○ THE EFFECT OF CELL RESOLUTION ON DEPRESSIONS IN DIGITAL ELEVATION MODELS

*Paul A Zandbergen, Department of Geography, University of South Florida
Correspondence to Paul A Zandbergen: zandberg@cas.usf.edu*

A proper understanding of the occurrence of depressions is necessary to understand how they affect the processing of a Digital Elevation Model (DEM) for hydrological analysis. While the effect of DEM cell resolution on common terrain derivatives has been well established, this is not well understood for depressions. The more widespread availability of high resolution DEMs derived through Light Detection and Ranging (LIDAR) technologies presents new challenges and opportunities for the characterization of depressions. A 6-meter LIDAR DEM for a study watershed in North Carolina was used to determine the effect of DEM cell resolution on the occurrence of depressions. The number of depressions was found to increase with smaller cell sizes, following an inverse power relationship. Scale-dependency was also found for the average depression surface area, average depression volume, total depression area and total depression volume. Results indicate that for this study area the amount of depressions in terms of surface area and volume is at a minimum for cell sizes around 30 to 61 meters. In this resolution range there will still be many artificial depressions, but their presence is less than at lower or higher resolutions. At finer scales, the (small) vertical error of the LIDAR DEM needs to be considered and introduces a large number of small and shallow artificial depressions. At coarser scales, the terrain variability is no longer reliably represented and a substantial number of large and sometimes deep artificial depressions is created. The results presented here support the conclusion that the use of the highest resolution and most accurate data, such as LIDAR-derived DEMs, may not result in the most reliable estimates of terrain derivatives unless proper consideration is given to the scale-dependency of the parameters being studied.

INTRODUCTION

WHAT ARE DEPRESSIONS?

Digital Elevation Models (DEMs) are widely used in hydrological analysis. Most DEMs contain numerous topographic depressions, which are defined as areas without an outlet; they are also referred to as sinks or pits. In regular-grid DEMs, topographic depressions are represented by an area of one or more contiguous cells that is lower than all of its neighboring cells. Single-cell depressions are often referred to as pits, but not all literature is consistent in this terminology. Determining hydrologically relevant terrain attributes in most cases requires the removal of these depressions; the DEM needs to be made “hydrologically correct”, i.e. water running over the surface must continue to flow downstream. A proper understanding of the occurrence of depressions is necessary to understand how their presence affects the processing of a DEM for hydrological analysis.

With the more widespread use of higher resolution and more accurate DEMs using technologies like Light Detection and Ranging (LIDAR), there is an expectation that the presence of artificial depressions will be reduced. However, research using high resolution LIDAR-derived DEMs has indicated that in fact high resolution DEMs have a very large number of (mostly small) depressions because of greater surface roughness and finer resolutions (MacMillan et al., 2003). Despite the availability of interpolation techniques such as ANUDEM (Hutchinson, 1989), which

effectively reduce the presence of depressions, most DEMs available today and those currently being created using LIDAR and related technologies contain a substantial amount of topographic depressions. Depression removal, therefore, will remain a necessary step in the use of DEMs for hydrological analysis.

Topographic depressions can be artificial or real. Artificial depressions are introduced primarily because of input-data errors, interpolation defects during DEM generation, truncation or rounding of interpolated values to lower precision, averaging of elevations within cells, or smoothing effects caused by resampling (Martz and Garbrecht, 1998; Tribe, 1991; Florinsky, 2002). The occurrence of artificial depressions is also linked to the vertical and horizontal resolution of the original elevation data and the variability of the landscape being modeled. For example, artificial depressions are very common in low-relief areas (Martz and Garbrecht, 1998; Liang and Mackay, 2000), which can in part be attributed to the limited vertical accuracy of DEMs.

Real or natural depressions represent areas of natural storage or man-made modifications to the landscape. In certain geomorphological settings, like in karst environments, such real depressions are very common. Normally, however, real depressions are much less common than artificial ones and may be almost absent in most terrain types (Mark, 1984; Goodchild and Mark, 1987). This is because fluvial erosion processes do not normally produce such features; exceptions are karst or recently glaciated terrains, where potholes, sinkholes, ponds, lakes and other natural depressions may prevail. Man-made structures, such as detention basins, ponds, and quarries are also real depressions and can be very common in urbanized landscapes.

Without supplementary information or field investigation it is not possible to determine with certainty using only the DEM whether a topographic depression is artificial or real, although promising modeling approaches exist which can assist in distinguishing artificial from real depressions (Lindsay and Creed, 2006). Whatever the nature of the depressions, they all artificially truncate flow and prevent the analysis of downstream flow paths. All hydrologic models ultimately rely on some form of overland flow simulation to define drainage courses and watershed structure (Garbrecht and Martz, 2000). To create a fully connected drainage network, water outflow for every DEM grid cell needs to be routed to an outlet; a topographic depression prevents simulated water flow from being routed to an outlet, resulting in disconnected stream-flow patterns and interior subwatersheds with no outlets. For some real depressions this may in fact be a correct representation of the actual hydrology, but even most real depressions ultimately overflow into a downstream hydrologic system. Due to the undesirable effects resulting from the occurrence of depressions, most hydrological modeling that employs DEMs has as its first step the identification and removal of depressions.

A number of depression removal techniques are available, which fall into two main categories: filling and breaching. Depression filling raises the elevations of the cells in a DEM up until the elevation of the lowest neighboring cell – this process continues until the entire depression is removed. This approach to depression removal effectively floods the depression until an outlet is reached and the depression “overflows”. The result is a filled DEM whose cell elevation values are either the same or higher than the original DEM, never lower. Depression filling assumes all depressions are caused by elevation underestimation. Depression filling has become by far the most widely used approach, in part because of its relative simplicity. Several different algorithms have been developed to fill depressions (Marks et al., 1984; O’Callaghan and Mark, 1984; Jenson

and Domingue, 1988; Martz and Jong, 1988; Planchon and Darboux, 2001). Some of the differences between these algorithms include the size of scan window used, the scan direction, and the degree to which flow direction in the resulting filled depressions is characterized (Wang and Liu, 2006). However, depression filling can sometimes result in substantial modifications of the original DEM (Jenson and Domingue, 1988; Planchon and Darboux, 2001), in particular in areas with very low relief.

