Sensitivity of Digital Landscapes to Artifact Depressions in Remotely-sensed DEMs

John B. Lindsay and Irena F. Creed

Abstract
Depressions are often removed from digital elevation models (DEMs) used in hydro-geomorphic applications. Light detection and ranging (lidar) and interferometric synthetic aperture radar (INSAR) DEMs of flat to mountainous landscapes were used to evaluate the occurrence of artifact depressions caused by the representation of surfaces using grids and random elevation error. The number of depressions in DEMs that result from grid representation was inversely related to grid spacing; however, normalizing for the number of grid cells in a DEM demonstrated that coarser grids were relatively more vulnerable to depressions. Flat landscapes containing extensive lakes experienced more depressions related to grid spacing and placement than high-relief areas. Stochastic modelling showed that error magnitude controlled the extent of vulnerability within a landscape to depressions caused by random error. Nevertheless, certain areas were likely to experience depressions regardless of the magnitude of random error, including flat areas, valley bottoms, and highly convergent topography.

Introduction
Technological advances in topographic data acquisition have contributed to the current trend towards increasingly accurate and finer resolution DEMs. DEMs are now routinely generated from remotely-sensed data sources, including digital photogrammetry from stereo aerial photography and satellite imagery, INSAR, and lidar (Wilson and Gallant, 2000). These data sources have become increasing important for DEM generation. Although fine-resolution DEMs derived from remotely-sensed data can provide an opportunity to study geomorphic and hydrological processes at appropriate scales (Lance and Chandler, 2003), they also present unique challenges for analysis. Specifically, remotely-sensed DEMs often contain far more depressions (including single-cell depressions (i.e., pits) and multi-cell depressions) than DEMs generated from contours, because of greater surface roughness and finer resolutions (MacMillan et al., 2003).

Topographic depressions possess ambiguous flow directions, thereby interrupting modelled overland flow paths and making the automated extraction of stream networks and basins from DEMs impossible (O’Callaghan and Mark, 1984; Band, 1986; Jenson and Domingue, 1988; Tarboton et al., 1991). Digital depressions should be removed prior to using DEMs in hydrological and geomorphic applications that require fully integrated overland flow networks (Fairfield and Leymarie, 1991; Martz and Garbrecht, 1998; Rieger, 1998; Lindsay and Creed, in press). As fine-resolution DEMs derived from remotely-sensed data sources become widely available, pre-processing for depressions can be expected to become a serious problem. Therefore, there is a need to eliminate, or at least to minimize, the occurrence of depressions in DEMs.

Although in some landscapes many of the depressions in fine-resolution DEMs may represent actual landforms, the majority of depressions in DEMs are spurious (Band, 1986; Mark, 1988; Hutchinson, 1989; Tarboton et al., 1991; Martz and Garbrecht, 1998; Rieger, 1998). Thus, the term depression will henceforth refer to an artifact depression. Rieger (1998) recognized two general causes of depressions in grid-deposited DEMs: the consequences of finite representation of surfaces using grids and errors in grid elevation data.

Depressions associated with finite grid representation can result from (a) interpolation procedures, this is more of a problem with contour-derived DEMs (Hutchinson, 1989; Wise, 2000; Florinsky, 2002), and (b) because of inadequate grid spacing to resolve flow paths (Mark, 1988; Tarboton et al., 1991; McCormack et al., 1993; Wilson and Gallant, 2000). A raster grid of elevations can obscure flow paths owing to its inability to explicitly represent breaks in slope, particularly flow paths along valley bottoms. For example, researchers have observed how inadequate grid spacing can cause chains of depressions along narrow valley bottoms (Rieger, 1998; Wise, 2000), a phenomenon that Burrough and McDonnell (1998) argued would occur at all levels of resolution. The obstruction of flow paths, and therefore occurrence of depressions, resulting from representation of surfaces by regular grids is related to both grid spacing and grid placement. Grid placement refers to the positioning of the grid relative to the terrain. Artifact depressions occur when a grid spacing is too coarse to capture important landscape features or when the grid is positioned such that it does not coincide with a local flow path. No previous research has examined the sensitivity of various landscapes to depressions caused by grid spacing and placement.

