COMPARISON OF TWO ALGORITHMS FOR REMOVING DEPRESSIONS AND DELINEATING FLOW NETWORKS FROM GRID DIGITAL ELEVATION MODELS

by

Anurag Srivastava

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirement for the degree of

Master of Science
in
Biological Systems Engineering

Dr. Conrad D. Heatwole, Chair
Dr. Mary Leigh Wolfe
Dr. Theo A. Dillaha
Dr. John V. Perumpral, Department Head

February 10, 2000
Blacksburg, VA

COMPARISON OF TWO ALGORITHMS FOR REMOVING DEPRESSIONS AND DELINEATING FLOW NETWORKS FROM GRID DIGITAL ELEVATION MODELS

by

Anurag Srivastava

(ABSTRACT)

Digital elevation models (DEMs) and their derivatives such as slope, flow direction and flow accumulation maps, are used frequently as inputs to hydrologic and nonpoint source modeling. The depressions which are frequently present in DEMs may represent the actual topography, but are often the result of errors. Creating a depression-free surface is commonly required prior to deriving flow direction, flow accumulation, flow network, and watershed boundary maps. The objectives of this study were: 1) characterize the occurrence of depressions in 30m USGS DEMs and assess correlations to watershed topographic characteristics, and 2) compare the performance of two algorithms used to remove depressions and delineate flow networks from DEMs.

Sixty-six watersheds were selected to represent a range of topographic conditions characteristic of the Piedmont and Mountain and Valley regions of Virginia. Analysis was based on USGS 30m DEMs with elevations in integer meters. With few exceptions watersheds fell on single 7.5 minute USGS quadrangle sheets, ranged in size from 450 to 3000 hectares, and had average slopes ranging from 3 to 20 percent. ArcView (3.1) with the Spatial Analyst (1.1) extension was used to summarize characteristics of each watershed including slope, elevation range, elevation standard deviation, curvature, channel slope, and drainage density. TOPAZ (ver 1.2) and ArcView were each used to generate a depression-free surface, flow network and watershed area. Characteristics of the areas 'cut' and 'filled' by the algorithms were compared to topographic characteristics of the watersheds. Blue line streams were digitized from scanned USGS 7.5 minute topographic maps (DRGs) then rasterized at 30 m for analysis of distance from the derived flow networks.

The removal of depressions resulted in changes in elevation values in 0 - 11% of the cells in the watersheds. The percentage of area changed was higher in flatter watersheds. Changed elevation cells resulted in changes in two to three times as many cells in derivative flow direction, flow accumulation and slope grids. Mean fill depth by watershed ranged from 0 to 10 m, with maximum fill depths up to 40 m. In comparison with ArcView, TOPAZ, on average affected 30% fewer cells with less change in elevation. The significance of the difference betweenArcView and TOPAZ decreased as watershed slope increased. A spatial assessment of the modified elevation and slope cells showed that depressions in the DEM's occur predominantly on or along the flow network. Flow networks derived by ArcView and TOPAZ were not significantly different from blue line streams digitized from the USGS quadrangles as indicated by a paired t test. Watershed area delineated by ArcView and TOPAZ was different for almost all watersheds, but was generally within 1%.

Conclusions from this study are: 1) The depressions in 30 mDEMs can make up a significant portion of the area especially for flatter watersheds; 2) The TOPAZ algorithm performed better than ArcView in minimizing the area modified in the process of creating a depressionless surface, particularly in flatter topography; 3) Areas affected by removing depressions are predominantly adjacent to the stream network; 4) For every elevation cell changed, slopes are changed for two to three cells, on average; and 5) ArcView and TOPAZ derived flow networks closely matched the blue line streams.
Keywords: *ArcView, Topaz, depressions, stream network, DEMs, watershed delineation, flow network, channel, filling sinks, DEM error, hydrology.*

In the memory of my uncle
Shri B. P. Srivastava

Acknowledgments

I would like to thank Dr. Heatwole for his knowledge, guidance and patience while serving as my major professor and also for the financial help provided for this study. I would also like to thank my committee members, Dr.Wolfe and Dr. Dillaha for their encouragement and support. Thanks to our department head Dr.Perumpral for the facilities and funding provided by the department for this project.

Thanks also goes to Dr. Yagow and Jennifer Miller for the help extended during various stages of the project. I would also like to thank all friends and fellow graduate student for their encouragement and support.

Finally, I would like to thank all my family members for their faith in me and all their encouragement.

Table of Contents

List of Figures

List of Tables

1. Introduction

Over the last several years there has been a dramatic increase in the use of GIS (Geographic Information System) in the field of hydrology including applications in hydrologic modeling. A GIS can be defined as a computer-assisted system for the capture, storage, retrieval, analysis, and display of spatial data (Clarke, 1986). The basic data required for hydrologic modeling such as elevation, land use, land cover, soil type, channel network, rainfall intensity, and runoff curve numbers (CN) can easily be stored and used for analysis within GIS. With the ability to compute hydrologic parameters required for models, a GIS helps in solving problems which are time consuming, repetitive, and error prone (DeVantier and Feldman, 1993).

Goodchild (1993) described three ways that GIS can assist with the spatial dimension of modeling:

(i) Preprocessing data into a form suitable for analysis (scale, coordinate system, data structure, data model, etc.).

(ii) Direct support for modeling, so the GIS carries out those tasks such as analysis, calibration and prediction itself.

(iii) Post-processing data through reformatting, tabulation, mapping and report generation.

A digital elevation model (DEM), as a representation of the elevation surface, is basic data required for hydrologic and non-point source modeling. Digital elevation models are commonly used in hydrological modeling for delineating drainage pathways and runoff contributing areas. DEMs are also used for computing curvature maps, slope maps, aspect maps, and slope profiles that can be used to prepare shaded relief maps, assist geomorphological studies, or estimate erosion and runoff. With the availability of DEM data, a GIS has been used for determining the drainage pattern in a watershed for nonpoint source modeling (Kao, 1992), and for simulating surface runoff from flash floods (Julien et al., 1995).

A basic problem encountered in the process of delineating a drainage network using a DEM is the presence of depressions, or sinks, in the DEM which pose problems in creating an accurate representation of flow direction and therefore accumulated flow. These depressions must be removed in order to process the DEM for delineating the stream network (Garbrecht and Martz, 1998).

Depressions may be accurate reflections of the terrain, or may be the result of input data errors, interpolation methods, or limitations in DEM resolution. Input data errors, which are often present in the DEM, are generally classified in two categories as random or systematic errors. These errors are introduced in the data set as a result of the DEM generation process.

The removal of depressions causes changes in the DEM data set which may have a significant impact on hydrologic and non-point source modeling parameters. Several algorithms have been developed to address the problem of depressions. Some algorithms assume that depressions result solely from underestimated elevations (Mark, 1988), whereas other algorithms assume that depressions can be the result of both under and overestimation of elevation values (Martz and Garbrecht, 1992). Thus, the choice of an algorithm to remove sinks from the DEM will affect the number of cells which have their elevation changed, and consequently will have an impact on derivative maps, such as slope and flow direction. Errors present in the DEM data and watershed terrain complexity may influence the performance of these algorithms, in terms of minimizing changes required in the elevation surface and in the accuracy of the stream network delineated as compared to the blue line streams. Watershed terrain complexity refers to the surface roughness of the watershed as defined by slope and measures of variation in elevation.

Many researchers have tried to address the problem of depressions. Jenson and Domingue (1988) and Martz and Garbrecht (1992) developed algorithms which treat depressions solely as errors in the DEM data set while Martz and De Jong (1988) treat depressions as features which reflect the terrain. One approach adopted by O'Callaghan and Mark (1984) was to use smoothing and filtering of the DEM data. While this method effectively reduced the number of sinks, it resulted in loss of information content of the DEM data. Another approach is to fill the depressions by increasing the values of cells in each depression by the value of the cell with the lowest value on the depression's boundary (Marks et al., 1984; Band, 1986; Jenson and Trautwein, 1987).

The ArcView[*] GIS (Version 3.1) with Spatial Analyst[*] extension (Version 1.1) incorporates the Jenson and Domingue (1988) algorithm which is based on the concept of "filling sinks" by increasing the elevation value to remove depressions. This approach assumes that depressions and flat areas are the results of underestimation of elevations. The ArcView Spatial Analyst extension provides the raster (grid) GIS representation and

[*]ArcView and Spatial Analyst are trademarks of ESRI, Inc., Redlands, CA

analysis tools. The Hydro1.1 extension for Spatial Analyst automates a number of raster-based analysis functions that assist in extracting watershed characteristics and hydrologic parameters based on an input DEM. These functions include those to derive stream networks from a DEM.

A recent approach for treatment of depressions and flat areas by Martz and Garbrecht (1998) is incorporated in the TOPAZ landscape analysis tool (TOPAZ User Manual, 1997). The algorithm used by TOPAZ assumes that depressions and flat areas are the results of both underestimation and overestimation of elevation. The algorithm simulates breaching of the outlet of closed depressions to reduce the elevation of those areas which were expected to have been the result of elevation overestimates. Any depression left after this is considered to be a sink-depression and is treated in the conventional manner by raising the elevations within the depression.

Objectives

The overall objective of this study is to determine if there is a significant difference in the outcomes of two algorithms used to remove sinks in a DEM. One algorithm (ArcView Spatial Analyst Ver 1.1) uses only fills, and the other (TOPAZ, Ver. 1.20) uses both cuts and fills to remove depressions from the DEM.

Specific objectives of the study were:

1. Characterize the frequency of occurrence and spatial distribution of “sinks” in USGS 30m DEMs for a range of small watersheds in Virginia that represent different physiographic regions and topographic characteristics.
2. Determine if there are watershed topographic characteristics that are correlated to the percentage of watershed area modified by the two algorithms in the process of removing depressions.
3. Determine which algorithm generates a flow network that best matches the ‘blue line stream’ from the USGS 7.5 min quadrangle maps.

2. Literature review

2.1 Introduction

Geomorphologists, hydrologists, engineers and specialists of other discipline related with earth sciences are often interested in deriving drainage networks. Using paper maps to measure spatially defined hydrologic variables is tedious, time consuming, error prone and cumbersome. However, much of this desired information can be extracted with the help of GIS because of its capability to integrate, visualize, and to derive spatial and non-spatial data (McDonnell, 1996).

Before the advent of computers, drainage network derivation took considerable effort and time to identify the drainage net and to measure its basic properties. The process involved delimiting the drainage network on maps or aerial photographs and then measuring properties such as link lengths and junction angles (McDonnell, 1996).

With the introduction of digitizers, the time and work required for the measurement phase could be reduced. The drainage network can be traced and coordinate encoded after the identification of the drainage net. Though this is an improvement over the previous method, still the digitization process is tedious, and the channel network needs to be identified.

Over the last two decades there has been growing interest in the use of digital elevation data in geomorphology and hydrology, specifically including the analysis of channel networks. O’Callaghan and Mark (1984) define a digital elevation model (DEM) as any numeric or digital representation of the elevations of all or part of a planetary surface. Regular square grids are the most commonly used data structure for DEMs. Grid elevations are available as a matrix of points equally spaced in two orthogonal directions. Triangulated Irregular Networks (TIN) and contour based grids have also been used for hydrological analysis.

A TIN is a terrain model that uses a sheet of connected triangular facets based on a Delaunay triangulation of irregularly spaced nodes or observation points. O'Loughlin (1986) and Moore, et al. (1988) used contour-based DEMs with some success. Contour based-DEMs offer the advantage of dividing the catchment into natural units related to water flow but require more computation (Moore et al., 1988).

Automated evaluation greatly simplifies and expedites the task of data preparation and organization and a number of algorithms have been developed for drainage delineation. Computer programs provide tools for rapid parameterization of drainage network and subcatchment properties from available DEMs for subsequent use in hydrologic surface runoff models. This is obtained by the direct measurement of network, subcatchment, and basin properties from the DEM.

The raster map of the drainage network and subwatersheds provides an interface to other data, which are available for the drainage basin. It can be registered to other data layers and can also be used to extract data for individual subwatersheds and network links using a variety of available GIS (Martz and Garbrecht, 1992).

2.2 Generation of DEMs from different sources

DEMs can be generated in several ways. Two of the most common techniques are interpolation from contour or point data and direct generation from stereo photos through image correlation devices.

2.2.1 Image correlation devices

Automated image correlation devices such as Defence Mapping's UNAMACE system and commercially available Gestalt Photomapper (GPM II) are used as an important source for digital elevation data (Elassal, 1978). Aerial photographs are first scanned at a pixel size of 182 μm at the scale of the photographs (Elassal, 1978). Elevations are calculated from the parallax and elevations are computed at every pixel. These local model grids are then resampled and a more smoothed grid with axes aligned generally with universal transverse mercator (UTM) are produced (Allam, 1978). United States Geological Survey (USGS) generally distributes DEMs resulting from GPM II method with grid spacing of 30 m and elevations are reported integers in meters. The data for forested area obtained from GPM II generally contains many sinks and it was observed that these were mainly located in relatively flat areas, mainly on the valley floors and on undissected plateau remnants (Allam, 1978).

Visual inspection of the analytical hill shading images of DEMs for forested areas revealed that elevations were generally those of a tree-top surface instead of the ground surface. The features such as road, powerline, and pipeline cuts were visible as grooves in the surface. These errors result in depressions even though they may be correct representations of the surface sensed by the GPM II (O'Callaghan and Mark, 1984).

2.2.2 Interpolation from contours

Contour maps are one of the forms representing elevation data. Elevation values are interpolated from irregularly spaced points to grid points (Briggs, 1981). Interpolated surfaces are generally smooth when the horizontal contour spacing is greater than the grid spacing but without elevation control, depressions are easily produced even when the source contour data does not have depressions. This problem necessitates the process of sink detection and their removal from elevation data derived from digitized contours (O'Callaghan and Mark, 1984).

2.3 Errors in DEM dataset

2.3.1 Accuracy

The accuracy of a DEM is dependent upon the level of detail of the source and the grid spacing used to sample that source. Scale of the source materials is the main factor which affects the level of detail of the source. The level of content that may be extracted from a given source during digitization is dependent on proper selection of grid spacing.

The DEM is visually inspected for the completeness on a DEM view and edit system for the purpose of performing final quality control and if necessary edits of the DEM. The physical format of each digital elevation model is validated for content completeness and logical consistency during production quality control and prior to archiving in the National Digital Cartographic Data Base.

2.3.2 Horizontal positional accuracy

The horizontal accuracy of the DEM is expressed as an estimated root mean square error (RMSE). The estimate of the RMSE is based upon horizontal accuracy tests of the DEM source materials which are selected as equal to or less than intended horizontal RMSE error of the DEM (USGS, 1998).

2.3.3 Vertical positional accuracy

The vertical positional accuracy incorporates both systematic and random errors introduced during production of the data. Linear interpolation of elevations are compared with the known elevation to determine the accuracy. Well distributed test points such as aerotriangulated test points, spot elevations, or points on contours from existing source maps with appropriate contour interval, which are representative of the terrain are selected. At least twenty-eight points composed of twenty interior points and eight edge points are needed per DEM to compute the RMSE (USGS, 1998).

The three types of vertical error in the DEM are blunder, systematic and random errors. These errors can be reduced in magnitude but they cannot be removed completely. Blunders occur in major proportions and can be easily identified and removed during editing. Errors following some fixed patterns are termed systematic errors and are the result of data collection procedures and systems such as misinterpretation of terrain surface due to trees, buildings and shadows, fictitious ridges, tops or benches. Random errors result from either unknown or accidental causes.

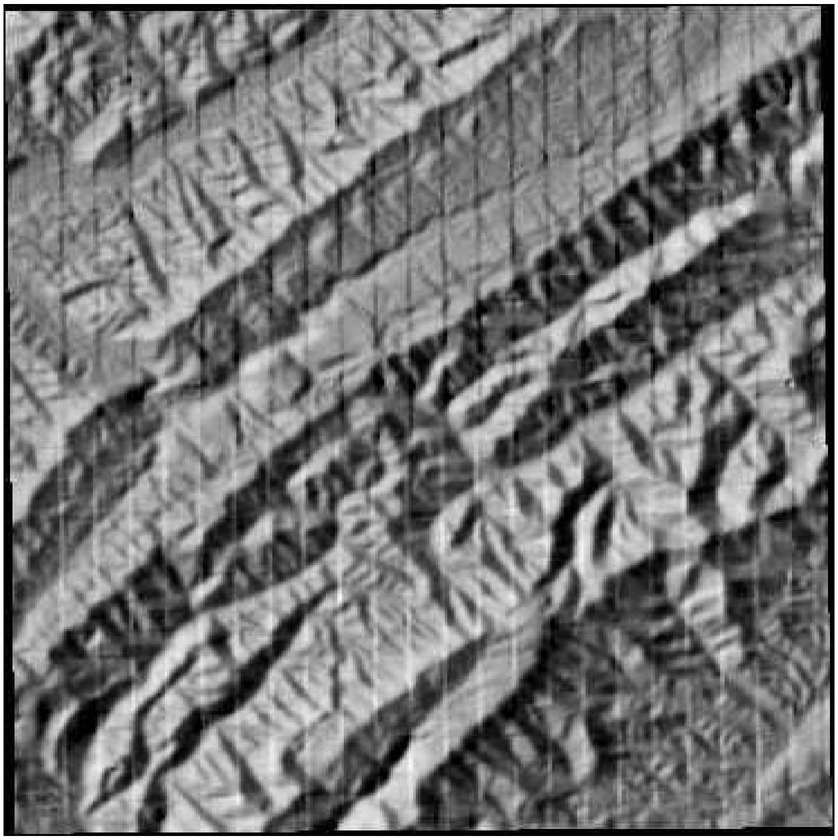

Figure 2.1. The vertical stripes in this hillshaded representation of a DEM reveal systematic error in the elevation data (Rucker Gap, Virginia, 7.5min USGS DEM).

Sometimes as a result of photogrammetric procedures, systematic and random errors are induced in the DEMs. The systematic errors are sometime evident in the form of stripping across the DEM dataset (Figure 2.1). Techniques such as filtering or smoothing are suggested by researchers to rectify these shortcomings in the dataset (Giles and Franklin, 1996; Brown and Bara, 1994).

2.4 Watershed terrain

The slope and topographic variation of a watershed are factors that may affect the drainage network derived from the DEM. The most common measures used to describe the variation in the relief of watershed is variance or standard error and standard deviation (Mark 1975). Other measures used are slope and curvature, the first and second derivatives of the elevation surface respectively (Brown and Bara, 1994).

Watershed slope and channel slopes are also used as measures to quantify the variation between watersheds. Watershed slope is simply the function of watershed length and the elevation difference between top and bottom of the watershed channel length whereas the channel slope is dependent on channel length. Channel length is defined by selecting a certain threshold and then channel slope is computed by dividing it by the elevation difference between the top and outlet of the channel.

2.5 Algorithms for drainage network identification

A channel is defined as the lines along which fluvial processes dominate over slope processes. Using this concept of dominance of fluvial process over slope process leads to two basic approaches in defining flow networks, those based on local surface concavity, and those based on flow accumulation. In addition, Palacios-Velez and Cuevas-Renaud (1986) proposed an algorithm based on geometrical observations that paths of steepest slope in a downstream direction from a point to another would lead back to the starting point then the path belongs to a pure streamline, otherwise it represents a river course. The method had limited use because of its sensitivity to low resolution matrices and DEM errors.

2.5.1 Local surface concavity

Local surface concavity is used to define the flow network based on the fact that any part of a topographic surface which is locally concave-upward will be a place where surface runoff will tend to be concentrated. Thus, it can be inferred that the areas of the DEM which are concave-upwards represent channels. A small window can be moved over the DEM to identify upwardly concave areas. Cells which are at the bottom of these concave-upward areas, are assigned to be part of the stream or channel network (Peucker and Douglas 1975).

An algorithm to identify areas which are upward concave was developed by Douglas (1986). This algorithm simply flags the highest point among each square of four adjacent points in the grid. When this flagging is over, all those points which are not flagged are drainage networks.

Band (1986) used this algorithm for the first stage of his topographic partition of watersheds to delineate stream channels. One of the most serious shortcomings of this method was that it generates discontinuous network segments which must be connected afterwards, requiring intensive postprocessing. For low relief and topographically complex landscapes this problem can become more acute and can prohibit the usefulness of the technique (O'Callaghan and Mark, 1984).

2.5.2 Flow accumulation

For the flow accumulation approach, O'Callaghan and Mark (1984) and Mark (1988) identified the following procedure for delineating channels:

1. Removal of sinks and calculation of drainage direction matrix.
2. Calculation of the accumulated area matrix.
3. Selection of accumulated area threshold to define channels as pixels exceeding the threshold.

The flow accumulation approach is based on the fact that the drainage network represents those points at which runoff is sufficiently concentrated that fluvial processes dominate over slope processes. The delineated drainage network is based on the flow accumulated across the landscape. This overland flow accumulation approach is widely used because of its ability to generate a fully connected network and it also represents basic hydrologic process largely responsible for initiating and maintaining stream channel. According to Tribe (1992), the method generates drainage networks that would result from a uniform rain over a barren and impermeable landscape. This approach causes the total amount of water summed in one grid cell during the whole draining process to be proportional to the size of drainage basin. Since there is no time concentration, water is assumed to drain at the same speed to all lower neighbors during the time unit, independently of slope and soil conditions (Tribe, 1992).

O’Callaghan and Mark (1984) were the first to use this approach. They suggested defining channels on a DEM as all points with accumulated area above some threshold. Since this approach is based on physical conditions related to processes, it holds considerable appeal to the geomorphologist. The method involves identifying the steepest downslope overland flow path between each cell of a raster DEM and its neighbors, and accumulating catchment area downslope along the flow paths connecting adjacent cells. After the assignment of flow direction on the elevation data set, the data sets are used to create the flow accumulation data set. In the flow accumulation data set, each cell is assigned a value equal to the number of cells that flow to it (O’Callaghan and Mark, 1984).

Cells having a flow accumulation value of zero (to which no other cells flow) generally correspond to the ridges. A threshold is selected to delineate the drainage network and all cells which are greater than the threshold value are classified as part of the drainage network. This approach is simple and reliable since it depends on accumulated flow to define the flow paths. The pattern formed by selecting cells with values higher than some threshold value delineates a fully connected drainage network since all cells in a depressionless DEM have an upslope path to the watershed boundary. As the threshold value is increased, the density of the drainage networks decreases.

The threshold used to delineate the network may vary from one watershed to another since it is dependent on factors such as land use and soil type. Some minimum drainage area is required to initiate a channel and any area smaller than the threshold value may not produce enough runoff to form and maintain a channel.

Chorowicz et al. (1992) generated a very dense drainage network which was an outcome most likely due to selection of inaccurate threshold value (Garbrecht and Martz, 1993). Arbitrary methods of threshold value selection without considering factors such as local terrain slope, soil properties, geology, infiltration capacity, surface cover and climatic conditions can lead to erroneous networks (Garbrecht and Martz, 1993).

2.6 Problems in drainage delineation

The presence of sinks and depression pose problems in the drainage delineation process. Since depressions in the DEM create obstacles during the calculation of flow direction, it is a common practice to fill all the sinks to the level of the next lowest cell of the perimeter. Chorovicz et al. (1992) pointed out that sinks might represent actual characteristics of the landscape. He suggests an approach, which combines both geometrical and hydrological concepts for flow network extraction. Mark (1988) suggested the use of digitized stream channels for the removal of sinks or use of a local flooding procedure where sinks are made to drain towards the point at which water would overflow from the sink.

Mark (1988) and Jenson and Domingue (1988) defined depressions as areas surrounded by neighbors having higher elevation, and introduced preprocessing procedures to remove them before the application of automatic algorithms to define the flow network. Jenson and Domingue (1988) used the approach of filling the depression as the first step and then modifying flow directions in all flat areas so that the flow is directed from each inflow cells to the nearest outflow cell on the perimeter of the sink.

Various researchers have followed different approaches to treat sinks. Martz and De Jong (1988) assumed all sinks to be real topographic features which represent ponds or reservoirs whereas Jenson and Domingue (1988) follow the assumption that sinks are primarily data errors or artifacts. While the Martz and De Jong (1988) algorithm works well for low relief landscapes with a horizontal resolution of 15 m and a vertical resolution of 0.1 m, it was observed that the method did not give satisfactory results for relatively lower resolution DEM (Martz and Garbrecht, 1992). Martz and De Jong (1988) filled the depressions after accumulating catchment areas along flow paths derived from the DEM. Catchment area was modified after filling depressions to simulate overflowing at the lowest point of the perimeter.

2.7 Treatment of depressions in DEMs

Each cell of the depressionless DEM is part of the cells leading to the edge of the data set. A path of cells, which leads to edge, is defined as a structure composed of cells that are horizontally, vertically or diagonally adjacent in the raster since they have eight-way connectedness. Depressions present in DEMs hinder flow routing or flow connectivity. A depressionless DEM is commonly generated using one of two approaches: a) filling all depressions and b) a combined use of breaching (reducing the elevation of cells that form an obstruction), along with filling depressions.

2.7.1 Filling depressions

Mark et al. (1984) developed a depression filling procedure in which depression cells are raised to the elevation of the lowest elevation neighbor and are encoded as a flat area. One assumption in this approach is that depressions are the result of underestimation of elevation values.

Morris and Heerdegen (1988) found that the procedure of treating sinks as genuine features yielded more realistic flow directions. The approach they used is similar to the procedure of filling sinks. It works even if, while unblocking one sink, it encounters another of higher elevation. The procedure allows the water to flow in sinks and the water level is raised to the level of the sink’s lowest neighbor. The neighboring points of this location are evaluated to determine if there are any points at the same or lower elevation, the drainage directions are then used to determine the possibility for water to flow to a level lower than that of the sink. If no outlet is found, the water level is raised to that of next lowest point and the procedure is repeated. In Figure 2.2(a) drainage directions along steepest gradients are indicated by arrows, and point A is a sink. The water level is raised to 49 to incorporate point B, but no outlet is found. The level is then raised to 52 to incorporate point C and an outlet via point D to point E which is lower than the original sink. The final network with the altered directions indicated by broken lines is shown in Figure 2.2(b).

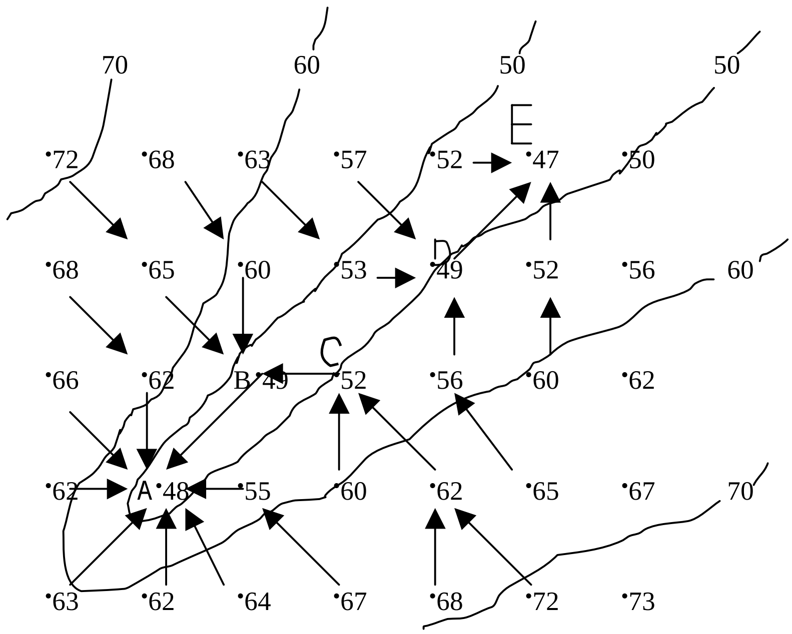

Figure 2.2(a)

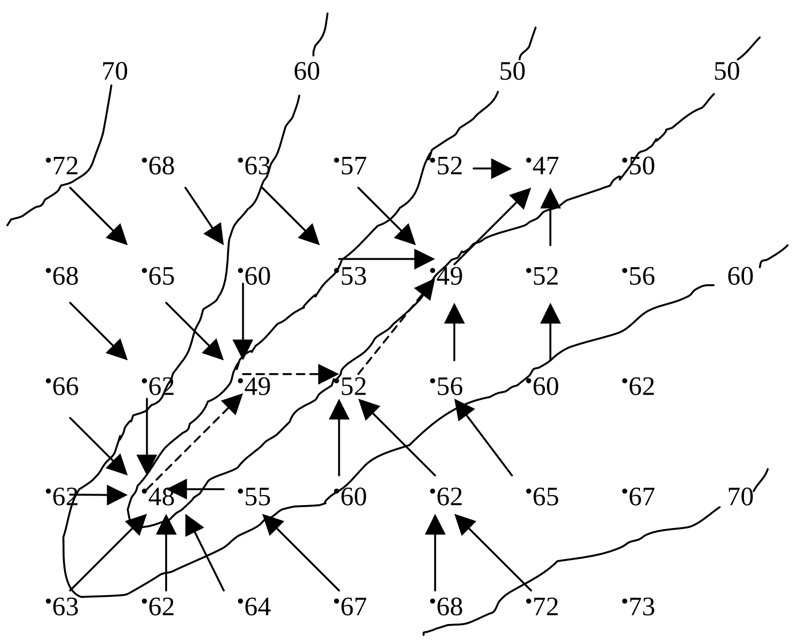

Figure 2.2(b)

Figure 2.2. Procedure for unblocking a sink: (a) initial condition; (b) the final network (Morris and Heerdegen, 1988).