Depression breaching lowers the elevations of the cells in a DEM along a breach channel, which is analogous to creating a trench through the “dam” or obstacle in front of the depression. The result is a breached DEM whose cell elevation values are either the same or lower than the original DEM, never higher. Depression filling and breaching represent two opposite approaches, and several techniques have been established to combine the two into a single approach. Constrained breaching (Martz and Garbrecht, 1999) limits the breach channel length to a maximum of two grid cells, while all other depressions are filled. The Impact Reduction Approach (Lindsay and Creed, 2005a) selects filling or breaching depending on which method results in the least amount of modification of the DEM.

Despite the existence of several alternative depression removal techniques, a better understanding of depressions is needed. Early work on digital terrain modeling commonly used DEMs of fairly coarse resolution and limited vertical accuracy and it was therefore justified to assume that all depressions were artifacts (O’Callaghan and Mark, 1984; Jenson and Domingue, 1988; Hutchinson, 1989; Fairfield and Leymarie, 1991). However, recent DEMs derived from digital photogrammetry, LIDAR and related technologies are of much higher resolution and vertical accuracy, allowing for a more reliable representation of terrain variability, including real depressions. Real depressions play a role in the storage of water, sediment and nutrients, contribute to evaporation and groundwater recharge, and can provide critical habitat (Hayashi and van der Kamp, 2000; Rosenberry and Winter, 1997). As a result, the practice of depression removal is being challenged (McCormack et al., 1993), and a more thorough characterization of depressions is needed prior to the use of a DEM for hydrological analysis.

EFFECT OF DEM CELL RESOLUTION ON TERRAIN ATTRIBUTES

The ability to derive an understanding of watershed processes depends on reliability of the landscape input data, which is strongly influenced by the DEM scale or cell resolution (Kenward et al., 2000; Thompson et al., 2001; McMaster, 2002). Advances in numerical models to monitor and predict hydrology and geomorphology rely heavily on DEMs and their integrity. Many studies have examined the effect of different DEM cell sizes on the ability of a DEM to accurately and reliably represent form and function of landscapes (e.g. Zhang and Montgomery, 1994; Walker and Willgoose, 1999; Thieken et al., 1999; Zhan et al., 1999; Schoorl et al., 2000; Wolock and McCabe, 2000; Thompson et al., 2001; McMaster, 2002; Kienzie, 2004). The underlying purpose of this body of research is to determine at what DEM cell resolution it is appropriate to examine watershed behavior and landscape features, and how dependent analysis results are to DEM cell resolution.

Most research supports the notion that a smaller DEM cell size produces a more accurate representation of the actual terrain and therefore results in more reliable terrain derivatives. In general the effects of DEM cell resolution on common terrain derivatives are fairly well understood. For example, as the cell size gets smaller, the mean slope for a given area will increase.

This is simply a reflection of the fact that with smaller cell sizes terrain variability is better represented. In theory, as cell size decreases further, the cell size would become smaller than the spatial variability of the actual terrain, and no further increase in mean slope would be observed. Several studies have pointed out that a DEM cell resolution of 10 meters appears to represent sufficient terrain detail to produce very reliable terrain derivatives (Zhang and Montgomery, 1994; Hancock, 2005) suggesting that little additional information is gained from even smaller DEM cell sizes. However, the empirical evidence for this is very sparse, in part because up until recently very few DEMs at such high resolutions with appropriate vertical accuracy were available.

The effect of DEM cell resolution is not well understood for all terrain derivatives, in particular the more complex ones. Calculating terrain derivatives is a procedure in which new variables describing the properties of the surface are computed from the elevation points of a DEM. These derivatives are commonly divided into primary topographic attributes, such as slope, aspect, curvature, and catchment area; and secondary topographic attributes, such as topographic wetness index and stream power index. Primary topographic attributes are calculated directly from the elevation data or from one of its derivatives, while secondary topographic attributes are calculated from two or more primary ones. From this perspective, determining the presence of depressions is a primary topographic attribute. While this distinction is useful, from the perspective of understanding the effect of cell resolution a more useful classification of terrain derivatives is based on their spatial properties rather than their source of calculation. Derivatives based on a fixed neighborhood can be considered as *constrained*, while derivatives that are based on far-reaching spatial interactions can be considered as *unconstrained*. Derivatives such as slope and aspect would be considered constrained, while derivatives such as catchment area and the presence of depressions would be considered unconstrained. The behavior of constrained derivatives is fairly predictable since they are commonly determined by analyzing a 3x3 cell window around the cell for which the derivative is calculated and this behavior can to some degree be described analytically. For unconstrained derivatives the behavior is much less predictable since it may vary across multiple scales and this behavior requires empirical characterization. Due to their complexity, the body of research on the effect of DEM cell resolution on unconstrained terrain derivatives is not very extensive, but does include a characterization of watershed boundaries (Hancock, 2005; Oksanen and Sarjakoski, 2005) and stream networks (Clarke and Burnett, 2003; McMaster, 2002; Wang and Yin, 1998). Only one study (Lindsay and Creed, 2005b) has considered the effect of DEM cell resolution on depressions, suggesting that the number of depressions sharply increases with finer scale DEMs. The research presented here builds upon the work by Lindsay and Creed (2005b) by expanding the parameters used to characterize depressions and by developing more detailed explanations for the scale-dependency of depressions in DEMs.

RESEARCH OBJECTIVE

Despite the body of research on the effect of DEM resolution on terrain derivatives, the effect on the presence, shape and size of depressions has received very little attention and is not well understood. Studies that make reference to depressions and DEM resolution are quick to point out that most depressions are artificial, introduced by errors in the original data and/or the DEM creation, and that these are expected to disappear when a DEM of higher vertical accuracy is employed. Higher resolution DEMs, however, have been shown to contain a very large number of depressions. The purpose of this study, therefore, is to determine the effect of DEM cell resolu-

ution on the nature of depressions, including their number, surface area, volume and spatial distribution in the landscape. Statistical relationships are developed to characterize the scale-dependency of the occurrence of depressions. A case-study is used to develop empirical relationships between cell resolution and quantitative characteristics of depressions. While the exact empirical findings are specific to the general morphological characteristics of the case-study area, the general patterns identified are expected to be applicable to different regions.

