The Effect of Grid Cell Size on Major Terrain Derivatives

Stefan W Kienzle, Department of Geography, University of Lethbridge,
4401 University Drive, Lethbridge, Alberta, T1K 3M4, Canada
Tel.: 403 380 1875
Fax: 403 329 2016
stefan.kienzle@uleth.ca

Abstract

This paper examines the quality of DEMs derived from the interpolation of photogrammetrically derived elevation points. DEMs with grid cell sizes ranging from 100 to 5 m were interpolated from 100 m regularly spaced elevation points and many surface-specific point elevations using the ANUDEM interpolation method. In order to identify the grid resolution that matches the information content of the source data, three approaches were applied: density analysis of point elevations, an analysis of cumulative frequency distributions using the Kolmogorov-Smirnov test and the root mean square slope measure. Results reveal that the optimum grid cell size is between 5 and 20 m, depending on terrain complexity and terrain derivative. Terrain variables based on 100 m regularly sampled elevation points are compared to an independent high-resolution DEM used as a benchmark. Correlation analysis, root mean square errors, and relative root mean square errors further quantify the quality of terrain derivatives.

1 Introduction

The proliferation of digital elevation sources and terrain analysis tools enables researchers and operators in environmental science, agriculture, hydrology, biology and engineering to compute terrain-dependent variables and indices easier than ever before. They extend from slope and aspect analyses using a variety of algorithms (e.g. Evans 1980, Horn 1981, Zevenbergen and Thorne 1987) to the estimation of more sophisticated terrain attributes such as the drainage network (e.g. Jenson and Domingue 1988, Tarboton et al. 1991, Garbrecht and Martz 1997), or the topographic wetness index (Beven and Kirkby 1979). By using one or several terrain variables and indices, a catchment area can be characterized in terms of geomorphology (e.g., Jenson and Domingue 1988, Moore et al. 1993), stream network patterns (e.g. Tarboton et al. 1991, Band 1993) or landform classification (Blaszcynski 1997, Macmillan et al. 2000). Subsequently, ecosystem modelling, such as modelling the distribution of energy, water, sediments, nutrients (e.g. Mitasova et al. 1996, Kienzle et al. 1997, Klaghofer et al. 1993) and pollutants (e.g. Kern and Stednick 1993), depends on realistic terrain representation.

For the province of Alberta, Canada, elevation data with the highest spatial resolution are provided by AltaLIS, the agent for Spatial Data Warehouse, which is a not_for_profit organization maintaining and promoting Alberta's digital mapping. DEMs are available for the so-called “white area”, which is the surveyed area of Alberta. These elevation data sets are available in 100 m regular grids with many additional surface-specific point elevations to define the framework of the terrain, including spot heights and points along ridges, streams and saddles. DEMs and DTMs (digital terrain models, containing terrain variables such as slope, aspect or curvature rather than elevation values) based on this data set are referred to as DEM_{100} or DTM_{100}.

It is well known that most terrain attributes derived from a DEM change with a variation in the underlying grid cell size. **Figure 1** illustrates the impact that the interpolated grid cell size has on terrain representation. Elevations were interpolated to grid cell sizes of 100, 50, 25 and 5 m, using the same input data and interpolation algorithm (ANUDEM, Hutchinson 1989, 1991). It is evident from **Figure 1** that terrain features such as slope, aspect or upslope area will depend on DEM resolution. **Figure 2** shows the location of the profile line used for this demonstration. Previous studies have shown that the DEM grid cell size significantly effects derived terrain parameters (Kienzle 1994; Zhang and Montgomery 1994; Elsheikh and Guercio 1997). Kienzle (1994) has used terrain variables to distinguish between terrain units and estimate soil erosion potential for catchments in South Africa. He determined that in the 179 km^2 Henley Dam catchment in the Midlands of KwaZulu-Natal, South Africa, both slope and soil erosion estimations increase with a decrease in interpolated grid cell size (Table 1).

Table 1 Geomorphological parameters estimated from DEMs with varying grid cell size in the Henley catchment, KwaZulu-Natal, South Africa (from Kienzle, 2004, with permission)

Variable	Grid cell size	Mean	Max
Slope [%]	250 m	8.8	33
	100 m	13.4	85.5
	50 m	13.9	143.2
Soil loss [t ha^{-1} $year^{-1}$]	250 m	21.4	164
	100 m	31.3	303
	50 m	35.4	603

Saulnier et al. (1997) investigated the analytical compensation between DEM grid resolution and hydrological terrain derivatives and found, for example, that the topographic index, which combines the local slope and the associated upslope area, increases with grid cell size. This means that catchments are modelled to be wetter using a coarse grid cell size and drier using a finer grid resolution. Other investigators, such as Zhang and Montgomery (1994), used photogrammetrically sampled elevation data and compared grid cell sizes ranging from 2 to 90 m for two small catchments in Oregon (Mettman Ridge catchment) and California, USA (Tennessee Valley catchment). They derived a mean slope of 65% and 34% for the 2 m grid and 41% and 29% for the 90 m grid and found that the grid size significantly effects the cumulative frequency distributions of the specific catchment area, the topographic index and, consequently, hydrological simulations.

