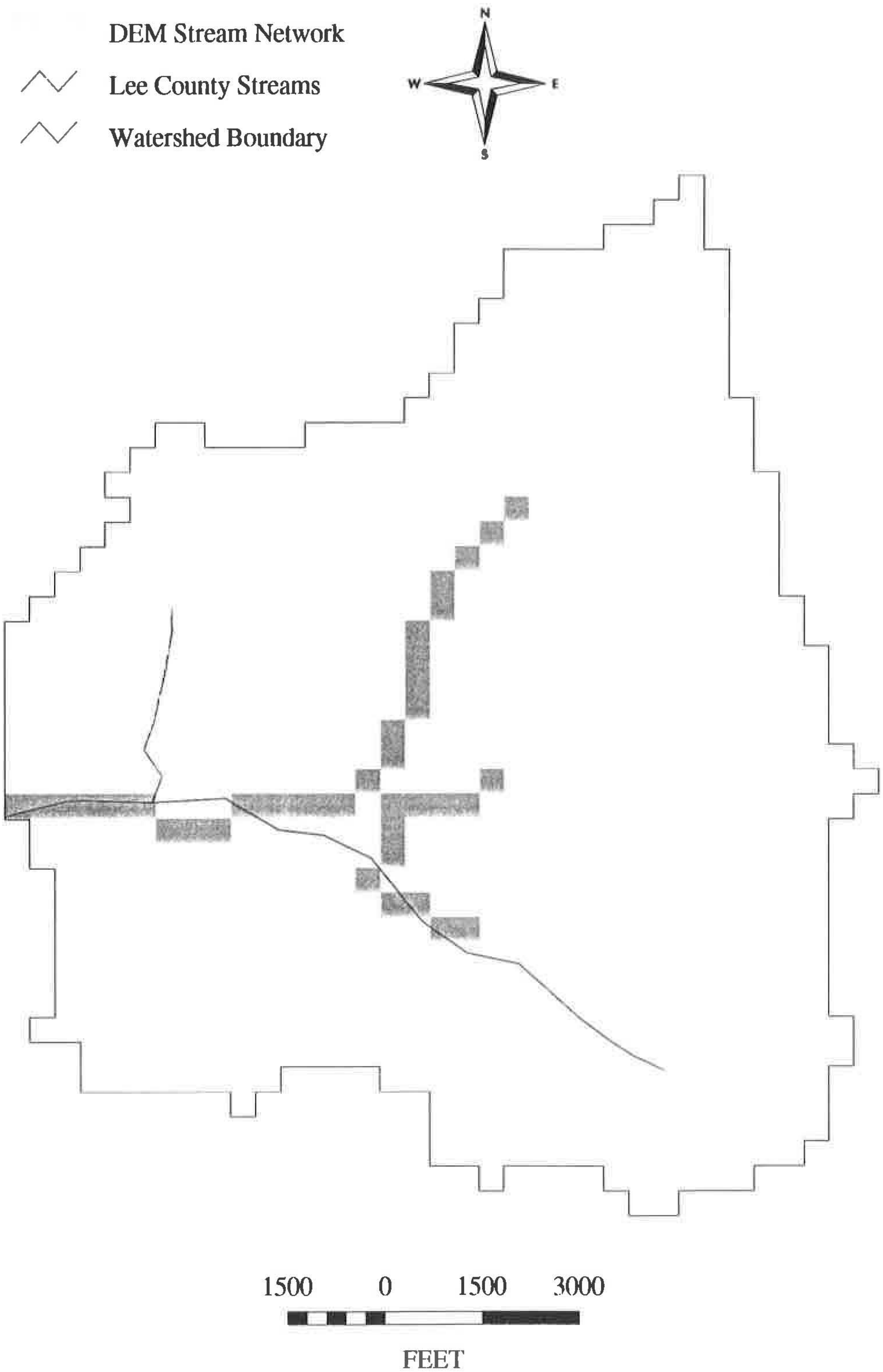

Figure 6. TIGER File Streams Overlaying Stream Network Derived from a DEM.

point as the lowest point of the watershed and moved upslope to include all of the area contributing flow to that pour point. This continued to the ridge lines determined in the *FLOWDIRECTION* grid. The new watershed grid was converted to a polygon coverage named *wshedp*, and the planimetric surface area is listed in a file under a heading named *area*.

Figure 7 is an approximation of the West Opelika, Alabama, 7.5 minute topographic map with 25 foot elevation contour lines and the TIGER streams and roads, and the watershed is highlighted in yellow. Figure 8 is a three dimensional view of the same map, and Figure 9 is the same view without the watershed to demonstrate the slopes around the Pepperell Creek watershed. Figure 10 is a close up of the watershed with 25 foot contours.

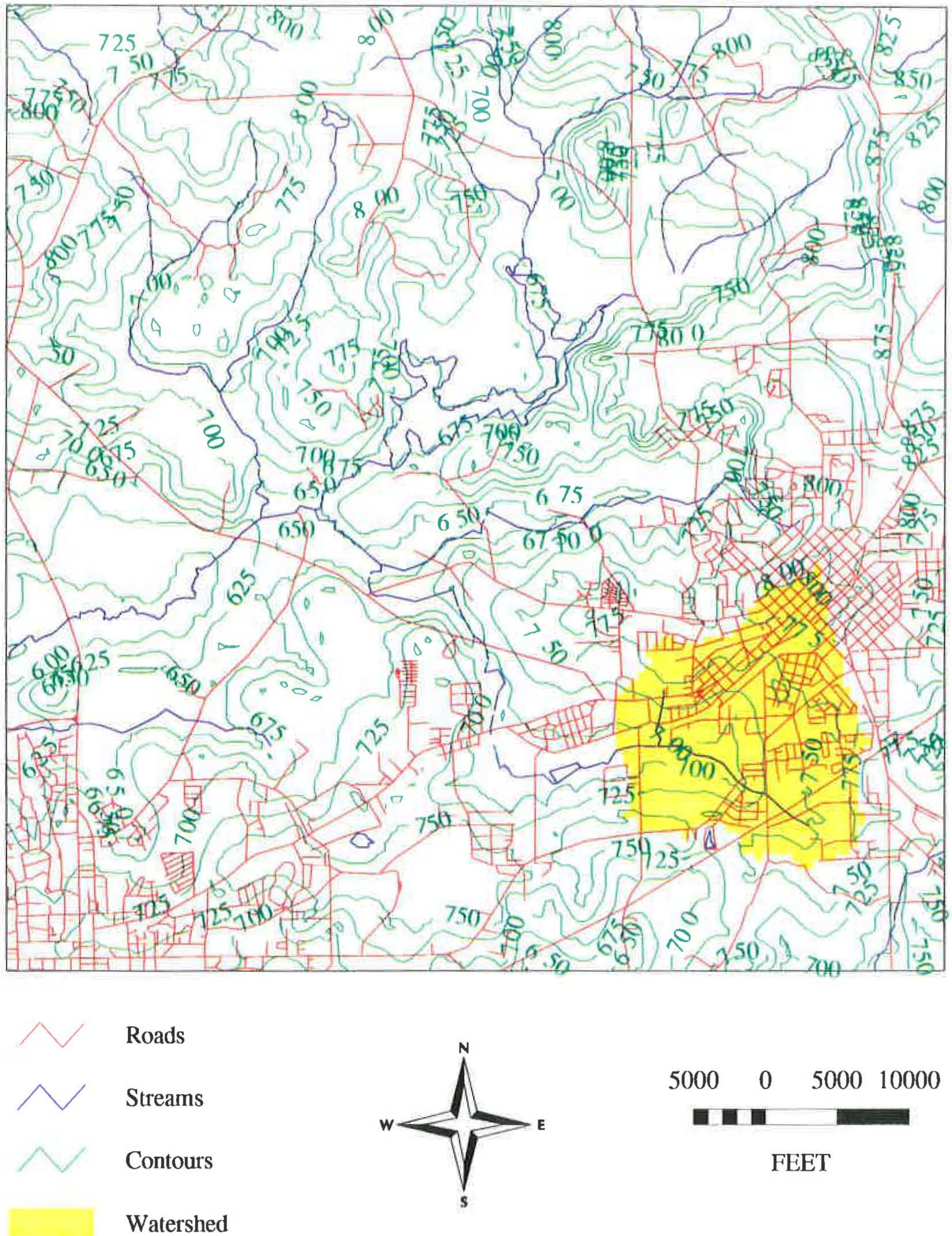

Figure 7. Two-Dimensional, 7.5 Minute Topographic Representation with 25 Foot Contour Lines, TIGER Streams and Roads, and the Basin Boundary.