2.7.2 Combined cuts and fills

TOPAZ has the ability to differentiate between two types of depressions: sink-depressions and impoundment-depressions. When a group of raster cells are at a lower elevation as compared to their surroundings then it is termed a sink-depression. Impoundment-depressions are the result of a narrow band of raster cells of higher elevation across drainage paths, similar to an obstruction or dam across a stream. TOPAZ treats the impoundment-depressions by lowering the selected DEM elevation values to simulate breaching of the obstruction or dam across the drainage path. This process reduces the size of the impoundment-depression.

Initially, contributing area is defined for each sink-depressions and then the cells within the contributing areas are evaluated to determine if they are the potential outlets. TOPAZ then evaluates these potential outlets of the sink-depressions to determine whether reduction of elevation value in one or two cells at the outlet will effectively simulate breaching of the outlet. Potential outlets are those cells within the contributing area which are adjacent to a cell outside the contributing area and also at higher elevation than a cell outside the contributing area (Figure 2.3).

All cells which are potential outlets are also potential breaching sites. When more than one breaching site is encountered, then the one with the steepest slope to a cell outside the contributing area is selected. The elevation of cell at the breaching site is reduced to the elevation of the cell outside the contributing area to which it flows. For a two-cell breaching length, the same procedure of reduction of the next cell along the most direct path to the inside of the contributing area is repeated.

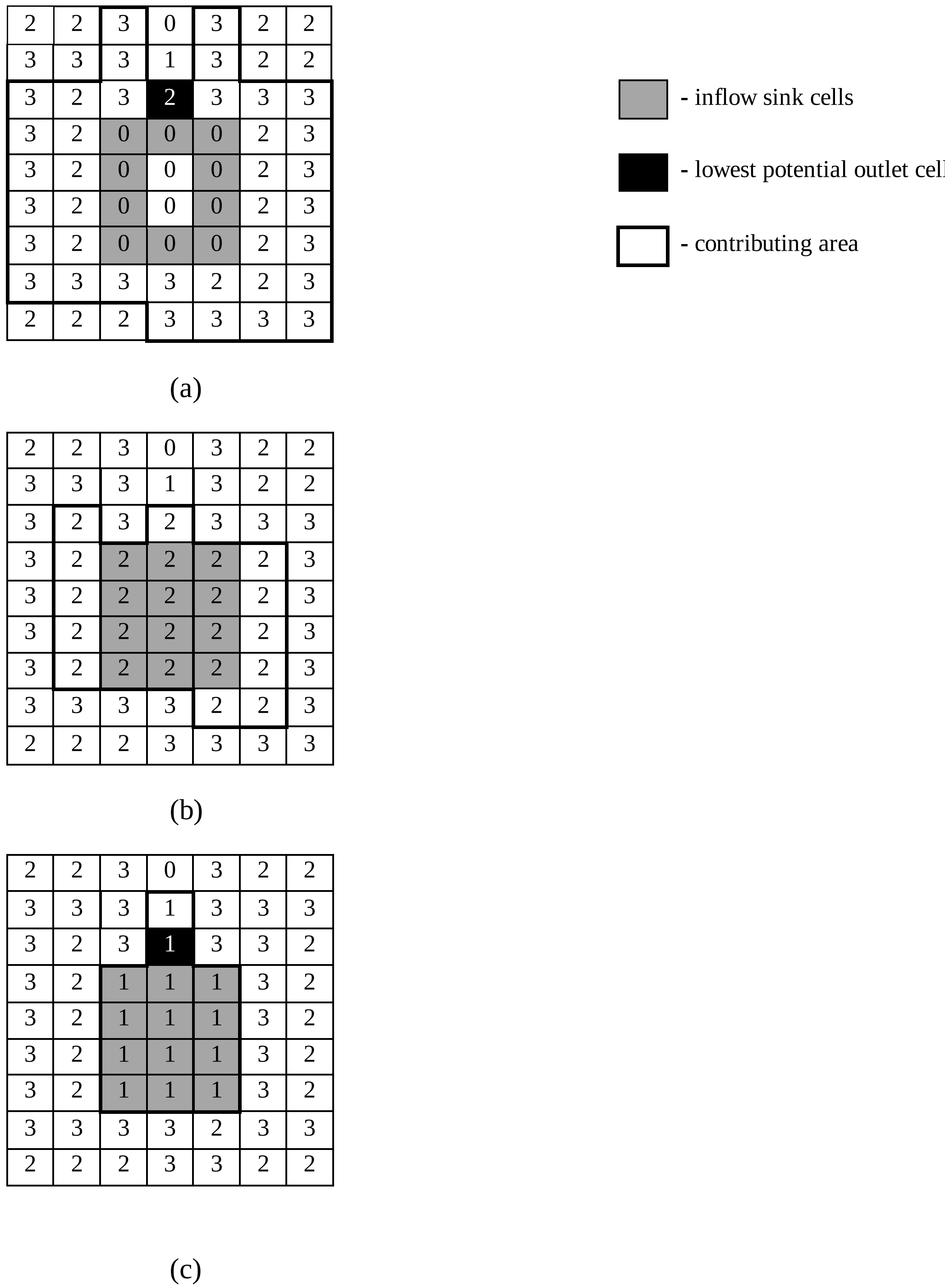

Figure 2.3. Example of depression filling approaches: (a) the original elevation grid; (b) conventional depression filling; (c) breaching and filling.

This procedure is effective and successfully breaches the obstruction responsible for the closed depression. Even after breaching there is the chance that a closed depression may still exist, though reduced in spatial extent. Therefore, the next step involves raising the elevation of all cells that are both inside the contributing area and lower than the selected outlet, to the elevation of the outlet. After this step all the closed depressions are filled resulting in a continuous flat surface at the elevation of the outlet.

Breaching is restricted to one or two cells assuming that major flow obstructions due to elevation overestimates are of limited spatial extent. The algorithm does not differentiate between overestimation and underestimation errors in a DEM, instead it reduces the elevation values of those cells, which it reasonably expects to have resulted from overestimation errors.

2.8 Assignment of flow direction

In the present study, TOPAZ and ArcView were used to assign flow directions. The difference lies in the approach these two algorithms adopt in treating flat areas. Elevation values are altered under the assumption that flow directions are oriented both towards the lower elevation and away from higher elevations (TOPAZ), while in ArcView, flow directions across flat areas are solely affected by adjacent cells of lower elevation.

2.8.1 ArcView

Once the depressions are treated, flow directions are assigned to each cell. The flow direction of the cell is the direction in which water will flow out of the cell. It is encoded to correspond to the orientation of one of the eight cells that surrounds the center cell z, using the following integer values:

64	128	1
32	z	2
16	8	4

If cell z flows towards the right, it will be assigned the flow direction of 2 or if cell z flows vertically upwards then it will be assigned the flow value of 128.

Jenson and Domingue (1988) describe the computation and assignment of flow direction as follows:

1. A cell is assigned a negative value when all the eight neighboring cells have higher elevations. This condition represents a single cell depression. Single cell depressions get eliminated by the depression filling procedure, but are included in the flow direction procedure for completeness.
2. The distance weighted drop in elevation is computed for a cell's eight neighbors by subtracting the neighbor's elevation from the center cell's elevation and dividing by the distance from the center cell, $\sqrt{2}$ for a diagonal cell and one for a non-diagonal cell.
3. Drop values are examined to determine the neighbors with the largest drop and then one of the following operations is performed, also illustrated in Figure 2.4.

    a). A negative flow direction is assigned when the largest drop is less than zero. (There is no such situation in a depressionless DEM).

    b). When the largest drop is greater than or equal to zero and takes place at only one neighbor then the flow direction is towards that neighbor.

    c). When the largest drop is greater than zero and occurs at more than one neighbor then the flow direction is assigned using a table look-up. For example, if there are three adjacent cells with equal drops then the center cell will be selected for the flow direction; while, if two cells on opposite sides have equal drops then one is selected arbitrarily.

d). When the largest drop is equal to zero and occurs at more than one neighbor then the sum of the direction values to those neighbors is assigned to the center cell. The cell is part of flat area and the flow directions of these cells are resolved in an iterative process.

	Elevation Grid			Weighted Drops Grid			Flow Direction of center cell		
Condition a	102	104	102	-10	-12	-10			
	97	92	94	-5		-2		-4	
	98	96	93	-6	-4	-1			
Condition b	102	94	91	-10	-2	1			
	97	92	92	-5		0		1	
	98	96	92	-6	-4	0			
Condition c	91	94	91	1	-2	1			
	92	92	92	0		0		1	
	92	96	92	0	-4	0			
Condition d	102	94	92	-10	-2	0	Temporarily		
	97	92	92	-5		0	encoded as		
	98	96	92	-6	-4	0	1 + 2 + 4 = 7		

Figure 2.4. Assignment of flow direction (Jenson and Domingue, 1988).

2.8.2 TOPAZ

2.8.2.1 Assignment of flow direction over flat areas

TOPAZ defines the drainage direction in flat areas by imposing two gradients on the surface; one towards lower elevations and another away from higher elevations. This approach produces a convergent drainage pattern. It is assumed that gradients can be computed independent of each other and that their linear addition is sufficient in assigning the drainage pattern over the flat surface.

Modifications in elevation values are made by successive addition of an elevation increment to the initial cell values. This incremental value is very small relative to cell elevation and vertical DEM resolution and does not significantly alter the elevation values in the digital elevation model, at the same time it is sufficient to help in assignment of flow direction over the flat surface. At least one cell of lower elevation is present at the edge of the flat area so that downslope drainage off the flat surface is possible. The cells, which are outside the flat surfaces, are not altered since the elevation increment is applied only to DEM cells within flat surfaces (Martz and Garbrecht, 1992).

The elevation values are incremented by 0.00002 units of the vertical DEM resolution. For example, if the vertical resolution is 2 m, then one elevation increment is 40 μm. Thus, such a small increment produces an insignificant change in actual cell elevation, though it is sufficient to numerically define a flow direction. The value of 0.00002 is chosen as incremental value rather than the incremental value of 0.00001 units, because the latter increment is required for the treatment of exceptional situations.

In reality, input DEM values are handled as 32-bit integers to prevent any rounding errors related with using such small elevation increments. Input DEM values are required to be in units of meters and limited to the range 1.0 to 9999.0. These values are read to a precision of 1 decimeter and multiplied by 100000. Therefore, an input elevation value of 32.33 m would be stored internally as 3230000 and the basic elevation increment would be 2.

2.8.2.2 Gradient towards lower terrain

To impose the gradient towards the lower terrain, the elevation increment is added to the elevations of all cells on the flat surface which are not adjacent to a cell with a lower elevation. This procedure is repeated until all the cells have a downstream gradient (Figure 2.5a). Using the above approach, a flow gradient towards lower terrain is constructed as a backward growth from the outlets into the flat surface and at the same time all the boundary conditions imposed by the higher and lower terrain surrounding the flat surfaces are satisfied. Figure 2.5b shows the resulting drainage from the imposed gradient along with flow directions which are represented with arrow sizes proportional to the upstream drainage area. The method sometimes results in parallel flow patterns which are commonly observed in the elevation data sets where this approach is used.

2	2	3	0	3	2	2
3	3	3	**1**	3	3	3
3	2	3	**1.02**	3	3	2
3	2	**1.04**	**1.04**	**1.04**	3	2
3	2	**1.06**	**1.06**	**1.06**	3	2
3	2	**1.08**	**1.08**	**1.08**	3	2
3	2	**1.10**	**1.10**	**1.10**	3	2
3	3	3	3	2	3	3
2	2	2	3	3	2	2

(a)

2	2	3	0	3	2	2
3	3	3	↑	3	3	3
3	2	3	↑	3	3	2
3	2	↗	↑	↖	3	2
3	2	↑	↑	↑	3	2
3	2	↑	↑	↑	3	2
3	2	↑	↑	↑	3	2
3	3	3	3	2	3	3
2	2	2	3	3	2	2

(b)

Figure 2.5. Example of imposition of a gradient on a flat surface (using a 0.02 increment): (a) addition of cumulative elevation increments; (b) the resulting drainage.

2.8.2.3 Gradient away from higher terrain

In this method, the gradient away from higher terrain is imposed by first incrementing the elevation values of all those cells in the flat area which are adjacent to higher terrain and have no adjacent cell at a lower elevation. A downward gradient away from higher terrain results from the imposed increment for all cells immediately adjacent to higher terrain. In subsequent passes: a) those cells which have been incremented previously are incremented again; and b) those cells that have not been incremented earlier and that are adjacent to cells which have been incremented and that are not adjacent to a cell lower than the flat surface, are incremented.

This procedure is repeated until all cells in the flat surface have been incremented (Figure 2.6a). This approach develops a gradient away from higher terrain by inward growth from the edges of the flat surface that are adjacent to higher terrain. Figure 2.6(b) exhibits the drainage pattern resulting from this gradient. Flow directions are represented with arrow and sinks as circles and most flow direction point away from higher terrain because flat areas are surrounded by higher elevations.

2	2	3	0	3	2	2
3	3	3	**1**	3	3	3
3	2	3	**1.04**	3	3	2
3	2	**1.04**	**1.04**	**1.04**	3	2
3	2	**1.04**	**1.02**	**1.04**	3	2
3	2	**1.04**	**1.02**	**1.04**	3	2
3	2	**1.04**	**1.04**	**1.04**	3	2
3	3	3	3	2	3	3
2	2	2	3	3	2	2

(a)

2	2	3	0	3	2	2
3	3	3	↑	3	3	3
3	2	3	↑	3	3	2
3	2	↘	↓	↙	3	2
3	2	→	**O**	←	3	2
3	2	→	**O**	←	3	2
3	2	↗	↑	↖	3	2
3	3	3	3	2	3	3
2	2	2	3	3	2	2

(b)

Figure 2.6. Illustration of imposition of gradient away from higher elevations: (a) addition of cumulative elevation increments; (b) the resulting drainage.

2.8.2.4 Combined gradient and final drainage pattern

In the final step, the increments resulting from the previous two steps are added linearly for every cell. This total increment is then added to the initial elevation of each cell producing a surface that is no longer flat and comprises a gradient away from higher terrain and a gradient towards lower terrain. Since the increments are small the changes in cell elevation are insignificant.

In some conditions, certain adjustments are required. Sometimes the imposed gradients towards the lower terrain and away from higher terrain are in exactly opposite directions and of the same magnitude. Since this cancels out each other, it leaves a cell without drainage direction. To deal with these situations, an additional gradient from the affected cell towards lower terrain is imposed (i.e., an additional one-half basic increment is added to the problem cell). The magnitude of the incrementation is 0.00001 elevation units since full incrementation is not feasible because it can sometime lead to another flat surface in some situations with upstream cells that are exactly one increment higher than the problem cell. If several cells are affected by the above situation then the one-half increment is added repeatedly following the same procedure as is used in imposing the gradient towards lower terrain.

The results of combining the gradients imposed in the first two steps of the analysis are shown in Figure 2.7(a). The cell positioned at row 5, column 4 proved to be a problem cell with an initial modification elevation of 1.08. One-half increment was added to produce an elevation value of 1.09 since without this increment there will be no downslope path to an adjacent cell.

2	2	3	0	3	2	2
3	3	3	**1**	3	3	3
3	2	3	**1.06**	3	3	2
3	2	**1.08**	**1.08**	**1.08**	3	2
3	2	**1.10**	**1.09**	**1.10**	3	2
3	2	**1.12**	**1.10**	**1.12**	3	2
3	2	**1.14**	**1.14**	**1.14**	3	2
3	3	3	3	2	3	3
2	2	2	3	3	2	2

(a)

2	2	3	0	3	2	2
3	3	3	↑	3	3	3
3	2	3	↑	3	3	2
3	2	↗	↑	↖	3	2
3	2	↑	↑	↑	3	2
3	2	↗	↑	↖	3	2
3	2	↗	↑	↖	3	2
3	3	3	3	2	3	3
2	2	2	3	3	2	2

(b)

Figure 2.7. Illustration of combination of gradients away from higher elevations and towards lower elevations: (a) addition of cumulative elevation increments for both gradients; (b) the resulting drainage.

The final flow pattern is depicted in Figure 2.7 (b). The flow pattern shown in this figure is different from that shown in Figure 2.5 (b) because from the combined gradient approach as evident in Figure 2.7 (b) the flow tends to converge and concentrate along the central axis of the flat area.

2.9 Use of DEM's in obtaining flow network information

The knowledge of distributed landscape topography provided by a DEM (Ebner and Edner, 1992) provides the basis for a series of automatic procedures able to derive drainage network structures at the proper scale of information (Brath et al., 1989) and to support hydrological modeling. Information about landscape morphology can be obtained from watershed parameters derived from these procedures (Quinn et al., 1991).

Jenson (1991) applied automated information extraction analysis to a DEM of the Kasei Valles area of Mars having a resolution of 200 meters to provide surface information. This efficiently delineated craters that could be further processed to derive depth, area and volume information.

Jenson and Trautwein (1987) presented an analogous example in pothole terrain on the Earth. The systematic pattern of depressions was present in some flat areas. The automatic formation of a realistic drainage network in the valley region, compared with non-fluvial features elsewhere, supported the fact that valley was created by hydrologic processes.

Band (1986) used the Glendora, California 7.5 min quadrangle DEM to extract networks. The algorithms used were assembled into a package of FORTRAN and C subroutines called the stream and ridge edge analysis and mapping system (STREAMS).

The extracted networks when compared with those in the topographic map showed a close correspondence (Band,1986).

Hutchinson and Dowling (1991) performed drainage basin analysis of the Australian 1/40th degree DEM. The automated drainage basin analysis was found to be in close agreement with the existing basin analyses for the surveyed coastal areas with well-defined relief and also for large areas of low relief.

To delineate a drainage network, Speight (1968) manually applied the concept that when runoff exceeds some minimum threshold then it represents a channel network. A square grid of 30.5 m was drawn on a contour map. Next, a slope line perpendicular to the contours was traced downslope from each grid point. Finally, a line segment 30.5 m in length was centered, parallel to the contours, on each grid point, and the number of slope lines crossing it was determined. If this count was greater than 100 lines, or if the line had a density over a narrow zone of flow concentration of more than one slope line per foot of sampling line, the point was declared to lie on a watercourse (Speight, 1968).

Mark (1984) used both a local concavity approach and a runoff simulation approach to delineate the drainage network. He used the program HILO to implement a local concavity approach for representing the topography of the Keating Summit, Pennsylvania, quadrangle whereas the FORTRAN program DRAIN was used to process the DEM with a runoff simulation approach. In his finding, the local concavity approach was more space and time efficient while the overland flow method was extremely slow and required more than an hour to process a single DEM.

O’Callaghan and Mark (1984) used DEMs derived from various sources to extract channel networks. Three test areas were used: Broken Hill (Australia), which was interpolated from contour data after sampling (Briggs, 1981); Keating Summit, Pennsylvania; and Bighorn, Wyoming. The DEM of Keating Summit was generated by USGS using Gestalt Photomapper (GPM II) while the Wild B-8 method was used to generate the Bighorn DEM. The test areas were selected to represent a range of terrain types.

The channels derived by O'Callaghan and Mark (1984) were in agreement with the channels inferred from the contours. Areas with low gradient and where the distinct channels did not exist in the DEM surface showed unrealistic patterns. For Keating Summit, the depressions lie along the channels and it was interpreted that those with positive errors are dams across the channel. A large number of sinks and depressions were found in flat areas.

Tarboton et al. (1991) emphasized that for a channel network to be useful, it should be extracted at a correct length scale or drainage density. They extracted channel networks for 21 digital elevation data sets well distributed throughout the U.S. They found reasonable agreement between drainage densities. Douglas (1986) gave a good description of methods developed to define ridges, channels, watersheds, and other hydrological features from DEMs.

Martz and Garbrecht (1992) used a set of ten algorithms to extract the drainage network and corresponding subwatersheds from a DEM combined in the computer program DEDNM (Digital Elevation Drainage Network Model). They concluded that while visual inspection of first order channel showed some variation, the higher order channels were well produced.

3. Methodology

3.1 Study overview

In this study, two algorithms were used to remove depressions from watersheds, and the magnitude (depth of cuts and fills) and spatial distribution of the modifications determined. The watersheds used represent different physiographic regions including the mountain, valley, and Piedmont regions of Virginia (Figure 3.1). The DEM data used was the 7.5 minute USGS DEMs that correspond to the 1:24,000 USGS topographic "quad" maps at a grid resolution of 30 meters. Elevation values were integer meters for the analysis. Figure 3.2 shows an example of hill shaded image of a DEM.

ArcView (Version 3.1, Spatial Analyst 1.1) and TOPAZ (Version 1.2) were used for watershed delineation, depression removal and extraction of the flow network. In ArcView, an AVENUE script was used to delineate watersheds (Table A1 in Appendix A). DEMs were filled to remove depressions, and maps of flow direction and flow accumulation were derived using the script (Table A2 in Appendix A). Streams were delineated for the DEMs using the command "Stream network as line shape" within the Hydro menu.

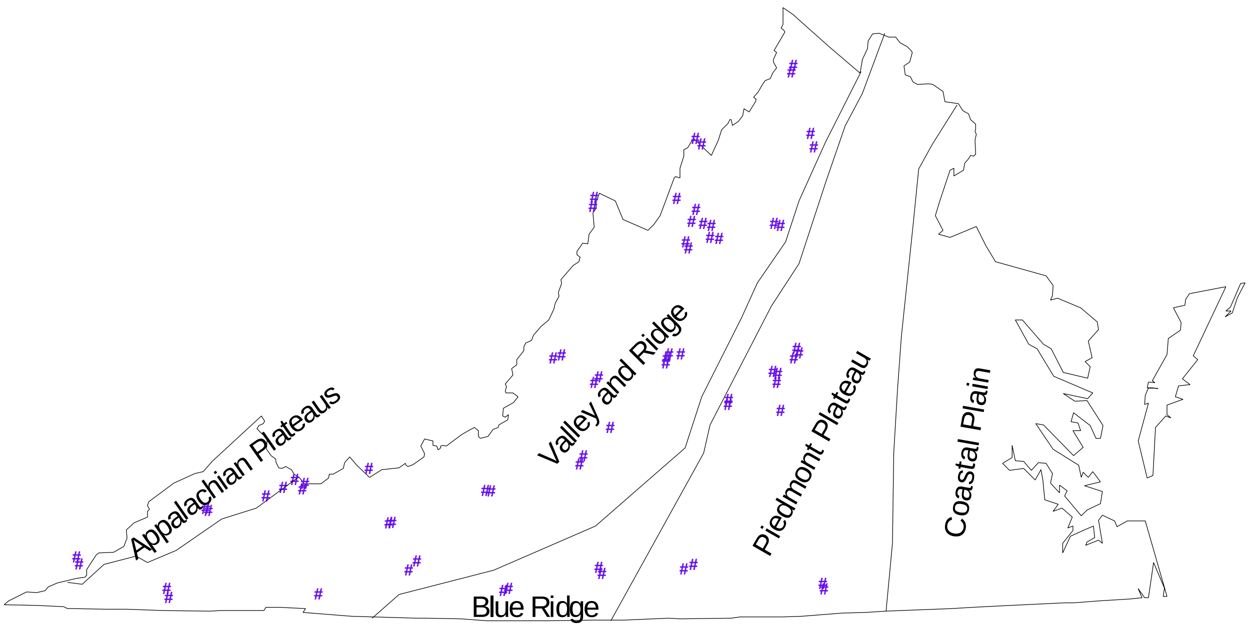

Figure 3.1. Location of watersheds used in this study.

Figure 3.2. The USGS DEM quadrangle for Blacksburg (Hillshade image of the DEM).

Since TOPAZ was also used to delineate watersheds from the quadrangles, the procedure is briefly described as follows:-

a). DEM data was first converted to grid theme.

b). The grid theme was exported from ArcView in ASCII format.

c). Header information was removed from each ASCII file to make it readable for the DEDNM module within TOPAZ.

d). A C++ program was used to arrange the rows and columns of the ASCII file in a single column for input to the DEDNM module.

e). Parameters required by DEDNM are:

i) Number of rows and columns in the DEM.

ii) Maximum and minimum elevation values within the watershed.

iii) Row and column coordinate defining the outlet of the watershed.

iv) Resolution or cell size of the DEM.

After using DEDNM, the RASFOR module was used to reformat the output. The header information required by ArcView was added to the output ASCII files. Output ASCII files contained the corrected DEM (i.e., the DEM representing a hydrologically continuous surface), the stream network, and the flow accumulation maps. These ASCII files were imported in ArcView for further analysis. The derived channel network obtained by the two methods was compared with the blue line streams digitized from the USGS quadmaps to ascertain which method better matches the digitized streams.

The number of cells in a watershed, minimum and maximum elevation, and area (hectares) were computed for each watershed using the “Statistics” function in the ArcView Legend editor. Elevation range for every watershed was computed by subtracting the minimum elevation from the maximum elevation value. Other properties such as watershed slope, channel slope, drainage density and curvature were determined for all the watersheds.

The watersheds were categorized in four groups according to their area, to evaluate any possible influence of watershed size on various watershed characteristics. The categorization was done on the basis of natural breaks in the data.

3.2 Measures of watershed characteristics

3.2.1 Deviation, variance and standard deviation

To interpret the variability of a single set of measurements, a measure of deviation is widely used. A data set with low variability will have few measurements located away from the center of the distribution while the majority of the measurements will be located near the center of the distribution. Thus, the deviations from the mean value will be large for the data set having a more variable set of measurements. Deviation is represented by $a - \bar{a}$ where a refers to the measurements and $\bar{a}$ is the mean of the set of measurements.

Many different measures of variability using the deviations $a - \bar{a}$ can be constructed, but the most common and easily interpreted function is the sum of the squared deviations, known as the variance. The definition for the variance of a set of measurements depends on whether the data is regarded as a sample or population of measurements. Since population measurements are generally difficult to find, it is assumed that a sample is being used. The variance of a set of n measurement $a_1, a_2, \ldots, a_n$ with mean $\bar{a}$ is the sum of the squared deviations divided by n-1 (Ott, 1992):

$$\text{Variance } (s^2) = \Sigma_i (a - \bar{a})^2/n\text{-}1$$

Standard deviation is another useful measure of variability and is defined as the positive square root of the variance (Ott, 1992). Thus, the sample standard deviation(s) is $\sqrt{s^2}$.

The arithmetic mean, or mean of a set of measurements is an important measure of central tendency and is defined by the total number of measurements. A sample of measurements is a subset of measurements selected from the population of interest. If $a_1, a_2, \ldots, a_n$ denote the measurements observed in a sample of size n, then the sample mean $\bar{a}$ can be written as $\bar{a} = \Sigma_i a_i/n$.

Where the symbol $\Sigma_i a_i$, is the notation used to designate a sum of n measurements, a_i:

$$\Sigma_i a_i = a_1 + a_2 + \ldots + a_n.$$

3.2.2. Watershed slope, channel slope and drainage density

A number of watershed characteristics were computed to establish the relationship between watershed terrain complexity and the area affected as a result of creating a depressionless surface. The measures used were watershed slope, range, channel slope, drainage density, slope and curvature of elevation surface.

Watershed slope was calculated by using the formula:

$$\text{Watershed slope} = \text{watershed length}/\ (E1\text{-}E2)$$

Watershed length is the length of drainage channel from watershed outlet to the drainage divide.

E1 = elevation at the watershed divide where the channel meets when extended.

E2 = elevation at the watershed outlet.

Similarly,

$$\text{Channel slope} = \text{Channel length}/\ (Ce1\text{-}Ce2)$$

Channel length refers to the length of main drainage channel. Appropriate selection of threshold is essential for proper delineation of network.

Ce1 = elevation at the top of channel network.

Ce2 = elevation at the watershed outlet.

Drainage density was another measure, which was used to characterize watershed surface.

$$\text{Drainage density} = \text{Length of drainage channel}/\ \text{total watershed area}$$

3.2.3 Slope of watershed

Slope is defined as the maximum rate of change in its elevation value from each cell to its neighbors. An output slope can be computed in percent or degree of slope. The slope for the cell is computed by using 3x3 neighborhood across the center cell. Using the following notation to indicate cell location,

A	D	G
B	E	H
C	F	I

The slope of cells is calculated as follows:

$$\text{Slope of E} = \sqrt{\left(\frac{(A+2B+C)-(G+2H+I)}{8*resolution}\right)^2+\left(\frac{(A+2D+G)-(C+2F+I)}{8*resolution}\right)^2}$$

3.2.4 Curvature of surface

Curvature indicates the extent to which the surface is curved. If the surface is upwardly convex at that cell curvature is positive. Curvature is negative when the surface is upwardly concave. If the surface is flat then it has a value of zero. The curvature of a surface is computed on a cell-by-cell basis using fourth-order polynomial of the form:

$$Z = Ax^2y^2 + Bx^2y + Cxy^2 + Dx^2 + Ey^2 + Fxy + Gx + Hy + I$$

for every cell is fitted to a surface consisting of a 3x3 window.

3.3 The spatial distribution of affected elevation and slope cells

The elevation cells changed by ArcView and TOPAZ were clumped in clusters to determine any spatial patterns which might exist. The elevation cells which were filled and those that were reduced in value by TOPAZ were separated to determine the spatial pattern of cuts and fills for each type.