One can summarize that the ability to carry out realistic terrain analyses is limited primarily by the quality of the DEM applied in terms of

- the accuracy and distribution of the elevation points used to interpolate the DEM
- the interpolation algorithm to generate a continuous DEM, and
- the chosen grid cell size

and will effect subsequent modelling of surface processes, such as erosion, deposition, slope stability, hydrological or water quality processes.

A grid cell size is often selected not with a specific subsequent terrain analysis in mind, but to overlay a DEM with other raster data such as satellite imagery, where the grid cell size is predetermined. While this is an important consideration, one should be aware of the consequences that the grid cell size has on terrain analyses. For example, there may be merit in choosing a fraction of the given grid resolution such as 12.5 or 6.25 m grid cell sizes for overlay with 25 m resolution Landsat imagery. It is important to know the exact grid cell size and projection of other raster data before one begins the creation of a DEM.

In this paper, the author systematically assesses how grid cell size effects terrain derivatives using elevation points available in Alberta, Canada, by applying the same interpolation algorithm and two sets of photogrammetrically derived elevation points at two different scales. Terrain derivatives of the first and second order are investigated as well as the wetness index representing compound terrain derivatives. This analysis also allows one to statistically define the grid resolution that corresponds to the content of the source data. Issues associated with the quality of sampled elevation points or various DEM interpolation techniques are not addressed.

2 Study Area

In order to represent areas typical for the Rocky Mountain foothills and Great Plains regions of western North America (**Figure 3**), three areas were chosen with steep, moderately sloped and flat relief (**Figure 4**). Mountainous areas with very steep slopes are not investigated. Study Area 4 lies within the city boundaries of Lethbridge and is used to compare the DEM derived from the provincial elevation database to one with a higher resolution. Table 2 lists some key terrain parameters to characterize the different study areas.

Table 2 Key terrain characteristics of the four study areas. Slope values are based on a 10 m DEM using Horn's (1981) method (from Kienzle, 2004, with permission)

Study Area	Scale of Aerial Photo	Size [km^2]	Density of sampled elevation points [points km^{-2}]	Mean slope [%]	Max Slope [%]	Elevationr ange [m]
1	1:60,000	15.21	484	24.1	78.4	326
2	1:60,000	15.21	251	12.3	55.9	235
3	1:60,000	15.21	178	1.1	18.1	66
4	1:60,000	2.25	345	3.9	34.7	107
4	1:10,000	2.25	11162	4.1	33.8	105

Study Areas 1 to 3 are square areas, containing a small watershed of about 7 km^2, and are 3.6 by 3.6 km in size. This size was chosen because it is a multiple of all grid cell sizes that are to be investigated, thus eliminating possible errors in calculating terrain derivatives along the edges. Study Area 4 is 1.5 by 1.5 km in size and lies to the west of the Oldman River within the city limits of Lethbridge. The area is not urbanized and includes very flat areas as well as steep slopes along the 60 m deep river valley.

Study Areas 1, 2 and 3 are used to determine significant differences in terrain derivatives from a variety of grid cell sizes under terrain conditions ranging from moderately steep to very flat. Study Area 4 is used to compare DEMs derived from two independent sources in order to reveal the quality of terrain attributes derived from the DEM_{100}, which is available for most of Alberta.

3 Analysis

3.1 Terrain Derivatives

Instead of analyzing elevation data directly, terrain derivatives are analyzed to reveal statistically significant differences. Slope and aspect (slope direction) are the first derivatives of the mathematically continuous surface that constitutes a raster DEM. Both Horn's and Zevenbergen and Thorne's algorithms are widely used, although numerous other slope and aspect estimators exist (Burrough and McDonnell 1998). Horn's (1981) third-order finite difference estimator is used by the SLOPE command in ArcView and ArcInfo (both products of Environmental Systems Research Institute Inc., 2000), whereas Zevenbergen and Thorne's (1987) second-order finite difference method is incorporated in ArcInfo's CURVATURE command. Jones (1998) investigated eight algorithms for calculating slope and aspect using both actual and manufactured DEMs. His research revealed that Zevenbergen and Thorne's method ranked best and Horn's method ranked second best.

The second order derivatives, the rates of change of slope in down slope direction (profile curvature) and perpendicular to the down slope direction (plan curvature) are investigated because of their importance in geomorphological and hydrological analyses. Finally, terrain indices that combine two or more terrain attributes (compound derivatives) are analyzed because of their wide applications and expected sensitivity to grid resolution. The curvature, which is the product of plan and profile curvature, is important in modelling erosion and runoff processes. The second compound derivative chosen is the wetness index (also called the topographic index). It is extensively used in hydrology, agriculture, geomorphology and vegetation studies because it represents the spatial distribution of soil moisture, surface saturation, groundwater recharge and discharge areas, as well as potential runoff generation (variable source areas). Table 3 lists the terrain derivatives tested.

Table 3 Terrain derivatives tested

Order of Derivative	Terrain attribute
First order derivatives	Slope (Horn)
	Slope (Zevenbergen and Thorne)
	Aspect (Horn)
	Aspect (Zevenbergen and Thorne)
Second order derivatives	Profile curvature
	Plan curvature
Combined derivatives	Curvature
	Wetness index

3.2 Comparison of Two Independent DEMs

In addition to identifying significant differences between important terrain derivatives for a variety of grid resolutions, the terrain derivatives are compared to an independently sampled, high resolution DEM. Major cities in Alberta have DEMs that are sampled photogrammetrically at 10 m regular intervals, referred to here as DEM_{10}. High resolution DEMs with grid cell sizes of 5 and 1 m are used to evaluate the errors of terrain derivatives based on the DEM_{100}. Correlation analysis is applied to reveal the strength of relationships between terrain derivatives based on the two independent data sources. Using the DEM_{10} as a benchmark, RMS and relative root mean square (RRMS) errors are computed for elevation and terrain derivatives to indicate the quality of terrain derivatives based on DEM_{100} .