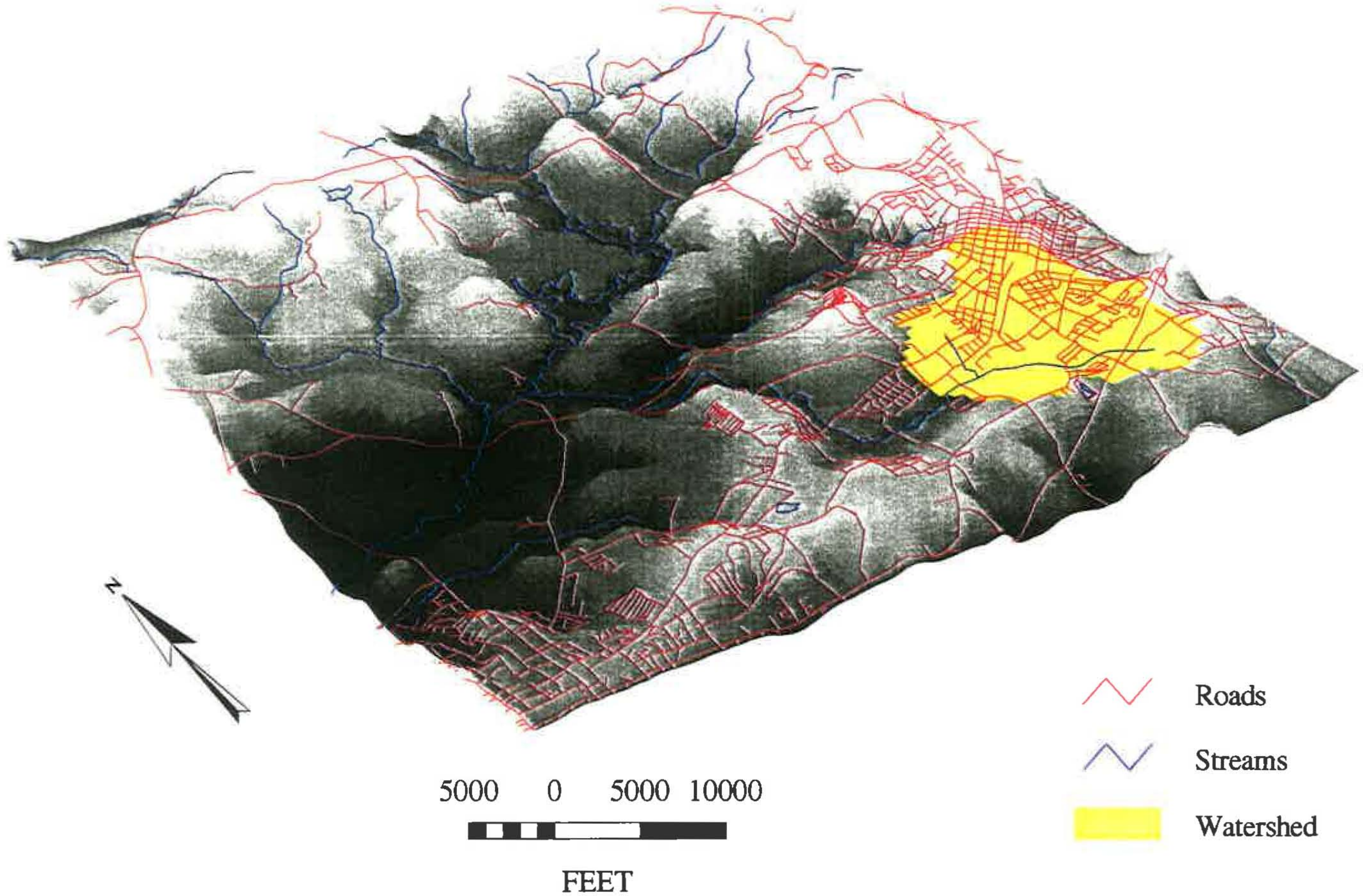

Figure 8. Three-Dimensional, 7.5 Minute Topographic Representation with the TIGER Streams and Roads and the Basin Boundary.

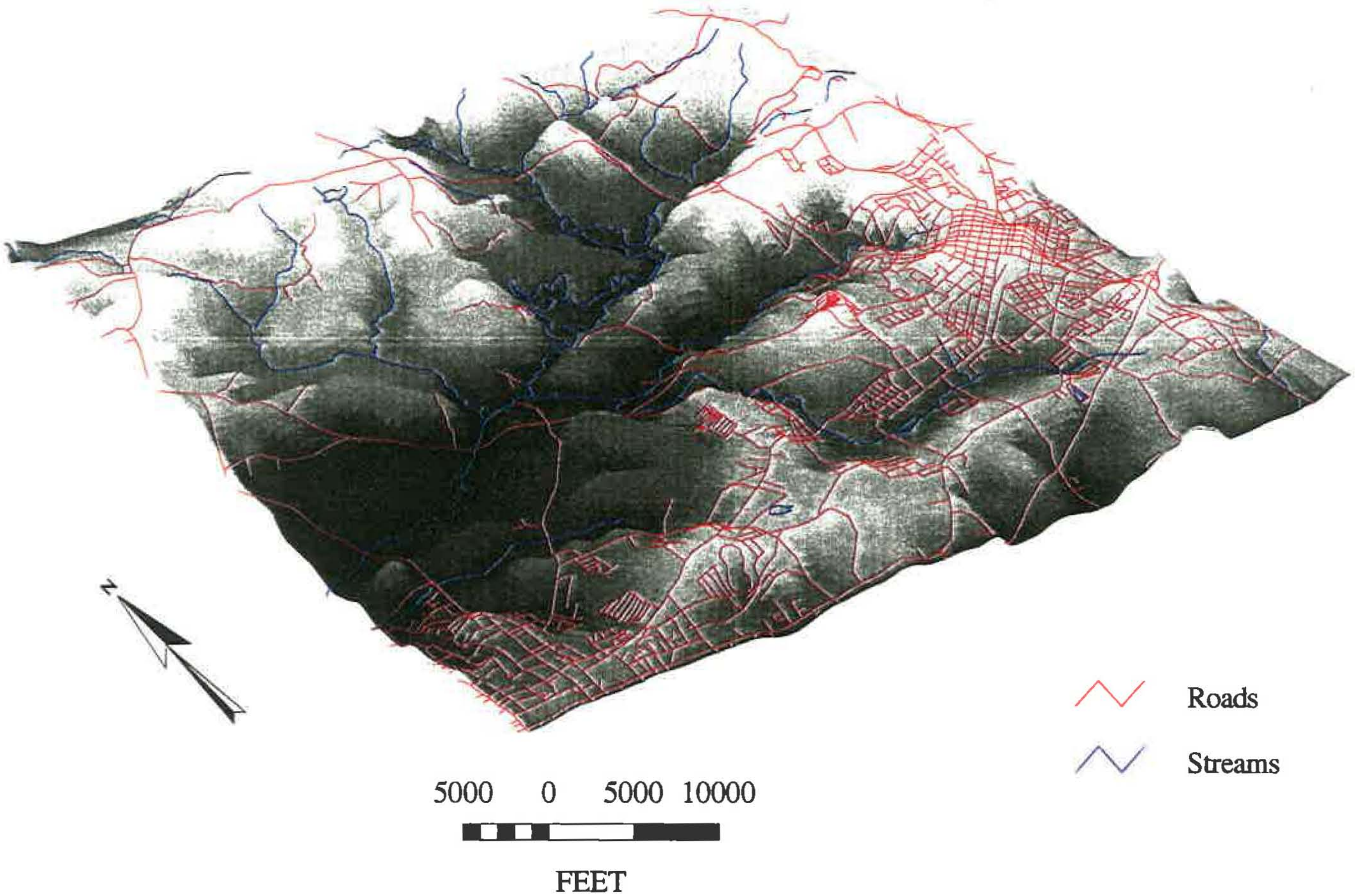

Figure 9. Three-Dimensional, 7.5 Minute Topographic Representation with the TIGER Streams and Roads but without the Basin Boundary.

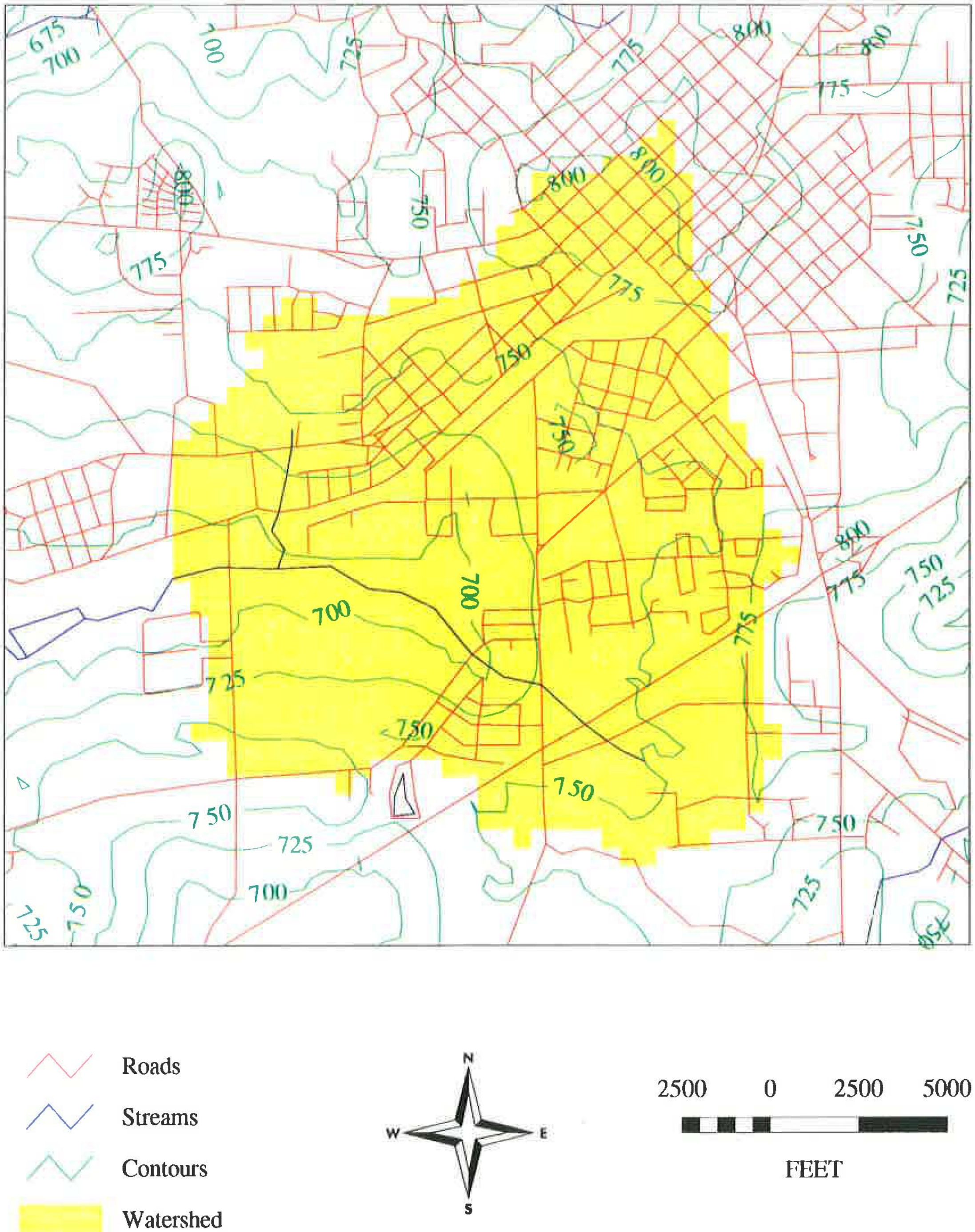

Figure 10. Close-up of the Basin with 25 Foot Contours
and TIGER Streams and Roads.