The elevation cells affected by ArcView and TOPAZ were classified in different number of clusters and the number of cells in each cluster. The diagonal cells were included in the procedure. AVENUE script was used for this purpose and then map query and statistics functions were used on the attribute table to find the number of cells in each cluster.

The elevation and slope cells changed by ArcView and TOPAZ were further classified according to their distance from the stream. All the changed cells were classified in three categories: cells which were on the stream, cells at a distance of 1-100 m from the stream and cells which were more than 100 m away from the stream.

The cells changed (cuts and fills) from the TOPAZ method were classified in the above three categories. The depth of the fills was calculated for each watershed for ArcView along with the depth of cuts and fills for the TOPAZ using map calculator.

Mean and standard deviation of curvature and mean slope of mean slope for all the watersheds were computed by using AVENUE script and “Derive slope” command within the Surface menu of ArcView. The number of cells which are within the categories of greater then 1, 2, 3, 4 and 5 (Absolute Curvature) were computed by reclassifying the map of absolute value of curvature.

3.4 Comparison of digitized USGS blue line stream with streams derived from ArcView and TOPAZ

3.4.1 Digitization

Blue line streams were digitized on-screen in ArcView from the USGS DRGs (Digital Raster Graphics), which are georeferenced images (TIFF format) made by scanning the quadrangle maps. The DRGs have a pixel size of 2.4 m. The TIFF images were zoomed at an approximate scale of 1: 20,000 at which the blue line stream was clearly evident and then the blue line streams were digitized using “Add new line theme”.

Snapping tolerance was set at 40 meters. When the snapping environment is set, ArcView moves the vertices or line segments of the new features added to align with the vertices or line segments of other features that are within a specified distance, the snap tolerance. This way, for line themes, all line features coming together at an intersection share the same endpoint, and there are no overshoots or undershoots, and for polygon themes, there are no gaps or overlap between adjacent polygon features. The digitizing of the entire blue line network was done in relative (mouse) mode. The digitized stream was generally within 30 m i.e., within one cell distance of the blue line stream.

3.4.2 Accuracy of flow networks derived by TOPAZ and ArcView

The accuracy of the flow network derived by TOPAZ and ArcView was assessed through comparison with the USGS blue line stream. This analysis included both continuous and intermittent streams. The raster drainage networks derived from the ArcView and TOPAZ were converted to a shape file. The channel networks derived by ArcView and TOPAZ were more detailed than the blue line network (Figure 3.3a). To provide a reasonable comparison, generated flow networks were pruned to match the extent of the network with the blue line streams (Figure 3.3b). The vector flow network of ArcView and TOPAZ were pruned by editing on-screen based on a display of the flow network and blue line stream. The pruned ArcView and TOPAZ generated streams are shown with the digitized USGS blue line stream network for all watersheds in Appendix B. After the cleaning process, the flow networks were converted back to a raster form. Thus, there were three-vector networks, which were rasterized using the ArcView

“Theme” menu function “Convert to Grid”. A 30 m cell resolution was used for the analysis, with care taken to ensure that the grids were registered to the DEM grid.

The “Find Distance” command was used to generate a grid theme containing the distance from the blue line stream. The “Summarize zone” command was used to compute the number of cells in the ArcView and TOPAZ stream grids which were in categories of 0-1, 1-50, 50-100 and greater than 100 m from the blue line stream network.

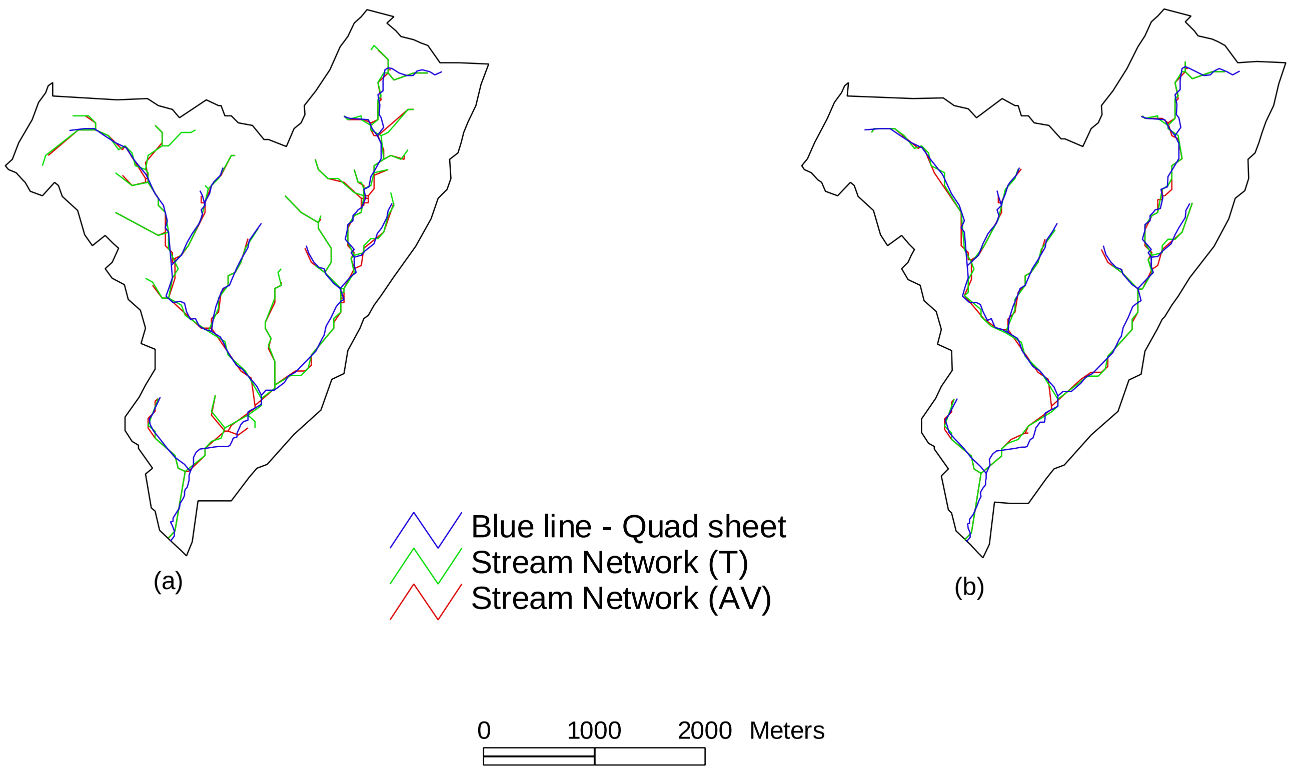

Figure 3.3. Blue line stream from the quad sheet along with the stream networks derived by TOPAZ and ArcView: a) ArcView and TOPAZ streams as generated; b) ArcView and TOPAZ streams after pruning to match the extent of the blue line stream.

3.4.3 Two-sample paired t test for means

A paired t test was used to analyze the percent of ArcView and TOPAZ stream cells matching with the rasterized blue line stream. In this case, each measurement in one sample is matched or paired with a particular measurement in the other sample. The test can also be used to determine whether the proportions of two populations giving a certain response are equal to each other (or differ by a specified amount).

3.4.3.1 Null and alternative hypotheses

A value for the hypothesized difference between the proportions must be specified. The difference of zero results in the test that the true proportions in the sample are equal. Typically, each observation has a value of 0 or 1 for the specified variables. Then, if 0 and 1 represent "failure" and "success," then it can be used to analyze whether the proportions of failures or successes are equal.

The hypothesis test is performed in this method to decide which of two theories about the difference between the proportions of certain responses of the selected variables is supported by the data. The first theory, known as the null hypothesis, represents the status quo. This hypothesis is retained unless the data gives strong evidence to the contrary. The alternative hypothesis represents a change in the status quo. Choice of the alternative hypothesis should be based on the expectations of the experimenter, not on the data. When the p-value is smaller than a certain threshold (typically 0.05 or 0.01), then null hypothesis is rejected and it is concluded that the evidence strongly favors the specified alternative hypothesis.

A proper analysis of the proper data makes use of the difference in the measurement to test the null hypothesis that the mean difference, μ_d, is D_o. This hypothesis is equivalent to H_o: $\mu_1 - \mu_2 = D_o$.

Where μ_1 and μ_2 are population means.

The null hypothesis (H_o), alternative hypotheses (H_a) and rejection region (R.R.) are summarized below:

H_o: $\mu_d = D_o$

H_a: 1. $\mu_d > D_o$

2. $\mu_d < D_o$

3. $\mu_d \neq D_o$

Test Statistics (T.S.): $$t = \frac{\bar{d} - D_o}{s_d / \sqrt{n}}$$

where $\bar{d}$ and s_d are the sample mean and standard deviation of the n differences.

Rejection Region (R. R.): For a specified value of Type I error[1] (α) and degree of freedom (df) = n –1

1. reject H_o if $t > t_{\alpha}$
2. reject H_o if $t < - t_{\alpha}$
3. reject H_o if $|t| > t_{\alpha/2}$.

[1] A Type I error is committed when null hypothesis which is true is rejected.

4. Results and discussion

4.1 DEM modification

The DEM modification was carried out on watersheds to remove depressions and to delineate the stream network. The changed elevations in the modified DEMs show different spatial distributions depending upon the watershed characteristics and the type of errors in the DEMs. Figures 4.1 to 4.10 show the spatial distribution of elevation cells modified by ArcView and TOPAZ along with the generated flow network for 5 watersheds. Figure 4.2 shows an example where cuts are more or less uniformly distributed all over the watershed whereas Figure 4.6 shows a case with few cuts which are confined near the stream. Similarly, the other figures represent a sampling of the types of responses found. The figures also show depth of the fills and cuts classified in four categories: 1-2, 2-5, 5-9, and greater than 9 meters.

It can be observed from Tables 4.1 and 4.2 that for ArcView the percent area changed ranges from 0 to 11.36% with mean of 3.1% and median of 2.47%. The percent area changed by TOPAZ (fills and cuts) ranges from 0 to 8.88% for fills and 0 to 2.8% for cuts. The majority of the watershed (58 of 66) have fewer cuts than fills. The mean for percent area changed by TOPAZ fills is 1.55%, which is lower than the mean of percent area changed by ArcView fills. TOPAZ has less impact (less percent area modified) than ArcView for 64 watersheds. This shows that fewer cells in the DEM

Table 4.1 Characteristics of watersheds and percent of area modified in removing depressions.

	Area (ha)	Std. Dev of Elev. (m)	Range of Elevation (m)	Abs. Mean Curvature	Mean Slope (%)	Watershed slope (m/m)*	Channel slope (m/m)	Drainage Density (per meter)	Percentage of area changed ArcView Fills	TOPAZ Fills	TOPAZ Cuts
Av2	1224	65	338	0.58	10.5	0.043	0.027	0.0025	4.59	1.78	1.92
Av3	641	58	339	0.70	15.7	0.037	0.025	0.0024	2.53	0.98	0.87
Bb2	1071	27	184	0.54	8.8	0.031	0.032	0.0023	3.34	1.29	1.88
Bb3	2051	27	183	0.55	7.3	0.007	0.006	0.0023	9.01	4.27	2.34
Be2	2097	75	386	0.66	13.0	0.039	0.024	0.0028	0.00	0.00	0.00
Be3	792	76	334	0.72	17.5	0.056	0.057	0.0019	0.00	0.00	0.00
Bl2	543	54	250	0.31	9.9	0.056	0.045	0.0024	1.76	0.88	0.48
Bl3	536	64	283	0.32	10.4	0.057	0.040	0.0021	1.45	0.64	0.47
BR	4954	29	188	0.23	3.9	0.006	0.005	0.0023	1.25	0.33	0.36
CCr	6428	30	227	0.93	4.2	0.006	0.004	0.0022	2.41	1.30	0.35
CG2	1902	121	511	0.42	13.4	0.084	0.084	0.0024	0.00	0.00	0.03
CG3	1076	193	688	0.36	11.6	0.092	0.094	0.0027	0.03	0.00	0.03
Cn10	462	52	230	0.56	16.6	0.057	0.052	0.0020	2.92	1.17	0.92
Cn11	807	77	354	0.56	18.2	0.068	0.060	0.0019	2.26	1.28	0.52
Cn12	984	74	348	0.50	17.3	0.062	0.054	0.0017	2.80	1.17	0.70
Cn14	591	72	334	0.67	19.8	0.074	0.063	0.0020	1.81	0.93	0.53
Cn17	1127	69	356	0.78	19.9	0.057	0.045	0.0020	0.78	0.32	0.34
Cn6	466	51	242	0.73	15.6	0.048	0.028	0.0023	1.91	0.64	1.06
Ct2	1299	210	789	0.34	14.2	0.087	0.081	0.0021	2.04	2.11	0.17
Ct3	2442	218	879	0.37	16.2	0.070	0.070	0.0022	0.94	0.45	0.30
DRl	2959	34	217	0.20	3.1	0.012	0.011	0.0025	3.78	1.53	0.74
DRu	19959	174	889	0.40	17.2	0.021	0.020	0.0019	0.72	0.37	0.13
Fle2	4391	208	915	0.49	18.0	0.069	0.058	0.0021	1.00	0.35	0.24
Fle3	867	123	786	0.43	14.4	0.097	0.074	0.0021	3.06	1.65	0.47
Ft10	630	15	77	0.19	3.2	0.011	0.011	0.0022	2.63	1.07	0.83
Ft14	571	18	87	0.20	3.8	0.016	0.015	0.0022	3.42	1.75	0.66
Ft2	767	11	72	0.21	3.6	0.013	0.011	0.0023	3.38	1.38	0.93
Ft7	556	15	72	0.21	3.9	0.013	0.013	0.0023	5.94	3.58	0.79
GC2	1155	56	271	0.33	12.4	0.030	0.027	0.0020	3.19	2.16	0.32
GC3	513	37	203	0.32	10.5	0.036	0.033	0.0019	1.93	1.10	0.56
Gl2	1169	46	307	0.54	10.7	0.031	0.025	0.0023	3.83	1.94	1.25
Gl3	667	55	362	0.50	9.1	0.057	0.029	0.0027	4.33	1.17	2.11
Gn10	606	133	675	0.45	18.5	0.120	0.112	0.0018	0.16	0.06	0.21
Gn16	1051	144	760	0.46	17.4	0.097	0.088	0.0023	0.24	0.03	0.27

…. continued….

*m/m = meter/meter

Table 4.1 (continued)

	Area (ha)	Std. Dev of Elev. (m)	Range of Elevation (m)	Abs. Mean Curvature	Mean Slope (%)	Watershed slope (m/m)*	Channel slope (m/m)	Drainage Density (per meter)	Percentage of area changed ArcView Fills	TOPAZ Fills	TOPAZ Cuts
Gn4	752	158	675	0.29	15.8	0.118	0.112	0.0018	0.02	0.00	0.05
Gn5	584	156	687	0.31	16.2	0.143	0.123	0.0019	0.39	0.35	0.05
Gn6	1567	131	681	0.30	14.8	0.086	0.066	0.0021	0.44	0.20	0.10
Hf2	952	26	203	0.55	7.0	0.027	0.016	0.0023	9.08	3.41	2.80
Hf3	823	22	134	0.54	7.0	0.015	0.014	0.0022	5.78	1.55	2.72
Longgl	2745	28	171	0.51	5.7	0.012	0.010	0.0024	8.02	3.53	2.54
MC	8153	82	598	0.25	6.9	0.021	0.013	0.0022	0.74	0.24	0.23
Mill	3896	38	219	0.23	4.4	0.012	0.010	0.0022	5.80	2.23	0.88
MossyCr	1948	21	128	0.53	6.6	0.009	0.008	0.0024	8.81	3.78	2.37
MV2	2685	68	322	0.24	7.5	0.024	0.024	0.0022	5.24	3.21	0.68
MV3	1881	25	133	0.24	5.3	0.014	0.012	0.0023	7.83	5.18	0.98
No2	797	63	280	0.59	18.9	0.044	0.039	0.0021	1.02	0.58	0.28
No3	546	55	242	0.59	18.2	0.045	0.042	0.0018	1.76	1.06	0.43
PIRun	2091	28	185	0.24	4.7	0.013	0.010	0.0021	4.44	2.04	0.66
Pl2	1396	14	76	0.20	3.9	0.010	0.009	0.0024	2.73	0.83	0.72
Pl3	943	14	78	0.19	3.6	0.008	0.008	0.0022	8.76	4.64	0.91
Pl4	1260	17	91	0.19	3.7	0.009	0.008	0.0022	4.91	2.65	0.78
SoH2	2134	11	70	0.23	3.3	0.006	0.005	0.0024	11.36	5.92	1.56
SoH3	775	11	60	0.23	3.5	0.009	0.009	0.0023	9.23	5.19	1.28
St1	1337	152	632	0.40	22.7	0.100	0.073	0.0017	0.14	0.07	0.05
St3	2576	159	816	0.53	14.2	0.059	0.058	0.0026	0.61	0.10	0.70
St4	1353	147	605	0.40	21.9	0.078	0.070	0.0019	0.16	0.08	0.05
Stu2	1493	148	523	0.64	16.9	0.050	0.051	0.0023	2.91	1.27	0.92
Stu3	1169	104	511	0.64	15.5	0.069	0.066	0.0024	3.22	1.31	1.22
Su2	2883	132	681	0.63	17.8	0.050	0.041	0.0024	0.00	0.00	0.00
Su3	685	100	531	0.56	17.8	0.078	0.049	0.0021	0.00	0.00	0.00
Sv2	2009	130	516	0.44	10.5	0.039	0.036	0.0024	6.80	2.99	0.95
Sv3	1616	96	451	0.44	12.3	0.044	0.042	0.0023	3.05	1.23	0.75
Tw2	1928	68	380	0.52	16.6	0.041	0.037	0.0019	0.01	0.01	0.00
Tw3	1683	67	391	0.50	16.0	0.047	0.032	0.0020	0.00	0.00	0.00
VH2	881	17	97	0.22	4.8	0.015	0.011	0.0021	4.74	2.12	0.85
VH3	479	21	86	0.26	5.4	0.018	0.016	0.0022	10.27	8.88	0.88

Table 4.2 Summary statistics of watersheds.

	Area (ha)	Std. Dev of Elev. (m)	Range of Elevation (m)	Abs. Mean Curvature	Mean Slope (%)	Watershed slope (m/m)*	Channel slope (m/m)	Drainage Density (per meter)	Percentage of area changed ArcView Fills	TOPAZ Fills	TOPAZ Cuts
Min	462	11	60	0.19	3.1	0.006	0.004	0.0017	0	0	0
Max	19959	218	915	0.93	22.7	0.143	0.123	0.0028	11.36	8.88	2.8
Mean	1869	76	370	0.43	11.5	0.045	0.038	0.0022	3.15	1.55	0.74
Med	1162	64	328	0.44	11.9	0.042	0.032	0.0022	2.47	1.17	0.66
S. D.	2655	57	245	0.18	5.8	0.033	0.03	0.0002	3.01	1.72	0.71

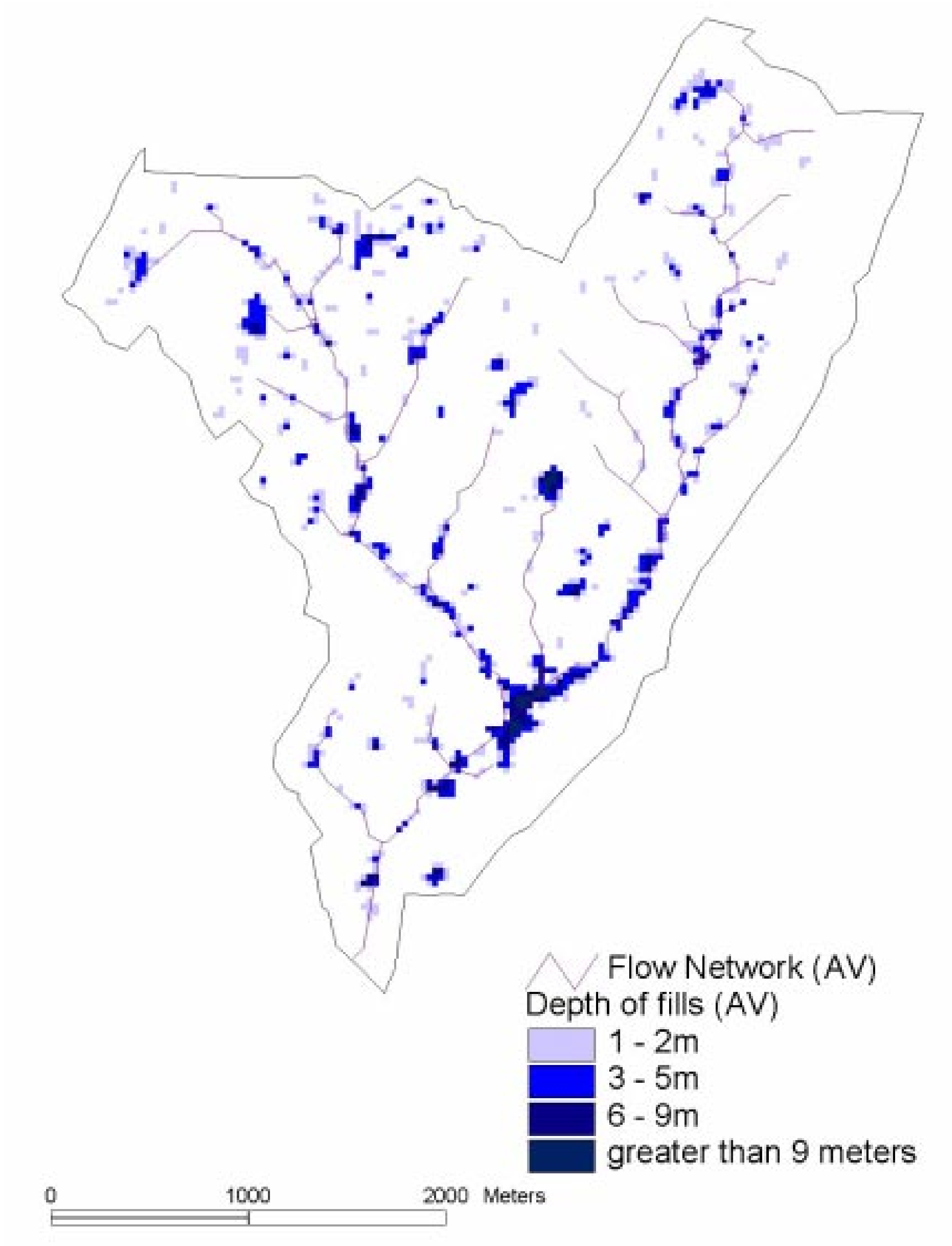

Figure 4.1. Depth of fills along with flow network generated by ArcView for the Hf2 watershed (Cell size = 30m ; Area = 10,579 cells).

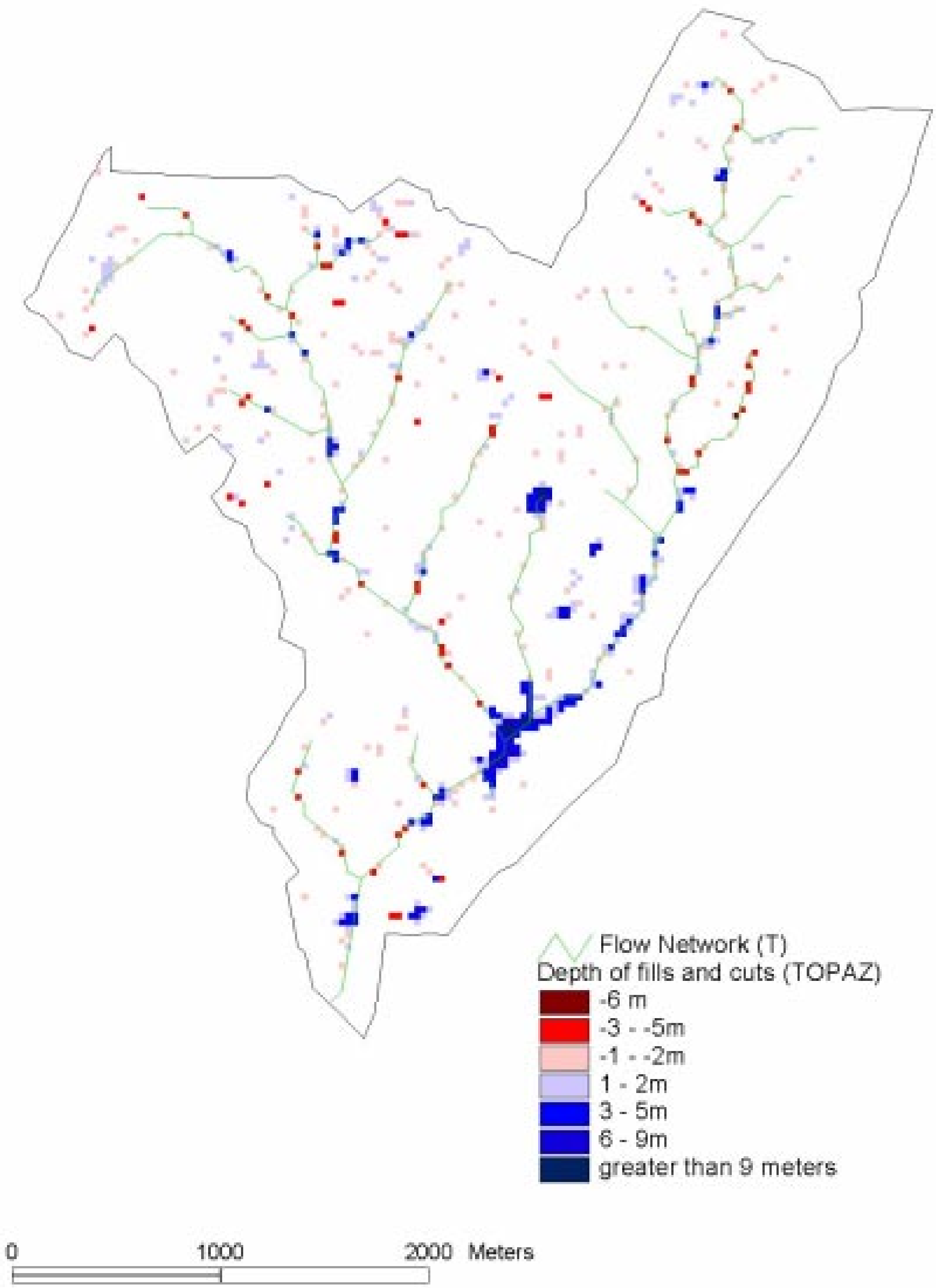

Figure 4.2. Depth of fills and cuts along with flow network generated by TOPAZ for the Hf2 watershed (Cell size = 30m ; Area = 10,579 cells).

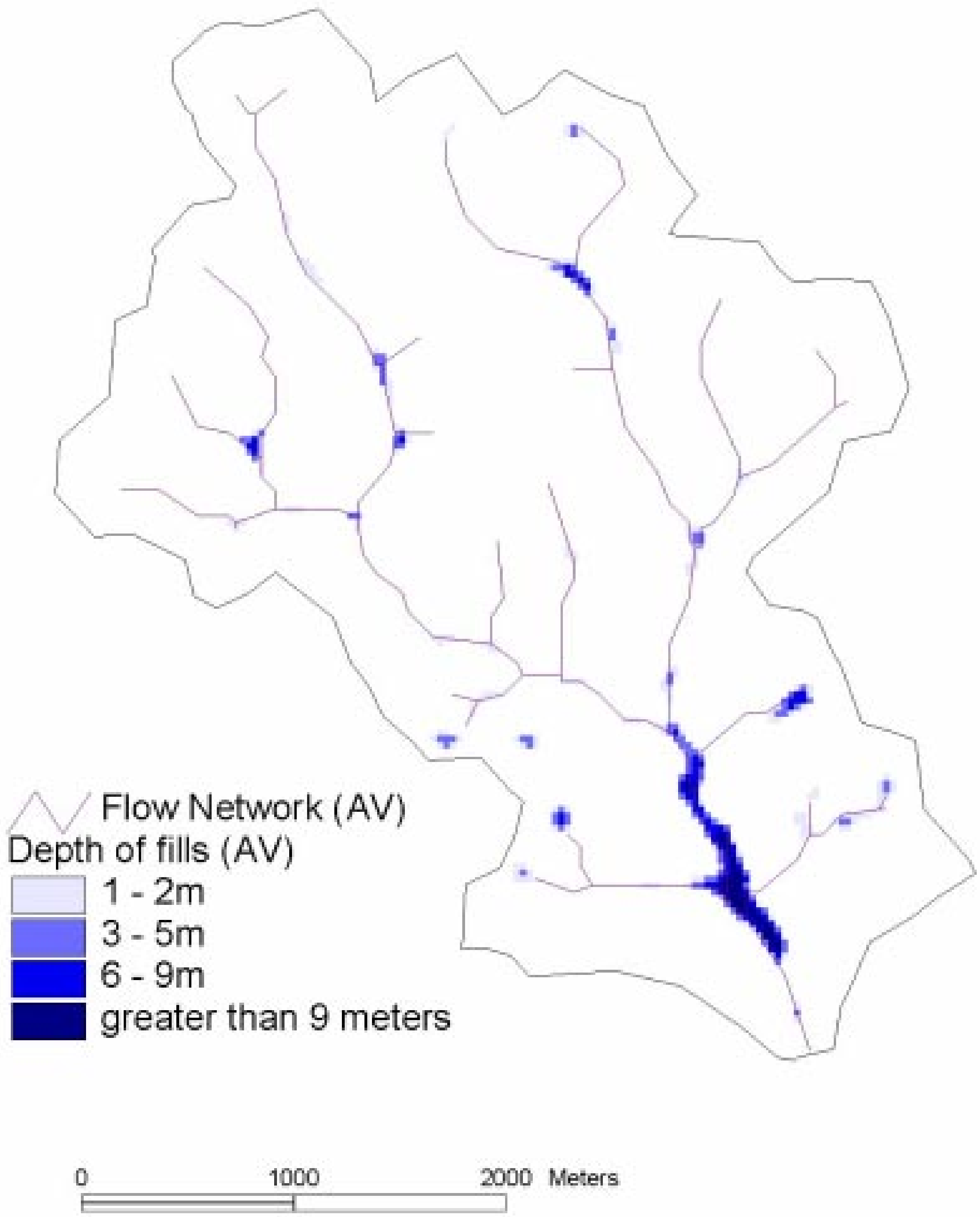

Figure 4.3. Depth of fills along with flow network generated by ArcView for the GC2 watershed (Cell size = 30m ; Area = 12,833 cells).