Data errors that cause depressions include missing data (Liang and Mackay, 2000; Lindsay, 2003), measurement of off-terrain objects (e.g., buildings, roads, forest canopy), systematic errors (e.g., imperfect joints between mosaiced DEMs), and most commonly random elevation errors (Burrough and McDonnell, 1998). Random error is extremely common in most DEMs (Hunter and Goodchild, 1997) and particularly
evident in those derived from remotely-sensed data (Endreny et al., 2000). Every grid cell in a DEM has the potential to belong to a depression given a large enough random error distribution. However, some landscapes are more likely to experience a depression. To our knowledge, the probability of occurrence of depressions caused by random error has never been quantified.

Depressions in DEMs have been recognized as a serious problem in terrain analysis research for a long while (O’Callaghan and Mark, 1984; Jenson and Domingue, 1988). However, there has never been a quantitative analysis of the sensitivity of different terrain to these errors. Such a study is needed now more than ever given the current availability of fine-resolution DEMs derived from remotely-sensed data sources. The objective of this research is to quantify the sensitivity of digital terrain to depressions caused by grid representation, including grid spacing and placement, and random errors in the elevation of grid cells.

Study Areas
Three DEMs were utilized in this study, representing a range in terrain from flat and hummocky to mountainous. The first DEM was a 1 m grid resolution lidar DEM of a 2.3 km² area of the Utikuma Uplands on the Boreal Plains of northern Alberta, Canada. The DEM was based on the last return of the laser altimeter, thereby estimating the ground surface elevation rather than the elevation of the vegetation canopy. Hopkinson et al. (in review) describe the collection and processing of the lidar data used to create this DEM in detail. The hummocky terrain of the Utikuma Uplands Site (UUS) is underlain by glacial moraines and possesses only 40 m of topographic relief. The UUS contains extensive wetlands complexes.

The second DEM was a 2.5 m grid resolution lidar DEM (based on the last return) of the Turkey Lakes Watershed (TLW), a 10.5 km² experimental watershed located approximately 60 km north of Sault Ste. Marie, Canada. The TLW is located in the Abitibi Uplands of the Canadian Shield, and contains a chain of headwater lakes that drain into Lake Superior through the Batchawana River (Semkin and Jeffries, 1983). Local slopes are moderate to steep with a maximum relief of 410 m. The TLW contains wetlands under open and closed canopy, i.e., cryptic wetlands (Creed et al., 2003), which occupy both upland and bottomland slope positions.

The third DEM used in this study was a 917 km² portion of a Shuttle Radar Topography Mission (SRTM) DEM of the Mad River Watershed (MRW) in Vermont. The MRW is located within the Green Mountains and has a local relief of 1200 m. Topography is dominated by fluvial processes, which is evident by the well-defined network of stream valleys. Valley bottoms are flat at lower elevations and the MRW contains one small lake. The SRTM topographic dataset was derived from spaceborne INSAR and was recently released into the public domain as part of a collaborative effort by U.S., German, and Italian government agencies (Farr and Kobrick, 2000). The original MRW DEM was in geographic coordinates with a resolution of 1 arc-second and was converted to a Universal Transverse Mercator (UTM) 30 m resolution elevation grid using a projection conversion utility and cubic-convolution resampling. The original integer meter vertical resolution of the data was also converted to floating point during resampling. The MRW DEM did not contain any patches of missing data, which are commonly observed in SRTM data that contain large water bodies or steep topography (Farr and Kobrick, 2000).

We would like to have compared lidar and INSAR data across landscapes. However, lidar data were not available for the MRW and the SRTM data that is available for the UUS and TLW is at a 3 arc-seconds (approximately 90 m) resolution, which was not deemed adequate for the current study because both sites were covered by too few grid cells. Henceforth, the UUS, TLW, and MRW DEMs will be referred to as flat, moderate, and steep.