METHODS

A high resolution LIDAR DEM was obtained for the Middle Creek watershed in Wake County, North Carolina. This study area was selected in part because of the availability of a high resolution LIDAR DEM and a highly accurate small-scale stream network. This study area also represents a range of low to moderate slopes where many depressions are likely to occur.

The LIDAR DEM was obtained from the North Carolina Flood Mapping Program. The raw LIDAR data for this area was collected in 2002 and processing of the data was completed in 2004. A 6-meter bare earth DEM was created by the North Carolina Flood Mapping Program. *[Note: the original DEM was created in the State Plane Coordinate System in US Survey Feet and the resolution was exactly 20 feet. The actual DEM cell size was therefore approximately 6.096 meters, but is referred to here as the 6-meter DEM. Resampled versions of the original 20-foot DEM are also reported to the nearest meter, while the actual DEM processing was accomplished in the non-metric State Plane Coordinate System.]* Details on the collection, processing and accuracy assessment of the LIDAR data are provided in a series of Issue Papers produced by the North Carolina Flood Mapping Program (2006); a brief summary follows. The original LIDAR data was collected with a ground spacing of sampling points of approximately 3 meters. To produce the bare earth DEM a combination of manual and automated cleaning techniques were employed. These post-processing techniques included the use of automated procedures to detect elevation changes that appeared unnatural to remove buildings, as well as the use of last returns to remove vegetation canopy. The accuracy specifications for the collection of the LIDAR data report a vertical accuracy requirement of 25 cm for all the inland Counties; field testing of the vertical accuracy for Wake County resulted in a vertical accuracy assessment of 13.2 cm (North Carolina Flood Mapping Program, 2002). This estimate represents the 95% Root Mean Square Error (RMSE) of 125 surveyed checkpoints across a range of land cover classes. The original accuracy assessment was completed for all of Wake County. The Middle Creek Watershed falls almost completely within Wake County and both landform and land cover of the rest of County are very similar to that of the study watershed; therefore, the accuracy assessment for Wake County can be considered a reliable estimate for the accuracy of the LIDAR data used in this study. The 95% RMSE was adopted instead of the 100% RMSE as the most reliable accuracy statistic in the North Carolina Floodplain Mapping Program due to non-normal distribution of observed errors, resulting in skewed 100% RMSE values due to a small number of outliers (North Carolina Flood Mapping Program, 2001)

Individual tiles of the LIDAR DEM covering the Middle Creek watershed were obtained and mosaiced together to form a single continuous DEM. This DEM was processed using an automated stream and watershed delineation procedure to generate 103 subwatersheds. This procedure consisted of filling all depressions using the Planchon and Darboux (2001) algorithm, followed by determining flow direction using the D-8 method (Mark, 1984; O'Callaghan and Mark,

1984). Subwatersheds and stream networks were delineated using a constant stream threshold of 0.016 km^2 which produced first order streams which closely approximated the stream network derived by Wake County through photogrammetry.

The entire Middle Creek Watershed covers an area of 173 square kilometers with a total relief of 98 meters and an average slope of 3.56 degrees. Figure 1 shows the location of the Middle Creek Watershed as well as the 103 subwatersheds.

Figure 1 Location of Middle Creek Watershed in Wake County, North Carolina

The original 6-meter LIDAR DEM was resampled by factors of 3, 5, 10, 25, 50 and 100 to produce DEMs with resolutions of 6, 18, 30, 61, 152, 305 and 610 meters. This set of resampling factors is very similar to those employed in other recent efforts to examine the scale-dependency of terrain derivatives using high to medium-resolution DEMs (Brasington and Richards, 1998; Kienzie, 2004; Usery et al., 2004; Claessens et al., 2005). The chosen set of resampling factors was also based on the fact that some degree of non-linear behavior could be expected (Lindsay and Creed, 2005b) and therefore the sequence of DEM cell sizes used is also non-linear. Nearest neighbor resampling was used, which in effect takes every 3rd, 5th, 10th, 25th, 50th and 100th point, respectively, in both X and Y direction from the original 6-meter DEM. This approach

was deemed appropriate since it maintains the exact same elevation values for those selected locations without any additional smoothing effects. The information loss in each resampling can therefore be attributed entirely to the effect of using a coarser grid, and not to the smoothing effect inherent in other resampling techniques such as bilinear interpolation or cubic convolution. Nearest neighbor has also been the technique of choice in most other recent studies which used resampling to produce coarser DEMs for the purpose of determining the effect of DEM resolution (Barber and Shortridge, 2005; Lindsay and Creed, 2005b; Wolock and McCabe, 2000).

The depression characteristics for each of the seven DEMs were determined in the following manner. First, depressions were filled using the Planchon and Darboux (2001) algorithm. Second, by comparing the original DEM with the filled DEM, the amount, extent and depth of depressions were determined. The following parameters were obtained for each DEM: number of depressions, number of depression cells, combined surface area of all depressions, combined volume of all depressions, number of cells of each depression, surface area of each depression, and volume of each depression. For visual comparison, slope grids were derived from each DEM using the method developed by Horne (1981).

RESULTS

Prior to a more quantitative analysis of the depressions' parameters, results for two selected study areas are presented to characterize the effect of DEM cell resolution in a visual manner. Figure 2a shows an orthophoto and the stream network for the first study area. This area consists of a meandering riverbed in a wide channel in a relatively undeveloped and vegetated area. Slopes vary between 0 and 5 degrees through most of this area, but there is one steep streambank on the south-side of the main channel, as shown in Figure 2b. Figures 2c, 2d, and 2e show the extent and depth of the depressions based on DEMs of 6, 30 and 61 meters, respectively.