In order to determine the smallest grid cell size that can be used for the interpolated grids, the nearest distances between sampling points were calculated for the four study areas. **Figure 5** shows the spatial distribution of sampled point elevation densities. The cumulative frequency distributions of the distances of the nearest sampled elevation points are shown in **Figure 6**. This graph shows that 50% of the point elevations in the study areas, with the exception of the very flat Study Area 3, have a nearest distance to each other of less than 24 m (Study Area 1), less than 30 m (Study Area 4) and less than 35 m (Study Area 2). All study areas have sampled elevation points with a nearest distance of between 10 and 15 m for 10% of their area. Nearest distances can be as close as between 3 and 6 m. At locations where the sampling density is high (close distances between sampled elevation points) the terrain is typically complex, which will locally effect slope, aspect, flow direction, erosion potential, radiation budgets etc. Based on the local density, it can be justified that the smallest grid cell size that can be used from the DEM_{100} may be less than 10 m, and therefore grid cell sizes as low as 7.5 and 5 m were interpolated.

The fourth study area is situated in the City of Lethbridge. The mean sampling density of the AltaLIS data set is 345 samples km^{-2}, while the city data set provided 11,162 samples km^{-2}.

3.3 Interpolation of DEMs

The TOPOGRID command in ArcInfo, which incorporates the 1996 version of ANUDEM, requires that stream network data have all arcs pointing down slope. Consequently, an algorithm was applied to ensure that this condition was met. After proper preparation, the TOPOGRID command was executed to produce an initial DEM with a 25 m resolution. Using this DEM, Study Areas 1, 2 and 3 were selected to represent three terrain types commonly found in southern Alberta. The size of the sub-sets was set to be 3,600 by 3,600 m, chosen to be a multiple of all grid cell sizes to be produced for the final analyses. During the subsequent execution of the TOPOGRID command the size of the sub-sets was increased in all directions by at least 20 times the chosen grid size to avoid potential edge problems of the resulting DEMs. The DEMs with resolutions of 100, 80, 75, 50, 30, 25, 20, 15, 12.5, 10, 7.5 and 5 m were then clipped to their original size of 3,600 by 3,600 m.

For Study Area 4 one section was made available by the City of Lethbridge. To avoid edge problems during the interpolation process, a central portion with the size 1,500 by 1,500 m was used as the final DEM.

3.4 Computation of Terrain Derivatives

The terrain derivatives were calculated for each DEM using the SLOPE (Horn's method), ASPECT (Horn's method) and CURVATURE (Zevenbergen and Thorne's method for slope, aspect, profile and plan curvatures, and compound curvature) commands in ArcInfo. The wetness index was calculated according to the equation:

$$w = \ln(A_s / \tan\beta) \qquad (1)$$

where A_S is the specific catchment area (catchment area divided by the cell width in slope direction) and is the local slope of the terrain in degrees (Beven and Kirkby 1979).

3.5 Statistical Analyses and Results

3.5.1 Comparison of terrain variables derived from different grid resolutions

For all terrain derivatives and all grid resolutions cumulative frequency distributions (CFDs) were computed. As an example, for the general patterns found, CFDs for five grid resolutions were calculated for slope **(Figure 7**), profile curvature **(Figure 8)** and the wetness index **(Figure 9)** for Study Area 1. The patterns presented in **Figures 7, 8 and 9** are the same for all four study areas. Generally, the larger the grid cell size the smaller the derived slope values become (**Figure 7**). The two different slope algorithms employed also reveal the same frequency distribution patterns. Differences in slope estimation are larger in terrain with high relief (Study Area 1) than in terrain with low relief (Study Area 3). Grid cell sizes over 25 m are not able to identify steep slopes successfully. This has a particular impact where slope stability or erosion is being estimated from relatively coarse DEMs. Calculated differences of slope generally become smaller with grid cell sizes under 25 m. It becomes evident from this analysis that the 100 and 50 m DTMs, and to a lesser extent the 25 m DTM, result in a considerable underestimation of slope values, particularly along hill slopes.

The profile curvature is strongly underestimated using larger grid cell sizes. A very similar pattern exists for plan curvature. Again, the same patterns are found in all study areas. A characterization of an area in terms of concavity or convexity information fails to an increasing degree with higher DTM grid cell sizes. The impact of underestimating plan or profile curvatures is to underestimate dispersion and convergence areas, which are particularly important for erosion, sedimentation and hydrological analysis.

The wetness index, used as an example for compound terrain variables in **Figure 9**, varies strongly with different grid resolutions. The general rule for all study areas is that with increasing grid cell size there is a more or less parallel shift towards higher wetness indices, with higher values indicating the potential for higher soil saturation. Therefore, when comparative studies between a number of watersheds are undertaken, the wetness would be overestimated with a coarser DEM. High relief terrain (Study Area 1) has lower mean wetness indices (6.3), representing drier soils, than terrain with moderate relief in Study Area 2 (7.2) and the flat Study Area 3 (9.4).