Methods
Depressions Caused By Grid Representation
A series of tests were conducted to evaluate the occurrence of depressions caused by grid spacing (d) and grid placement. All depressions and flat areas were first removed from the three study DEMs using the depression-filling algorithm of Planchnon and Darboux (2001). Depression filling may have decreased the degree of surface roughness in the lidar and INSAR DEMs slightly; however, by using depressionless DEMs, all of the resulting depressions could be attributed to flow path obstructions related to grid spacing and placement. The depressionless DEMs served as the basis for deriving several elevation grids with coarser grid spacing using a nearest-neighbor resampling algorithm (Mather, 1987). Three DEMs were created at each of the tested resolutions to consider the effect of grid placement. The first grid coincided with the coordinates of the edges of the original DEM. The second grid was shifted in the vertical direction by 1/2d, and the third grid was shifted the same distance in the horizontal direction: 1/2d represents the largest shift distance that can be made when displacing two equally spaced grids. The number and area of depressions were then measured in each of the derived DEMs. The number of depressions was considered to be the number of groups of contiguous grid cells without downslope neighbors and not occurring along a DEM edge. This measure takes into account single-cell pits and nested or composite depressions (McCor- mack et al., 1993). The analysis included normalization, by comparing the relation between grid spacing and the density of depressions rather than simply the number of depressions, so that comparisons could be made among the three landscapes.

Depressions Caused By Random Error in Elevation
The probability that a grid cell belongs to a depression (p_{dep}) is a function of the topography and the error distributions associated with the cell, its neighbors, and downslope cells. This is a complex problem that is ideally suited to investigation by stochastic modelling. Therefore, the Monte Carlo method was used to evaluate the occurrence of depressions resulting from random error. A depression-filled DEM (i.e., containing no depressions or flats) was considered to be the true topographic surface, Z, so that all depressions generated during the simulation could be attributed entirely to random elevation error. An error matrix, E, a realization of a population of random errors, was then added to Z. The error-added DEM was once again depression-filled. A temporary matrix, T, was created and initialized with zeros. A value of one was added to each cell in T that corresponded to a cell in the error-added DEM that was modified by the depression-filling algorithm. This process was then repeated, adding a different E (drawn from the same error distribution) to Z each time. Dividing T by the realization number, n, yielded the depression probability matrix, P_{dep}, an image that describes the spatial pattern of the probability of a depression occurring in a landscape given a specified error distribution in the DEM. The simulation ended when the root-mean-square-difference between realizations was lower than a threshold of 0.001, indicating a stable solution. Depending on the DEM and error magnitude, convergence typically occurred between 200 and 400 realizations.
Random errors in DEMs are often assumed (a) to form a Gaussian (i.e., normal) distribution, with a mean of zero and standard deviation ($\sigma$) equal to the root-mean-square-error (RMSE), and (b) to be spatially autocorrelated (Hunter and Goodchild, 1997). Each $E$ in the Monte Carlo procedure was generated by first creating a stationary error field (i.e., a Gaussian field with very low spatial autocorrelation). The spatial autocorrelation of the stationary error field was increased using a $3 \times 3$ moving average filter, based on the neighborhood autocorrelation method described by Wechsler (2000). This is an efficient method of adding a small degree of spatial autocorrelation to an error field, although it has the unfortunate effect of decreasing $\sigma$ of the resulting error distribution. To overcome this problem, the filtered error matrix was rescaled such that $E$ had the desired $\sigma$. Note that the root-mean-square-difference between realizations, used to determine when the simulation was complete, is different than the RMSE, which was used to define random surfaces for the simulation.

The stochastic simulation procedure is computationally intensive. Therefore, the goal was to conduct the simulation procedure on an area that represented topographic variation in the landscape, but possessed a grid size (i.e., number of grid cells) that was computationally possible. For the flat DEM an expanded region of 145,536 grid cells was used at a 10 m resolution, by resampling the original DEM using a cubic convolution algorithm. For the moderate DEM a representative sub-region of 134,302 grid cells was used at a 25 m resolution. For the steep DEM a representative sub-region of 123,152 grid cells was used at a 30 m resolution.