The pattern in the depressions derived from the original 6-meter LIDAR DEM in Figure 2c reveals a large number of depressions, with many of the small ones being very shallow; most of these are expected to be artificial depressions. Larger depressions are also deeper and occur within or immediately adjacent to the stream; some of these are expected to be real depressions, although some will be artificial. As the cell size increases to 30 and 61 meter, the total number of depressions decreases substantially. Most of the very small and shallow depressions in the 6-meter DEM have disappeared, but a number of shallow single-pixel depressions remain in both the 30 and 61 meter DEM. Most of the larger and deeper depressions are still identified, but with much fewer cells. No depressions appeared in the 152 meter DEM.

While the decrease in the number of depressions is evident based on a visual comparison of Figures 2c, 2d and 2e, trends in surface area and volume are not immediately clear and require further quantification.

Figure 2a Orthophoto and stream network for selected study area 1

Figure 2b Slope grid derived from 6-meter original DEM for selected study area 1

Figure 2c Extent and depth of depressions in 6-meter original DEM for selected study area 1

Figure 2d Extent and depth of depressions in 30-meter DEM for selected study area 1

Figure 2e Extent and depth of depressions in 61-meter DEM for selected study area 1

Figure 3a shows another type of study area, consisting of a more developed part of the watershed with low density residential development. Several man-made re/detention ponds are visible which represent real depressions. [Note: this study area was chosen on purpose to show examples of real depressions. Obviously, real depressions do not have to be ponds and can naturally occur in undeveloped areas, but these cannot always be identified easily on an orthophoto.] The slope grid in Figure 3b confirms the presence of some well-defined banks around ponds, roads and buildings, but otherwise the area is characterized by very smooth topography resulting from the grading of the residential sub-division and the agricultural fields. Figures 3c, 3d, and 3e show the extent and depth of the depressions based on DEMs of 6, 30 and 61 meters, respectively. The pattern in the depressions derived from the original 6-meter LIDAR DEM in Figure 3c reveals

several well defined depressions which correspond very closely to the ponds observed in the orthophoto. Relatively few other depressions occur; where they do, they consist mostly of shallow single-pixel depressions. As the cell size increases to 30 and 61 meter, most of the shallow single-pixel depressions have disappeared as well as some of the larger ones. Two of the largest depressions maintain their outline fairly well in the 30-meter DEM, but at 61 meters the outlines become more unreliable. No depressions appeared in the 152 meter DEM. As with the first study area, a drop in the total number of depressions is evident, and in this case total surface area also appears to have decreased.

Figure 3a Orthophoto and stream network for selected study area 2

Figure 3b Slope grid derived from 6-meter original DEM for selected study area 2

Figure 3c Extent and depth of depressions in 6-meter original DEM for selected study area 2

Figure 3d Extent and depth of depressions in 30-meter DEM for selected study area 2

Figure 3e Extent and depth of depressions in 61-meter DEM for selected study area 2

The qualitative results so far provide a number of insights: 1) The number of depressions appears to decrease with increasing cell size; 2) Many of the small and shallow depressions, which are most likely artificial, disappear with increasing cell size, but a number of single-pixel depressions persist; 3) Trends in surface area and volume require more detailed quantification; and 4) Interpretation of results will require a consideration of artificial and real depressions.

Following this introductory visual characterization, a more rigorous quantitative description is presented. First, Figure 4 shows the spatial extent of the depressions in the entire watershed for selected resolutions. It confirms that the number of depressions decreases with increasing cell size, but the changes in total surface area are less obvious. Statistical analyses presented in what follows all pertain to the entire watershed.

Figure 4 Extent of depressions in Middle Creek Watershed for selected DEM cell resolutions

Moving on to the more quantitative characterization of depressions, Figure 5 shows the relationship between the number of depressions and cell resolution; a log scale is used for both axes. Table 1 reports the statistical summary for this data. Results reveal that the number of depressions decreases with cell resolution. Curve-fitting regression revealed an inverse power relationship with a very high degree of correlation. Table 2 reports the curve-fitting results – the results for surface area and volume are included here and will be discussed in later sections.

The relationship between the number of depressions and cell resolution can be described as: number of depressions = $476,146 * (\text{cell size in meters})^{-1.5098}$. The R-square for this regression is 0.9988 and the relationship is significant with $p < 0.0001$. The power value for this relationship is -1.5098.

Figure 5 Relationship between number of depressions and DEM cell resolution

DEM Resolution (meters)	# of depressions	Total # of depression cells
6	34,421	210,680
18	5,913	18,055
30	2,568	5,801
61	874	1,470
152	225	372
305	89	136
610	32	54

Table 1 Descriptive statistics for depressions

Variable	N	A	B	R ²	F	Sign F	SE
Number of Depressions	7	476,146	- 1.5098	0.9988	5007.26	< 0.0001	0.0851
Average Depression Size	7	6.9382	+ 1.7331	0.9907	639.46	< 0.0001	0.2736
Average Depression Volume	7	0.509	+ 2.3061	0.9726	213.82	< 0.0001	0.6290

Table 2 Non-linear regression results for depression characteristics
Regression model is based on an inverse power function: variable = A*(cell size)^B

At first glance it may not seem very intuitive that the number of depressions for fine-scale DEMs is much larger than for coarse-scale DEMs for the same area, since it is expected that fine-scale DEMs are more accurate in their representation of morphological features. However, it should be recognized that the number of grid cells for a given area decreases by a power of 2

with increasing cell size. For example, a square parcel of land of 90,000 square meters would be represented by 2500 cells in a 6-meter raster and by 25 cells in a 60-meter raster. For every 10-fold increase in cell size, the number of cells required to represent the exact same area is reduced 100-fold, which is represented by a power factor of 2. The inverse power relationship between the number of depressions and cell size, therefore, is not unexpected. What is somewhat surprising, however, is the fact that the power factor is much lower than 2. If the relation between the number of depressions and cell size were scale-independent, the power value would be exactly 2, and it can therefore be inferred that the relationship is indeed scale-dependent. The fact that the power value is much less than two (i.e. 1.5098) is of significance, since this suggests that the number of depressions relative to the total number of cells in fact increases with large cell sizes.