Determining the Slopes Within the Watershed

To determine the slopes within the basin, the *LATTICETIN* command created a Triangular Irregular Network (TIN) of the *FILL*ed grid. A TIN was necessary to conduct surface analyses that require the third dimension of elevation such as slope and true surface area. The TINs stored topological relationships between triangles and their adjacent neighbors. The *TINARC* function converted the TIN into a point coverage called *pcarc*.

To account for elevation trends beyond the watershed boundary, an area of at least 10% greater than the watershed boundary was required. For the extra area, the *MEASURE WHERE* function was used to determine the set of four (x,y) points around the watershed by trial and error.

Finally, in the *CREATETIN* environment, the *pcarc* combined with the *wshedp* coverage. The *MASS* option specified that the *pcarc* features were points, and the *SOFTCLIP* option included only the points and the interpolated values up to the edge of the watershed polygon. The *TINARC* command converted the new TIN to a polygon coverage named *shedpoly*. Figure 11 shows the TIN of the watershed.

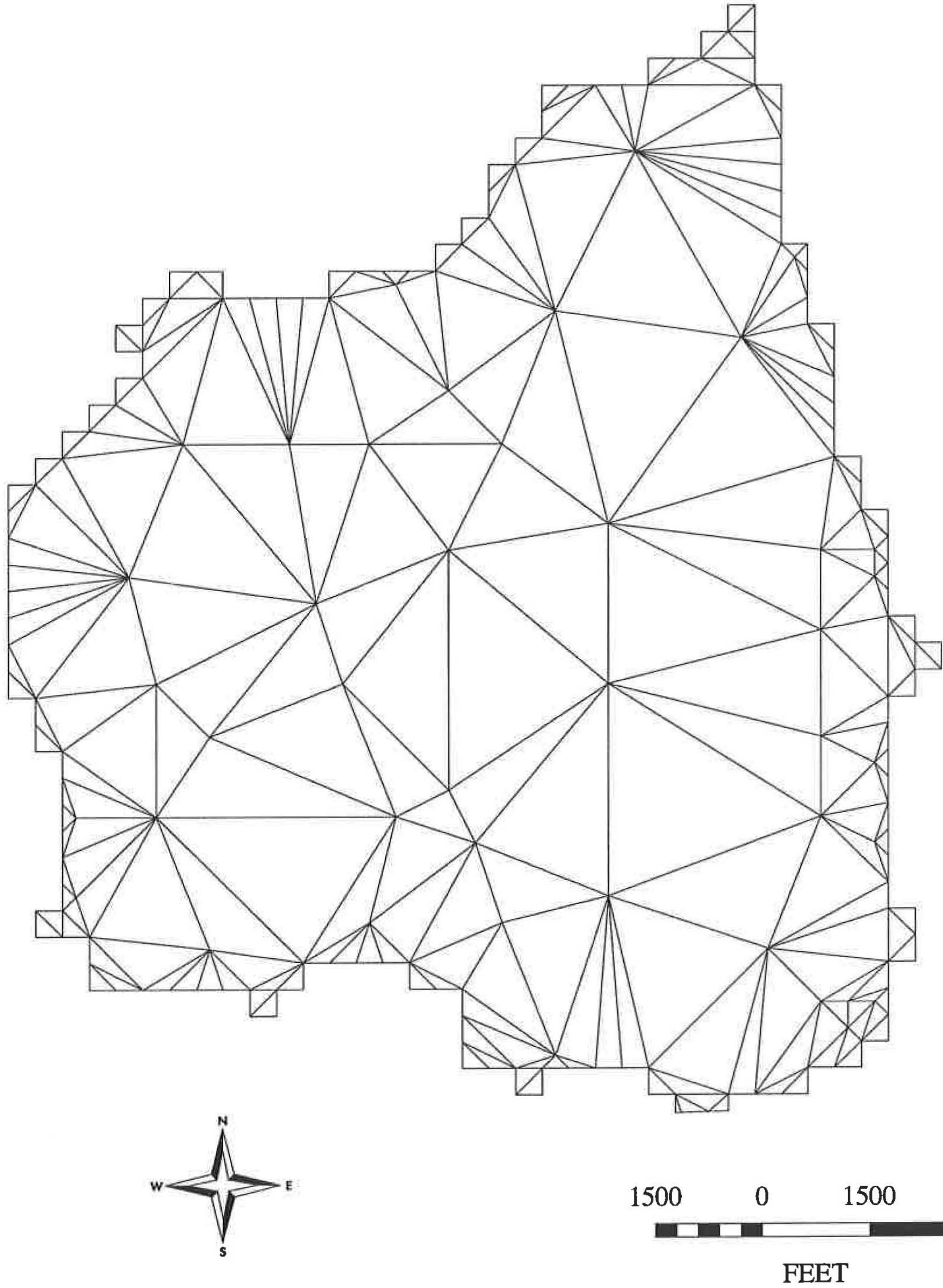

Figure 11. TIN of the Watershed.

Defining the Watershed Curve Number

To find the areal averaged curve number (CN) for the watershed, an application of the SCS runoff method similar to Warwick and Haness was used (Warwick and Haness, 1994). The first step was to *CLIP* the Phenix City land use polygon data to the watershed polygon, *wshedp*, to create the coverage *shedlulc*. With the *CLIP* function, *shedlulc* only included the land use features within the watershed polygon. Figure 12 depicts this.

After *CLIP*ping the soils polygon data to *wshedp* to make *shedsoil* as in Figure 13, *shedlulc* and *shedsoil* were combined to make *shedcn* with the *UNION* command. *UNION* computed the geometric intersection of the two polygon coverages. This union is demonstrated in Figure 14. Each polygon contained only one soil type code and one land use code.

To calculate the areal averaged curve numbers for each polygon in the watershed, the land use and soils attributes were each assigned values to be combined (Warwick and Haness, 1994). The soil group identification numbers, *CN1*, were designated integers from 0 to 3 based on the hydrologic soil groups defined by the NRCS and then stored in a look up table called *leesoilp.lut*. Stored in *pclulcn.lut*, the land use identification numbers, *CN2*, were based on a level II antecedent moisture condition and skip to every fourth number to account for the combination with *CN1*. The *pclulcn.lut* and the *leesoilp.lut* are listed in Tables 3 and 4 respectively.

Then, *shedcn.pat* was joined with *pclulcn.lut* and *leesoilp.lut* to assign the correct hydrologic soil type identification number and land use identification number for each polygon containing one hydrologic soil type and land use combination. This was stored

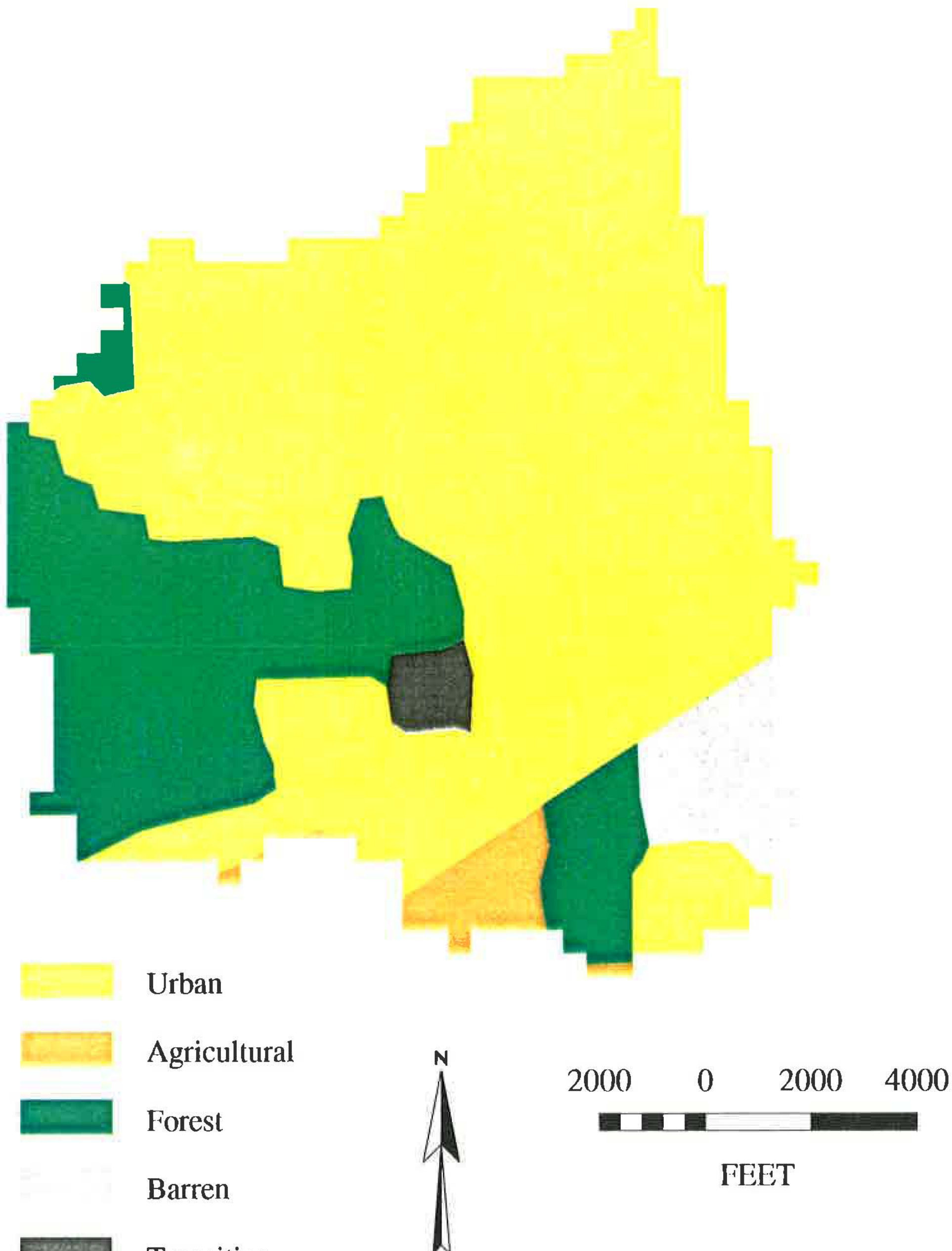

Figure 12. Land Use and Land Cover for the Watershed.