Several RMSE values were used in a series of simulations to test the sensitivity of $P_{dep}$ to variations in error magnitude. The RMSE values were selected to demonstrate the range of values for which each landscape was sensitive. The stochastic simulation procedure described above was developed as a module in the Terrain Analysis System (TAS), a freeware GIS software package that can be obtained from the authors.

**Results and Discussion**

**Depressions Caused By Grid Representation**

The relation between $d$ and the density of depressions ($D_{dep}$) resulting from grid representation was found to follow an inverse power function (Figure 1). Non-linear regression analyses were performed using the SPSS Curve Estimation utility. Tests showed that $d$ was a strong and significant predictor of $D_{dep}$ for each of the three landscapes (Table 1). The sensitivity of the $d$ versus $D_{dep}$ relation was inversely related to local relief, i.e., the exponent ($b$) in the inverse power function varied such that flat > moderate > steep (Table 1). For a given value of $d$, $D_{dep}$ also varied such that flat > moderate > steep (Figure 1), and thus, lower relief areas were more likely to experience depressions resulting from inadequate grid spacing.

The density of depressions ($D_{dep}$), and therefore the number of depressions ($N_{dep}$), was insensitive to variations in grid placement. This was evident in Figure 1 by the minimal scatter observed among the three grids that were created for each value of $d$, i.e., many of the data points corresponding to three grid placements overlapped to the extent that they were indistinguishable. It should be noted that each of the three series in Figure 1 do not start with the finest resolution DEMs because these models were depressionless. Furthermore, the upper value in the range of $d$ for each landscape (Figure 1) was determined by a need to have a reasonable number of grid cells in the coarsened DEM to identify depressions.

The inverse power function relating $D_{dep}$ and $d$ has serious implications for the common practice of removing depressions from DEMs as a pre-processing step for modeling overland flow networks. Most of the algorithms that are commonly used to remove depressions from DEMs rely on identifying individual depressions prior to removal (e.g., Jenson and Domingue, 1988). These algorithms are already at the limit of DEM size that they can reasonably handle, even with increased computing power (Planchon and Darboux, 2001; MacMillan et al., 2003). Because even moderate increases in DEM resolution can be expected to significantly increase $N_{dep}$, it is likely that these pre-processing algorithms will be of limited use in the future. Planchon and Darboux (2001) describe a method of filling depressions that does not require prior identification of individual features, and thus, is very efficient. Although this algorithm is insensitive to data size and complexity, Lindsay and Creed (in press) found that depression-filling based algorithms impact DEMs and their derivatives significantly more than other methods. Therefore, there is a need to develop alternative methods for efficiently removing depressions from DEMs that contain numerous depressions.

At first, it may appear illogical that $N_{dep}$ is far greater in finer-resolution DEMs; this implies that although finer-resolution DEMs represent morphology more accurately than coarser-resolution DEMs, representation of flow paths is poorer. However, the non-linearity and direction of the $d$ versus $N_{dep}$ relation is not surprising given that the number of grid cells in a DEM ($N$) decreases by a power of two with

<table>
<thead>
<tr>
<th>Site</th>
<th>$n$</th>
<th>$a^*$</th>
<th>$b^*$</th>
<th>$r^2$</th>
<th>$F$ (df)</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>18</td>
<td>21391$^*$</td>
<td>1.6936$^*$</td>
<td>0.993</td>
<td>2414.67$^*$ (1,16)</td>
<td>0.1272</td>
</tr>
<tr>
<td>Moderate</td>
<td>27</td>
<td>2038$^*$</td>
<td>1.5392$^*$</td>
<td>0.993</td>
<td>3872.36$^*$ (1,25)</td>
<td>0.1204</td>
</tr>
<tr>
<td>Steep</td>
<td>18</td>
<td>32.11$^*$</td>
<td>0.9327$^*$</td>
<td>0.989</td>
<td>1471.54$^*$ (1,16)</td>
<td>0.0554</td>
</tr>
</tbody>
</table>