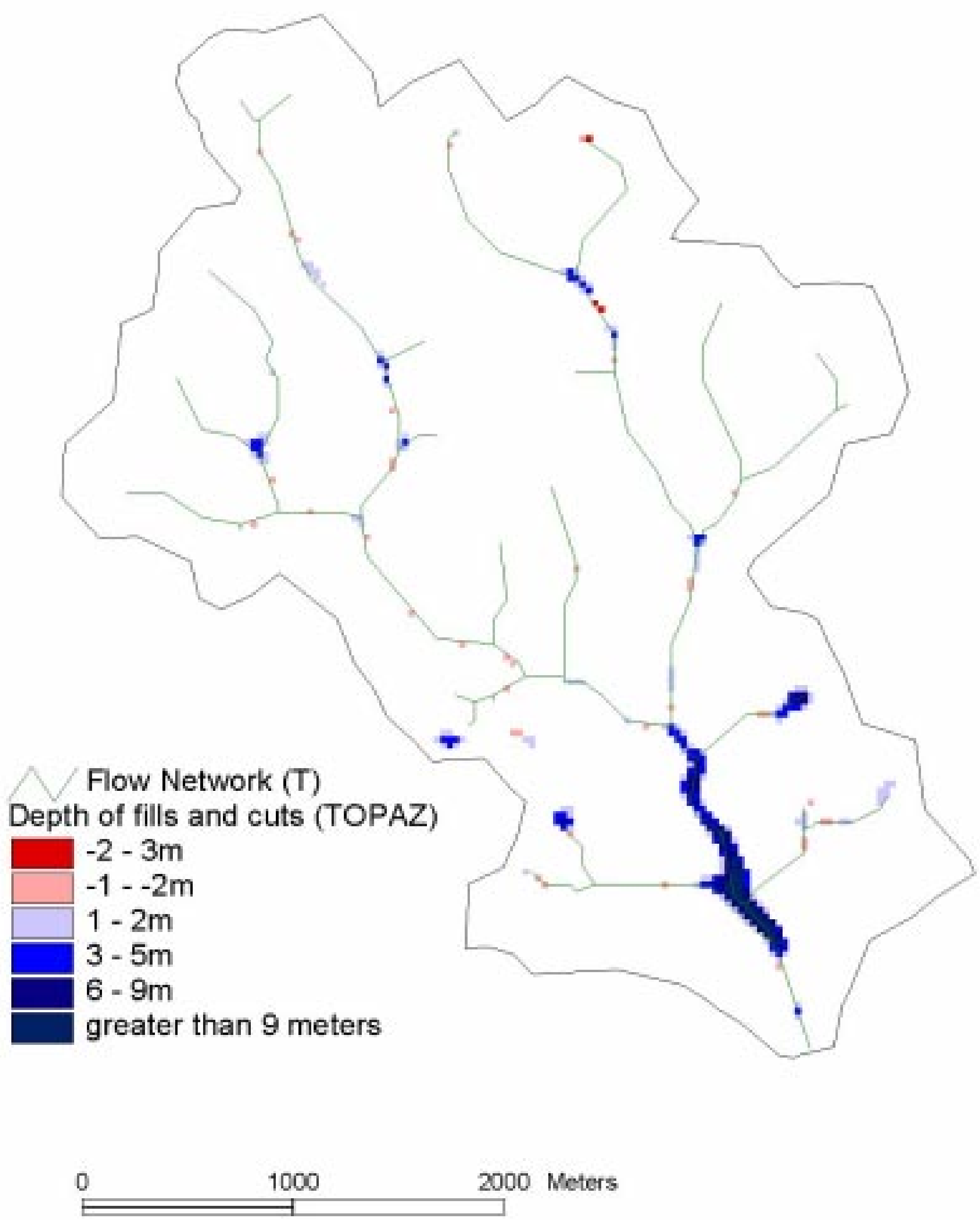

Figure 4.4. Depth of fills and cuts along with flow network generated by TOPAZ for the GC2 watershed (Cell size = 30m ; Area = 12,833 cells).

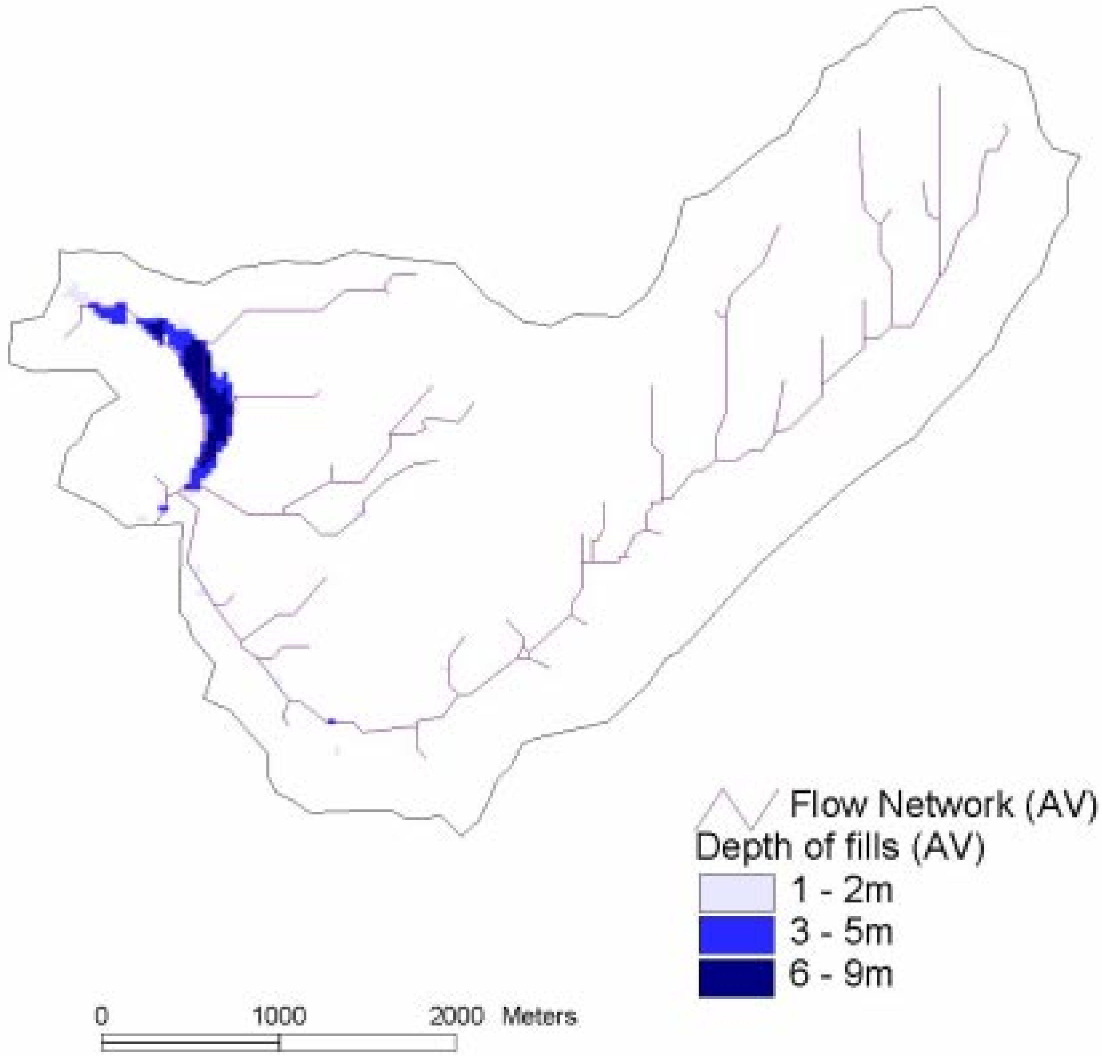

Figure 4.5. Depth of fills along with flow network generated by ArcView for the Ct2 watershed (Cell size = 30m ; Area = 14,429 cells).

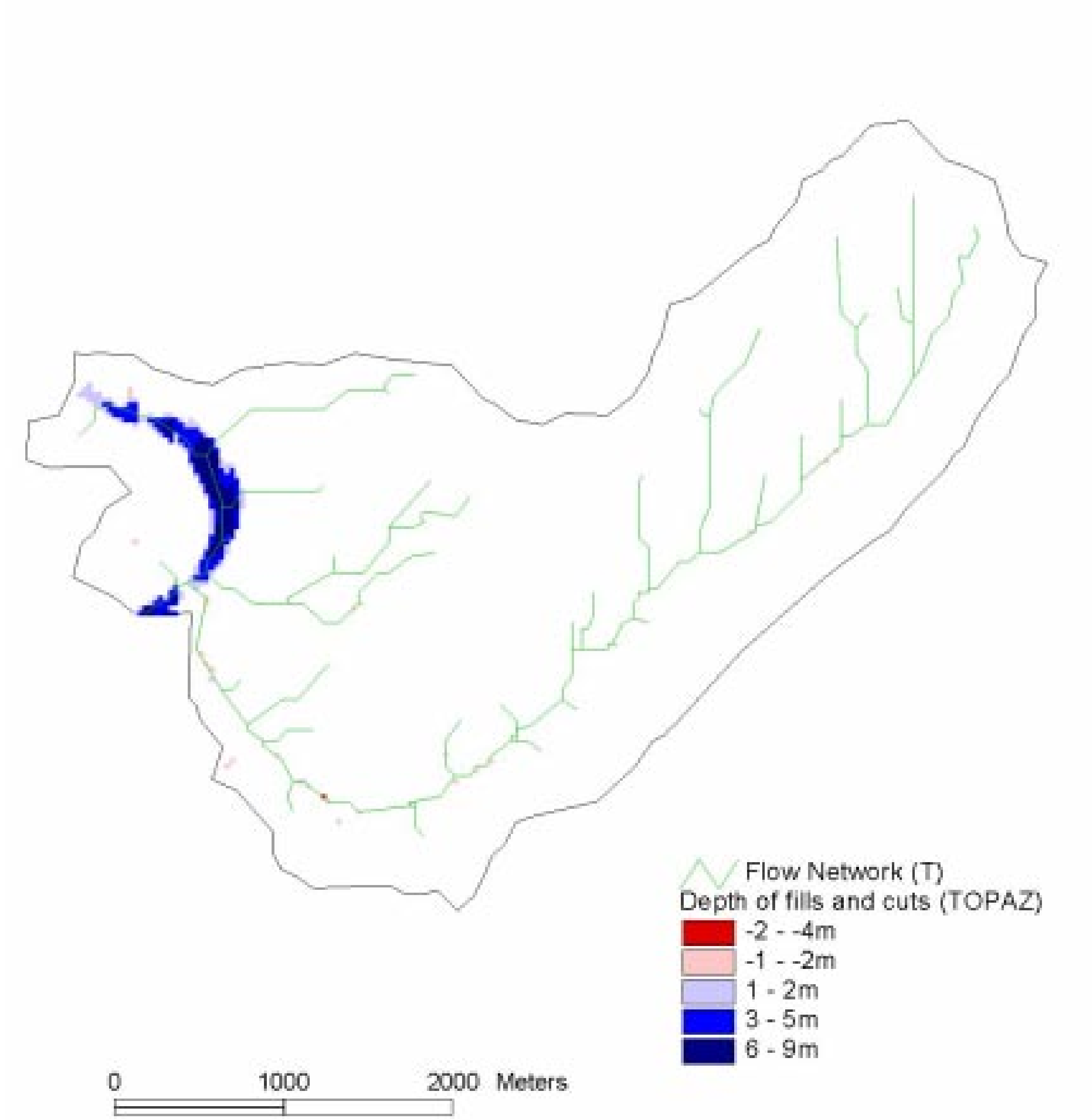

Figure 4.6. Depth of fills and cuts along with flow network generated by TOPAZ for the Ct2 watershed (Cell size = 30m ; Area = 14,429 cells).

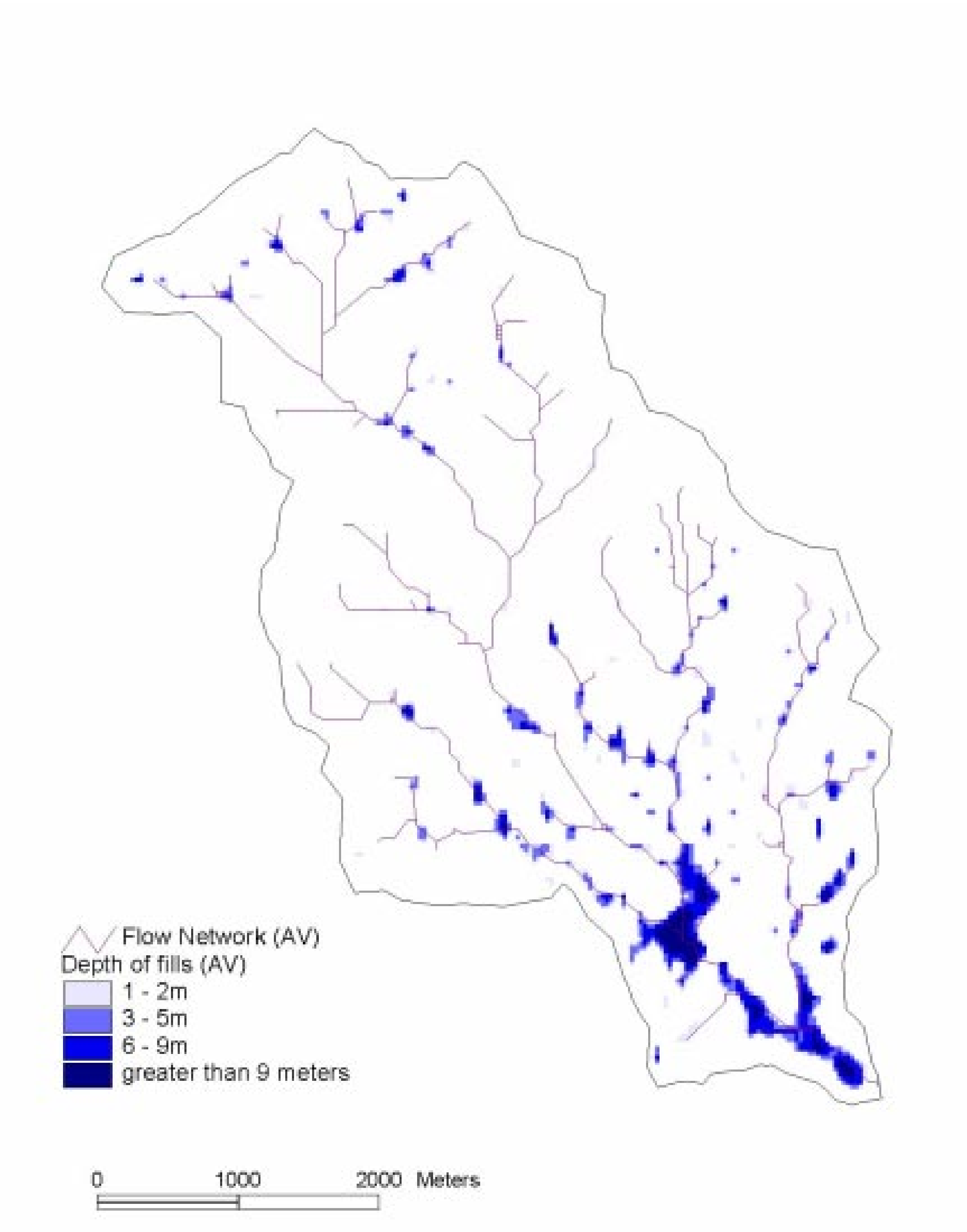

Figure 4.7. Depth of fills along with flow network generated by ArcView for the SV2 watershed (Cell size = 30m ; Area = 22,325 cells).

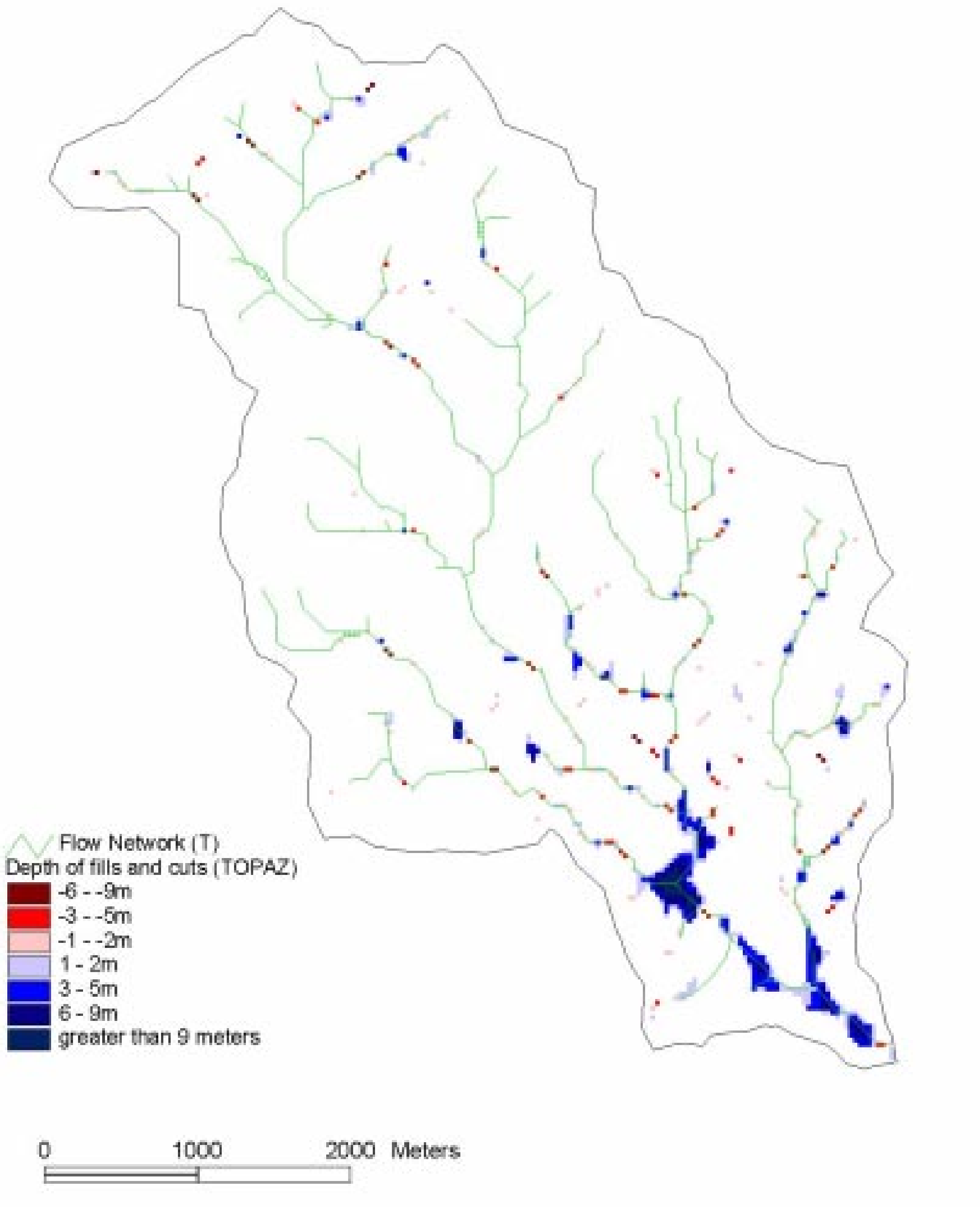

Figure 4.8. Depth of fills and cuts along with flow network generated by TOPAZ for the SV2 watershed (Cell size = 30m ; Area = 22,325 cells).

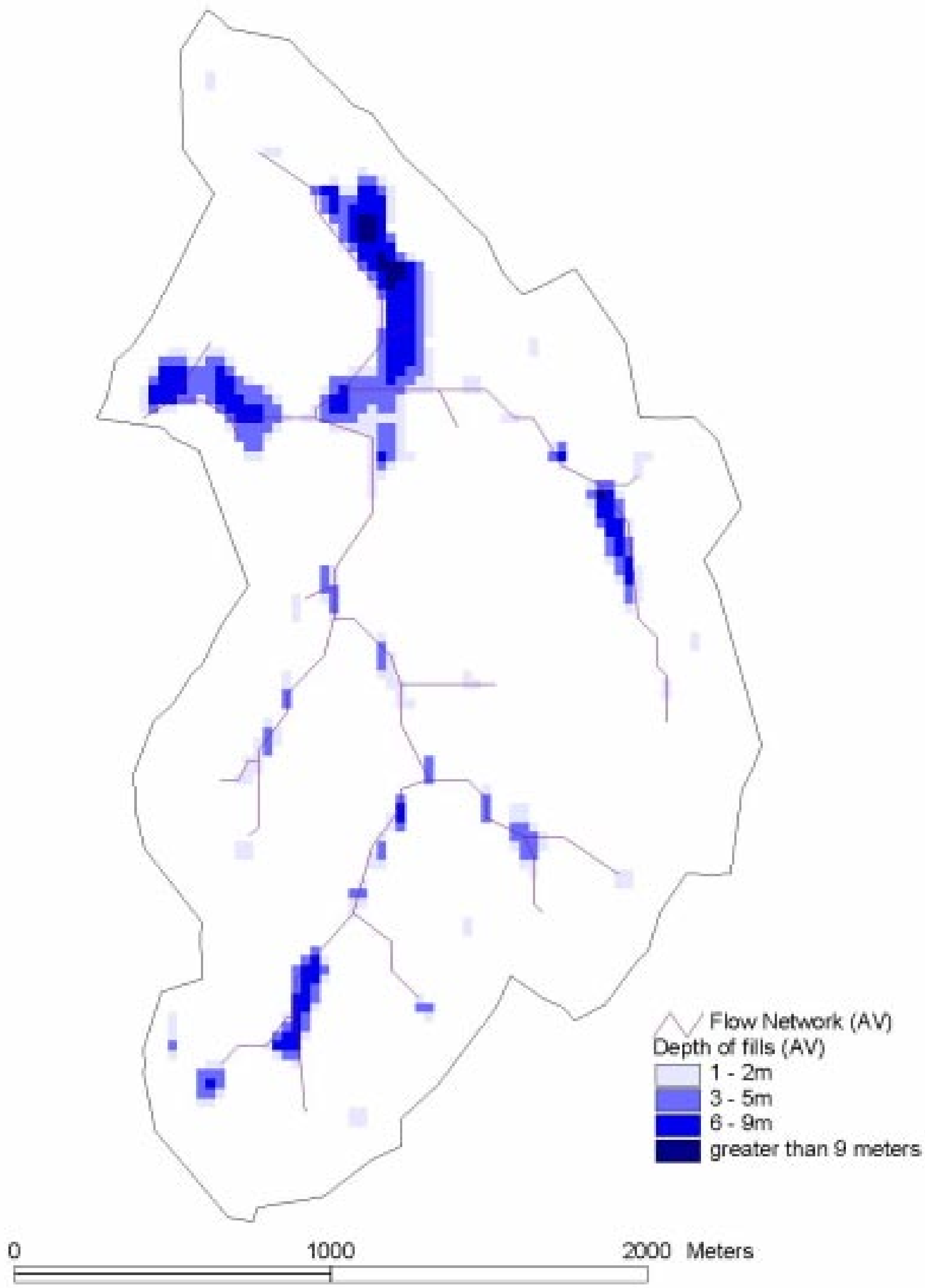

Figure 4.9. Depth of fills along with flow network generated by ArcView for the VH3 watershed (Cell size = 30m ; Area = 5,325 cells).

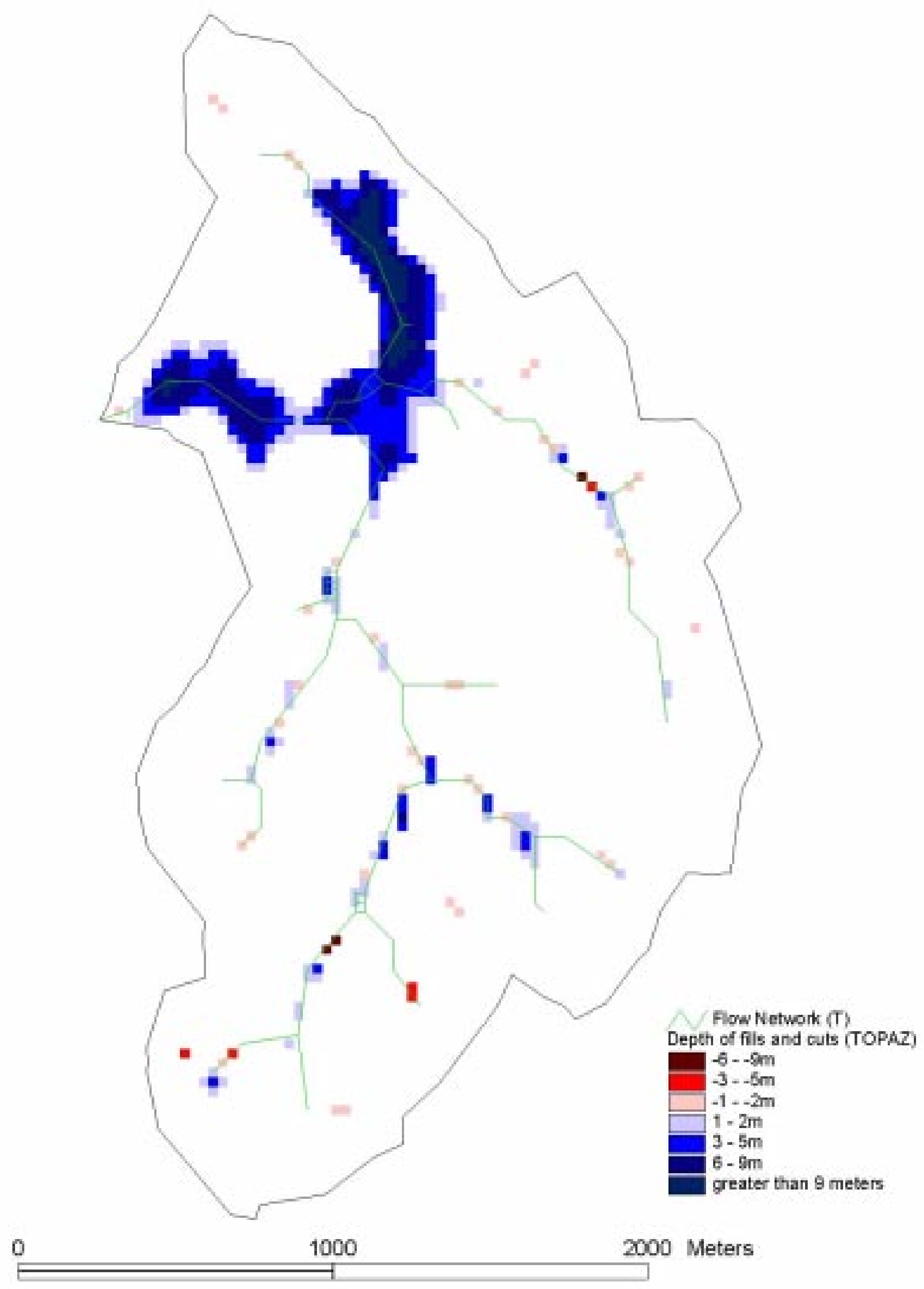

Figure 4.10. Depth of fills and cuts along with flow network generated by TOPAZ for the VH3 watershed (Cell size = 30m ; Area = 5,325 cells).

data set are modified when TOPAZ is used as compared to ArcView. The variability in the percent area changed for ArcView which is represented by standard deviation is 3.01% whereas it is 1.72% for TOPAZ fills and 0.71% for TOPAZ cuts. It can also be observed from Table 4.1 that there are 7 watersheds where no cells are modified, irrespective of whether ArcView or TOPAZ is used. For such watershed it can be concluded that they did not contain sinks or depressions.

4.2 Relationship between terrain complexity and percent of area changed

Several measures were used to represent the topographic variability in a watershed: standard deviation of elevation, range of elevation, mean surface slope, curvature of elevation surface, drainage density, watershed slope and channel slope. The relationship between these measures and percent area (elevation) changed was examined to establish any correlation that might exist.