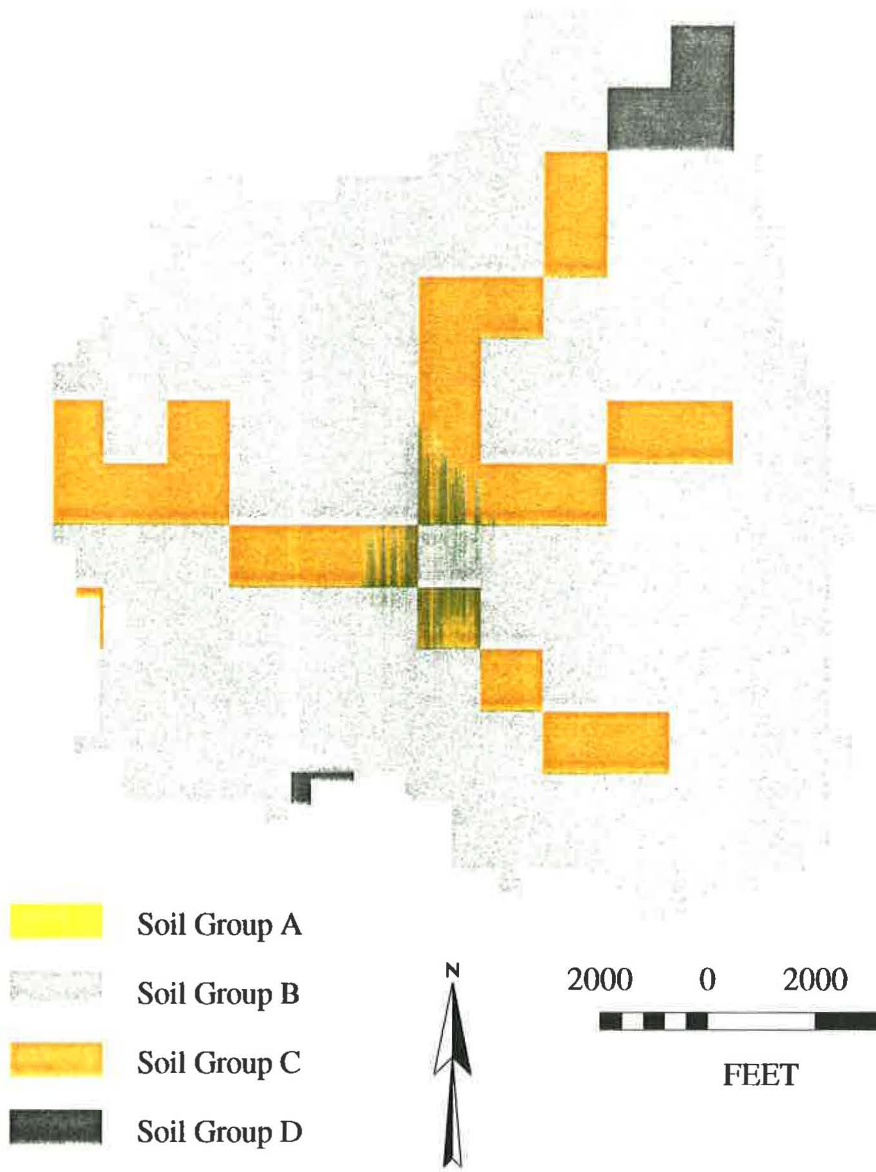

Figure 13. Watershed Hydrologic Soil Groups.

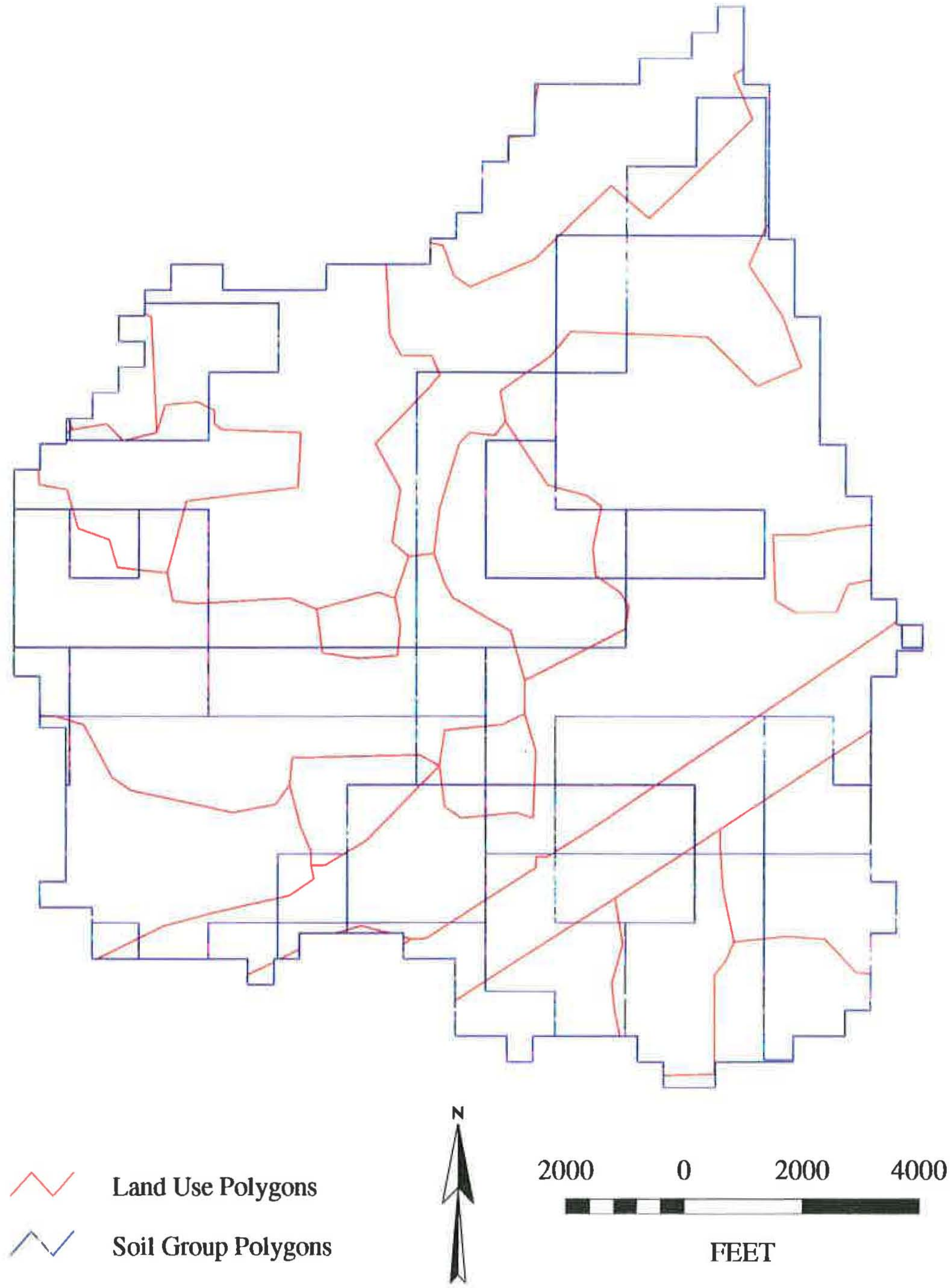

Figure 14. Combination of Watershed Land Use and Soils Polygons.

Table 3.Assigning an Identification Number to the Hydrologic Soil Groups.

CN1	Hydrologic Soil Groups
0	A
1	B
2	C
3	D

Table 4. Assigning an Identification Number to the Land Use Codes.

CN2	Land Use Codes
1	11
5	12
9	13,15,16,17
13	14
17	21
21	22
25	23
29	24,32
33	41,42,43
37	51,52,53,54,61,62
41	75
45	76

in *shedcn.pat* where *CN1* and *CN2* were added to give *CN3*. To convert each *CN3* into a SCS curve number (*CN*) for each land use and hydrologic soil type polygon, *shedcn.pat* was *JOIN*ed with *cn3.lut*. This conversion is listed in Table 5. Item *CNAVE* was then added to *shedcn.pat* to calculate the average curve number for each polygon area by the equation

$$CNAVE = CN \times AREA \,/\, TOTAL\ AREA.$$

Summing every *CNAVE* for the watershed resulted in the average areal curve number, *CN*, for the watershed.

Table 5. Relating the Land Use and Soil Group Combination to a SCS Number.

CN3	CN	CN3	CN	CN3	CN	CN3	CN
1	61	13	98	25	49	37	100
2	75	14	98	26	69	38	100
3	83	15	98	27	79	39	100
4	87	16	98	28	84	40	100
5	89	17	64	29	35	41	76
6	92	18	74	30	56	42	85
7	94	19	81	31	70	43	89
8	95	20	85	32	77	44	91
9	81	21	43	33	43	45	72
10	88	22	65	34	65	45	82
11	91	23	76	35	76	47	87
12	93	24	82	36	82	48	89

Hydrology Equations

The curve number, abbreviated *CN*, determined *S*, the potential maximum retention, in inches, by the SCS empirical equation

$$S = \frac{1000}{CN} - 10.$$

With *S*, the storm runoff for the watershed was computed by the SCS runoff equation

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S}$$

where *Q* was the storm runoff in inches and *P* was the 24 hour rainfall in inches of a specified return period. From TR-55 (SCS, 1986), the peak flow rate, q_p, in cfs, was calculated by

$$q_p = QAq_u$$

where q_u was the unit peak discharge in cfs/mi^2/in, *Q* was determined above, and *A* was the drainage basin area in square miles. The unit peak discharge, q_u, was determined from the peak flow chart in TR-55 (SCS,1986) using the parameters I_a/P and T_c. I_a, the initial abstraction in inches, was all of the losses before runoff begins and empirically determined by the SCS to be equal to 0.2*S*. T_c, the time of concentration in hours, was the time required by the runoff to travel from the hydraulically most distant point of the watershed to the pour point.