$^*$Unstandardized co-efficient; $p < 0.001.
increased \( d \) Therefore, it is necessary to consider the ability of finite grid representation to cause depressions relative to \( N \). If the relation between \( d \) and \( N_{\text{dep}} \) were scale-independent, \( b \) would equal two. However, we infer that the relation between \( d \) and \( N_{\text{dep}} \) is scale-dependent because \( b \) was <2 for each DEM (Table 1). This scale dependence is evident when \( d \) is plotted against \( N_{\text{dep}}/N \) (Figure 2). Simple linear regression analyses showed that each site exhibited a significant direct relation between \( d \) and \( N_{\text{dep}}/N \) (Table 2). Thus, while the absolute number of depressions decreases with increased grid spacing, the relative number of depressions actually increases.

Figure 2 demonstrates similar trends with respect to local relief as those noted for the \( d \) versus \( D_{\text{dep}} \) relation (Figure 1). The sensitivity of the \( d \) versus \( N_{\text{dep}}/N \) relation (i.e., the slopes in Table 1) and the likelihood of occurrence of depressions resulting from inadequate grid spacing (demonstrated by the relative position of lines in Figure 2) varied such that flat > moderate > steep. The \( d \) versus \( N_{\text{dep}}/N \) relation exhibited greater sensitivity to grid placement than the \( d \) versus \( D_{\text{dep}} \) relation, particularly at larger \( d \) values and for the flat and moderate DEMs. The difference in the sensitivity to grid placement between the flat and moderate DEMs and the steep DEM may partially reflect differences in data sources (i.e., lidar versus INSAR), although it is more likely a result of landscape features. The steep DEM contained one small lake and the flat and moderate DEMs each contained several lakes and wetlands. Whether a lake or wetland appeared as a depression in a DEM depended on how well represented the feature’s outlet was in the elevation grid, which was largely dependent on grid placement. This phenomenon is further illustrated by examining the relation between \( d \) and the total area of depressions, \( A_{\text{dep}} \) (Figure 3). A very large variation in \( A_{\text{dep}} \) among the three grids created for each \( d \) value is apparent in the flat and moderate DEMs, indicative of the relatively extensive wetlands and lakes in these regions. This high sensitivity to grid placement in the flat and moderate DEMs reduces the predictive strength and reliability of the \( d \) versus \( A_{\text{dep}} \) linear regression models substantially (Table 3).

To test the sensitivity of these findings to the method used to coarsen the resolution of the DEMs (i.e., nearest neighbor resampling), we repeated the analyses using bilinear interpolation and cubic convolution resampling algorithms. In terms of both \( N_{\text{dep}} \) and \( A_{\text{dep}} \), nearest neighbor < bilinear interpolation < cubic convolution. However, each of the above relations held true despite these differences.

### Table 2. Linear Regression Results for Relations Between the Number of Artifact Depressions Resulting From Grid Obstructions Per Grid Cell and Grid Spacing for Three Study Landscapes

<table>
<thead>
<tr>
<th>Site</th>
<th>( n )</th>
<th>( \text{Constant}^* )</th>
<th>Slope*</th>
<th>( r^2 )</th>
<th>( F ) (df)</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>18</td>
<td>( 3.145 \times 10^{-2} )</td>
<td>9.990</td>
<td>( 10^{-4} )</td>
<td>0.609</td>
<td>24.97 (1, 16)</td>
</tr>
<tr>
<td>Moderate</td>
<td>27</td>
<td>3.090 \times 10^{-3}</td>
<td>8.467</td>
<td>( 10^{-3} )</td>
<td>0.890</td>
<td>203.04 (1, 25)</td>
</tr>
<tr>
<td>Steep</td>
<td>18</td>
<td>(-1.620 \times 10^{-4} )</td>
<td>4.608</td>
<td>( 10^{-3} )</td>
<td>0.986</td>
<td>1118.05 (1, 16)</td>
</tr>
</tbody>
</table>

*Unstandardized co-efficient; † \( p < 0.001 \); ‡ not significant ( \( p > 0.05 \)).