Based on the CFDs for Study Area 4, key percentiles were extracted and plotted against the grid resolution to reveal the sensitivity of the distribution parameters on grid cell size (**Figures 10 to 14**). In order to evaluate the distribution of terrain variables derived from DEM_{100}, the distribution values for terrain variables derived from a DEM_{10} are also displayed.

Figures 10 to 14 show selected percentile values of the CFDs for Study Area 4 to reveal differences of terrain values as a function of both grid cell size and DTM source. In these Figures, "10 m" signifies the DEM_{10} and "100 m" the DEM_{100}. There is a wealth of information that can be derived from this series of graphs. Slope grids with grid cell sizes of 15, 12.5, 10, 7.5 and 5 m have very similar distributions (**Figure 10**) and show similar values even for the 99^{th} percentile slope values, which represent the steepest one percent of the study area. Median slope values are also quite similar, although the larger grid cell sizes seem to slightly overestimate the median slope. This is due to the relatively greater number of grid cells associated with the valley slopes within the coarser DTMs. The slope distribution based on 5 m grid cell sizes of DTM_{100} and DTM_{10} compare favourably, indicating that an overall characterization of slope over a terrain similar to Study Area 4 can be represented by a 5 m DTM_{100}. Both the Horn and Zevenbergen and Thorne methods result in similar slope distribution values over the range of grid cell sizes investigated, with slope values derived using Zevenbergen and Thorne's method consistently being slightly higher. Using the 10m grid cell sizes based on DEM_{10} as an example, median slope values are 2.11% (Horn) and 2.18 (Zevenbergen and Thorne), while the 95 percentile slope values are 35.98 and 36.53 respectively.

Since a correct representation of aspect is crucial for the derivation of hydrological variables such as flow direction, flow accumulation, stream networks and catchment boundaries or environmental variables such as solar radiation or the wetness index, the potential misrepresentation of aspect values based on a DEM_{100} at any interpolated resolution may have significant consequences. In order to show the effects of grid resolution and DEM source, the differences in aspect angles were calculated, using the 5 m grid resolution based on DEM_{10} as a benchmark. An algorithm was developed to calculate the true angle differences between two grids in order to avoid challenges stemming from the fact that aspect angles of 359 and 1 are only 2 apart. The results in **Figure 11** show that, virtually independent from the grid cell size chosen, the aspect angles based on DEM_{100} are out by 35 to 40 for 50% of the study area, out by 70 to 80 for 25% of the study area (75^{th} percentile) and may be as much out as 150 and, in less than 0.5 % of cases, by a full 180 . No meaningful differences between the two algorithms used (Horn, 1981; Zevenbergen and Thorne, 1987) were observed.

Plan and profile curvature, as well as their product (curvature), show 95^{th} percentile values that increase in range with decreasing grid cell size (**Figure 12, 13**), showing an increase in convexity or concavity with decreasing cell size. Median values are close to zero, which means that approximately the same level of concavity and convexity occurs within the study area. Plan and profile values are, however, reaching a limit in their range when they are based on the DEM_{100}. As was pointed out earlier, this may significantly effect calculations of overland flow and its divergence or convergence and subsequent erosion or deposition simulations. The limitations of the DEM_{100} for detailed hydrological or environmental terrain analyses become evident.

Figure 14 shows key percentiles for the distribution of the wetness index. The wetness index values become smaller (signifying drier soils) with a decrease in grid cell size. The overall distributions of wetness index values based on DEM_{10} or DEM_{100} are quite similar, but median index values based on DEM_{100} tend to be slightly larger than those based on DEM_{10}. The grid cell size is therefore critical when one compares wetness index values from different regions which may be based on different grid cell sizes and DEM sources.

Figure 15 shows a visualization of Study Area 4 with the draped wetness index based on 5 m resolution derived from DTM_{100} and $DTM_{10.}$ It becomes evident that, particularly in flat terrain, the topographic position has a large impact on the wetness index, with low values (dry conditions) on small mounds and a concentration of potential water in shallow depressions, resulting in relatively high wetness index values there. Generally, within the flat part of Study Area 4, the wetness index values derived from DEM_{100} are considerably overestimated. This is largely due to problems in representing flow direction from DEM_{100}, which is imperative for the calculation of upslope areas, which, in turn, is used for the calculation of the wetness index. In steep terrain, the wetness index compares more favourably between DEM_{100} and DEM_{10}. A comparison of flow direction and aspect values, calculated using different algorithms, shows very similar spatial distribution over Study Area 4.

While the previous Figures offer visual impressions of the sensitivity of terrain derivatives on grid resolution, they do not offer a quantification of the level of significant differences between terrain variables based on different grid cell sizes, nor can they quantify any spatial correlations between respective terrain variables based on DEM_{100} and the benchmark DEM_{10}.