The inverse power relationship between the number of depressions and cell size has implications for how depressions are handled in terrain processing. As observed by Lindsay and Creed (2005b), the common practice of removing depressions prior to modeling flow networks uses algorithms which rely on the identification of individual depressions prior to removal. These algorithms are currently at their limit in terms of DEM size, and despite increases in computing power the depression removal techniques currently in use will not be able to process higher resolution DEMs.

The number of depressions in the original 6-meter LIDAR DEM was 34,421 for the entire watershed, or approximately 198 depressions per square kilometer. An even finer scale DEM of for example 3 meters would result in a depression density of approximately 500 depressions per square kilometer. While 3 to 6 meters presents the current limit of most widely used LIDAR DEMs, it does not seem impossible that with continued improvements in LIDAR collection and processing technology a cell resolution of 1 meter or even less might be used in the future. A 30 cm DEM would have a depression density of approximately 16,500 depressions per square kilometer, which presents serious challenges for current computing power.

While a number of depression filling algorithms exist, the one developed by Jensen and Domingue (1988) is by far the best known and has been implemented in nearly all GIS and hydrologic software packages. Despite the widespread adoption of the Jensen and Domingue (1988) algorithm, the identification and filling of depressions remains the most time-consuming step in the extracting stream networks and subwatersheds (Wang and Liu, 2006). Due to the relatively complex nature of the Jensen and Domingue (1988) algorithm and the large file size of the LIDAR DEM used in this study, none of the commercial GIS software with terrain analysis functions was able to process the LIDAR DEM for depression characterization within an acceptable time-frame. Processing of the LIDAR DEM for this study was therefore accomplished using the Planchon and Darboux (2001) algorithm, which is insensitive to data size and complexity. Both the Planchon and Darboux (2001) algorithm and the more recently developed Wang and Liu (2006) algorithm are expected to perform well with the large file sizes of LIDAR. These algorithms have not been implemented in commercial GIS and hydrologic software, but that can be expected as processing of LIDAR data becomes more widespread. However, both these algorithms only address depression filling and do not consider alternative approaches to handle depressions such as breaching.

In addition to the number of depressions, the effect of cell size on the characteristics of depressions was determined by looking at the average area and the average volume of all depressions. Figure 6 shows the relationship between average depression area (in square meters) and cell size,

while Figure 7 shows the relationship between average depression volume (in cubic meters) and cell size.

Figure 6 Relationship between average area of depressions and DEM cell resolution

Figure 7 Relationship between average volume of depressions and DEM cell resolution

As could be expected based on the discussion of the number of depressions, both average area and average volume increase with larger cell sizes, revealing a power function. Curve-fitting regression results are reported in Table 2. Both relationships are very strong (R-square value of 0.9907 and 0.9726, respectively) and highly significant ($p < 0.0001$ for both).

This general pattern was to be expected: as the DEM cell size increases, there are fewer depressions, but they are larger in size. The power relationships found reflects the same phenomena described earlier: the number of grid cells for a given area decreases by a power of 2 with increasing cell size. Again, it is significant to note the power values: +1.7331 for average area and 2.3061 for average volume. If the relationship was scale-independent these power values would be 2, and it can therefore be inferred that the relationships between cell resolution and average area and volume are indeed scale-dependent.

In addition to exploring the effect of cell resolution on the average area of depressions in terms of actual surface area, it is useful to explore the average size of depressions in terms of the number of cells. Figure 8 shows the relationship between the average size of depressions in number of cells and cell resolution. Summary statistics are provided in Table 3.

Figure 8 Relationship between average size of depressions in number of cells and DEM cell resolution

DEM Resolution (meters)	Minimum	Average	Maximum	Standard Dev
6	1	6.120	2716	36.194
18	1	3.053	878	13.474
30	1	2.259	373	8.122
61	1	1.682	111	4.118
152	1	1.653	75	4.963
305	1	1.528	10	1.333
610	1	1.688	8	1.630

Table 3 Summary statistics for the size distribution of depressions in number of cells

In general, many DEMs have a large number of very small depressions consisting of only one or several cells, and a much smaller number of depressions consisting of a large number of cells. The DEMs used here are no exception to this, and the values for average size are fairly low. However, Figure 3 reveals a very surprising trend: the average number of cells per depression for the original 6-meter LIDAR DEM is 6.1 cells, but this drops off quickly to less than half that with increasing cell size, and levels out at around 1.6 cells at a cell size of around 61 meters. The descriptive statistics in Table 3 confirm that the maximum and standard deviation also decrease substantially with increasing cell size. These results suggest that the fine-scale DEMs are much less dominated by single-cell depressions.

The pattern in Figure 3 requires some further consideration. The fact that for most of the DEM cell resolutions considered the value is between 1 and 2 is not surprising; the theoretical minimum is one and single-pixel artificial depressions are often very dominant in DEMs. However, the much higher values for fine-scale DEMs suggest there is another phenomena at work. The hypothesis presented here is that the original 6-meter LIDAR DEM contain a substantial number of real depressions of relatively small size (i.e. several dozen to several hundred cells); some typical examples are shown in Figure 3c, and may include man-made ponds as well as natural depressions. As the cell size increases, these real depressions are represented by a much lower number of cells or may disappear altogether, making them more difficult to distinguish from artificial depressions. Therefore, the fine-scale DEMs are less dominated by single-cell depressions and the average number of cells per depression is higher.

While this hypothesis explains the pattern observed in Figure 8 it should be pointed out that proving this hypothesis requires that artificial and real depressions can be separated. At present, no such technique using only the DEM exists and intensive field verification would be required to accomplish this. At present, therefore, it remains a hypothesis for further research. It does however, present some interesting new perspectives relevant for future research on depressions. For example, it suggests that many real depressions in landscapes might be relatively small, making them impossible to recognize in coarse-scale DEMs of 30 meters or larger. This emphasizes the need to use high resolution DEMs to characterize depressional storage, particular in areas of moderate to low slopes. It also suggests that a technique to separate real from artificial depressions in a DEM should consider depression surface area as one of its criteria, but that the ability to accomplish this is limited at coarse resolutions.