To determine the time of concentration according to TR-55 (SCS, 1986), all of the estimated travel times, T_t, also in hours, of the water to flow down gradient from the furthest point on the watershed to the pour point were summed. Sheet flow, overland flow and channel flow were used in the TR-55 (SCS, 1986) travel time calculations, and several runoff starting points were chosen to determine the hydraulically most distant point of the watershed. For the SCS Lag Equation (Viessman and Lewis, 1996), one time of concentration was calculated for the whole basin.

For the first 300 feet of flow according to TR-55 (1986), sheet flow was assumed and a simplified form of the kinematic-wave equation was applied:

$$T_t = \frac{0.007(nL)^{0.8}}{(P_2)^{0.5} s^{0.4}}$$

where n was Manning's roughness coefficient, L was the flow length, P_2 was the 2 year rainfall for 24 hours, and s was the slope in ft/ft. This was based on steady shallow uniform flow, constant intensity of a 24 hour rainfall available for runoff, and minor infiltration affects on travel time (SCS, 1986).

After about 300 feet, sheet flow becomes shallow concentrated flow (SCS,1986). For unpaved surfaces, the velocity, V, was given in feet per second as

$$V = 16.1345\ (s)^{0.5},$$

and for paved surfaces, V was

$$V = 20.3282\ (s)^{0.5}.$$

T_t was then calculated by the SCS graphical peak flow method (1986) to be

$$T_t = \frac{L}{3600V} \ .$$

Where streams are listed on a USGS quadrangle sheet, open channel flow was assumed (SCS, 1986). Manning's equation for open channel flow was

$$V = \frac{1.49 r^{2/3} s^{1/2}}{n}$$

where V was the average velocity in feet per second; r was the hydraulic radius in feet and was equal to a/p_w; a was the cross sectional flow area in ft^2; p_w was the wetted perimeter in ft; s was the slope of the channel in ft/ft; n was the Manning's roughness coefficient for open channel flow. This velocity was substituted into

$$T_t = \frac{L}{3600V}$$

to determine the T_t (SCS, 1986).

The SCS Lag Equation (Viessman and Lewis, 1996) applied one equation for the whole basin. This was

$$t_c = \frac{1.67 L^{0.8} [1000/CN - 9]^{0.7}}{1900 S^{0.5}}$$

where t_c was in hours, L was the hydraulic length or longest flow path of the watershed in feet, CN was the SCS runoff curve number, and S was the average basin slope in percent.

IV. RESULTS AND DISCUSSION

Peak Flow Rates Determined Using the SCS Method

Time Of Concentration

To find the time of concentration for the basin, three starting points were chosen at the basin boundary to estimate the hydraulically most distant point from the basin outlet. The overland flow lengths, *L*, and the slopes, *S*, were necessary to calculate the times of concentration by the method in TR-55 (SCS, 1986). The flow lengths were estimated by applying an Arc/Info interactive function that measured distance. Slopes were determined for the paths using a similar tool on the TIN.

Using the *STREAMNET* function and varying the threshold number of cells emptying into each cell, the growth of the flow network within the basin was created. Figure 15 shows the three starting points used to estimate the time of concentration and the stream network for the cells in the basin that received flow from more than 15 cells, and Figure 16 shows the cells that receive flow from over 4 cells in relation to the three starting points. The flow paths used to calculate the travel times were overlaid on the stream networks in Figures 15 and 16. These were compared to Figure 17, which has the watershed streams and roads coverages overlaid, to estimate the flow lengths, slopes, and the type of overland velocity along the path.

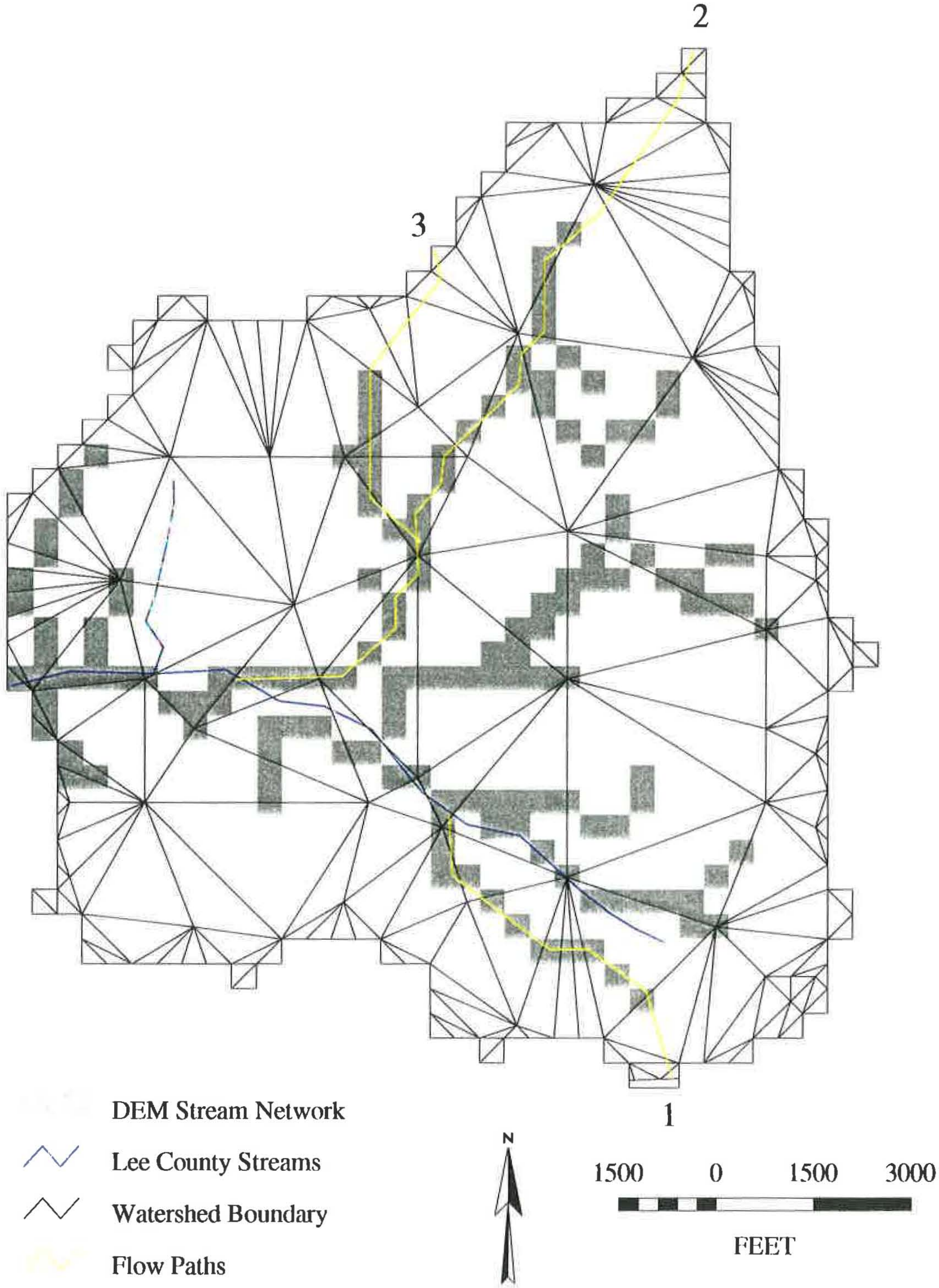

Figure 15. Stream Network for Cells Receiving Flow from 15 Cells or More with the TIN and Lee Streams Overlaid.