4.2.1 Standard deviation of elevation

Standard deviation is a common measure used to represent watershed characteristics. The watersheds have values of standard deviation ranging from 11 to 218 (Table 4.2). In general, it can be observed from Table 4.1 that watersheds with lower standard deviation have higher percent of area changed by ArcView and TOPAZ (Figures 4.9 and 4.10) as compared with watersheds having comparatively higher standard deviation. Due to errors present in the DEMs, it can also be observed from Table 4.1 that some of the watersheds have high percent of area changed at higher standard deviation

(Figure 4.7 and 4.8). Figure 4.5 and 4.6 are an illustration of a watershed which has low percent of area changed at higher standard deviation.

The percent of area changed by ArcView and TOPAZ for each watershed is plotted against standard deviation in Figure 4.11. The graph shows a general trend that with increasing standard deviation the total percent of area changed decreased. It can be inferred from the figure that there is generally a decreasing trend for ArcView and TOPAZ, with percent of area changed for TOPAZ generally lower than ArcView. The percent of area changed and standard deviation of elevation are correlated negatively but it is not a clearly defined relationship (Figure 4.11).

The size of the watersheds ranged from 462 hectares to 19,959 hectares (Table 4.2). Figures 4.12, 4.13 and 4.14 show the relationship between different sizes of watershed, standard deviation and percent of area changed for ArcView and TOPAZ. The watersheds were categorized in four groups as follows:

1). Watersheds greater than 900 hectares

2). Watersheds greater than 900 hectares but less than 1400 hectares

3). Watersheds greater than 1400 hectares but less than 2200 hectares

4). Watersheds greater than 2200 hectares.

This categorization was done to evaluate any possible effect which watershed size might have on percent of area changed or on standard deviation. Figure 4.11 along with Figure 4.13 and Figure 4.14 show that there is little influence of watershed size on either percent of area changed or on standard deviation. There is no evidence of correlation between watershed size and percent of area changed or standard deviation, and this conclusion holds for both ArcView and TOPAZ.

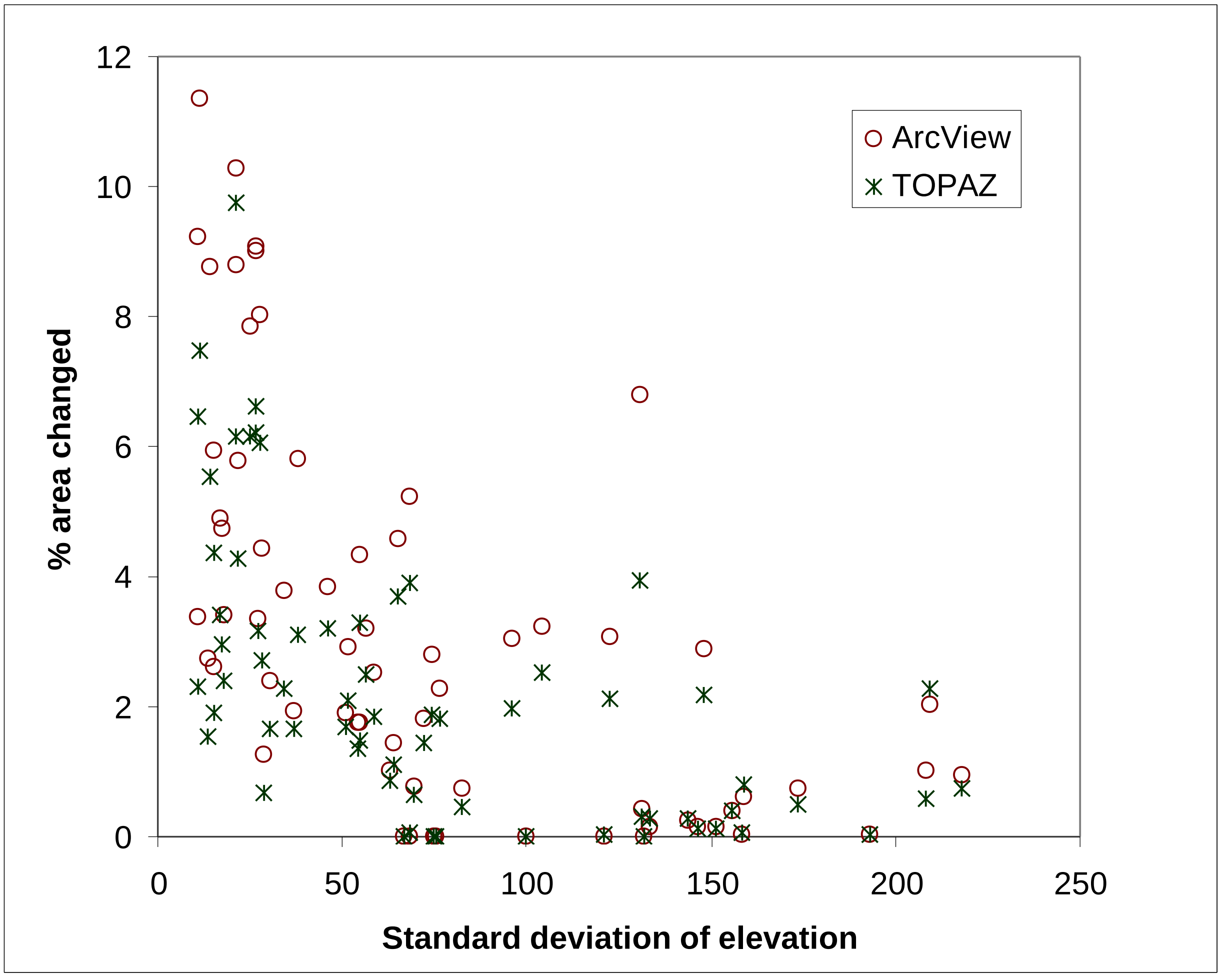

Figure 4.11. Relationship between standard deviation of elevation and percent of area changed.

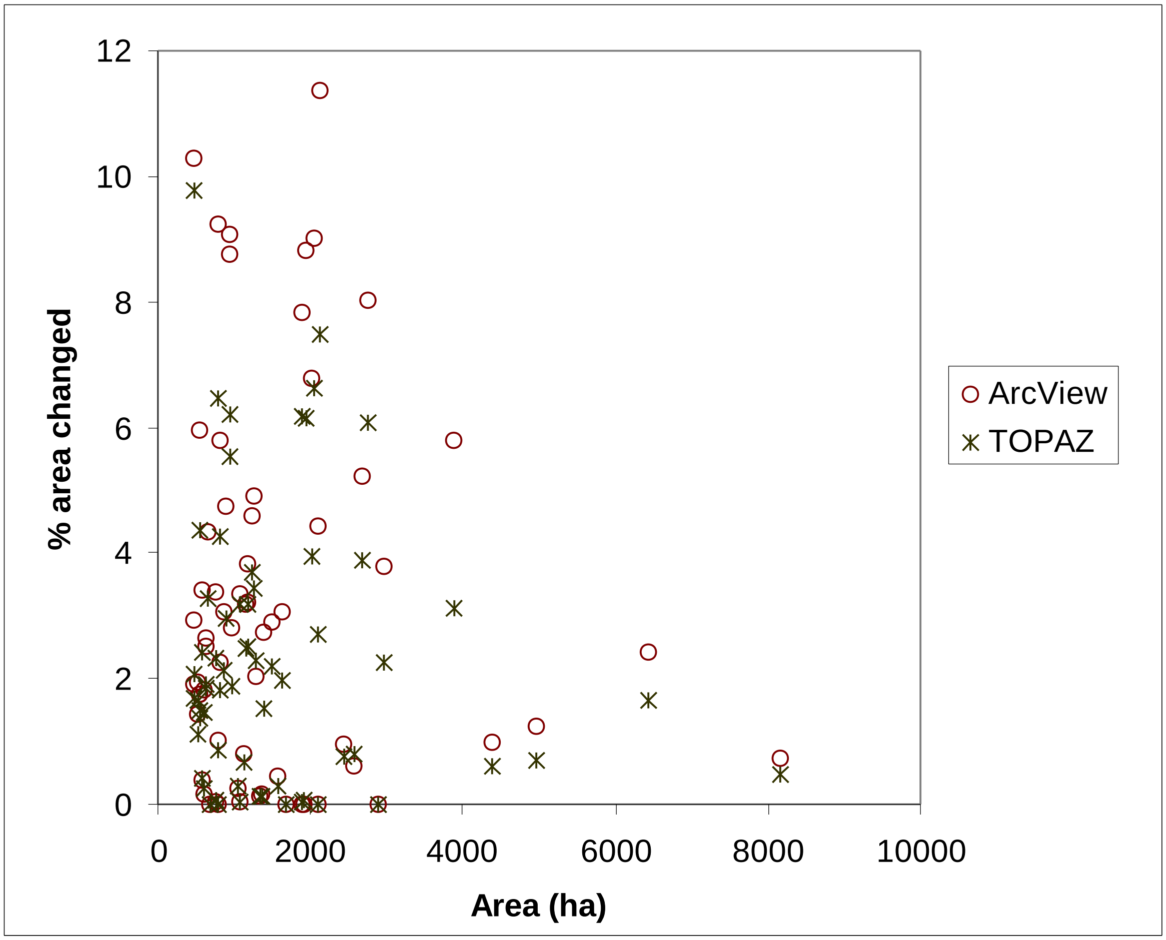

Figure 4.12. Relationship between watershed area (hectares) and percent of area changed in removing depressions.

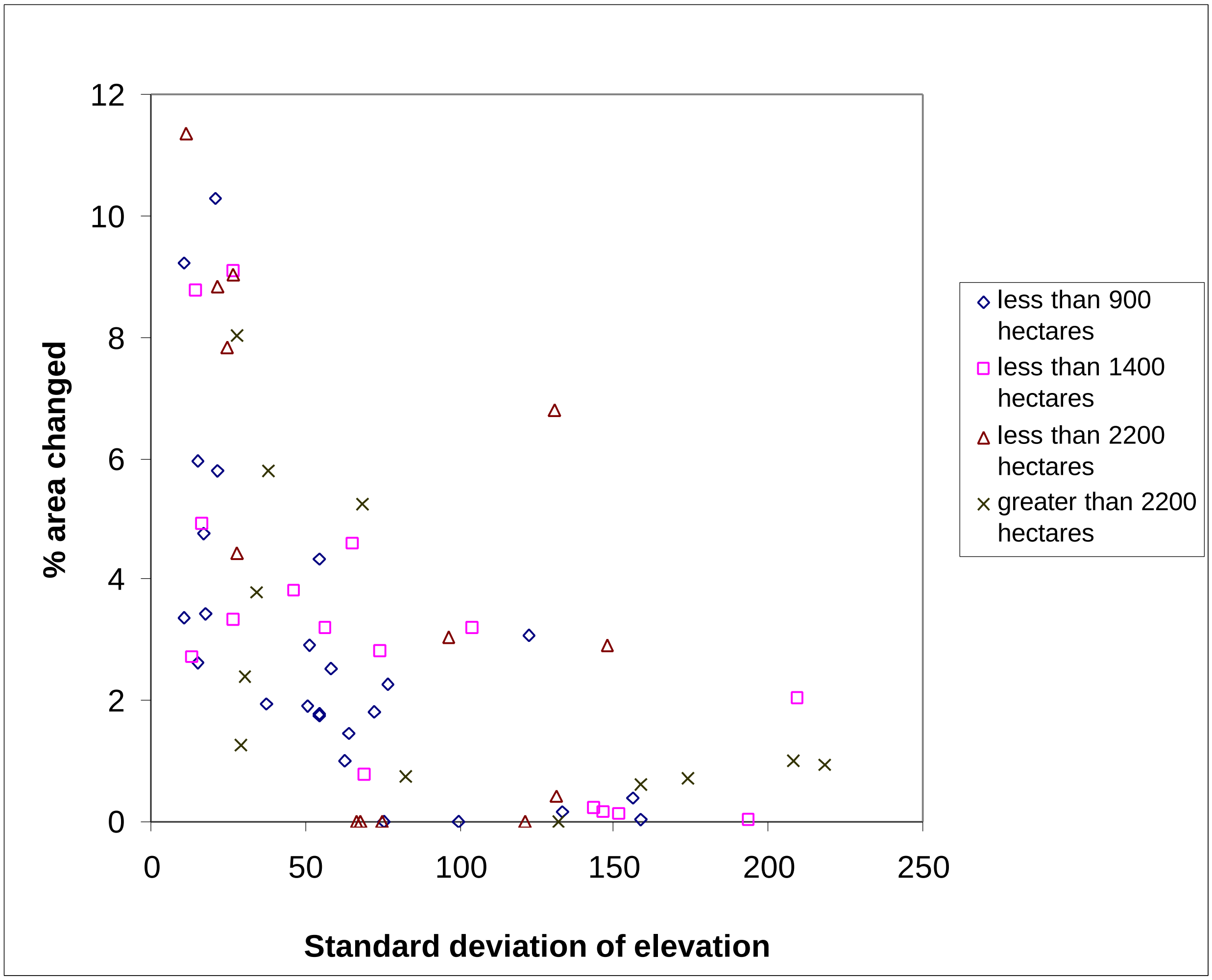

Figure 4.13. Percent of area changed (ArcView) as a function of standard deviation of elevation for four watershed size groups.

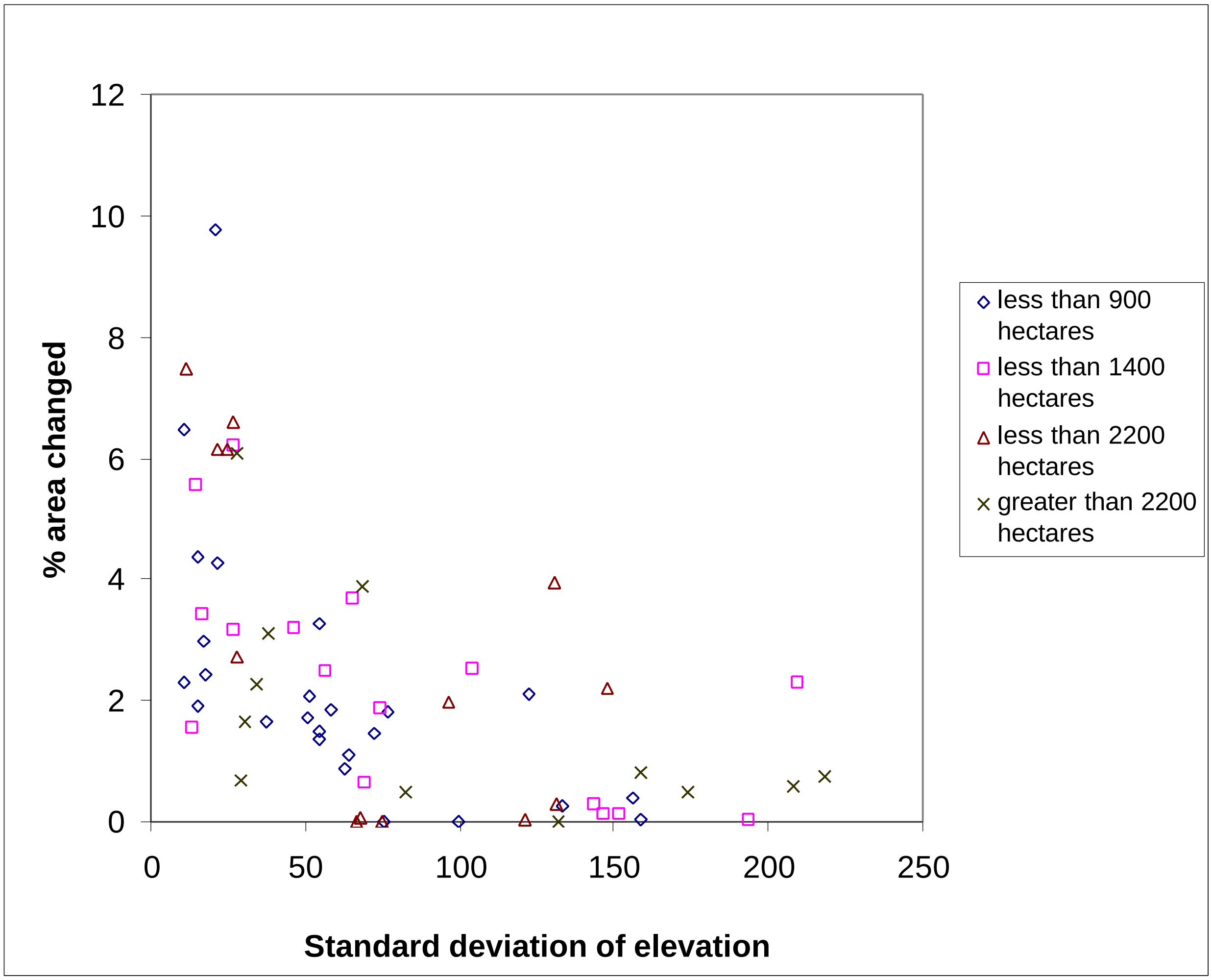

Figure 4.14. Percent of area changed (TOPAZ) as a function of standard deviation of elevation for four watershed size groups.

4.2.2 Range of elevation

Elevation range was used as a measure to characterize watershed topography. Sample range may not be a better measurement of variability as compared to standard deviation but it is easily understood (Ott, 1992).

Table 4.1 shows the values of the range of elevation used for the watersheds. The values ranged from 60 m to 915 m with the mean of 370 m and median of 328 m (Table 4.2). The high standard deviation of 245 m suggests that there was wide variability between the values of range of elevation. It can be observed from Figure 4.15 that there is a definite correlation between range of elevation and percent of area changed for both ArcView and TOPAZ.

4.2.3 Mean slope

Mean slope provides a quantitative measure of landform (O'Callaghan and Mark, 1984) and is calculated here as the mean of the slope of all grid cells in the watersheds. The lowest value of mean slope was 3.1% and the highest was 22.7% with a mean value of 11.5%. The median was 11.9% and standard deviation of the mean slope was 5.8% (Table 4.2). Mean slope for each watershed is plotted against percent area changed in Figure 4.16. It can be observed from the figure that as the mean slope increases the percent area changed decreases. The overall distribution of data points is similar to previous figures relating percent area changed to standard deviation of elevation. The values of mean slope range from 3% (flat areas) to 23% (steep areas). There is more fluctuation in the values of percent area changed at low slope. One possible reason for this is that in areas of low slope the effects of error which may be present in the data set

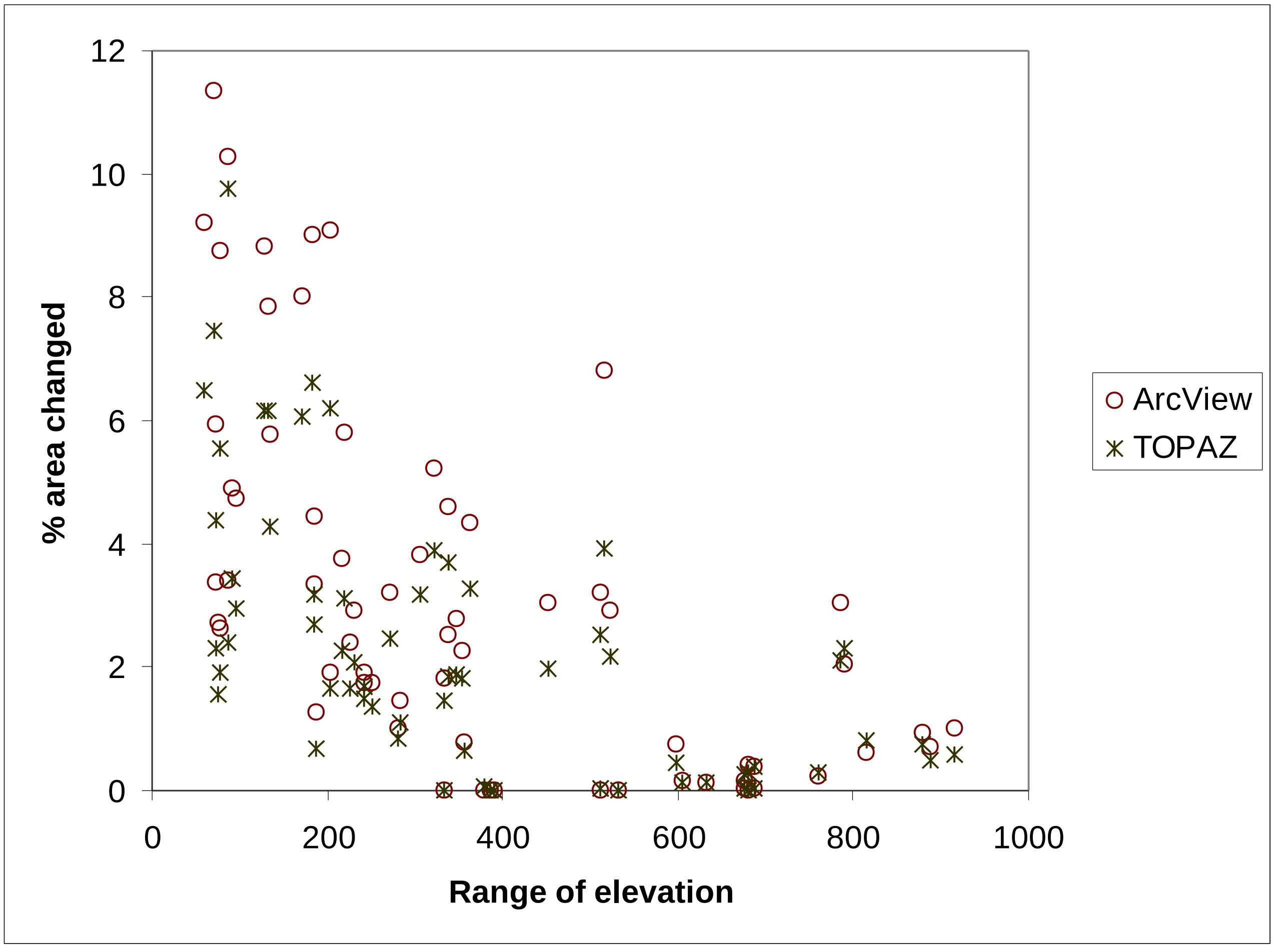

Figure 4.15. Relationship between range of elevation and percent area changed.

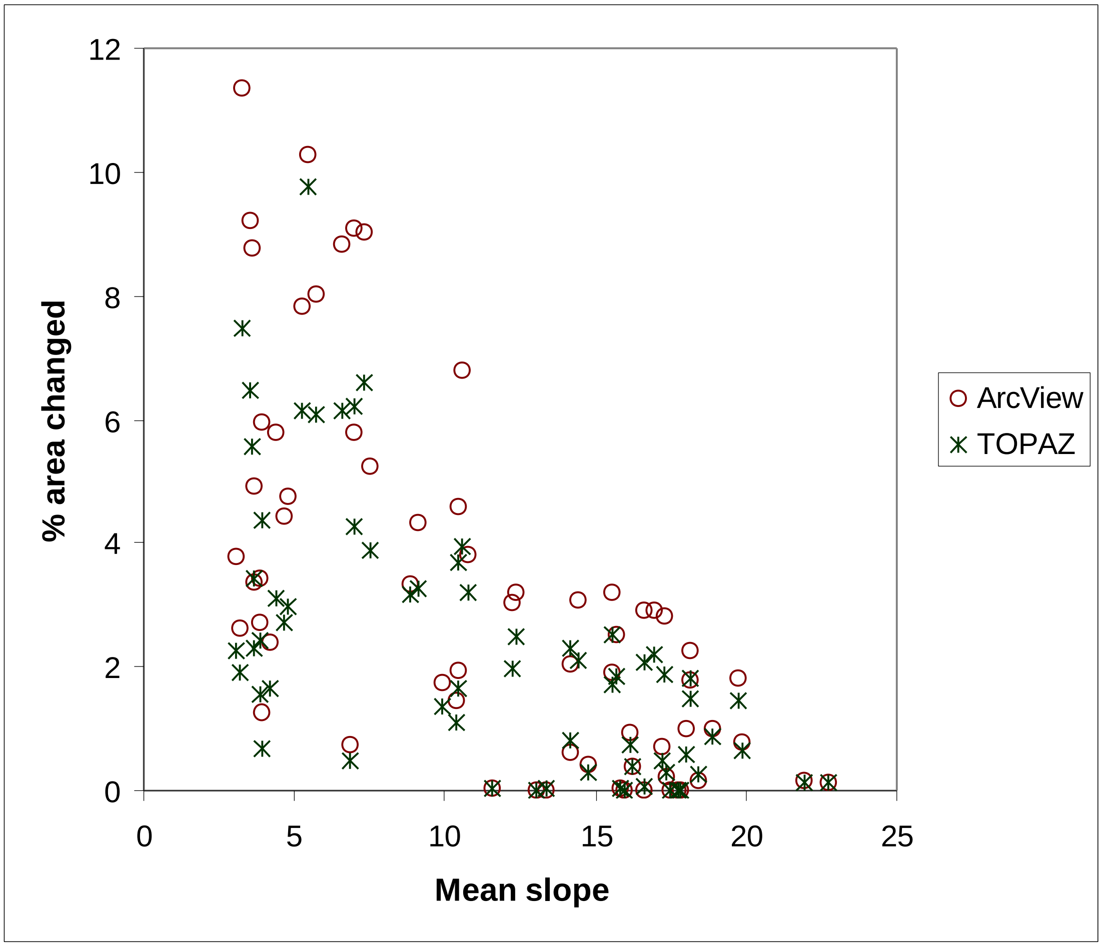

Figure 4.16. Relationship between mean slope and percent area changed.

may be greatest (Veregin, 1997). In general, the trends are similar to those for standard deviation (Figure 4.11) and range (Figure 4.15). It can also be observed that the general trend is similar for both ArcView and TOPAZ although there are fewer cells changed for TOPAZ in comparison with ArcView.

4.2.4 Curvature of elevation surface

Curvature of elevation, like slope and aspect, is a derivative of the elevation surface. Curvature is the derivative of slope and indicates areas where slope is changing. Figure 4.17 does not suggests strong evidence of any correlation between curvature and percent of area changed. It can be observed from the figure that there is high variability in curvature across a range of values, and that the general pattern of decreasing variability seen with the other variables is not evident here.

The presence of random and systematic error might be one reason for the lack of correlation between mean curvature and percent of area changed. The effect of errors in DEM data generally become more acute for higher order derivation topographic surfaces (Giles and Franklin, 1996). Another possible explanation for the absence of a clear relationship between mean curvature and percent of area changed might be that the curvature values are too small in magnitude. Thus, any error which might be present will have greater impact on these values.

As an alternative basis for analysis, the absolute value of curvature was classified in 5 classes (Table 4.3). The 5 categories in which the absolute curvature map was classified were absolute values greater than 1, 2, 3, 4 and 5. The number of cells belonging to each category were computed and generally the steep watersheds had higher

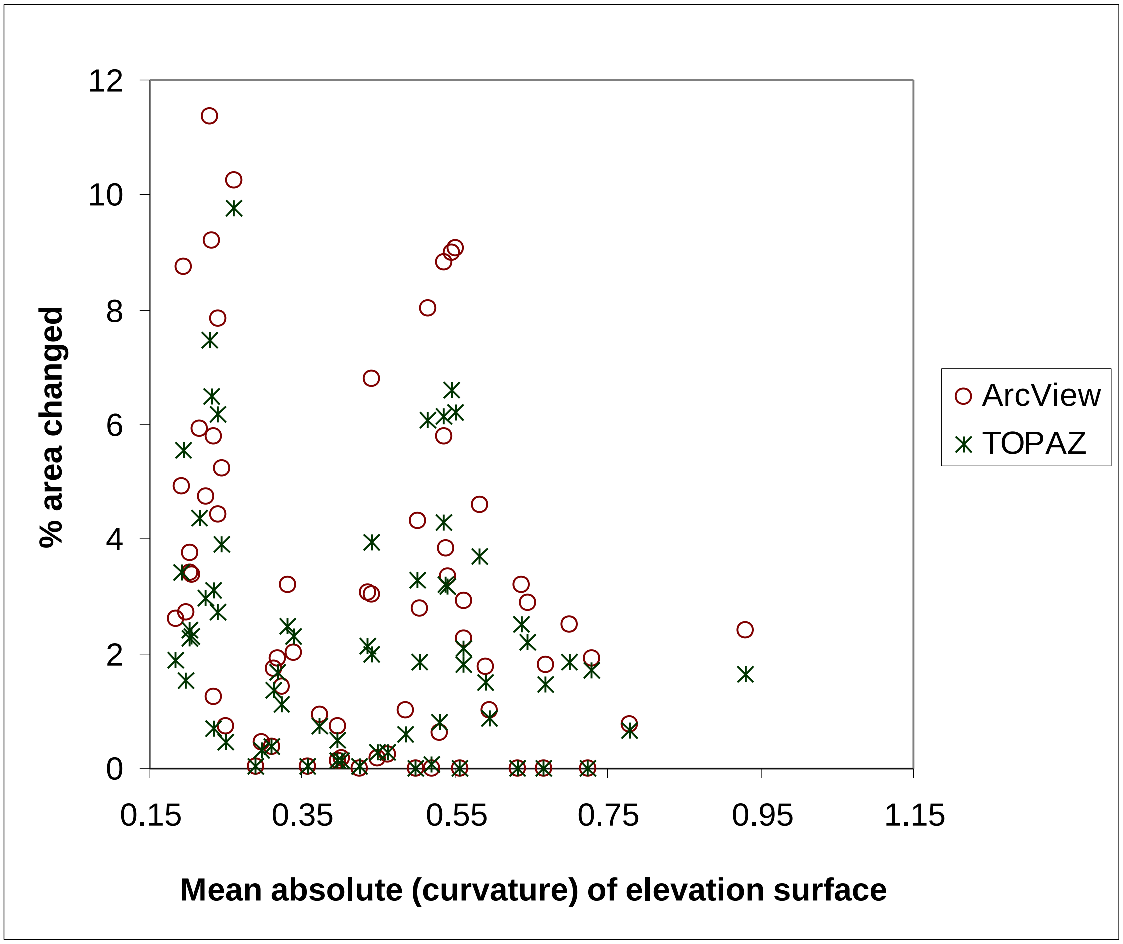

Figure 4.17. Relationship between mean absolute value of curvature and percent area changed.