3.7.2 Optimization of Grid Cell Size Using the Kolmogorov-Smirnov Test

The question arises: Which grid resolution corresponds to the information content of the source data ? The first method used compares CFDs of terrain derivatives based on a number of different grid resolutions. In order to determine significant differences of terrain derivatives with decreasing grid cell size, each CFD was compared to the associated 5 m CFD by carrying out the two-sample Kolmogorov-Smirnov test. To be as unbiased as possible, only the grid cell sizes 80, 50, 30, 20, 12.5, 7.5 and 5 were used, because the ratio of two sequential cell sizes is consistently in a narrow range between 1.5 and 1.67. This test is the non-parametric equivalent to the two-sample t-test with unequal variances. It is typically applied when data sets to be compared are neither normally distributed nor meet the assumption of equal variances. The Kolmogorov-Smirnov test determines its test statistic by finding the point at which the two samples have the greatest difference between the cumulative proportions of the two samples. Consistently using 100 values representing all integer percentiles, the cumulative frequency distributions are considered different at the 95% confidence level when the maximum difference between the curves is greater than 0.1367. The differences must be greater than 0.1638 for the 99% confidence level. In cases where the difference between a 5 m grid and a 7.5 m grid was insignificant, and the difference between the 7.5 m grid and the 12.5 m grid was also insignificant, but the difference between the 5 m grid and the 12.5 m grid was significant, the 5 m grid would have been selected as the maximum grid cell to be used.

All results are listed in Table 4. Results from this test show at what resolution all relevant terrain information has been extracted from the source point data, and further refinement of the DTM resolution would not significantly change the tested terrain derivatives. Results for DEM_{100} show that first order derivatives can be based on DEMs with a 7.5 m grid cell size for steep terrains, while flat terrains can be represented with a 20 m resolution. A rule for slope becomes apparent in that the smaller the sampling density (the flatter the terrain), the coarser the slope grid can be. For Study Area 1 slight differences exist depending on the slope algorithm used.

Distributions of second order derivatives (plan and profile curvature) show a reverse trend: the higher the sampling density (the steeper the terrain) the coarser the grid cell size can be. For plan curvature, the cell sizes range from 12.5 for Study Area 1 to 5 m for Study Area 3.

For both compound terrain derivatives investigated (curvature and the wetness index), the smallest grid resolution examined (5 m) shows a significantly different distribution to the 7.5 m and larger grid cell sizes. One exception is the representation of curvature in Study Area 2, where the 5 m grid cell size is not significantly different from the 7.5 m grid, and therefore the 7.5 m grid reveals the best information. Based

Table 4 Maximum grid cell sizes that are significantly different from the next larger grid cell size. Grid cell sized tested are: 80, 50, 30, 20, 12.5, 7.5 and 5 m.

Test Feature	Study Area (Sampling Density in points km^{-2})		
	1 (484)	2 (251)	3 (178)
Slope (Horn)	7.5 m	12.5 m	20.0 m
Slope (Zevenbergen and Thorne)	12.5 m	12.5 m	20.0 m
Plan curvature	12.5 m*	5.0 m*	5.0 m
Profile curvature	12.5 m*	7.5 m	7.5 m
Curvature	5.0 m	7.5 m	5.0 m
Wetness Index	5.0 m	5.0 m	5.0 m

Test results of the Kolmogorov-Smirnov test for tested features and three study areas. The grid cell size reported is that grid cell size where values derived from smaller grid cell sizes are not significantly different. The next larger grid cell size would result in a value that is significantly different at the 95% confidence level (* marks a 99% confidence level).

on these results, researchers using the province-wide Alberta elevation points should be aware that grid cell sizes reported in Table 4 contain the best information that can be extracted from the original source data. It must be kept in mind, however, that the use of the reported grid cell sizes is no guarantee for realistic spatial representation of these terrain variables, as was shown earlier.

3.7.3 Optimization of Grid Cell Size Using the Root Mean Square Measure

Another measure to qualify the optimum grid cell size is to calculate the RMS of terrain derivatives and plot them against the grid cell size. This method was applied by Hutchinson (1996) and Hutchinson and Gallant (2000). The criterion for an optimum grid cell size is found at the point in the graph where the RMS value starts to flatten. RMS slope values are plotted against grid cell size for all four study areas in **Figure 16**. For Study Area 1, a minor flattening is apparent at a grid cell size of 7.5 m, which conforms with the results of the Kolmogorov-Smirnov test shown in Table 4. However, an outlier exists at a grid cell size of 20 m, which could either be the result of the localized grid cells (where a shift in grid cell location could result in different RMS slope values), or it could be the indication that the information contained in the source point elevations has reached an optimum at 20 m. Two such outliers are present in Study Area 2 at grid cell sizes 75 and 30 m. No area of flattening of the RMS slope can be found here, so the results from the Kolmogorov-Smirnov test cannot be confirmed. No clear optimum grid cell size can be derived from **Figure 16b**. A different pattern is shown in **Figure 16c** for Study Area 3. RMS slope values for grid cell sizes below 20 m seem to be erratic, suggesting that an optimum grid cell size has been reached at a grid cell size of 20 m. This value again coincides with the test results reported in Table 4. In **Figure 16d** results for both DEM_{100} and DEM_{10} are plotted for Study Area 4. While a clear flattening on RMS slope values for DEM_{100} at a grid cell size of 7.5 m is evident, no clear flattening can be found for DEM_{10}, indicating that the optimum grid cell size for DEM_{10} for slope is 1 m or less. Since the density of source elevation points in Study Areas 1 and 4 are quite comparable (484 and 345 point km^{-2} respectively), it is interesting that the optimum grid cell size for slope analysis of 7.5 m can be qualified for both study areas.