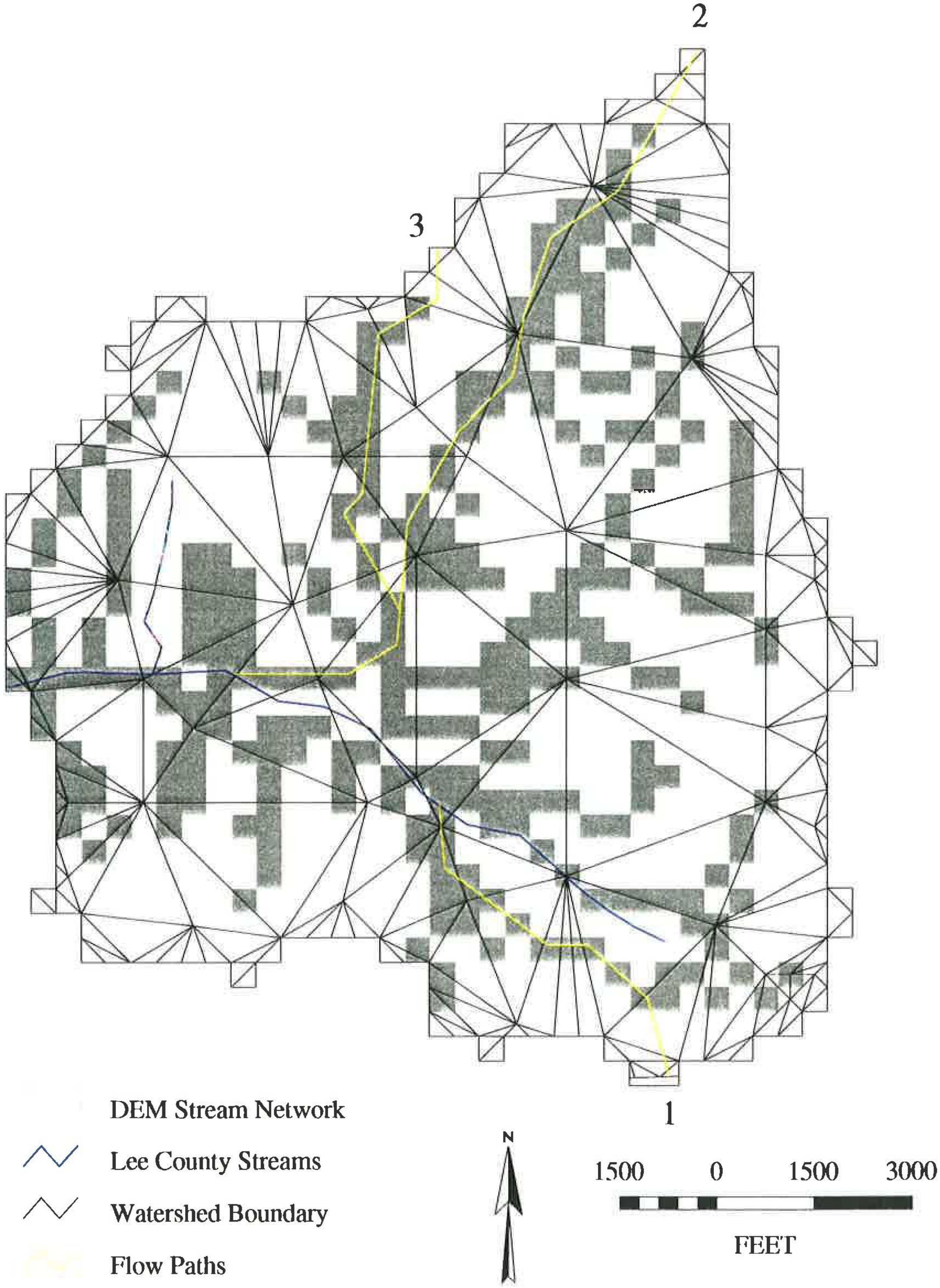

Figure 16. Stream Network for Cells Receiving Flow from 4 Cells or More with the TIN and Lee Streams Overlaid.

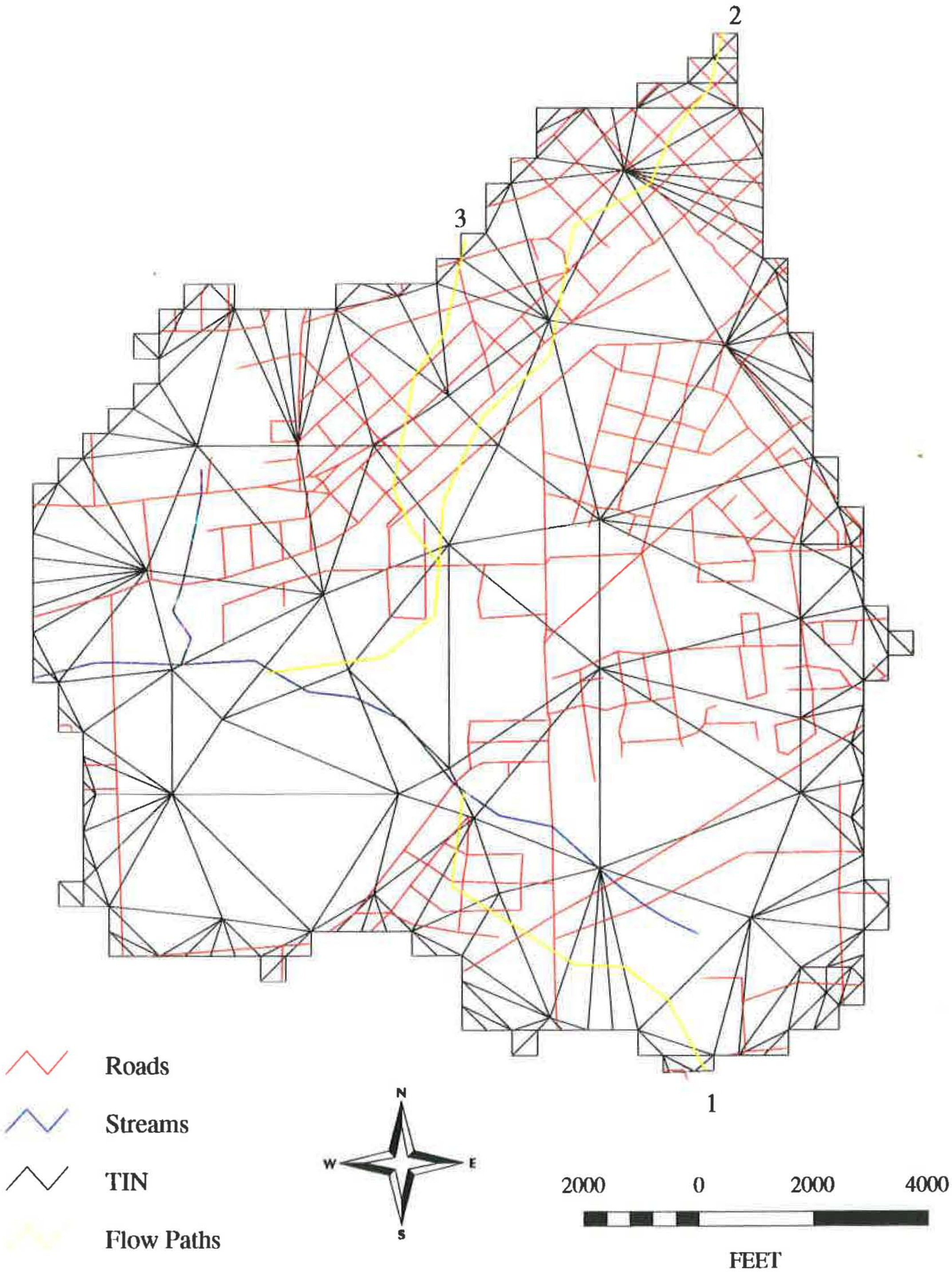

Figure 17. Basin Streams and Roads Combined with the TIN.

Appendix IV lists the overland flow travel time computations for the paths 1, 2, and 3, respectively, from the edge of the watershed to the channel. Because path 1 went through some wooded and residential sections, a Manning's roughness coefficient, n, of 0.6 for sheet flow and the unpaved velocity for shallow flow were applied. Paths 2 and 3 went through residential sections, so an approximate roughness coefficient, n, of 0.1 for sheet flow and the paved velocity for shallow flow was used (SCS, 1986).

Figure 18 shows the stage-discharge relationship calculated by the USGS for Pepperell Creek (ALDOT, 1993). From this chart, the 25 year storm discharge shows a stage of 706.8 feet, and Figure 19 shows that this stage is about an average of 1.9 feet above the berm and into the flood plain. This depth was used for the 50 and 100 year floods because their depths would not vary much. Since the flow in the flood plain would be very slow compared to the channel giving a larger travel time than the channel, the travel time for the flow in the flood plain was used for the travel time calculations.

Figure 19 shows the flood plain cross section with some of the hydrologic characteristics of the channel for Pepperell Creek as recorded in the report by the Alabama DOT (ALDOT, 1993). Assuming constant shape along the length of the flood plain, the area of flow in the flood plain for 1.9 feet above the berm was about 1,180 square feet, and the perimeter was about 620 feet. Manning's roughness coefficients were estimated by the ALDOT to be 0.175 for the flood plain. Appendix III contains the travel times for the flow in the flood plain for paths 1, 2 and 3.

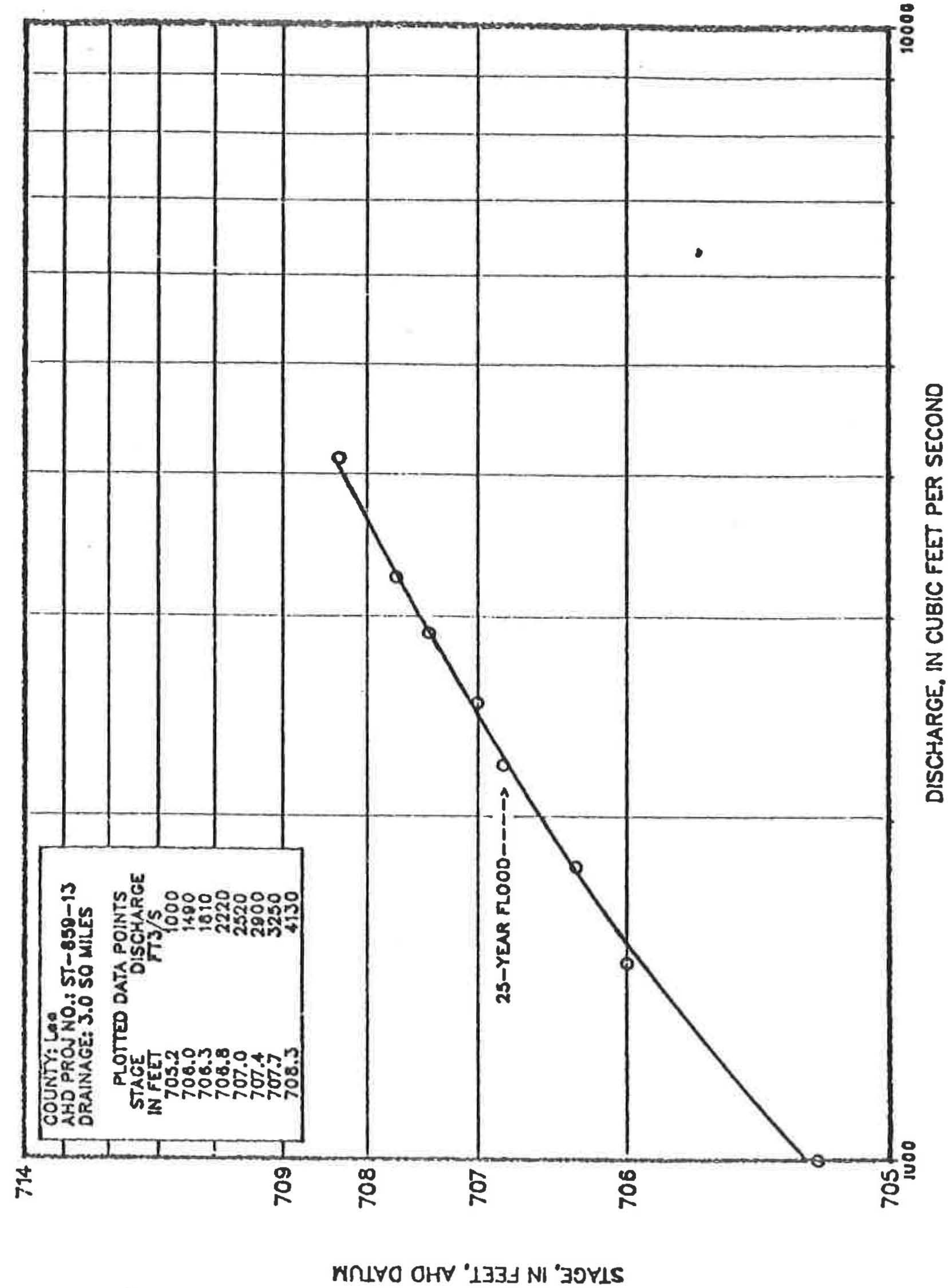

Figure 18. Stage-Discharge Relationship for the Pepperell Creek.