To further explore the size distribution of depressions, Figure 9 shows the cumulative distribution function of the number of depressions by depression size in cells. To maintain legibility, only the result for the original 6-meter LIDAR DEM and the 30-meter DEM are shown. Figure 9 reveals that the coarser-scale DEM is indeed more dominated by smaller depressions. The number of depressions of 1 cell as a percentage of the total number of depression is 56.6% for the 6-meter DEM and 68.6% for the 30-meter DEM. Figure 9 reveals that this difference is fairly consistent across all depression sizes.

Figure 9 Cumulative distribution function of the number of depressions by depression size in cells for 6-meter and 30-meter DEMs

Another meaningful and perhaps more insightful way to quantify the dominance of small depressions is to plot the cumulative distribution function of the depression area by depression size in cells. This is shown in figure 10, again only using the results for the original 6-meter LIDAR DEM and the 30-meter DEM. The area of depressions of 1 cell as a percentage of the total area of depressions is 9.2% for the 6-meter DEM and 30.4% for the 30-meter DEM. Figure 10 reveals that this difference is fairly consistent across all depression sizes; however, at the upper end of the distribution the two curves approach each other as a result of the occurrence of one very large depression. It should be noted that the curves in Figure 10 are not as smooth as in Figure 9; while in Figure 9 each depression is counted equally in determining the proportion, in Figure 10 each depression is weighted by its surface area, resulting in “jumps” in the distribution for larger depressions.

Figure 10 Cumulative distribution function of the area of depressions by depression size in cells for 6-meter and 30-meter DEMs

While the number of depressions and their size distribution provide a meaningful characterization of the occurrence of depressions in DEMs of varying resolutions, ultimately of most interest are the total surface area and the total volume of depressions. These two parameters are meaningful in their own right for hydrological modeling since they provide information on the potentially available depressional storage in the landscape. They are also relevant for terrain processing since they provide an indication of how much the original DEM needs to be modified in order to become hydrologically correct. The scale-dependency of the relationship between cell size and the number of depressions, as well as the scale-dependency of the size distribution of depressions, have already provided some indication that the total surface area and the total volume are not likely to be scale-independent either. Figure 11 shows the relationship between the total surface area of depressions and cell size and Figure 12 shows the relationship between the total volume of depressions and cell size.

Figure 11 Relationship between total surface area of depressions and DEM cell resolution

Figure 12 Relationship between total volume of depressions and DEM cell resolution

Both Figure 11 and 12 reveal scale-dependency of total surface area and total volume, respectively, but in somewhat surprising patterns. In Figure 11 the total surface area of depressions is 7.8 square kilometers for the original 6-meter DEM, drops to 6.03 for 18 meters, appears to reach a minimum of 5.4 for 30 and 61 meters, and then gradually increases up to 20.1 square

kilometers for the 610-meter DEM. In Figure 12 the total volume of depressions is 2.6 million m^3 for the original 6-meter DEM, drops slightly to 2.1 and 2.0 million m^3 for the 18 and 30-meter DEM, respectively, and then starts to gradually increase to 2.7 million m^3 for the 61-meter DEM and ultimately to 76.7 million m^3 for the 610-meter DEM. It should be recognized that if the surface area and volume were scale-independent, the curves in both Figures 11 and 12 would be flat (i.e. straight lines with a slope of zero). The fact that they are not flat presents strong evidence of the scale-dependency of total depression area and volume.

In trying to find possible explanations for the scale-dependency of depression area and volume, it should be remembered that for many areas most depressions are expected to be artificial. Considering how artificial depressions are formed will provide some insights into the patterns found in Figures 11 and 12.

The following explanation is presented for the observed scale-dependency of total area and total surface of depression. At the very highest resolution of 6 meters most of the small sinks (1 to several cells) are likely artifacts – the DEM can be characterized by a very large number of very small and very shallow artificial depressions and a small number of larger and deeper true depressions. As cell size increases to 30 or 61 meters, many of these small artificial depressions disappear since they are smaller than the new cell size; many (but fewer) small artificial depressions remain (and some are newly created) as well as most of the larger and deeper real depressions. As a result, the total surface area of depressions drops. Since these small and artificial depressions are very shallow, the effect on volume is much less strong; there is a small drop in depression volume going from 6 to 30 meters, but it is much less dramatic than for depression area.

As cell size increases from 61 to 152 meters and higher, small artificial depressions and some larger real depressions disappear, but new large artificial depressions are created. In many areas the cell resolution becomes insufficient to properly represent the variability in the terrain within the river channels. Minor channels which were well defined at the 30 or 61 meters resolution become much less well defined, and many artificial depressions are created in and around the stream channel. As cell size increases to 305 meters and higher, even larger channels lose their definition and very large artificial depressions are created.

This explanation suggests that there are two effects at work, both related to how artificial depressions are formed: 1) For a very small cell size of 6 meters, the elevation difference between adjacent cells becomes very small and in many areas starts to approximate the vertical error of the elevation values. For this DEM the estimated vertical error is around 15 centimeters, which can be considered quite small for elevation data. Nevertheless, this small vertical error appears to result in a large number of small and shallow artificial depressions at the original resolution. As the DEM gets coarser, this effect becomes less and artificial depressions start to disappear. 2) For larger cell sizes (100 meters and higher) much of the spatial variability in elevation is lost and major channels lose their definition. This results in some very large and deep artificial depressions which because of their size have a strong influence on the total surface area and total volume of all depressions combined.

The first effect has been illustrated already in Figures 2c, 2d and 2e. The second effect is illustrated in Figure 13a, 13b and 13c, which show the extent and depth of depressions as well as the slope derived from DEMs of 6, 30 and 152 meters, respectively. The purple line represents the cross section used later in Figure 14.