### Table 3. Linear Regression Results for Relations Between the Area of Artifact Depressions Resulting From Grid Obstructions Per Grid Cell and Grid Spacing for Three Study Landscapes

<table>
<thead>
<tr>
<th>Site</th>
<th>( n )</th>
<th>( \text{Constant}^* )</th>
<th>Slope*</th>
<th>( r^2 )</th>
<th>( F ) (df)</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>18</td>
<td>25.775</td>
<td>(-1.280 \times 10^{-14} )</td>
<td>0.210</td>
<td>4.27 (1, 16)</td>
<td>2.1229</td>
</tr>
<tr>
<td>Moderate</td>
<td>27</td>
<td>3.642</td>
<td>4.843</td>
<td>( 10^{-2} )</td>
<td>0.519</td>
<td>26.96 (1, 25)</td>
</tr>
<tr>
<td>Steep</td>
<td>18</td>
<td>0.538</td>
<td>9.147</td>
<td>( 10^{-3} )</td>
<td>0.971</td>
<td>533.90 (1, 16)</td>
</tr>
</tbody>
</table>

*Unstandardized co-efficient; † \( p < 0.001 \); ‡ not significant ( \( p > 0.05 \)).
presence of hotspots within landscapes that are highly sensitive to depressions caused by random elevation error. Areas of high probability of depression occurrence were predominantly located in flatter areas and convergent topography along valley bottoms (Plate 1). The steepest areas within each landscape were the least likely to experience a depression during the simulations.

Several researchers have also asserted that depressions are abundant in areas of gentle relief, owing to lower signal-to-noise ratios (Mark, 1988; Rieger, 1998; Burrough and McDonnell, 1998; Liang and Mackay, 2000). The results of the stochastic simulations allow this relation to be evaluated quantitatively. The slope between a grid cell and its steepest downslope neighbor ($\beta_{\text{SDN}}$) can be used as a
measure of the signal. Based on the observations from the simulations, the probability of depression occurrence (i.e., $p_{\text{dep}}$) is inversely related to $\beta_{\text{SDN}}$ and directly related to the RMSE, i.e.,

$$p_{\text{dep}} = \frac{\text{RMSE}}{\beta_{\text{SDN}}}$$

(1)

Because the distance in the steepest downslope direction ($d_{\text{SDN}}$) is largely dependent on $d$, so is $\beta_{\text{SDN}}$. This complicates comparisons of $p_{\text{dep}}$ among DEMs of different resolutions. To remove this effect, a dimensionless measure of error, $\text{RMSE}/d_{\text{SDN}}$, was used. Therefore, Equation 1 reduces to

$$p_{\text{dep}} = \frac{\text{RMSE}}{\Delta z_{\text{SDN}}}$$

(2)

where $\Delta z_{\text{SDN}}$ is the elevation drop in the direction of the steepest downslope neighbor. It should be noted that $\Delta z_{\text{SDN}}$ can also be expected to be somewhat related to $d$ in certain landscapes. However, Equation 2 is clearly less dependent on $d$ than is Equation 1. Finally, expressing Equation 2 in terms of a signal-to-noise ratio,

$$p_{\text{dep}} = \left( \frac{\Delta z_{\text{SDN}}}{\text{RMSE}} \right)^{-1}$$

(3)

Figure 4 plots $p_{\text{dep}}$ values from the flat DEM (RMSE = 0.50 m), the moderate DEM (RMSE = 0.25 m), and the steep DEM (RMSE = 3.00 m) simulations against this measure of signal-to-noise ratio. It is interesting how similar these relations are given the contrast in terrain of the three landscapes. Each relation demonstrates that for signal-to-noise ratios greater than approximately two, the probability of a depression occurring is negligible. Furthermore, $p_{\text{dep}}$ increases exponentially as the signal-to-noise ratio tends to zero with significant scatter. This scatter in the relation at lower signal-to-noise ratios is likely the result of other factors of the topographic setting. For example, a location may have a high signal-to-noise ratio but highly convergent topography (i.e., a flow path bottleneck) directly downslope can cause upslope blockages in flow. Proper representation of terrain at flow path bottlenecks is important because of their sensitivity to interruptions by random error. For example, Figure 5 plots the average $p_{\text{dep}}$ against the number of downslope neighbors for the moderate DEM simulations (RMSE = 0.15 m, 0.25 m, and 0.50 m). Grid cells with only one downslope neighbor (i.e., flow path bottlenecks) demonstrated relatively high average values of $p_{\text{dep}}$ regardless of the RMSE. It is likely that there is upslope carryover of this effect, by which any flatter areas behind these potential dams also have locally elevated $p_{\text{dep}}$.