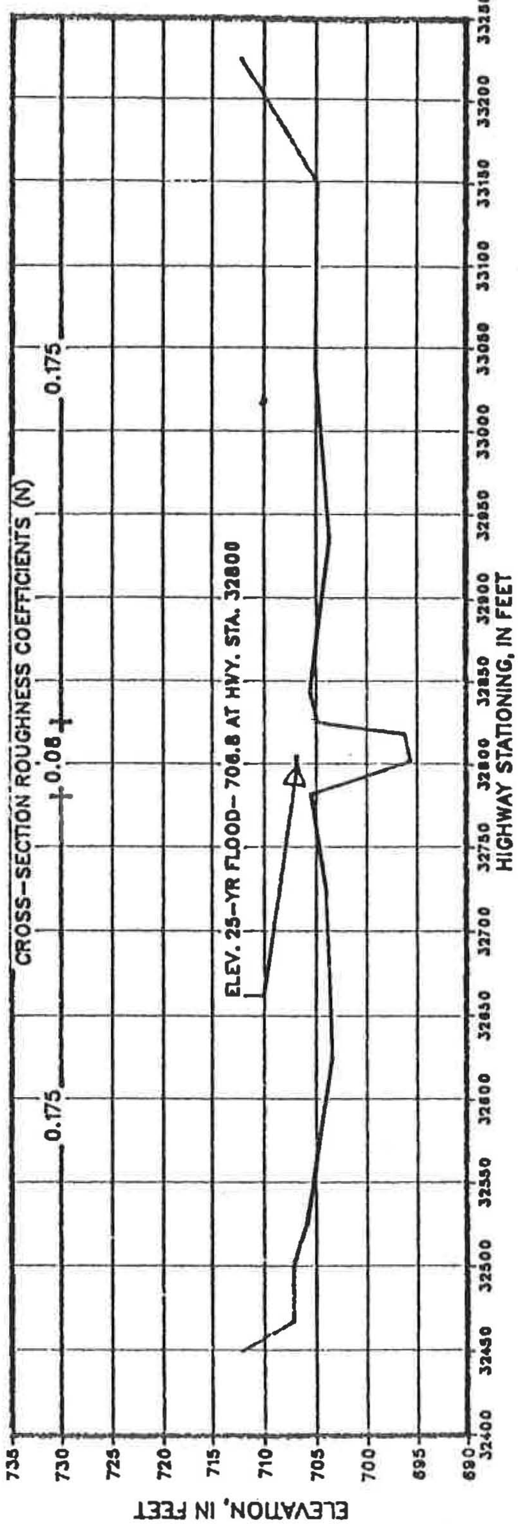

Figure 19. Pepperell Creek Flood Plain Cross Section.

Table 6 summarizes the times of concentration for the SCS graphical peak flow method after adding the travel times for the overland, channel flow and overflow for each path. The time of concentration to be used for the hydrologic analysis was the largest travel time computed which was 3.26 hours (SCS, 1986).

For the SCS Lag Equation, the longest flow path, the curve number and the average watershed slope were needed. Entering these into the equation resulted in a lag time of 2.68 hours for the watershed.

Table 6. Summary of the Travel Times for the Three Paths.

Path	**t_t (hrs)**
1	2.77
2	3.02
3	3.26

Peak Discharges

Arc/Info computed the area, *A*, of the watershed to be 3.24 square miles and the areal average Curve Number for the watershed to be 79.3 for an antecedent moisture condition of II. Using this Curve Number for the watershed, the potential maximum retention after runoff, *S,* in inches, by the SCS empirical equation was 2.61 which gave an initial abstraction, I_a, of 0.52. From the isohyetal precipitation chart in TR-55 (SCS, 1986), the 2 year, 24 hour precipitations for the 25, 50, and 100 year storm interval were identified. Table 7 summarizes this information.

Table 7. Summary of the SCS Parameters for the 25, 50 and 100 Year Storms.

	25 Year Storm	50 Year Storm	100 Year Storm
P (inches)	7.33	8.0	8.8
I_a/P	0.0712	0.0653	0.0593
Q (inches)	4.92	5.54	6.29

The time of concentration for the basin was 3.26 hours from the Velocity Method and 2.68 hours from the Lag Equation. The peak flow rates, q_p, for the 25, 50 and 100 year return interval storms are outlined in Table 8 after using the SCS graphical peak discharge method for a type III rainfall distribution (SCS, 1986).

Table 8. Peak Flow Rates for the 25, 50 and 100 Year Storms.

	T_c (velocity) = 3.26 hrs	**T_c (lag) = 2.68 hrs**
q_u (cfs/mi^2/in)	120	165
25 Year q_p (cfs)	1910	2630
50 Year q_p (cfs)	2160	2960
100 Year q_p (cfs)	2450	3360

Peak Flow Rates Determined Using the DOT Regression Equations

For the traditional hydrologic analysis, the Alabama Department of Transportation estimated the peak flow rate, Q(u), by inputting the watershed area, A, in square miles and impervious area percentage, IA, into a regression equation for urban streams in Alabama. The equations applicable to Opelika for various recurrence intervals and their standard errors are reproduced in Table 9 (Atkins, 1996).

Table 9. Peak Flow Regression Equations for Opelika, Alabama.

Recurrence Interval (years)	Regression Equation for Streams in Urban Area	Standard Error of Estimate (percent)
2	$Q(u)=150A^{0.70}IA^{0.36}$	26
5	$Q(u)=210A^{0.70}IA^{0.39}$	24
10	$Q(u)=266A^{0.69}IA^{0.39}$	24
25	$Q(u)=337A^{0.69}IA^{0.39}$	24
50	$Q(u)=396A^{0.69}IA^{0.38}$	25
100	$Q(u)=444A^{0.69}IA^{0.39}$	25

The regression equations were derived by analysis of peak discharges obtained from calibrated basins for 23 urban stations in Alabama. In applying the regression equations, the area of the watershed was estimated after the ridge lines surrounding the basin were outlined on a USGS 7.5 Minute Topographic Map, and the impervious area of the watershed was measured by gridding the topographic map or aerial photographs (Atkins, 1996).

On a USGS 7.5 Minute Topographic Map, the Alabama Department of Transportation estimated the area of the watershed to be 3.0 square miles with 15% impervious area (IA). Table 10 lists the peak flow rates estimated by inputting these parameters into the regression equations for Alabama for the 25, 50, and 100 year storm intervals. Table 11 contains the peak flow rates using the watershed area computed by Arc/Info.

Table 10. Peak Discharges Estimated Using the Regression Equations and Basin Area of 3.0 Square Miles.

Recurrence Interval (years)	**Peak Discharge (cfs)**
25	2070
50	2360
100	2720

Table 11. Peak Discharges Estimated Using the Regression Equations and Basin Area of 3.24 square miles.

Recurrence Interval (years)	**Peak Discharge (cfs)**
25	2180
50	2490
100	2870

Discussion

Using the GIS to acquire the watershed area, curve number, overland and channel flow lengths, and slopes, the estimated peak flow rates did not deviate much from that estimated by the Alabama DOT who used the regression equations. In fact, the peak flow estimated for this study from the Velocity Method was essentially the same as that predicted by the DOT. However, the maximum peak flow estimated from the lag equation was 30% larger than that from the regression equations which could become significant during large storms. Entering the watershed area of 3.24 square miles determined from Arc/Info into the regression equations, the peak discharges were 5% larger than with 3.0 square miles.

The digital data used in this project had many pros and cons. All of the data came at no cost. The DEM and the Land Use and Land Cover were available from the USGS ftp site. The soils data was available at no cost from the Opelika, Alabama, NRCS office. The Tiger files usually cost $40 per data set, but this set was obtained from the Auburn University Alabama Land Resource Information Center of the Alabama Agricultural Experiment Station for no cost.

The DEM, which created cells that were 75 meters by 90 meters, was very coarse which could over or under estimate areas such as a watershed. The USGS was in the process of digitizing all of their 7.5 minute topographic maps which would have a 30 meter spacing between data points. This finer resolution DEM would mean that each cell was much smaller and would give a closer representation of the actual terrain. The basin boundaries would have less error due to the smaller cell size. The cost of a 7.5 minute data set that had already been digitized was $100 per DEM.

The Land Use and Land Cover data was also of a very small scale. Since the data was compiled in the late 1970's, some of the data would not necessarily be current. As a solution to this, the old data could be updated or a new land cover could be digitized.

The soils data was compiled in the late 1970's also, and not all of counties in Alabama were completed. Even though each 15.44 acre cell contained only the dominant soil type, it was good enough to give the overall characteristics of the soils in the area. The NRCS has plans to digitize new soil type coverages for all of the counties in Alabama with Arc/Info.