Figure 13a Depressions and slope derived from 6-meter original LIDAR DEM

Figure 13b Depressions and slope derived from 30-meter DEM

Figure 13c Depressions and slope derived from 152-meter DEM

Figure 13a shows the somewhat typical pattern for the depressions in the original 6-meter LIDAR DEM: a large number of small and relatively shallow depressions, as well as some larger and deeper ones close to the stream. The results in Figure 13b for 30 meters represent the first effect mentioned above: fewer depressions but covering a similar total area. As can be seen in

the results for the slope grid, at the highlighted cross section the stream channel is very narrow (1 or 2 cells) but still clearly defined. The slope grid for 152 meters in Figure 13c shows that the stream channel definition has been lost – it appears as if the channel has “collapsed”. As a result, a very large and deep artificial depression is created.

Figure 14a and 14b further illustrate this effect. A cross sectional elevation profile was created for both the 30 and 152-meter DEM; the location of the cross section is shown in Figure 13. The profile for the 30-meter DEM reveals a well defined channel, with only a minor depression due to slightly higher elevation values in the streambed downstream. The profile for the 152-meter DEM shows a much less defined channel, and while the actual elevation at the stream bed is almost identical at 70 meters, the collapse of the stream channel downstream has resulted in a large and deep depression, up to 5 meters at its deepest along the cross section.

While the results in Figures 13 and 14 represent a relatively extreme example, many instances of this type of “channel collapse” were found within the watershed, each resulting in a depression of one or several cells in the 152, 305 and 610-meter DEMs. This strongly suggests the characterization of depressions becomes much less meaningful at cell sizes larger than 61 meters.

Figure 14a Cross sectional elevation profile for 30-meter DEM

Figure 14b Cross sectional elevation profile for 152-meter DEM

CONCLUSIONS

Results of this study indicate that DEM cell resolution has a very strong effect on the occurrence of depressions. The number of depressions was found to increase with smaller cell size following an inverse power relationship. The power factor of 1.51 strongly suggest that the relationship is scale-dependent and not simply determined by the increased number of cells in the DEM for the same study area. The inverse power relationship between the number of depressions and cell size has implications for how depressions are handled in terrain processing, and it is expected that most GIS software being used for terrain analysis will need to adopt revised algorithms for depression removal in order to process high resolution LIDAR DEMs.

Both average area and average volume of depressions were found to increase with larger cell size, revealing a power function. This supports the general expectation that coarse scale DEMs are dominated by fewer but larger depressions. The power factors of 1.73 and 2.31 for average area and average volume, respectively, again indicate strong scale-dependency.

Scale-dependency was also found for the total area and total volume of depressions, but the patterns are more complicated. Both decreased from the original 6-meter LIDAR DEM to 30 and 61 meters, and increased for coarser-scale DEMs. An explanation based on the formation of artificial depressions has been presented.

Based on the results for this study area, there appears to be “sweet spot” around 30 to 61 meters where the amount of depressions in terms of surface area and volume is at a minimum. In this resolution range there will still be many artificial depressions, but their presence is less than at finer or coarser scales. At finer scales, the (small) vertical error of the DEM needs to be

considered and introduces a large number of very small and very shallow artificial depressions. At coarser scales, the terrain structure is no longer reliably represented and a substantial number of large and sometimes deep artificial depressions is created.

The results presented here support the conclusion that the use of the highest resolution and most accurate data, such as LIDAR DEMs, may not result in the most reliable estimates of terrain derivatives unless proper consideration is given to the scale-dependency of the parameters being studied. The results also suggest that further research into depression removal techniques for high resolution DEMs is needed, as well as a characterization of the scale-dependency of artificial and real depressions for different landscape types.

REFERENCES

- Barber, C; Shortridge, A. 2005. 'LIDAR elevation data for surface hydrologic modeling: Resolution and representation issues'. *Cartography and Geographic Information Science* 32 (4): 401–410.
- Brasington, J; Richards, K. 1998. 'Interactions between model predictions , parameters and DTM scales for TOPMODEL'. *Computers and Geosciences* 24 (4): 299–314.
- Claessens, L; Heuvelink, G; Schoor, J; Veldkmap, A. 2005. 'DEM resolution effects on shallow landslide hazard and soil redistribution modeling'. *Earth Surface Processes and Landforms* 30: 461–477.
- Clarke, S; Burnett, K. 2003. 'Comparison of digital elevation models for aquatic data development'. *Photogrammetric Engineering and Remote Sensing* 69 (12): 1367–1375.
- Fairfield, J; Leymarie, P. 1991. 'Drainage networks from grid digital elevation models'. *Water Resources Research* 27 (5), 709–717.
- Florinsky, I.V.2002. 'Errors of signal processing in digital terrain modeling'. *International Journal of Geographical Information Science* 15 (5): 475–501.
- Garbrecht J.; Martz, L. 2000. 'Digital elevation model issues in water resources modeling'. In: *Hydrologic and hydraulic modeling support with Geographic Information Systems*, D. Maidment and D. Djokic (eds.), pp. 1–27. ESRI Press, Redlands, CA.
- Goodchild, M; Mark, D.M. 1987. 'The fractal nature of geographic phenomena'. *Annals of the Association of American Geographers* 77 (2): 265–278.
- Hancock, G.R. 2005. 'The use of DEMs in the identification and characterization of catchment over different grid scales'. *Hydrological Processes* 19: 1727–1749.
- Hayashi, M; van der Kamp, G. 2000. 'Simple equations to represent the volume-area-depth relationships of shallow wetlands in small topographic depressions'. *Journal of Hydrology* 237: 74–85.
- Horne B.K.P. 1981. 'Hill shading and the reflectance map'. *Proceedings of the IEEE* 69 (1): 14–47.
- Hutchinson, M.F. 1989. 'A new procedure for gridding elevation and stream line data with automatic removal of spurious pits'. *Journal of Hydrology* 106: 211–232.
- Jenson, S; Domingue, J.O. 1988. 'Extracting topographic structure from digital elevation data for geographic information system analysis'. *Photogrammetric Engineering and Remote Sensing* 54 (11): 1593–1600.
- Kenward, T; Lettenmaier, D; Wood, E; Fielding, E. 2000. 'Effects of digital elevation model accuracy on hydrologic predictions'. *Remote Sensing of the Environment* 74: 432–444.
- Kienzle, S. 2004. 'The effect of DEM raster resolution on first order, second order and compound terrain derivatives'. *Transaction in GIS* 8 (1): 83–111.
- Liang, C; Mackay, D.S. 2000. 'A general model of watershed extraction and representation using globally optimal flow paths and up-slope contribution areas'. *International Journal of Geographical Information Science* 14 (4): 337–358.
- Lindsay, J; Creed, I.F. 2005. 'Removal of artifact depressions from digital elevation models: Towards a minim impact approach'. *Hydrological Processes* 19: 3113–3126.