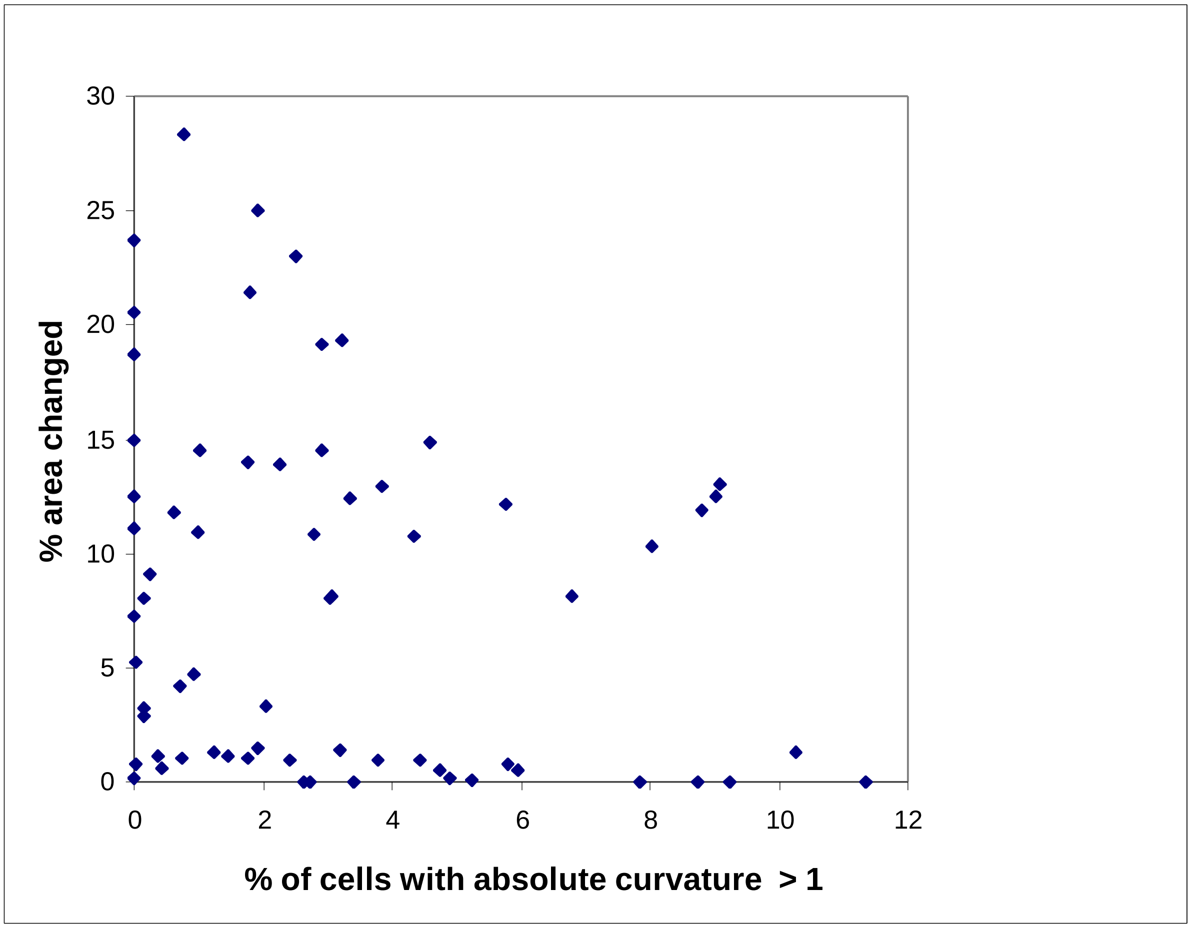

Figure 4.18. Relationship between percent of cells with absolute value of curvature greater than 1 and percent area changed (ArcView).

Table 4.3. The table shows 5 categories in which the absolute curvature map was classified. The table is sorted by standard deviation in ascending order.

		Percent of cells with absolute value of curvature greater than				
Watershed	Std. Dev of Elev.	1	2	3	4	5
SoH3	11.0	0.02				
Ft2	11.0	0.21	0.04			
SoH2	11.4	0.03				
Pl2	13.5	0.00				
Pl3	14.3	0.01				
Ft7	15.0	0.55	0.19	0.16	0.03	
Ft10	15.5	0.00				
Pl4	16.8	0.14				
VH2	17.1	0.50				
Ft14	18.0	0.02				
VH3	21.1	1.31	0.04			
MossyCr	21.4	11.90	0.33	0.03		
Hf3	21.8	12.20	0.35	0.01		
MV3	24.8	0.02				
Hf2	26.5	13.02	0.33	0.01		
Bb3	26.5	12.47	0.33	0.01		
Bb2	26.9	12.40	0.33	0.01		
Longgl	27.8	10.29	0.25			
PlRun	28.2	0.98	0.01			
BR	29.0	1.29	0.07	0.01	0.01	
CCr	30.3	0.94	0.01			
DRl	34.2	0.96	0.05			
GC3	37.1	1.51				
Mill	37.9	0.83	0.01			
Gl2	46.1	12.95	0.63	0.04	0.01	
Cn6	50.7	24.98	3.11	0.10	0.04	
Cn10	51.6	14.51	2.47	0.45	0.06	
Bl2	54.3	1.03				
Gl3	54.7	10.76	0.20			
No3	54.9	14.03	0.08			
GC2	56.3	1.39				
Av3	58.4	23.04	2.06	0.06		
No2	63.1	14.49	0.09			
Bl3	64.0	1.16				
Av2	65.1	14.90	0.75	0.11	0.03	0.02
Tw3	66.7	11.11	1.25	0.08	0.02	
Tw2	68.2	12.49	1.76	0.22	0.01	
MV2	68.3	0.05				
Cn17	69.5	28.31	3.10	0.20	0.05	0.02
Cn14	72.2	21.43	1.67	0.15	0.06	0.02
Cn12	74.4	10.87	0.79	0.04		
Be2	75.1	20.58	3.33	0.36	0.03	

……..continued………..

Table 4.3 ……….continued …….

Watershed	Std. Dev of Elev.	Percent of cells with absolute value of curvature greater than 1	2	3	4	5
Be3	75.6	23.74	5.09	1.06	0.22	0.02
Cn11	76.6	13.94	1.58	0.13		
MC	82.4	1.03	0.06	0.01		
Sv3	96.2	8.06	0.79	0.06	0.02	0.01
Su3	99.8	14.96	2.58	0.30	0.01	
Stu3	104.0	19.29	1.54	0.05		0.01
CG2	120.9	7.25	0.71	0.08	0.02	
Fle3	122.6	8.10	0.83	0.06		
Sv2	130.5	8.12	1.17	0.29	0.10	0.02
Gn6	131.2	0.57	0.01			
Su2	131.7	18.68	4.02	0.79	0.13	0.02
Gn10	133.4	8.01	0.34			
Gn16	143.6	9.11	0.28	0.01		
St4	146.6	3.23	0.01			
Stu2	148.1	19.19	1.96	0.22	0.09	0.05
St1	151.6	2.90				
Gn5	155.8	1.12				
Gn4	158.4	0.81				
St3	158.9	11.79	0.81	0.23	0.12	0.08
DRu	173.6	4.16	0.28	0.09	0.03	0.01
CG3	193.3	5.28	0.70	0.11	0.01	
Fle2	208.2	10.94	1.56	0.28	0.07	0.02
Ct2	209.5	3.31	0.08			
Ct3	218.0	4.77	0.31	0.03		

number of cells with absolute curvature values greater than 2 and 3, whereas the majority of the cells in flatter watersheds had values less than 1. Figure 4.18 shows a general declining trend in the percent area changed with increases in the percent of cells having absolute curvature greater than 1.

4.2.5 Channel slope, watershed slope and drainage density

The slope of the watershed and channel was also considered as a measure of the variability in relief. Channel slope is the rise of the main drainage channel over the length of drainage channel. The watershed slope is calculated by extending the channel to the watershed boundary and using that maximum flow path length and the elevation difference to calculate slope. The values of channel slope ranged from 0.004 (m/m) to 0.123 (m/m) and those of watershed slope had a range from 0.006 (m/m) to 0.143 (m/m) (Table 4.2). The mean of watershed slope was slightly higher at 0.045 (m/m) as compared to the mean of channel slope having a value of 0.038 (m/m). Figure 4.19 and Figure 4.20 show the relationship of percent area changed with channel slope and watershed slope respectively. Both the figures appear very similar in their relationship with percent of area changed. It can be observed that there is a decline in the percent of area changed as the slope value increases.

Drainage density was also explored for possible correlation with the number of cells with changed elevations. The drainage density had a minimum value of 0.0017 (m^2/m), maximum of 0.0028 (m^2/m), mean and the median were 0.0022 (m^2/m), and standard deviation of 0.0002 (m^2/m) (Table 4.2).

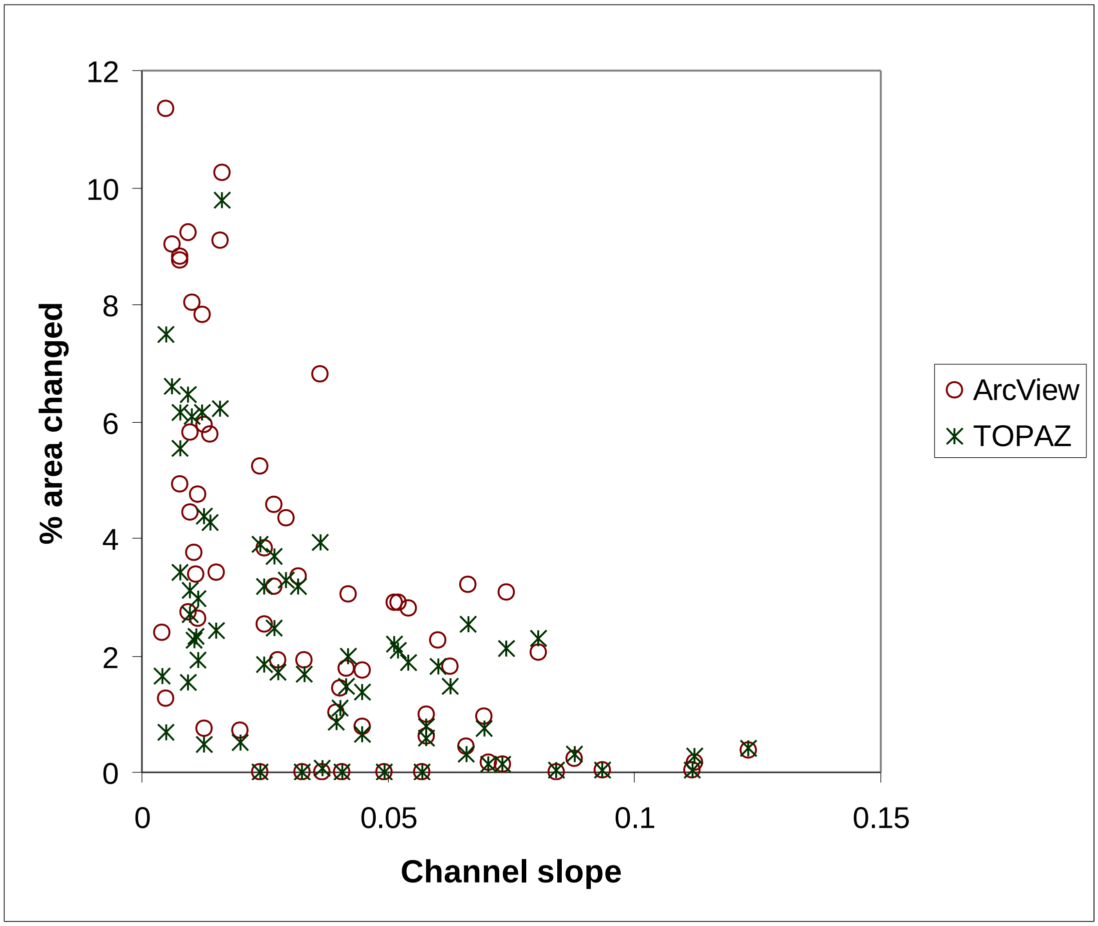

Figure 4.19. Relationship between channel slope and percent area changed.

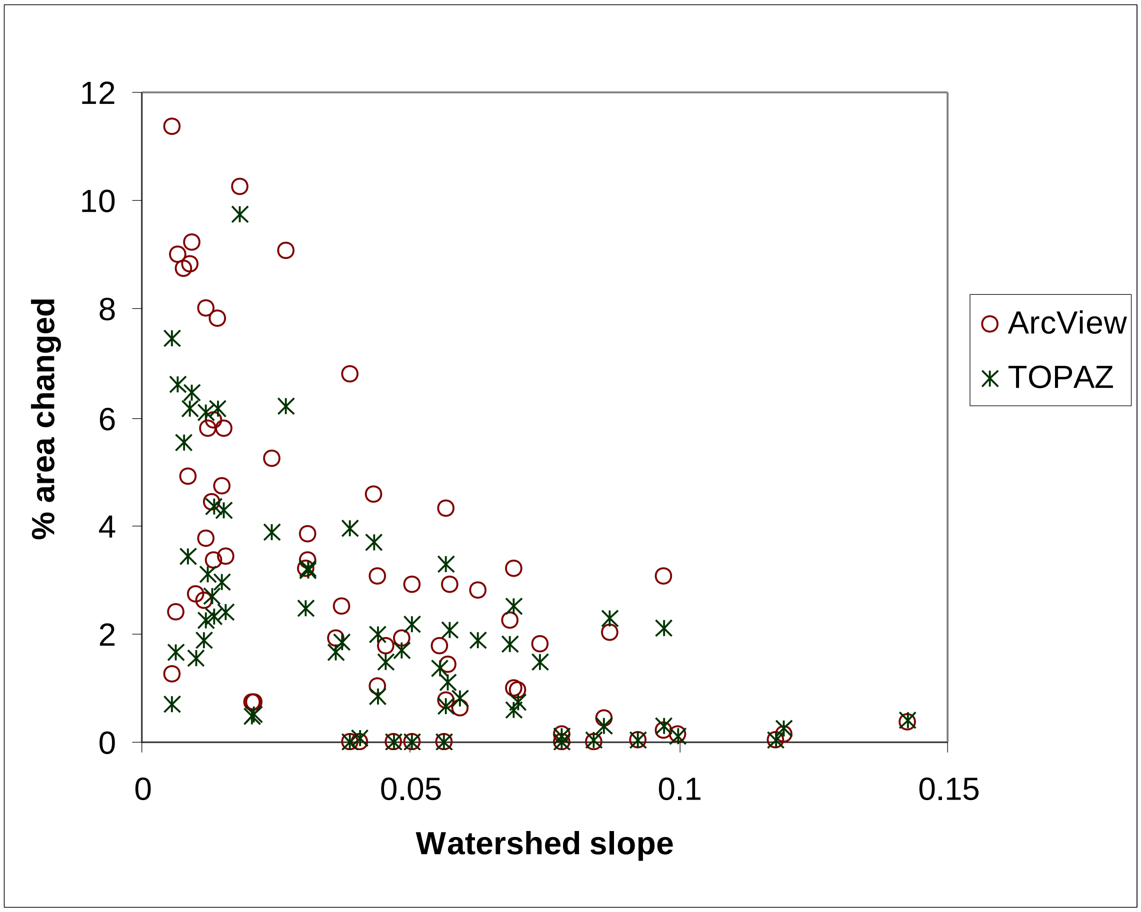

Figure 4.20. Relationship between watershed slope and percent area changed.

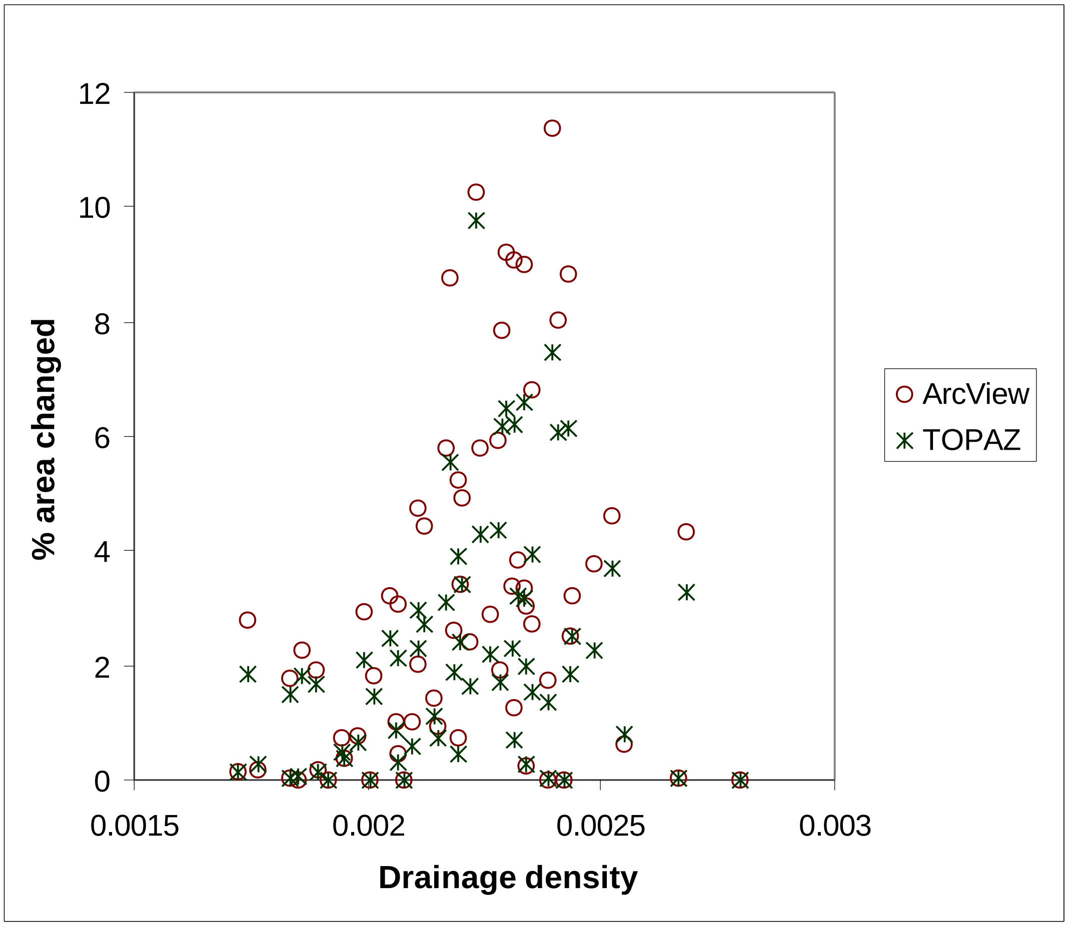

Figure 4.21. Relationship between drainage density and percent area changed.

Figure 4.21 shows the relationship between drainage density and percent of area changed and there seems to be no correlation between drainage density and percent of area changed. From the figure, it seems that there is high variation in the values of percent area changed within a very narrow range of drainage density but no definite trend is evident.

4.3 Spatial distribution of modified elevation cells (cuts and fills)

The changed elevation cells vary in their spatial distribution. The spatial distribution of changed elevation cells was evaluated to determine if there is any relationship with the manner in which ArcView and TOPAZ remove depressions.

4.3.1 Relationship between modified elevation cells

Cells were grouped in clusters (including diagonals) to assess the degree to which cuts and fills occur independently or in coincidence with other modified cells. Figures 4.1 to 4.10 show the spatial location of fill cells for ArcView along with TOPAZ (fills and cuts). Clusters formed by fill cells of ArcView and TOPAZ are larger in sizes. It can be observed from the figures that almost all the clusters of cut cells are composed of either one or two cells which is in contrast to the larger clusters formed by cells which were filled. The summary of clustering results for each watershed is included in Table 2 of Appendix C.

A detailed example for one watershed is shown in Table 4.4. It can be observed from the table that among the clusters of fill cells for ArcView and TOPAZ there are

Table 4.4. The table below shows the number of clusters by cluster size for watershed SoH2 for ArcView and TOPAZ.

Number of cells in cluster	Number of clusters			
	ArcView	TOPAZ		
		fills and cuts	cut cells	filled cells
1	12	144	131	32
2	44	98	112	20
3	22	18	2	12
4	24	18	2	15
5	16	6		8
6	11	10		9
7	7	5		2
8	8	9		7
9	9	2		1
10	2	2		2
>10	59	29		29

some clusters which consist of more than 10 cells whereas there are no more than 4 cells in the largest clusters of cut cells. The majority of ArcView and TOPAZ fill cells are composed of clusters having more than 2 cells and this is true for almost all watersheds (Table 2 in Appendix C).

4.3.2 Relationship between modified elevation cells with respect to proximity to channel network

From a hydrologic and non-point source modeling point of view, the distance at which the modified cells are located from the stream network is of considerable importance. It can be observed from Figures 4.1 to 4.10 that ArcView and TOPAZ fill cells are mainly located on or near the streams whereas the TOPAZ cut cells are distributed all over the watershed. The distributions of modified cells by distance from the flow network at 0, 1-100, and greater than 100 m distance are shown in Table 4.5. It can be observed from the table that for ArcView, fills within 100 m of the flow network with a mean of 2.68% ranges from 0 to 9.11% of the watershed area, while for TOPAZ, the equivalent range is from 0 to 7.98% with a mean of 1.43%. The lower mean for TOPAZ fills suggests that TOPAZ has less impact on the area within 100 m of the stream when compared to ArcView. The percent area changed that is greater than 100 m from the stream ranges from 0 to 2.59% with a mean of 0.47% for ArcView, and ranges from 0 to 0.90% with a mean of 0.12% for TOPAZ. The area changed by TOPAZ cuts ranges from 0 to 1.93% (mean of 0.57%) within 100 m of stream, and 0 to 1.07% (mean of 0.17%) for distances greater than 100 m from the stream.

Table 4.5. Proximity to stream of ArcView fills and TOPAZ fills and cuts.

	ArcView (fills)		TOPAZ (fills)		TOPAZ (cuts)	
	Distance to stream (m)		Distance to stream (m)		Distance to stream (m)	
Watershed	0-100	>100	0-100	>100	0-100	>100
	---------------------------------------(percent of watershed area)---------------------------------------					
Av2	3.69	0.90	1.57	0.21	1.38	0.54
Av3	2.23	0.29	0.95	0.03	0.67	0.20
Bb2	2.64	0.71	1.09	0.19	1.29	0.60
Bb3	7.15	1.86	3.77	0.50	1.27	1.07
Be2	0.00	0.00	0.00	0.00	0.00	0.00
Be3	0.00	0.00	0.00	0.00	0.00	0.00
Bl2	1.71	0.05	0.83	0.05	0.45	0.03
Bl3	1.45	0.00	0.64	0.00	0.40	0.07
BR	1.07	0.19	0.25	0.07	0.27	0.09
CCr	1.89	0.52	1.20	0.10	0.28	0.08
CG2	0.00	0.00	0.00	0.00	0.03	0.00
CG3	0.03	0.00	0.00	0.00	0.03	0.00
Cn10	2.43	0.49	1.05	0.12	0.72	0.19
Cn11	2.22	0.04	1.28	0.00	0.48	0.04
Cn12	2.73	0.07	1.16	0.01	0.65	0.05
Cn14	1.80	0.02	0.93	0.00	0.49	0.05
Cn17	0.73	0.05	0.32	0.00	0.28	0.06
Cn6	1.91	0.00	0.62	0.02	0.95	0.12
Ct2	1.93	0.10	2.04	0.08	0.14	0.03
Ct3	0.93	0.01	0.43	0.02	0.27	0.03
DRl	2.68	1.10	1.27	0.27	0.52	0.21
DRu	0.68	0.04	0.36	0.01	0.13	0.00
Fle2	0.96	0.04	0.34	0.00	0.24	0.00
Fle3	2.90	0.17	1.62	0.03	0.43	0.04
Ft10	1.74	0.89	0.86	0.21	0.51	0.31
Ft14	2.82	0.60	1.69	0.06	0.52	0.14
Ft2	2.81	0.56	1.30	0.08	0.73	0.20
Ft7	4.42	1.52	3.14	0.44	0.65	0.15
GC2	3.04	0.16	2.08	0.08	0.30	0.02
GC3	1.74	0.19	1.00	0.11	0.51	0.05
Gl2	3.69	0.14	1.90	0.04	1.12	0.13
Gl3	3.93	0.40	1.08	0.09	1.62	0.49
Gn10	0.16	0.00	0.06	0.00	0.19	0.01
Gn16	0.21	0.03	0.03	0.00	0.23	0.03
Gn4	0.00	0.02	0.00	0.00	0.02	0.02
Gn5	0.39	0.00	0.35	0.00	0.03	0.02
Gn6	0.44	0.00	0.20	0.00	0.09	0.01
Hf2	6.49	2.59	2.76	0.65	1.93	0.87
Hf3	4.50	1.28	1.18	0.37	1.78	0.94
Longgl	6.52	1.50	3.20	0.34	1.74	0.80

……..continued……..

Table 4.5…………..continued……

	ArcView (fills)		TOPAZ (fills)		TOPAZ (cuts)	
	Distance to stream (m)		Distance to stream (m)		Distance to stream (m)	
Watershed	0-100	>100	0-100	>100	0-100	>100
	-------------------------------------(percent of watershed area)-------------------------------------					
MC	0.70	0.04	0.24	0.00	0.21	0.02
Mill	4.83	0.97	2.08	0.15	0.75	0.13
MossyCr	7.55	1.27	3.51	0.27	1.74	0.63
MV2	4.96	0.27	3.13	0.08	0.56	0.12
MV3	6.64	1.20	4.92	0.26	0.77	0.21
No2	0.97	0.05	0.58	0.00	0.27	0.01
No3	1.76	0.00	1.06	0.00	0.43	0.00
PlRun	3.97	0.47	1.99	0.05	0.57	0.09
Pl2	2.34	0.39	0.79	0.04	0.55	0.17
Pl3	7.77	0.98	4.47	0.17	0.72	0.19
Pl4	4.41	0.49	2.54	0.11	0.63	0.15
SoH2	8.80	2.56	5.37	0.55	0.99	0.57
SoH3	7.04	2.18	4.50	0.70	0.82	0.45
St1	0.14	0.00	0.07	0.00	0.05	0.00
St3	0.53	0.08	0.06	0.04	0.59	0.11
St4	0.16	0.00	0.08	0.00	0.05	0.00
Stu2	2.71	0.20	1.22	0.05	0.79	0.13
Stu3	2.80	0.42	1.14	0.17	1.09	0.12
Su2	0.00	0.00	0.00	0.00	0.00	0.00
Su3	0.00	0.00	0.00	0.00	0.00	0.00
Sv2	5.83	0.96	2.81	0.18	0.76	0.19
Sv3	2.83	0.22	1.20	0.03	0.62	0.13
Tw2	0.01	0.00	0.01	0.00	0.00	0.00
Tw3	0.00	0.00	0.00	0.00	0.00	0.00
VH2	4.49	0.26	2.08	0.03	0.74	0.11
VH3	9.11	1.16	7.98	0.90	0.71	0.17
Summary Statistics						
Min	0.00	0.00	0.00	0.00	0.00	0.00
Max	9.11	2.59	7.98	0.90	1.93	1.07
Mean	2.68	0.47	1.43	0.12	0.57	0.17
Median	2.23	0.18	1.07	0.04	0.52	0.09
Std dev.	2.44	0.64	1.56	0.19	0.49	0.25

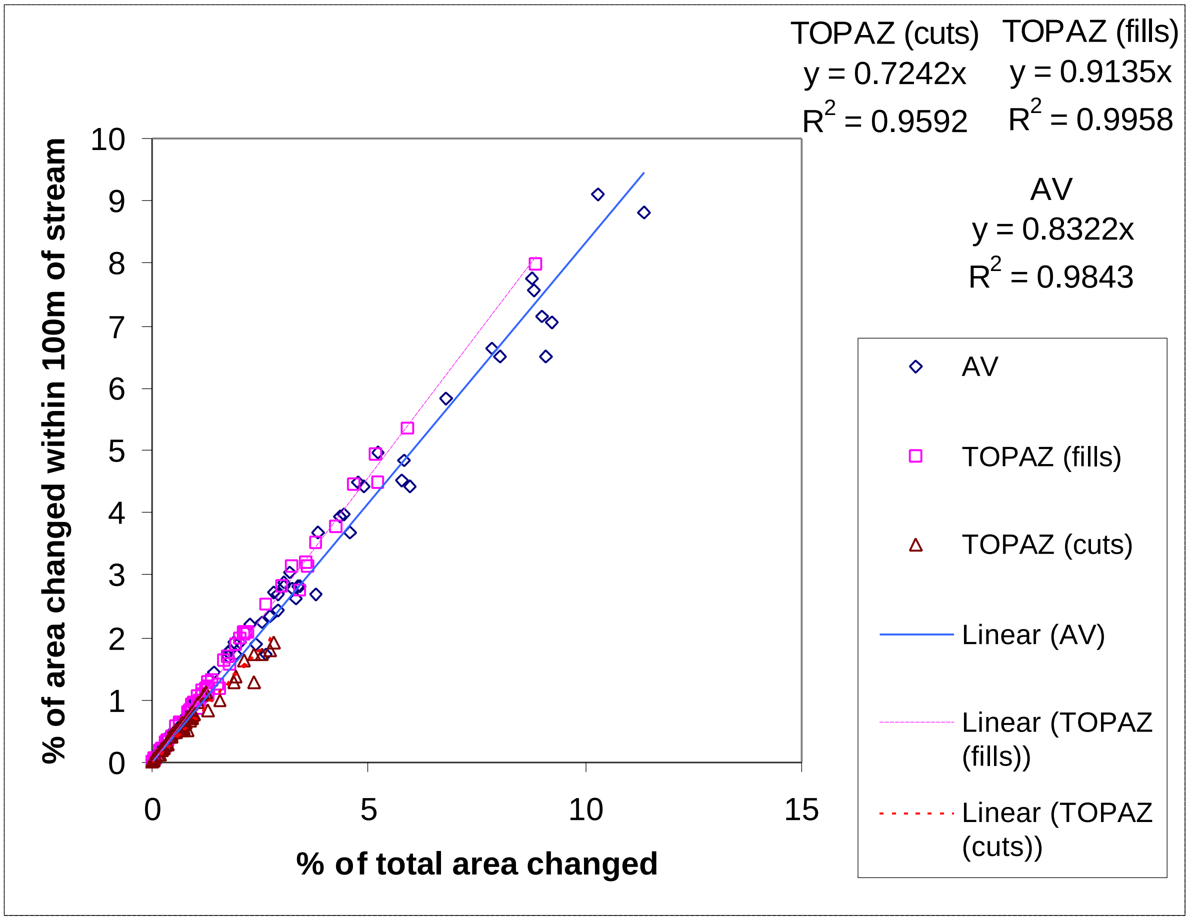

Figure 4.22. Relationship between percent of total area changed and percent of area changed within 100m of stream.