The Tiger files are updated periodically, especially after a census. This would compensate for changes in roads and streams. Users who do not want to digitize their own data would need to buy the new data.

V. SUMMARY AND CONCLUSIONS

For this study, a hydrologic analysis to determine the peak flow using Arc/Info was performed on the Pepperell Creek basin in Opelika, Alabama according to the Soil Conservation Service peak flow method in Technical Release 55 (SCS, 1986). The resulting peak flows were compared to an analysis using regression equations of the same watershed conducted by the Alabama Department of Transportation as a part of the plan to reroute US Highway 280 (ALDOT, 1993).

Based on studies of Alabama watersheds, the regression equation approach considered only the watershed area and the impervious area percentage for the peak flow calculations. Any variability in the slopes and the soil and land use and land cover, however, were not taken into account.

The Arc/Info GIS offers an alternative method of determining the peak flow rate for a watershed with the SCS peak flow method in TR-55. Arc/Info can convert the digital data into a format called a coverage which can be easily combined with other Arc/Info coverages. As presented in this study, Arc/Info provides many functions to delineate the watershed, define the watershed curve number, measure lengths and determine the planimetric surface area and slopes. Arc/Info could also consider subbasins within the watershed to determine their contribution to the flow.

Even though the peak flows from the SCS method and regression equations were fairly close, Arc/Info can handle more variability and future changes within the watershed. This additional information could aid the engineer since he would not have to rely on one set of data. Because the peak flow rate would increase due to more pavement, less natural cover, and a drainage system, the engineer could examine these changes in peak flow rates due to this development and ensure that any additional runoff would be properly handled by the drainage system.

REFERENCES

Alabama Department of Transportation. (1993) Project No. ST-859-13. Peperrell Creek-U.S. 280 Relocation Project File, Montgomery, Alabama.

Atkins, J. B. (1996) *Magnitude and Frequency of Floods in Alabama: U.S. Geological Survey Water-Resources Investigation Report 95-4199*, Tuscaloosa, Alabama.

Environmental Protection Agency. (1996) metadata.txt (Available from: ftp://ftp.epa.gov/pub /EPAGIRA/meta/general-metadata.txt), Washington, D. C.

Environmental Systems Research Institute. (1994) Arc/Info Data Management, Redlands, California.

Environmental Systems Research Institute. (1995) "Surface Concept," *Arc/Info Online Documentation*, Redlands, California.

Greene, R.G., and Cruise, J.F. (1995). Urban Watershed Modeling Using Geographic Information System, *Journal of Water Resources Planning and Management*, ASCE Vol. 121, No. 4, July/August, pp 318-325.

Krug, Mark R. (1994) Combining Thematic Databases Into a Single Integrated GIS Database, *Proceedings of the Fourteenth Annual ESRI User Conference*, Environmental Systems Research Institute, Redlands, California.

Ross, M. A., and Tara, P. D. (1993) Integrated Hydrologic Modeling with Geographic Information Systems, *Journal of Water Resources Planning and Management*, ASCE, Vol. 119, No. 2, March/April, pp 129-140.

Spratling, Larry. (1996) Personal Interview, National Resource Conservation Service, Auburn, Alabama.

U.S. Bureau of the Census. (1992) TIGER/Line Files (TM) 1992 Technical Documentation (Available from: http://www.census.gov/geo/www/tiger/content.html), Geography Division, Washington, D. C.

U.S. Geological Survey. (1996) Metadata Readme File (Available from: http://edcwww.usgs.gov/Webglis/glisbin/guide.pl?guide=/glis/hyper/guide/1_250_lulc), Earth Science Information Center, Reston, Virginia.

U.S. Geological Survey. (1993) U.S. Geodata Digital Elevation Models Factsheet, Earth Science Information Center, Reston, Virginia.

U. S. Soil Conservation Service. (1986) Urban Hydrology for Small Watersheds, Technical Release 55.

U. S. Soil Conservation Service. (1981) Soil Survey of Lee County, Alabama, Opelika, Alabama.

Viessman, W. and Lewis, G. (1996) Introduction to Hydrology, 4th Edition, Harper Collins, New York, NY.

Warwick, J. and Haness, S. (1994) Efficacy of Arc/Info GIS Applications to Hydrologic Modeling, *Journal of Water Resources Planning and Management*, ASCE, Vol. 120, No. 3, May/June, pp 366-381.

APPENDIX I

Projection File, dem.prj, Applied to the DEM.

```
input
projection geographic
units ds
zunits meters
datum wgs72
parameters
output
projection utm
zone 16
units meters
zunits meters
datum nas_c
parameters
end
```

APPENDIX II

This appendix lists the information gathered from each set of digital data.

Digital Elevation Model

- Contour Map
- Watershed delineation
- Stream network
- Slopes

Land Use and Land Cover and Hydrologic Soil Types

- Curve Number of the Watershed

TIGER files

- Streams and roads as references for location

APPENDIX III

This appendix lists the GIS steps used on the digital data.

Watershed Delineation: The DEM was converted to the UTM projection and delineated the watershed. Then the DEM was converted to a polygon coverage.

1. DEMLATTICE pcdem pclattice USGS
2. PROJECT GRID pclattice pcgrid dem.prj cubic
3. GRID environment was entered to manipulate the DEM grid.
4. FILL pcgrid pcfill sink
5. flow_dir = flowdirection(pcfill,NORMAL)
6. flowacc = flowaccumulation(flow_dir)
7. strmnet = con(flowacc>100,1)
8. Wshed = watershed(flowdir,selectpoint,(pcfill,x,y))
9. Exit grid
10. GRIDPOLY wshed wshedp

Slope and True Surface Area Calculation of the Watershed: The lattice of elevations was manipulated to determine the slopes and surface area of the watershed.

1.LATTICETIN pcfill ctin

2. TINARC pctin pcarc POINT

3. CREATETIN shedtin # # 1 649212.6 3609812 652660.9 3613915

4.COVER pcarc point SPOT MASS #

5. COVER wshedp POLY SPOT SOFTCLIP

6. END to exit Createtin.

7. TINARC shedtin shedpoly POLY

CN Determination with the Arc Command Menu in Arctools: The watershed soils data and the land use data were extracted from the Lee County Soils data and the Land Use and Land Cover data. Then, they were combined to result in the curve number for the watershed.

1. CLIP pclulcn shedpoly shedlulc POLY

2. CLIP leesoilp shedpoly shedsoil POLY

3. UNION shedsoil shedlulc shedcn POLY

4. JOIN shedcn.pat with pclulcn.lut using LUCODE to give shedcn.pat LINEAR

5. JOIN shedcn.pat with leesoilp.lut using SOILCODE to give shedcn.pat LINEAR

6. ADD ITEM to shedcn.pat: CN3 8 18 F5

7. CALCULATE shedcn.pat CN3 = CN1 + CN2

8. JOIN shedcn.pat with CN3.LUT using CN# to give shedcn.pat

9. ADD ITEM shedcn.pat: CNAVE 8 18 F5

10. CALCULATE CNAVE = CN * AREA / 8385158

11. STATISTICS shedcn.pat cnarea

12. SUM CNAVE

13. LIST CNAREA

APPENDIX IV

This appendix contains the hydrologic analysis.

Table IV-A. Overland Flow Time for Path 1.

Flow Type	S (Percent)	L (ft)	V (fps)	T_t (hrs)
Sheet	2.60	223	---	0.73
Shallow	3.77	75	3.13	0.01
Shallow	3.77	56	3.13	0.00
Shallow	0.47	607	1.11	0.15
Shallow	0.40	866	1.02	0.24
Shallow	0.49	190	1.13	0.05
Shallow	0.83	131	1.47	0.02
Shallow	0.92	236	1.55	0.04
Shallow	1.77	1083	2.15	0.14
Shallow	3.62	548	3.07	0.05
Shallow	3.62	699	3.07	0.06
			Total T_t =	1.50

Table IV-B. Overland Flow Time for Path 2.

Flow Type	S (Percent)	L (ft)	V (fps)	T_t (hrs)
Sheet	2.94	157	---	0.13
Sheet	2.59	131	---	0.11
Shallow	1.90	295	2.8	0.03
Shallow	3.21	446	3.64	0.03
Shallow	2.52	312	3.23	0.03
Shallow	0.14	1929	0.76	0.70
Shallow	0.69	837	1.69	0.14
Shallow	2.67	840	3.32	0.07
Shallow	2.10	1362	2.95	0.13
Shallow	0.08	1230	0.57	0.59
			Total T_t =	1.96

Table IV-C. Overland Flow Time for Path 3.

Flow Type	S (Percent)	L (ft)	V (fps)	T_t (hrs)
Sheet	2.25	292		0.23
Shallow	2.25	374	3.05	0.03
Shallow	0.43	246	1.33	0.05
Shallow	1.04	410	2.07	0.05
Shallow	0.04	843	0.41	0.58
Shallow	1.79	833	2.72	0.09
Shallow	2.42	262	3.16	0.02
Shallow	3.10	988	3.58	0.08
Shallow	0.08	2218	0.57	1.07
			Total T_t =	2.20

Table IV-D. Travel Times in the Flood Plain for Path 1.

Slope (%)	L (ft)	V(fps)	T_t (hrs)
3.04	1703	2.06	0.21
1.79	686	1.58	0.11
1.17	1421	1.28	0.28
0.58	725	0.90	0.20
1.07	656	1.22	0.13
0.82	489	1.07	0.11
2.51	1257	1.87	0.17
3.49	236	2.21	0.03
6.28	92	2.96	0.01
6.12	207	2.92	0.02
		Total T_t =	1.29

Table IV-E. Travel Times in the Flood Plain along Paths 2 and 3.

Slope (%)	L (ft)	V(fps)	T_t (hrs)
1.79	686	1.39	0.11
1.17	1421	1.21	0.28
0.58	725	0.79	0.20
1.07	656	1.07	0.13
0.82	489	0.94	0.11
2.51	1257	1.64	0.17
3.49	236	1.94	0.03
6.28	92	2.60	0.01
6.12	207	2.56	0.02
		Total T_t =	1.06