3.7.4 Testing for Correlations

The Pearson correlation coefficient is a measure of the strength of the linear relationship between two sets of data. A grid resolution of 5 m was chosen as the benchmark to compare the relationship of terrain variables derived from 10 m and 100 m regularly sampled elevation points respectively. Each 5 m grid cell containing a terrain value derived from the 10 m elevation points was correlated spatially to the relevant grid cell derived from the 100 m elevation points. Consequently, terrain values from a 5 m resolution were compared one-on-one with the benchmark grid, and terrain values with a 10 m resolution were compared to the four corresponding benchmark grid values and so forth. Only grid cell sizes ranging from 5 m to 30 m were analyzed. High correlation coefficient values, close to 1, indicate that the terrain values derived from 100 m elevation points represent the benchmark well, while low correlation coefficients indicate that a poor spatial correlation exists.

Table 5 Pearson correlation coefficients for terrain derivatives calculated from 5 m resolution based on the DEM_{10} vs terrain derivatives calculated from on a variety of grid resolutions based on the DEM_{100} (from Kienzle, 2004, with permission)

Variable	Grid cell size					
	5 m	10 m	15 m	20 m	25 m	30 m
Elevation	0.995	0.994	0.992	0.990	0.986	0.983
Slope (Horn)	0.926	0.927	0.911	0.896	0.867	0.857

Slope (Zevenbergen and Thorne)	0.928	0.928	0.915	0.899	0.879	0.857
Plan curvature	0.372	0.384	0.315	0.279	0.206	0.174
Profile curvature	0.437	0.511	0.491	0.450	0.404	0.350
Curvature	0.474	0.522	0.468	0.433	0.366	0.302
Wetness Index	0.140	0.133	0.128	0.107	0.102	0.086

All correlations are significant at the 0.01 level (2-tailed)

Results are summarized in Table 5. There exists a strong positive correlation for elevation between the 5 m DEM based on DEM_{10} and all DEM resolutions based on DEM_{100}. There also exists a strong positive correlation for slope values, however, with grid cell sizes larger than 10 m the correlation becomes increasingly weaker. This indicates that, in agreement with Table 4, a grid cell size of 10 m or smaller is sufficient to derive slope. The striking result presented in Table 5 is that all other terrain variables based on DEM_{100} are poorly correlated with respective terrain variables based on 5 m resolution DEM_{10}. Correlations of wetness index values are nearly nonexistent, even with the smallest grid resolution. The limitations of DEM_{100} for localized geomorphological, hydrological and environmental analyses are clearly evident.

3.7.5 Measuring the Quality of Terrain Variables Derived from DEM_{100}

RMS error (RMSE) is used as a measure to evaluate the quality of terrain derivatives based on DEM_{100}. It can be understood as the standard deviation of one surface against a benchmark surface. This analysis was only carried out for Study Area 4. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(y'_i - y_i)^2} \qquad (2)$$

where y is the cell value based on DEM_{10} with a 5 m grid resolution and y' is the respective cell value based on DEM_{100} with grid cell sizes ranging from 5 to 30 m and resampled to 5 m; and *n* is the number of compared points (here:90,000).

RMSEs for elevation values decrease from 4.3 m with a 50 m resolution to 2.0 m with the 5 m resolution. A threshold value of about 2 m is to be expected due to the large difference in sampling densities of elevation points between DEM_{10} and DEM_{100}. Slope values also reach an RMSE threshold. RMSEs for slopes derived using Horn's and Zevenbergen and Thorne's methods are similar, with threshold values of 4.8 and 4.9 % respectively. The wetness index is generally largely overestimated, with RMSEs ranging from 3.3 with a 50 m grid cell size to 2.7 with a grid cell size of 5 m.

The relative root mean square error (RRMSE) standardizes the RMSE calculated for each grid cell to the benchmark cell value (5 m grid cell size), and the derived terrain variables based on DEM_{10}. The resulting RRMSE is expressed as a percentage and represents the standard variation of the estimated terrain variable. The RRMSE is expressed as (explanation of elements see Equation 2):

$$RRMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}\frac{(y'_i - y_i)^2}{y_i}} \cdot 100$$

Results are shown in **Figure 17**. All RRMSEs decrease with a decrease in grid cell size. RRMSEs are small for elevation values, ranging from 4.8 to 2.3 %, and are large for all derivatives. RRMSEs for slope values reach a minimum of about 65% and can be larger than 100% for the coarse grid cell size of 50 m. These large values can be explained by the fact that the terrain is mainly very flat and a difference of a gentle slope such as 2% based on DEM_{100} versus a slope of 4% based on DEM_{10} will constitute a relative difference of 100%. The wetness index has RRMSEs ranging from between 50% and 42% for the grid cell sizes investigated.

3.7.6 Flow Direction Error

Flow direction, using the D8 method, were calculated for DEM100 and DEM10 for Study Area 4. Using a grid cell size of 5m, flow directions were compared against each other. **Figure 18** shows that about one third of the grid cells have the same flow direction, about another third has a difference of 45 degrees, and about one third has a deviation of 90 degrees or more. The spatial distribution of the error is shown in **Figure 19**, where flow directions for both DEMs are represented. It is clear from this Figure that in the flat areas, flow directions are unrealistic. The consequences are that watershed boundaries derived from the DEM100 would be false, and any other hydrological information, such as upslope area, stream power, wetness index or pollution sources would be false. The author is currently investigating this problem in greater detail.