The RMSE, and therefore signal-to-noise ratio, in remotely-sensed DEMs can be effectively reduced through data redundancy; however, some degree of random error will always be present. Averaging elevations from multiple measurements reduces the error by the square root of the number of data takes. Although this can significantly reduce the extent of areas within a landscape that are vulnerable to the occurrence of depressions resulting from random error (Plate 1), it will not reduce the probability of depressions in landscape hotspots (e.g., flats, valley bottoms, and flow path bottlenecks), and is dependent on data availability.

**Conclusions**

Evaluation of the ability of inadequate grid representation and random error in grid elevations to cause depressions in lidar and InSAR DEMs across a range of physiographic regions revealed the following major findings:

1. Both inadequate grid representation and random elevation error can be effective mechanisms for causing depressions in remotely-sensed DEMs, particularly in certain topographic settings.
2. The number of depressions in a DEM resulting from inaccurate grid representation is related to the grid spacing by an inverse power function, with observed exponents ranging...
It may not be prudent to remove actual depressions from scale, i.e., there are many more small depressions than large. The extent of areas with lower probabilities of occurrence of depressions in landscapes increases with decreased signal-to-noise ratios. The probability of a depression occurring in a digital landscape as a result of random error are sufficient to cause depressions in the DEM given signal-to-noise ratios exceeding two and therefore, it may be necessary in the future to develop alternative methods for removing depressions from DEMs that do not require identifying individual depressions. For similar grid spacings, flat landscapes contained far more depressions resulting from inadequate grid representation. Grid placement affected the occurrence of depressions in landscapes containing extensive wetlands or lakes, owing to the difficulty in adequately representing outlets with grids. Stochastic modelling showed that flat areas, both in extensive and narrow areas (e.g., valleys), are more likely to experience depressions resulting from random error because of their lower signal-to-noise ratios. The probability of a depression occurring in a digital landscape as a result of random error was negligible for signal-to-noise ratios greater than two and increased exponentially at lower values. Additionally, highly convergent topography was found to be very sensitive to the occurrence of depressions caused by random elevation errors. Reducing the magnitude of error in a DEM (e.g., by averaging multiple elevation measurements) can lessen the extent of areas with lower probabilities of occurrence of depressions within landscapes; however, even small levels of random error are sufficient to cause depressions in the landscape hotspots described above.

All of the depressions observed in the DEMs in this study were known to be artifacts; however actual depressions do occur in some natural landscapes. Furthermore, the number of actual depressions in landscapes increases with decreased scale, i.e., there are many more small depressions than large. It may not be prudent to remove actual depressions from DEMs used in hydro-geomorphic applications because of their potential control on environmental processes. Therefore, future research will focus on developing reliable methods of automatically discriminating between actual and artifact depressions in fine-resolution DEMs.

Acknowledgments
The authors would like to thank the three anonymous reviewers for their valuable feedback on early drafts of the manuscript. John Lindsay received financial support through the Natural Sciences and Engineering Research Council (NSERC) Postgraduate Scholarship B and the Ontario Graduate Scholarship program. Irena Creed acknowledges financial support through research grants including an NSERC Collaboration Research and Development (CRD) Grant with partners from Alpacia Ltd., Weyerhaesuer Ltd., and Ducks Unlimited-Canada for funding data acquisition for the lidar DEM in the Utikuma Highlands, P. Treitz (Queens University) and Geomatics for Informed Decisions (GEODE) for the lidar DEM in the Algoma Highlands, and her NSERC Discovery Grant.

References


(Received 21 January 2004; accepted 23 March 2004; revised 06 May 2004)