- Lindsay, J; Creed, I.F. 2005. 'Sensitivity of digital landscapes to artifact depressions in remotely-sensed DEMs'. *Photogrammetric Engineering and Remote Sensing* 71 (9): 1029–1036.
- Lindsay, J; Creed, I. 2006. 'Distinguishing actual and artifact depressions in digital elevation data'. *Computers and Geosciences* (in press).
- MacMillan, R; Marti, T; Earle, T; McNabb, D.H. 2003. 'Automated analysis and classification of landforms using high resolution digital elevation data: Applications and issues'. *Canadian Journal of Remote Sensing* 29 (5): 592–606.
- Mark, D.M. 1984. 'Automatic detection of drainage networks from digital elevation models'. *Cartographica* 21: 168–178.
- Marks, D; Dozier, J; Frew, J. 1984. 'Automated basin delineation from digital elevation data'. *Geo-Processing* 2: 299–311.
- Martz, L; Garbrecht, J. 1998. 'The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models'. *Hydrological Processes* 12: 843–855.
- Martz, L; Garbrecht, J. 1999. 'An outlet breaching algorithm for the treatment of closed depressions in a raster DEM'. *Computer and Geosciences* 25: 835–844.
- Martz, L; Jong, E.D. 1988. 'CATCH: A Fortran program for measuring catchment area from digital elevation models'. *Computers and Geosciences*, 14: 627–640.
- McCormack, J; Gahegan, M; Roberts, S; Hogg, J; Hoyle, B.S. 1993. 'Feature-based derivation of drainage networks'. *International Journal of Geographical Information Systems* 7 (3): 263–279.
- McMaster, K.J. 2002. 'Effects of digital elevation model resolution on derive stream network positions'. *Water Resources Research* 38 (4): 13-1–13-8. DOI: 10.1029/2000WR000150
- North Carolina Flood Mapping Program. 2006. Issues Papers of the North Carolina Cooperating Technical State Flood Mapping Program. Accessed online on April 6, 2006 at <http://www.ncgs.state.nc.us/floodmap.html>.
- North Carolina Flood Mapping Program. 2002. LIDAR Accuracy Assessment Report – Wake County. Accessed online on April 6, 2006 at <http://www.ncgs.state.nc.us/floodmap.html>.
- North Carolina Flood Mapping Program. 2001. Issues 7: LIDAR Specifications. Accessed online on April 6, 2006 at <http://www.ncgs.state.nc.us/floodmap.html>.
- O'Callaghan, J; Mark, D.M. 1984. 'The extraction of overland flow networks from digital elevation data'. *Computer Vision, Graphics and Image Processing* 28: 323–344.
- Oksanen, J; Sarjakoski, T. 2005. 'Error propogation analysis of DEM-based drainage basin delineation'. *International Journal of Remote Sensing* 26 (14): 3085–3102.
- Planchon, O; Darboux, F. 2001. 'A fast, simple algorithm to fill the depressions of digital elevation models'. *Catena* 46: 159–176.
- Rosenberry, D; Winter, T.C. 1997. 'Dynamics of water-table fluctuations in an upland between two prairie-pothole wetlands in North Dakota'. *Journal of Hydrology* 191: 266–289.
- Schoorl, J; Sonneveld, M; Veldkamp, A. 2000. 'Three-dimensional landscape process modelling: The effect of DEM resolution'. *Earth Surface Processes and Landforms* 25: 1025–1034.
- Thompson, J; Bell, J; Buttle, C.A. 2001. 'Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling'. *Geoderma* 100: 67–89.
- Thieken, A; Lucke, A; Dieckkruger, B; Richter, O. 1999. 'Scaling input data by GIS for hydrological modelling'. *Hydrological Processes* 13: 611–630.
- Tribe, A. 1991. 'Automated recognition of valleylines and drainage networks from grid digital elevation models: a review and a new method'. *Journal of Hydrology* 139: 263–293.
- Usery, E; Finn, M; Scheidt, D; Ruhl, S; Beard, T; Bearden, M. 2004. 'Geospatial data resampling and resolution effects on watershed modeling: A case study using the agricultural non-point source pollution model'. *Journal of Geographical Systems* 6: 289–306.
- Walker, J; Willgoose, G.R. 1999. 'On the effect of digital elevation model accuracy on hydrology and geomorphology'. *Water Resources Research* 35 (7): 2259–2268.

- Wang, L; Liu, H. 2006. 'An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modeling'. *International Journal of Geographical Information Science* 20 (2): 193–213.
- Wang, X; Yin, Z. 1998. 'A comparison of drainage networks derived from digital elevation models at two scales'. *Journal of Hydrology* 210: 221–241.
- Wolock, D; McCabe, G.J. 2000. 'Differences in topographic characteristics computed from 100 and 1000 meter resolution DEM data'. *Hydrological Processes* 14: 987–1002.
- Zhan, X; Drake, N; Wainwright, J; Mulligan, M. 1999. 'Comparison of slope estimates from low resolution DEMs: Scaling issues and a fractal method for their solution'. *Earth Surface Processes and Landforms* 24: 763–779.
- Zhang, W; Montgomery, D.R. 1994. 'Digital elevation model grid size, landscape representation and hydrologic simulations'. *Water Resources Research* 30 (4): 1019–1028.

Cite this article as: Zandbergen, Paul. 'The effect of cell resolution on depressions in Digital Elevation Models'. *Applied GIS* 2 (1). pp. 4.1–4.35. DOI: 10.2104/ag060004.