4 Summary

This paper presents vital information for users of DEMs generated from 100 m regularly sampled, or similar, elevation points such as those available in Alberta, Canada. For a more detailed analysis the reader is referred to Kienzle (2004). Major findings are summarised as follows:

- All computed terrain variables tested (slope, aspect, plan and profile curvature, curvature and the wetness index) vary with a change in grid cell size.
- Density or nearest distance analysis of point elevations, cumulative frequency distribution analysis and the calculation and plotting of root mean square derivatives are meaningful procedures to quantify the most meaningful grid cell size for given point elevation data sets.
- Using a DEM based on 100 m regularly sampled elevation points, only slope can be computed giving realistic results when using a grid resolution of 10 m or less. Larger grid cell sizes still give reasonable results, but an underestimation of slopes is to be anticipated along hills lopes.
- All other terrain derivatives have weak or no correlation to a tested benchmark DEM.
- Using a DEM based on 100 m regularly sampled points can, therefore, only be used to show the general range and distribution of terrain variables, but fails to spatially correctly represent all terrain variables, with the exception of slope.
- Flow directions in flat areas cannot be represented realistically from course DEMs, such as the DEM100 used here, which means that there are numerous restrictions in hydrological applications from relatively coarse DEMs.

Where possible and feasible, DEMs should be generated using elevation points sampled at significantly higher resolutions than 100 m regularly sampled elevation points, such as those provided by AltaLis.

Acknowledgements

The author would like to thank Alberta Agriculture, Food and Rural Development for the initial provision of the 100 m elevation points, as well as the City of Lethbridge for making the 10 m elevation points available for Study Area 4. This research was funded by the University of Lethbridge Research Fund. My research assistants Ms Amber Brown, Mr Guy Duke and Mr Bradley Higgins are thanked for undertaking parts of the analyses.

References relevant to this paper

Band L E 1993 Extraction of channel networks and topographic parameters from digital elevation data. In Beven K and Kirkby M J (eds) *Channel network hydrology*. Chichester, Wiley: 13-42.

Beven K J and Kirkby M J 1979 A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* 24: 43-69.

Blaszcynski J S 1997 Landform characterisation with geographic information systems. *Photogrammetric Engineering and Remote Sensing* 63(2): 183 – 191.

Blöschl G and Grayson R 2001 Spatial observation and interpolation. In Grayson R and Blöschl G (eds) *Spatial patterns in catchment hydrology*. Cambridge University Press: 17-50.

Burrough P E and McDonnell R A 1998 *Principles of Geographical Information Systems*. New York, Oxford University Press.

Elsheikh S and Guercio R 1997 GIS topographical analysis applied to unit hydrograph models: sensitivity to DEM resolution and threshold area. In *Proceedings of the Rabat Symposium on Remote Sensing and Geographic Information Systems for Design and Operation of Water Resources Systems*. International Association of Hydrological Sciences Publ. no. 242: 245-253

ESRI (Environmental Systems Research Institute) 2000 ArcGIS 8.0. Environmental Systems Research Institute, Redlands, CA.

Evans I S 1980 An integrated system of terrain analysis and slope mapping. *Zeitschrift für Geomorphology*. Suppl Bd 36: 274-95.

Fairfield J and Leymarie 1991 Drainage networks from grid digital elevation models. *Water Resources Research* 30(6): 1681-92.

Florinsky I V 1998 Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science* 12: 47-61.

Garbrecht J and Martz L W 2000 Digital elevation model issues in water resources modeling. In: Maidment D R and Djokic D 2000 *Hydrologic and Hydraulic Modeling Support with Geographical Information Systems*. ESRI Press: 1 - 28.

Gallant J C and Wilson J P 1996 TAPES_G: A grid_based terrain analysis program for the environmental sciences. *Computers and Geosciences*. 22(7): 713_722.

Gyasi-Agyei Y, Willgoose G and DeTroch F 1995 Effects of vertical resolution and map scale of digital elevation models on geomorphological parameters used in hydrology. *Hydrological Processes*. 9: 363-382.

Helios Environmental Modeling Institute (HEMI) 1999 http://www.hemisoft.com/topoview/manual/helpcontents1.htm

Hetrick W A, Rich P M, Barnes F J and Weiss S B 1993a GIS-based solar radiation flux models. *American Society for Photogrammetry and Remote Sensing Technical Papers*. Vol. 3, GIS Photogrammetry and Modeling: 132–143.

Hetrick W A, Rich P M and Weiss S B 1993b Modeling insolation on complex surfaces. Thirteenth Annual ESRI User Conference, 2: 447–458.

Horn B K P 1981 Hill shading and the reflectance map. *In Proceedings of the IEEE* 69(1):14-47.

Hofierka J, Paraijka J, Mitasova H and Mitas L 2002 Multivariate Interpolation of Precipitation Using Regularized Spline with Tension. *Transactions in GIS*, 6(2):135-150.