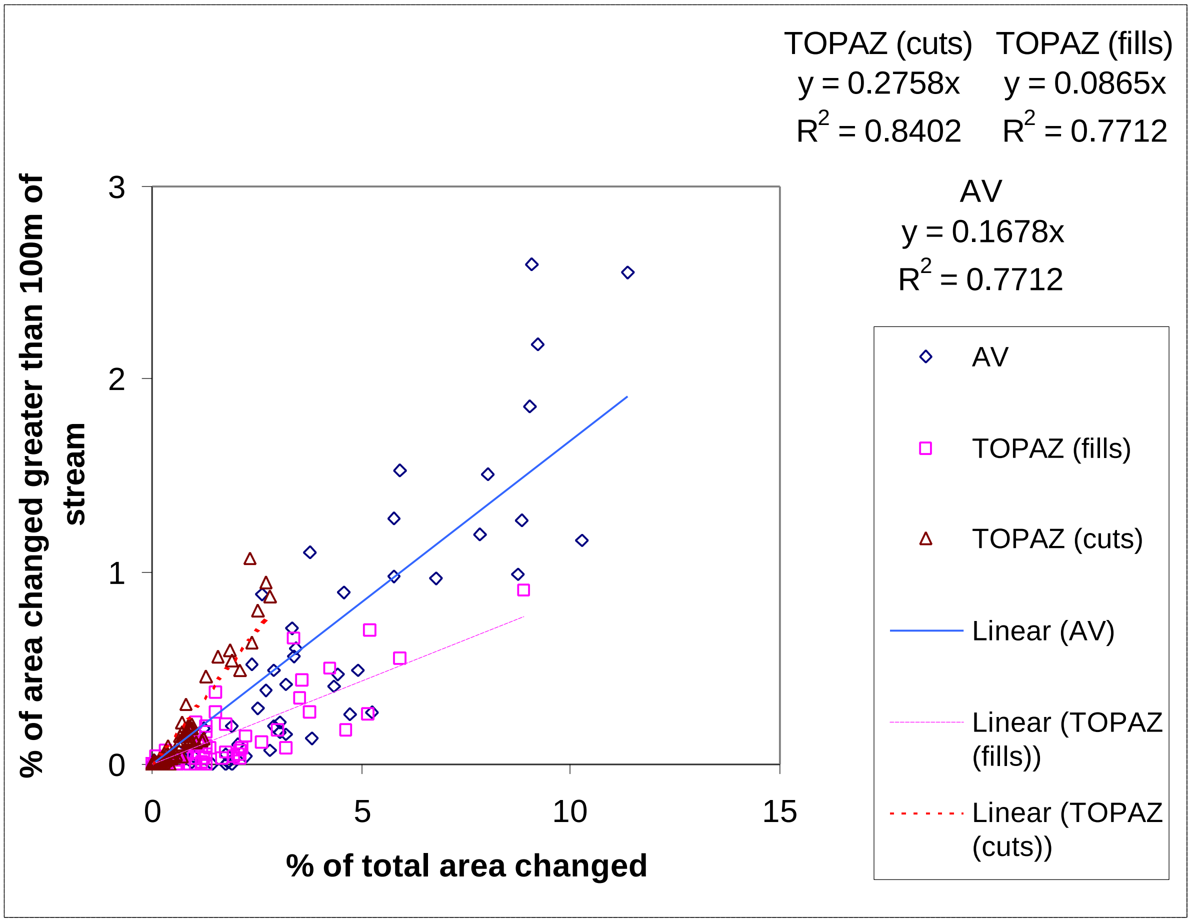

Figure 4.23. Relationship between percent of total area changed and percent of area changed which is greater than 100m from the stream.

Thus, it can be inferred that the depressions in the DEMs of this study are predominantly associated with the stream network.

Figure 4.22 shows high positive correlation between the total percent area changed and percent of area changed within 100 m of the stream. The percent of area changed within 100m of the stream increases with the increase in total percent area changed for ArcView (fills) and TOPAZ (cuts and fills). Figure 4.23 shows the relationship between total percent area changed and percent of area changed which is at a distance of greater than 100 m from the stream. The figure shows that for TOPAZ cuts the relationship is more linear when compared to TOPAZ fills and ArcView fills.

4.3.3 Proximity to stream and standard deviation of elevation

From Figure 4.11 it has already been established that percent area changed is correlated with standard deviation of elevation. It is also interesting to probe if similar relationship exists between standard deviation of elevation and percent of elevation cells changed which are on the streams. Figures 4.24, 4.25 and 4.26 show that with the increase in standard deviation there is a decrease in percent area changed which is similar to the relationship between standard deviation and percent area changed in Figure 4.11. It can also be observed from the figures that the magnitude of percent area changed is less for TOPAZ cuts (Figure 4.26) when compared to ArcView fills (Figure 4.24) and TOPAZ fills (Figure 4.25).

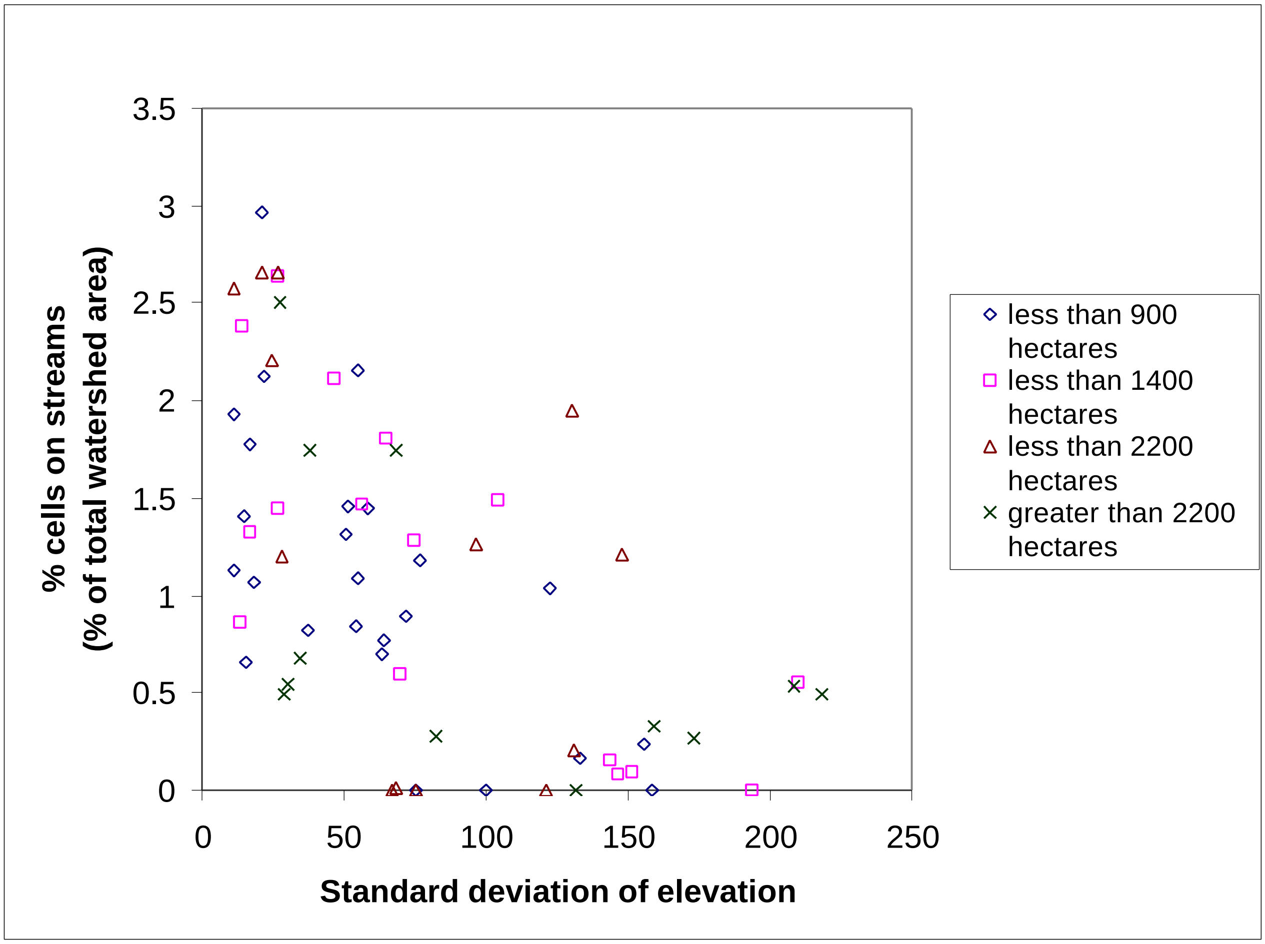

Figure 4.24. ArcView fills coinciding with stream cells as a function of standard deviation.

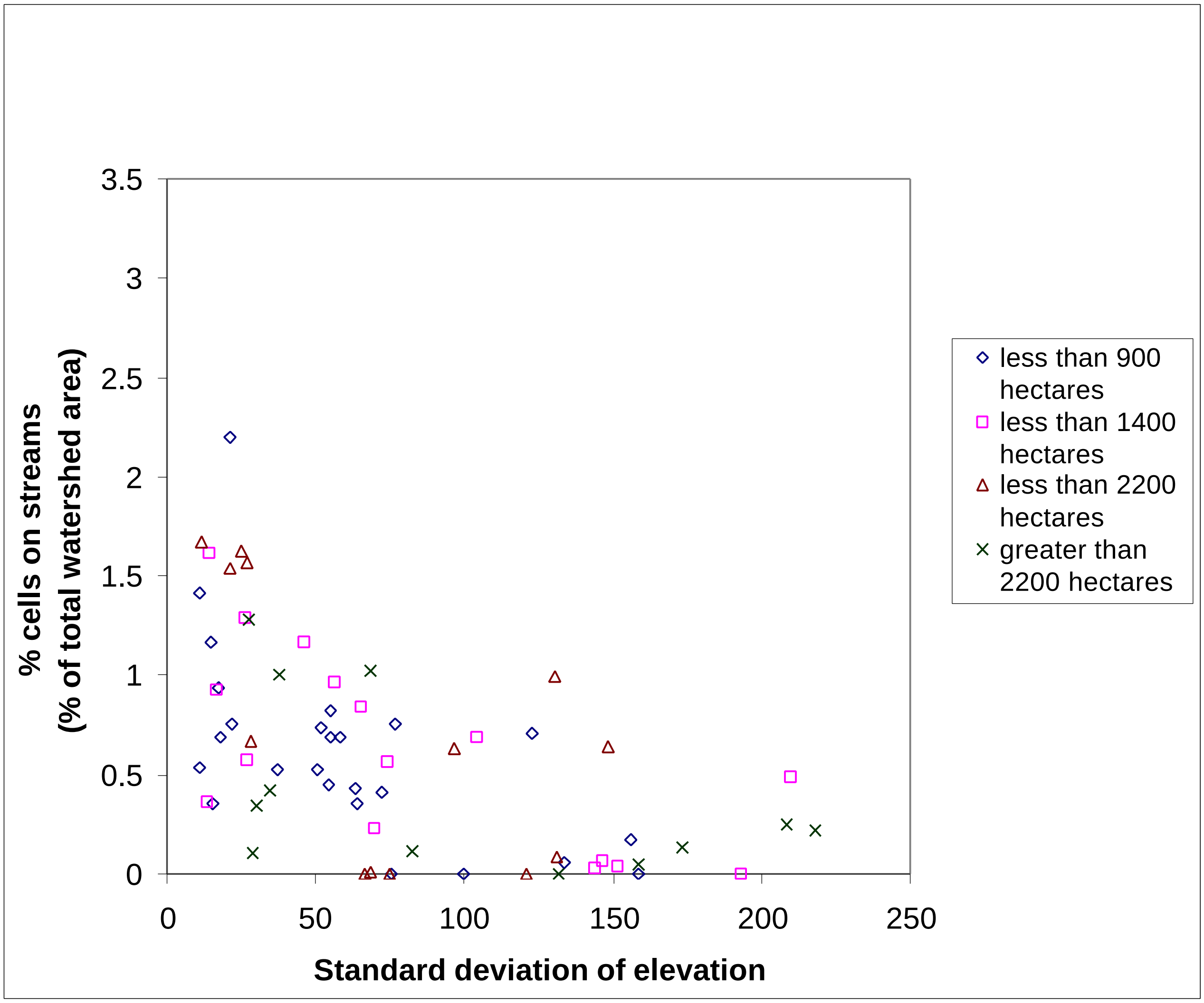

Figure 4.25. TOPAZ fills coinciding with stream cells as a function of standard deviation.

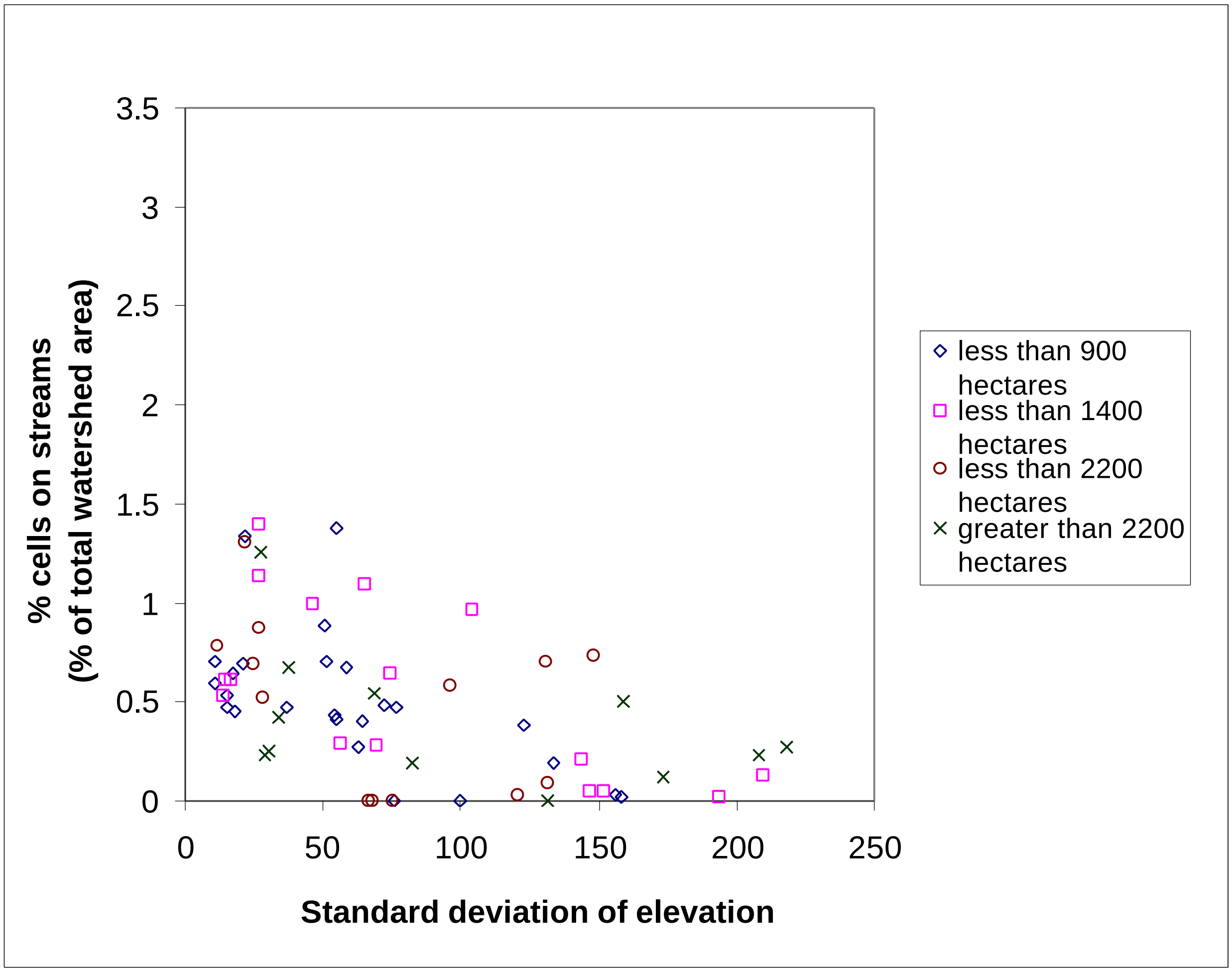

Figure 4.26. TOPAZ cuts coinciding with stream cells as a function of standard deviation.

4.4 Paired t test for changed elevation cells

The paired t test for means was used to evaluate if there was a significant difference in the percent of elevation cells changed by ArcView and TOPAZ. The watersheds were grouped in three categories (flat, average and steep) for this analysis. Twenty four watersheds belonged to the “flat” category, 24 to the “average” category, and 18 to the “steep” category.

Table 4.6 summarizes the result for flat watersheds. The p-value is much less than 0.05 and therefore, it can be concluded that mean number of elevation cells changed by ArcView is significantly greater than mean number changed by TOPAZ. For the “average” watersheds, (Table 4.7) the p-value is less than 0.05, it can be concluded that mean of ArcView affected elevation cells is significantly greater than that of TOPAZ affected cells.

For “steep” watersheds, the p-value is greater than 0.05 (Table 4.8), indicating insufficient evidence to suggest significant differences between ArcView and TOPAZ. Thus, we can conclude from the t test that TOPAZ performed better for flat and average watersheds in terms of percent of area changed because it induced fewer changes in the DEM data as compared to ArcView.

Table 4.6. Two sample paired t test for the means of ArcView and TOPAZ affected elevation cells for flat watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent area changed (ArcView)	5.79	2.97	23	0.0001
Percent area changed (TOPAZ)	4.12	2.32		

Table 4.7. Two sample paired t test for the means of ArcView and TOPAZ affected elevation cells for average watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent area changed (ArcView)	2.05	1.56	23	0.0001
Percent area changed (TOPAZ)	1.58	1.20		

Table 4.8. Two sample paired t test for the means of ArcView and TOPAZ affected elevation cells for steep watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent area changed (ArcView)	1.09	2.97	17	0.0627
Percent area changed (TOPAZ)	0.82	1.09		

4.5 Slope cells and their relationship with changed elevation cells

The values in the slope map are modified as a result of changes in the elevation map. Since slope maps are an important input to hydrologic and non-point source modeling, changes in the slope map may have serious implications on the output of a non point source model.

The modified slope cells due to changes in the DEM data set were identified and were reclassified in three categories i.e., cells which were on the stream, cells at a distance of 1-100 m from the stream and those which were at a distance of greater than 100 m from the stream.

4.5.1 Correlation between filled elevation cells and changed slope cells

Both ArcView and TOPAZ changed the cells in the slope map as a result of depression removal. Figures 4.27 and 4.28 show the relationship between changed elevation cells and modified slope cells for ArcView and TOPAZ respectively. The trend line has been fitted to both the figures in order to give concrete evidence of the existing relationship between changed slope and elevation cells.

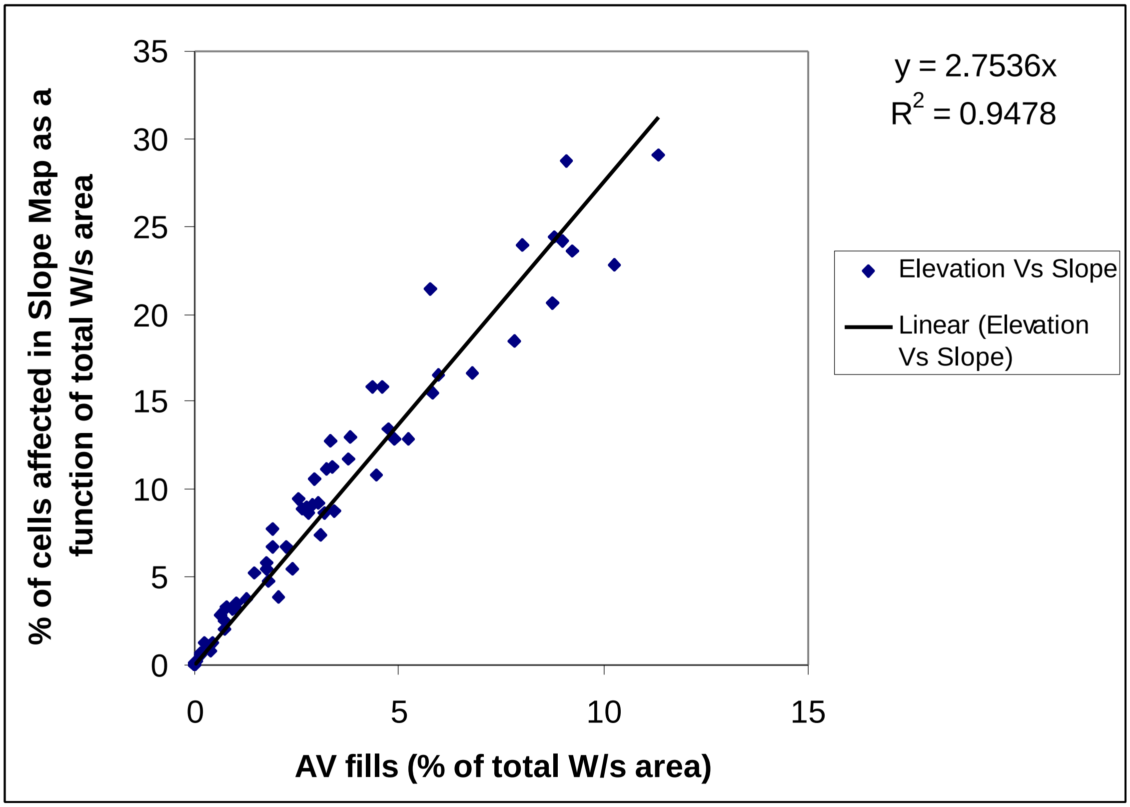

Figure 4.27. Relationship between number of elevation cells changed and number of slope cells changed (ArcView).

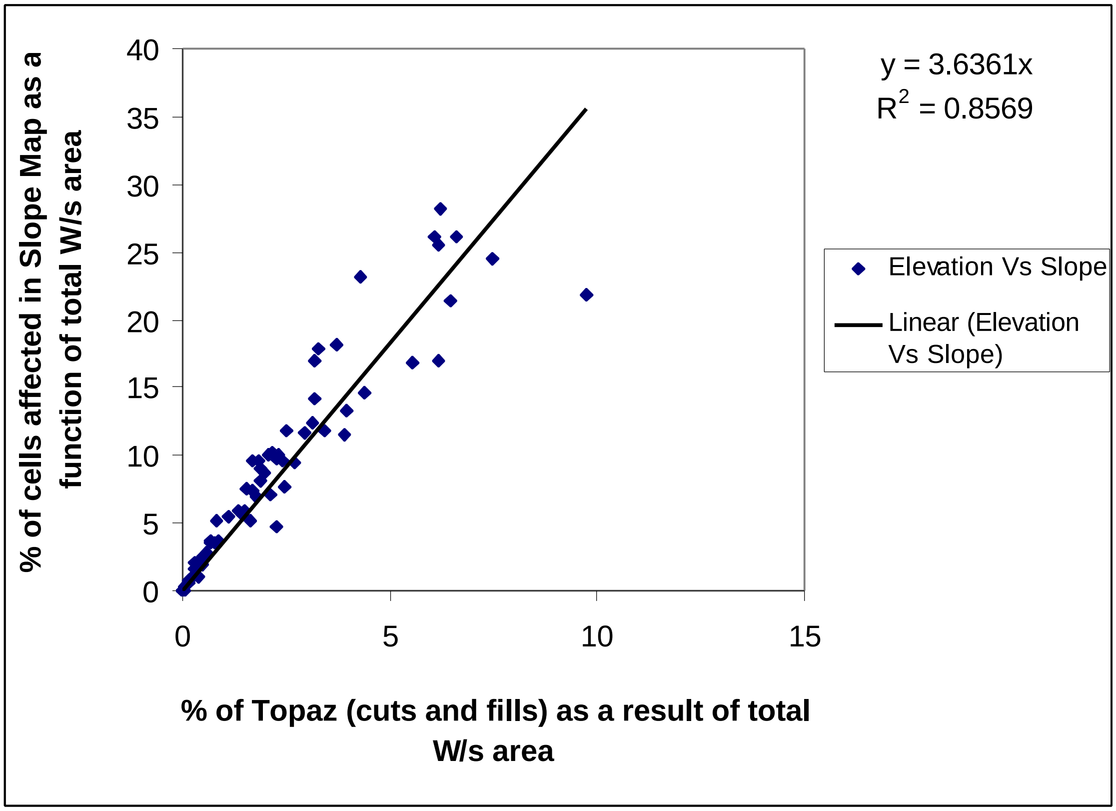

Figure 4.28. Relationship between number of elevation cells changed and number of slope cells changed (TOPAZ).

It can be observed from the figures that elevation cells modified by TOPAZ induces more changes in the slope map when compared to ArcView and this conclusion is supported by trend line equation shown in both figures. TOPAZ changes more slope cells because when modified elevation cells form large number of small clusters i.e., 1 or 2 cell cluster formed by cut cells then a greater number of slope cells are changed because the computation of slope is based on neighboring cells. Clusters formed by elevation cells changed by TOPAZ are greater in number but smaller in size whereas clusters formed by elevation cells changed by ArcView are comparatively smaller in number but larger in size affecting fewer neighboring cells, therefore less slope cells are modified by ArcView method of depression removal. There is high positive correlation between percent of changed elevation and slope cells for both ArcView and TOPAZ. It can be said that positive correlation between changed elevation and slope cells is marginally better for ArcView than for TOPAZ because the former has higher R^2 value (Figures 4.28 and 4.29).

4.6 Depth of modified elevation cells (fills and cuts)

The magnitude of cuts and fills was summarized based on the mean depth of the fills for ArcView and TOPAZ and cuts for TOPAZ. The modified elevation cells were classified according to the magnitude of the depth i.e., cells in the category of 1-2, 2-5, 5-9, and greater than 9 meters of depth (Table C3 of Appendix C).

Table 4.9 shows that mean depth for ArcView fill ranges from 0 to 10.29 m. Mean of TOPAZ fill depth ranges from 0 to 10.89 m and mean TOPAZ cut depth ranges from 0 to 4 m. The mean of mean fill depth for ArcView and TOPAZ is 2.59 m and

2.36 m whereas it is 1.54 m for TOPAZ cut. Thus, it can be observed that the mean depth of TOPAZ fills are similar to the mean depth of ArcView fill cells for most of the watersheds. The mean values of TOPAZ cuts are lower than the mean fill depths of ArcView and TOPAZ.

Table 4.9. Mean depth of ArcView fills and TOPAZ cuts and fills.