Hutchinson M F 1988 Calculation of hydrologically sound digital elevation models. In *Proceedings of the Third International Symposium on Spatial Data Handling*, Sydney, International

Geographical Union, Columbus, Ohio.
Hutchinson M F 1989 A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology* 106: 211-32.
Hutchinson M F and Dowling T I 1991 A continental hydrological assessment of a new grid based digital elevation model of Australia. *Hydrological Processes* 5: 45_58.
Hutchinson M F 1996 A locally adaptive approach to the interpolation of digital elevation models. In *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, National Center for Geographic Information and Analysis, University of California, Santa Barbara.
Hutchinson M F and Gallant J C 2000 Digital elevation models and representation of terrain slope. In Wilson J P and Gallant J C (eds) *Terrain Analysis: Principles and Applications*. John Wiley & Sons, Inc.
Jensen S K and Domingue J O 1988 Extracting topographic structure from digital elevation data for geographical information system analysis. *Photogrammetric Engineering and Remote Sensing* 54(11):1593-1600.
Jones K H 1998 A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers & Geosciences* 24 (4), 315–323.
Kern T J and Stednick J D 1993 Identification of Heavy Metal Concentrations in Surface Waters through Coupling of GIS and Hydrochemical Models. In *Proceedings of the Vienna Symposium on Application of Geographic Information Systems in Hydrology and Water Resources Management*. International Association of Hydrological Sciences Publication No. 211: 559-567.
Kienzle S W 1994 The application of a GIS-Grid system for hydrological and geomorphological analysis. *Fourth Symposium on Terrain Evaluation and Data Storage*, August 3 - 5, 1994, Midrand, South Africa: 35 pp. (not published)
Kienzle S W 1996 Using DTMs and GIS to define input variables for hydrological and geomorphological analysis. In: *Proceedings of the Vienna Symposium on the Application of Geographical Information Systems in Hydrology and Water Resources Management.* International Association of Hydrological Sciences Publication No. 235: 183-190.
Kienzle S W, Lorentz S A and Schulze R E 1997 Hydrology and water quality of the Mgeni catchment. *Water Research Commission Report* TT87/97, Pretoria: 1_88 (ISBN 1 86845 297 2)
Kienzle SW (2004) The effect of DEM raster resolution on first order, second order and compound terrain derivatives. Transactions in GIS, 8(1):83-111.
Klaghofer E, Birnbaum W and Summer W 1993 Linking sediment and nutrient export models with a geographic information system. In *Proceedings of the Vienna Symposium on the Application of Geographic Information Systems in Hydrology and Water Resources Management*. International Association of Hydrological Sciences Publication No. 211: 501-506.
Mackey B G, McKenney D W, Yang Y Q, McMahon J P and Hutchinson M F 1996 Site regions revisited: a climatic analysis of Hill's site regions for the province of Ontario using a parametric method. *Canadian Journal of Forest Research* 26: 333_354.
Macmillan R A, Pettapiece W W, Nolan S C and Goddard T W 2000 A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets and Systems*, 112: 81-109.
Martz L W and Garbrecht J 1992 Numerical definition of drainage network and subcatchment areas from digital elevation models. *Computers and Geosciences* 18(6): 747-761.
Mitasova H, Hofierka J, Zlocha M and Iverson L R 1996 Modeling topographic potential for erosion and deposition using GIS. *International Journal of Geographical Information Science* 10(5): 629_641.
Moore I D, Grayson R B and Ladson A R 1991 Digital terrain modelling: A review of hydrological, geomorphological and biological applications. *Hydrological Processes* 5:3_30.
Moore I D, Turner A K, Wilson J P, Jenson S K and Band LE 1993 GIS and land surface_subsurface process modeling. In M F Goodchild, B O Parks and L T Steyaert

(editors) Environmental Modeling with GIS. New York, Oxford University Press: 197_230
Parkinson B 2002 Personal communication. Manager, GIS and Data Management, PFRA, Agriculture and Agri-Food, Canada, Calgary, Alberta.
Quinn P, Beven K, Chevallier P and Planchon O 1991 The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrological Processes* 5:59-79.
Quinn P F and Beven K J 1993 Spatial and temporal prediction of soil moisture dynamics, runoff, variable source areas and evapotranspiration for Plynlimon, mid-Wales. *Hydrological Processes* 7:425-48.
Rees W 2000 The accuracy of digital elevation models interpolated to higher resolutions. *International Journal of Remote Sensing* 21:7-20.
Saulnier G, Obled C and Beven K 1997 Analytical compensation between DTM grid resolution and effective values of saturated hydraulic conductivity within the TOPMODEL framework. *Hydrological Processes* 11:1331-1346.
Skidmore K 1989 A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International Journal of Geographical Information Systems* 3:323-334.
Tarboton D G, Bras R L and Rodriguez-Iturbe I 1991 On the extraction of channel networks from digital elevation data. *Hydrological Processes*, 5(1):81-100.
Tarboton D G, Bras R L and Rodriguez-Iturbe I 1992 A physical basis for drainage density. *Geomorphology* 5(1/2):59-76.
Tarboton D G 1997 A new method for the determination of flow directions and contributing areas in grid digital elevation models. *Water Resources Research* 33(2):309-319.
Zevenbergen L W and Thorne C R 1987 Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12: 47-56.
Zhang W and Montgomery D R 1994 Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research* 30:1019-1028.
Zhang W, Drake N, Wainwright J and Mulligan M1999 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.