Watershed	Mean fill depth (AV)	Mean fill depth (TOPAZ)	Mean cut depth (TOPAZ)
Av2	2.83	2.57	-1.69
Av3	3.28	2.83	-2.15
Bb2	2.71	2.78	-1.58
Bb3	3.25	2.64	-1.61
Be2	0.00	0.00	0.00
Be3	0.00	0.00	0.00
Bl2	1.77	1.40	-1.24
Bl3	1.88	1.53	-1.18
BR	1.69	2.42	-1.20
CCr	2.26	2.23	-1.16
CG2	0.00	0.00	-1.00
CG3	1.25	0.00	-1.50
Cn10	4.24	3.68	-3.17
Cn11	4.24	4.03	-3.00
Cn12	4.91	4.35	-3.32
Cn14	10.29	10.89	-2.40
Cn17	2.80	3.63	-2.07
Cn6	2.12	1.64	-1.67
Ct2	4.05	4.24	-1.28
Ct3	2.24	2.21	-1.62
DRl	1.54	1.67	-1.06
DRu	2.81	3.37	-1.29
Fle2	3.96	3.36	-2.64
Fle3	5.15	3.87	-2.96
Ft10	1.41	1.33	-1.02
Ft14	1.96	1.59	-1.21
Ft2	1.52	1.57	-1.20
Ft7	2.02	1.96	-1.12
GC2	4.00	4.31	-1.37
GC3	2.55	2.56	-1.47
Gl2	4.20	4.18	-1.82
Gl3	2.38	2.17	-1.65
Gn10	2.55	2.75	-1.93
Gn16	1.79	1.67	-1.48
Gn4	1.50	0.00	-1.50
Gn5	4.48	3.78	-1.00
Gn6	2.64	2.35	-1.33
Hf2	2.98	2.84	-1.82
Hf3	2.62	2.65	-1.88
Longgl	2.72	2.42	-1.49
MC	1.41	1.42	-1.11
Mill	1.41	1.42	-1.11
MossyCr	2.67	2.11	-1.59
MV2	2.78	2.77	-1.34
MV3	3.16	3.36	-1.37
No2	2.51	2.25	-1.64

………continued………

Table 4.9……………continued……..

Watershed	Mean fill depth (AV)	Mean fill depth (TOPAZ)	Mean cut depth (TOPAZ)
No3	3.31	2.66	-1.69
PlRun	2.65	2.39	-1.49
Pl2	1.45	1.34	-1.12
Pl3	2.23	2.16	-1.20
Pl4	1.98	1.85	-1.07
SoH2	2.21	2.13	-1.28
SoH3	2.14	1.95	-1.35
St1	1.81	1.55	-1.00
St3	2.23	2.23	-1.73
St4	1.38	1.17	-1.00
Stu2	3.37	2.85	-2.02
Stu3	3.22	2.68	-2.03
Su2	0.00	0.00	0.00
Su3	0.00	0.00	0.00
Sv2	4.60	3.47	-2.62
Sv3	3.60	2.50	-2.55
Tw2	4.33	1.00	-4.00
Tw3	0.00	0.00	0.00
VH2	2.29	1.96	-1.77
VH3	3.79	4.84	-1.77
Summary Statistics			
Min	0.00	0.00	0.00
Max	10.29	10.89	4.00
Mean	2.59	2.36	1.54
Med	2.53	2.24	1.49
S. D.	1.57	1.6	0.76

4.7 Watershed delineation

The watershed boundaries used for analysis in this study were those delineated using ArcView. However, TOPAZ was also used to delineate the watersheds to assess differences in the closeness of the watersheds delineated by the two programs. Evaluated with reference to watershed roughness, Figure 4.29 shows that in general, the percent of unmatched area decreases with increasing standard deviation. In a few cases, there were differences of 2 to 3 percent of the watershed area which did not coincide between the spatial extent of ArcView and TOPAZ. Table 4.10 shows the extent to which ArcView and TOPAZ delineated areas were different. It can be observed that for the majority of the watersheds, less than 1 percent of the area is not common between ArcView and TOPAZ. Two watersheds are shown (Figures 4.30 and 4.31) to illustrate boundary differences.

Table 4.10. Differences in watershed areas as delineated byArcView and TOPAZ.

Watersheds	Area unique to ArcView	Area unique to TOPAZ	Combined Discrepancy
	----------------------------------(% of common area)	---------------------	
Av2	0.22	1.48	1.70
Av3	0.03	0.03	0.06
Bb2	3.39	0.43	3.81
Bb3	0.72	4.59	5.30
Be2	0.05	0.06	0.11
Be3	0.14	0.06	0.19
Bl2	0.18	0.45	0.63
Bl3	0.10	0.25	0.35
CG2	0.05	0.06	0.11
CG3	0.15	0.28	0.43
Ct2	0.14	0.32	0.46
Ct3	0.43	0.06	0.49
Fle2	0.03	0.04	0.07
Fle3	0.16	0.11	0.27
GC2	0.15	0.27	0.41
GC3	0.39	0.25	0.63
Gl2	0.19	0.12	0.30
Gl3	0.30	0.23	0.53
Hf2	1.86	0.31	2.17
Hf3	0.45	0.57	1.02
MV2	0.27	0.25	0.51
MV3	0.59	0.36	0.95
No2	0.08	0.08	0.16
No3	0.78	0.07	0.85
Pl2	0.71	1.94	2.65
Pl3	0.77	1.05	1.82
Pl4	0.78	0.92	1.71
SoH2	0.60	2.36	2.96
SoH3	2.10	1.44	3.53
Stu2	0.06	0.18	0.24
Stu3	0.09	0.03	0.12
Su2	0.02	0.12	0.14
Su3	0.13	0.08	0.21
Sv2	0.17	0.06	0.24
Sv3	0.17	0.70	0.88
Tw2	0.24	0.08	0.33
Tw3	0.11	0.13	0.24
VH2	0.32	1.79	2.11
VH3	1.27	0.53	1.81

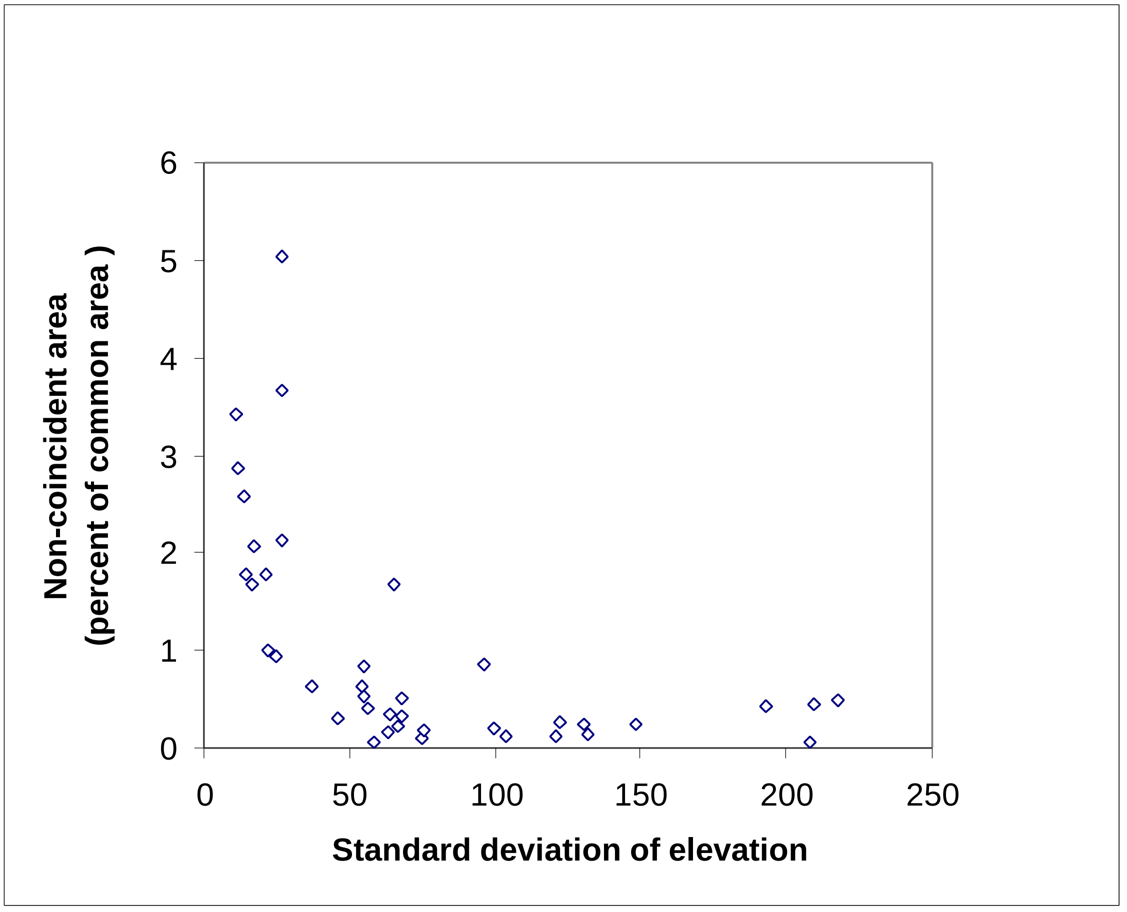

Figure 4.29. Percent of non-coincident watershed area delineated by ArcView and TOPAZ as a function of standard deviation of elevation.

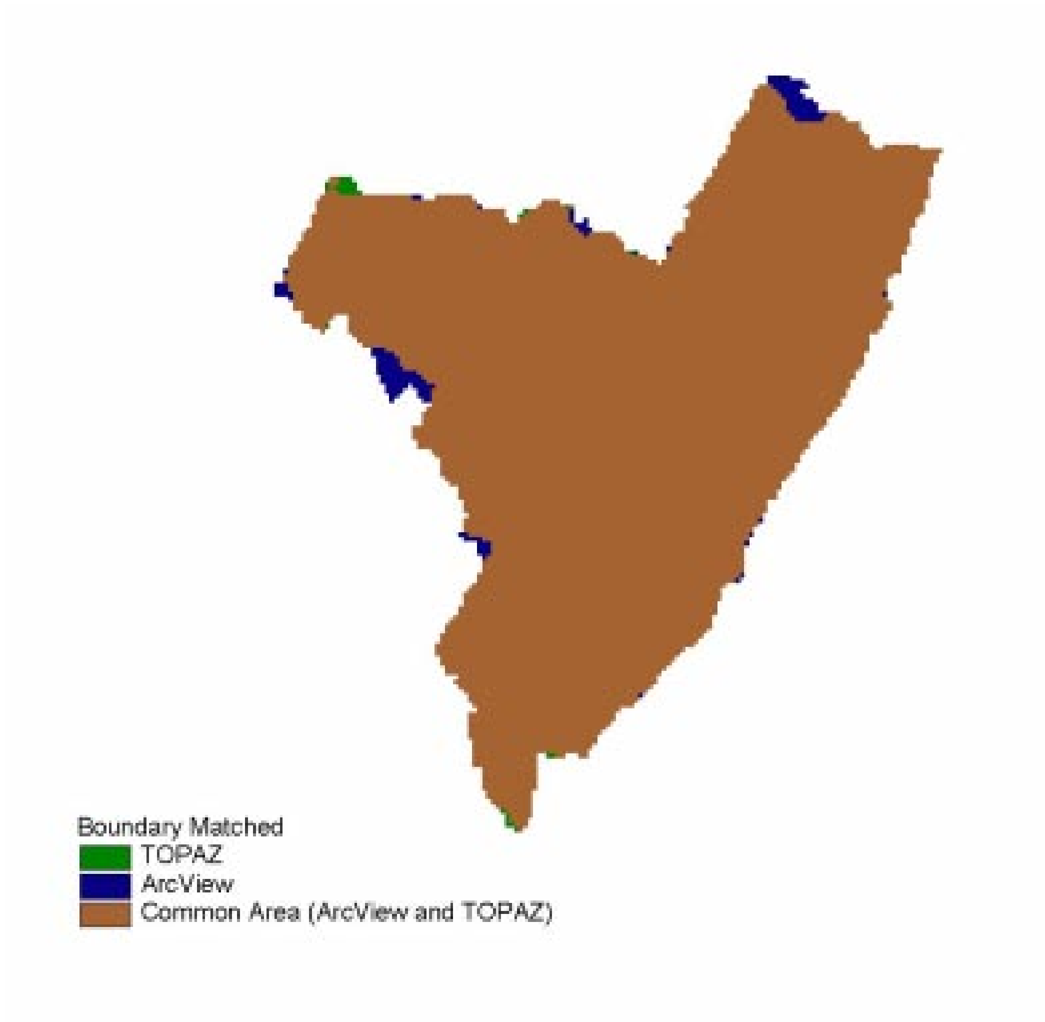

Figure 4.30. Comparison of watershed areas delineated by ArcView and TOPAZ given the same outlet point (Watershed Hf2).

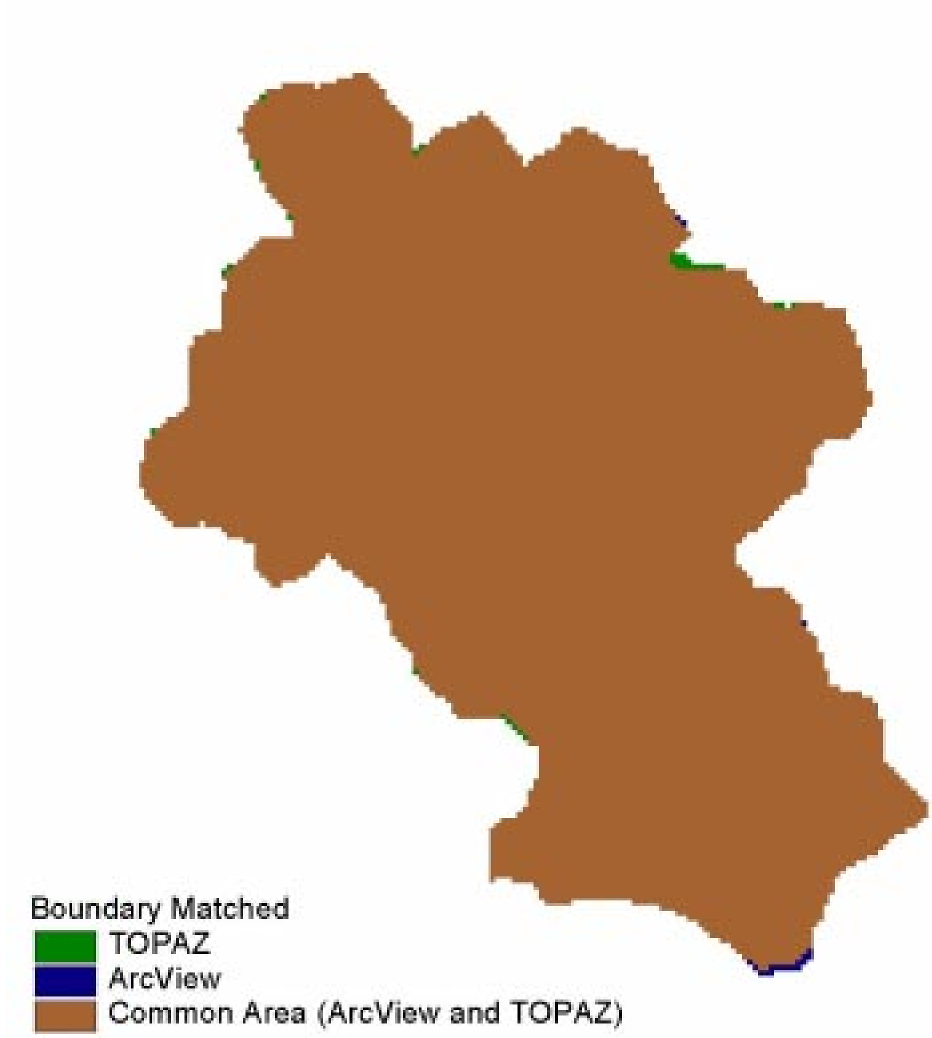

Figure 4.31. Comparison of watershed areas delineated by ArcView and TOPAZ given the same outlet point (Watershed GC2).

4.8 Stream network delineation accuracy

The percent of ArcView and TOPAZ stream cells matching with the rasterized blue line stream were analyzed using a paired t test to determine whether they are significantly different from each other. The cells of ArcView and TOPAZ derived streams were classified in four categories: cells which exactly coincided with the rasterized blue line stream, those within a distance of 1-50 meters, 50-100 meters, and greater than 100 meters (Table C1 in Appendix C).

Converting the number of cells to percent of stream cells normalized the data. The paired t test was used to identify if there were significant differences in how well stream networks derived by ArcView and TOPAZ matched the blue line stream for flat, average, and steep watersheds (Tables 4.11 to 4.22). The percentage distribution is very similar for both, but for ArcView there is a shift of about 2% of the stream cells from the coincident category to the 1-50 m category. The p-values in the tables indicate no significant differences between algorithms ($\alpha = 0.05$) except for flat watersheds in the 1-50 m category.

Thus, it can be concluded that in general there is no difference in ArcView and TOPAZ derived streams when compared to blue line stream. The graphical comparison of ArcView and TOPAZ derived flow network with the blue line stream (Appendix B) supports the above conclusion.

Table 4.11. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that coincides with blue line streams for flat watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	49.21	10.50	23	0.9999
Percent of TOPAZ stream	52.07	11.23		

Table 4.12. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that coincides with blue line streams average watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	60.36	11.86	23	0.9218
Percent of TOPAZ stream	60.82	12.05		

Table 4.13. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that coincides with blue line streams for steep watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	53.19	10.88	17	0.3045
Percent of TOPAZ stream	53.02	11.13		

Table 4.14. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of 1-50 m from the blue line streams for flat watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	41.49	7.64	23	0.0009
Percent of TOPAZ stream	39.46	7.15		

Table 4.15. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of 1-50 m from the blue line streams for average watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	35.89	10.11	23	0.0597
Percent of TOPAZ stream	35.41	10.30		

Table 4.16. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of 1-50 m from the blue line streams for steep watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	38.45	7.62	17	0.6667
Percent of TOPAZ stream	38.58	7.96		

Table 4.17. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of 50-100 m from the blue line streams for flat watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	7.73	4.40	23	0.1364
Percent of TOPAZ stream	7.12	4.89		

Table 4.18. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of 50-100 m from the blue line streams for average watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	3.38	2.51	23	0.5334
Percent of TOPAZ stream	3.39	2.47		

Table 4.19. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of 50-100 m from the blue line streams for steep watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	6.44	3.88	17	0.6728
Percent of TOPAZ stream	6.53	3.83		

Table 4.20. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of greater than 100 m from the blue line streams for flat watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	1.57	1.91	23	0.1681
Percent of TOPAZ stream	1.35	2.26		

Table 4.21. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of greater than 100 m from the blue line streams for average watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	0.38	0.85	23	0.4281
Percent of TOPAZ stream	0.38	0.83		

Table 4.22. Two sample paired t test for the means of the percent of ArcView and TOPAZ generated flow network cells that are at a distance of greater than 100 m from the blue line streams for steep watersheds.

Variable	Mean	Std. Dev.	Degree of freedom	p-value
Percent of ArcView stream	1.93	1.78	17	0.2343
Percent of TOPAZ stream	1.86	1.70		

5. Summary and Conclusions

5.1 Summary

Two algorithms to remove depressions and delineate drainage networks from grid elevation data (DEM) were evaluated to determine the magnitude and spatial distribution of modification in the elevation data, the effect on derived slope, and the accuracy of the stream network generated. The two algorithms represent different approaches used in ArcView(Version 3.1, Spatial Analyst 1.1) and TOPAZ (Version 1.2).ArcView removes depressions by raising the values of the elevation cells ("fills"), whereas TOPAZ uses cuts as well as fills to remove depressions.

Watersheds representing the varied topography of Virginia and adjacent regions were used for the analysis. USGS 7.5 minute DEMs (30 m grid) were used to delineate watersheds using ArcView. The changed elevation cells were analyzed to determine any existing spatial pattern. The USGS DEM was used to derive maps of slope, flow direction, flow accumulation and to delineate a channel network.

Several measures were used to characterize watershed topography: standard deviation of elevation, mean slope, watershed slope, channel slope, drainage density and elevation range. Watersheds were further classified in three categories of topographic "roughness" (flat, average and steep) based on values of the standard deviation of elevation. The number of changed elevation and slope cells were plotted against the above watershed topographic measures to assess the nature of any correlation. The changed elevation cells had influence on the derived maps of flow direction, flow accumulation and slope, impacting two to three times as many cells in these derivative layers. A spatial assessment of the modified elevation cells showed that depressions in the DEM's occur predominantly on or along the flow network.

The fill and cut cells were grouped in clusters of contiguous cells. The fill cells for ArcView and TOPAZ had larger cluster sizes as compared to TOPAZ cut cells. In reference to the stream network derived from the DEM by both ArcView and TOPAZ, a threshold value of 100 cells was selected to delineate the stream network. The majority of the changed elevation and slope cells were either on the channel network or within 100 m of a channel network.

A paired t-test was used to determine if there is any significant difference between the mean number of elevation cells changed by ArcView and TOPAZ for flat, average and steep watersheds. The statistical test indicated that the mean number of elevation cells changed by ArcView was greater for flat and average watersheds whereas for steep watersheds there was not a significant difference ($\alpha = 0.05$) in the mean number of elevation cells changed by TOPAZ and ArcView.

Blue line streams were digitized from the USGS 7.5 minute quadrangles and were used to assess the accuracy of ArcView and TOPAZ derived stream networks. The comparison was based on the percent of cells which fell in the following distance categories: coinciding with the rasterized blue line stream, within a distance of 1-50 meters, 50-100 meters, or greater than 100 meters from the rasterized blue line streams. A paired t test ($\alpha = 0.05$) indicated no significant difference in the mean number of ArcView and TOPAZ stream cells for flat, average and steep watersheds, except for flat watersheds in the zone of 1-50 meters.

5.2 Conclusions

1. The number of cells modified by removing depressions tended to be inversely correlated with various watershed topographic characteristics such as standard deviation of elevation, elevation range, mean slope, watershed slope and channel slope. Thus, a higher percentage of watershed area was affected by depressions in flatter/smoother watersheds.
2. The majority of filled cells were either on or adjacent to the channel network. The cut cells were to a greater extent distributed throughout the watersheds.
3. The majority of the fill cells were composed of groups which had 2 or more cells in the cluster whereas the cut cells occurred mostly as single cells or in a cluster of 2 cells.
4. In comparison with ArcView, TOPAZ generally changed fewer cells in the elevation data, with the difference statistically significant for the flatter watersheds but not for the steeper watersheds.
5. Channel networks derived by ArcView and TOPAZ were not significantly different from the blue line stream as suggested by a paired t test. Visual comparison of the streams also supported this conclusion.

5.3 Recommendations for further research

- Evaluate the effect of systematic and random error in the DEM on the presence of depressions in the DEM.
- Evaluate the effect that changes in the elevation data resulting from filling depressions has on prediction accuracy of hydrologic and non-point source models.
- Evaluate the relationship between DEM spatial resolution and the presence of depressions.
- Evaluate other algorithms and software, such as digital elevation model networks (DEMON), to ascertain which algorithm performs best in removing depressions and delineating flow networks.

REFERENCES

Allam, M. M. 1978. DTM's application in topographing mapping. *In Proc. Digital Terrain Models Symposium*, 1-15. St. Louis, Missouri., 9-11 May.

Band, L. E. 1986. Topographic partition of watershed with digital elevation models. *Water Resources Research* 22(1):15-24.

Bathurst, J. C. and P. E. O'Connell. 1992. Future of distributed modelling: the systeme hydrologique europeen. *Hydrological Processes* 6:265-277.

Brath, A. P., La. Barbera, M. Mancini, and R. Rosso. 1989. The use of distributed rainfall- runoff models based on GIS at different scales of information. ASCE Hydr. Div., 3rd Nat. Conf. on Hydraulic Engineering, New Orleans.

Briggs, I. C. 1981. Integration of elevation data with remotely sensed data. *In Proc. Landsat 81*, 4.1.1-4.1.4. Canberra.

Brown, D. G. and T. J. Bara. 1994. Recognition and reduction of systematic error in elevation and derivative surfaces from $7^1/_2$ – minute DEMs. *Photogrammetric Engineering and Remote Sensing* 60(2): 189-194.

Chorowicz, J., C. Ichoku, S. Riazanoff, Y. J. Kim, and B. Cervelle. 1992. A combined algorithm for automated drainage network extraction. *Water Resources Research* 28(5): 1293-1302.

Clark, M. J. 1998. Putting water in its place: a perspective on GIS in hydrology and water Management. *Hydrological Processes* 12:823-834.

Clarke, K. C. 1986. Recent trends in geographic information system research. *Geo-Processing* 3:1-15.

DeVantier, B. A. and A. D. Feldman. 1993. Review of GIS applications in hydrologic modeling. *Journal of Water Resources Planning and Management* 119(2): 246-261.

Douglas, D. H. 1986. Experiments to locate ridges and channels to create a new type of digital elevation models. *Cartographica* 23(4):29- 61.

Ebner, H. and K. Eder. 1992. State of art in Digital Terrain Modelling. *In Proc. EGIS'92*, Third European Conf. on GIS, Munich.

Elassal, A. A. 1978. USGS digital cartographic file management system. *In Proc. Digital Terrain Models Symposium,* 16-23. St. Louis, Missouri., 9-11 May.

Environmental Systems Research Institute, Inc. 1993. ArcView GIS version 3.1 online help. Redlands, CA.

Garbrecht, J. and L. W. Martz. 1993. Comment on "A combined algorithm for automated drainage network extraction" by J. Chorowicz, C. Ichoku, S. Riazanoff, Y. J. Kim, and B. Cervelle. *Water Resources Research* 29(2):535-536.

Giles, P. T. and S. E. Franklin. 1996. Comparison of derivative topographic surfaces of a DEM generated from stereoscopic SPOT images with field measurements. *Photogrammetric Engineering and Remote Sensing* 60(10): 1165-1171.

Goodchild, M. F. 1993. 'The state of GIS for environmental problem-solving', in Goodchild, M. F., B. O., and Steyaert, L. T. (eds), Environmental Modeling with GIS. Oxford University Press, Oxford. Pp. 8-15.

Hutchinson, M. F. and T. I. Dowling. 1991. A continental hydrological assessment of a new grid-based digital elevation model of Australia. *Hydrological Processes* 5: 45-58.

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

Jenson, S. K. and J. O. Domingue. 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing* 54(11):1593- 1600.

Julien, P. Y., B. Saghafian, and F. L. Ogden. 1995. Raster-based hydrologic modeling of spatially-varied surface runoff. *Water Resources Bulletin* 31(3):523-536.

Maidment, D. R. 1993. Developing a spatially distributed unit hydrograph by using GIS. *In HydroGIS 93: Application of Geographic Information Systems in Hydrology and Water Resources, Proceedings of the Vienna Conference,* eds. K. Dovar and H. P. Natchnebel, 181 - 192. Vienna: Int. Assoc. of Hydrological Sci.

Mark, D. M. 1975. Geomorphometric parameters: a review and evaluation. *Geografiska Annaler* 57A:165-177.

Mark, D. M. 1984. Automatic detection of drainage networks from digital elevation models. *Cartographica* 21(2/3):168-178.

Mark, D. M. 1988. Network models in geomorphology. In: modelling geomorphological systems, M. G. Anderson (Editor). John Wiley & Sons, New York, New York. pp. 73-96.

Martz, L. W. and E. DeJong. 1988. Catch: A FORTRAN program for measuring catchment area from digital elevation models. *Computers and Geosciences* 14(5): 627-640.

Martz, L. W. and J. Garbrecht. 1992. Numerical definition of drainage network and subcatchment areas from digital elevation models. *Computers and Geosciences* 18(6):747-761.

Martz, L. W. and J. Garbrecht. 1995. Comment on "Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method" by A. Tribe. *Journal of Hydrology* 167(1):393-396.

Martz, L. W. and J. Garbrecht. 1998. The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models. *Hydrological Processes* 12:843-855.

McDonnell, R. A. 1996. Including the spatial dimension: using geographical information systems in hydrology. *Progress in Physical Geography* 20(2):159-177.

Moore, I. D., E. M. O’Loughlin, and G. J. Burch. 1988. A contour-based topographic model for hydrological and ecological applications. *Earth Surface Processes and Landforms* 13(4): 305-320.

Morris, D. G. and R. G. Heerdegen. 1988. Automatically derived catchment boundaries and channel networks and their hydrological applications. *Geomorphology* 1: 131-141.

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

O’Loughlin, E. M. 1986. Prediction of surface saturation zones in natural catchments by topographic analysis. *Water Resources Research* 22: 794-804.

Ott, R. L. 1992. *An Introduction to Statistical Methods and Data Analysis*. California: Duxbury Press.

Palacios-Velez, O. L. and B. Cuevas-Renaud. 1986. Automated river-course, river and basin delineation from digital elevation data. *Journal of Hydrology* 86: 299-314.

Peucker, T.K. and D. H. Douglas. 1975. Detection of surface-specific points by local parallel processing of discrete terrain elevation data. *Computer Graphics and Image Processing* 4:375-387.

Quinn, P., K. Beven, P. Chevallier, and O. Planchon. 1991. The prediction of hillslope flow paths for distributed hydrologic modelling using digital terrain models. *Hydrological Processes* 5:59-79.

Speight, J. G. 1968. Parametric description of land form. In Stewart, G.A. (ed.), Land Evaluation. Macmillan, Melbourne, pp.239-250.

Tarboton D. G., R. L. Bras, and I. Rodriguez-Iturbe. 1991. On the extraction of channel networks from digital elevation data. *Hydrological Processes* 5:81-100.

Tribe, A. 1992. Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method. *Journal of Hydrology* 139(3): 263-293.

USGS. 1998. Digital Elevation Model. National Geospatial Data Clearinghouse. http://edcwww.cr.usgs.gov/nsdi/html/dem75/dem75.