A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution

Simon Wu\textsuperscript{a,}\textsuperscript{*}, Jonathan Li\textsuperscript{b}, G.H. Huang\textsuperscript{c}

\textsuperscript{a}Environmental Systems Engineering, University of Regina, Box 413, 4246 Albert Street, Regina, Saskatchewan, Canada S4S 3R9
\textsuperscript{b}Department of Geography, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
\textsuperscript{c}Department of Civil & Environmental Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Abstract

Primary topographic attributes play a critical role in determining watershed hydrologic characteristics for water resources modeling with raster-based digital elevation models (DEM). The effects of DEM resolution on a set of important topographic derivatives are examined in this study, including slope, upslope contributing area, flow length and watershed area. The focus of the study is on how sensitive each of the attributes is to the resolution uncertainty by considering the effects of overall terrain gradient and bias from resampling. Two case study watersheds of different gradient patterns are used with their 10 m USGS DEMs. A series of DEMs up to 200 m grid size are produced from the base DEMs using three commonly used resampling methods. All the terrain variables tested vary with the grid size change. It is found that slope angles decrease and contributing area values increase constantly as DEMs are aggregated progressively to coarser resolutions. No systematic trend is observed for corresponding changes of flow path and watershed area. The analysis also suggests that gradient profile of the watershed presents an important factor for the examined sensitivities to DEM resolution.

Keywords: Digital elevation model; Grid size; Topographic attributes; Hydrologic modeling

Introduction

Topography defines the pathways of surface water movement across a watershed, and thus is a major factor affecting watershed hydrologic response to rainfall inputs. Raster-based digital elevation models (DEMs) have been widely applied to efficiently derive topographic attributes used in hydrologic modeling such as slope and upslope contributing area. Any uncertainties in the topographic models will be propagated into the output of hydrologic model prediction, causing inaccuracies. One of such uncertainties stems from the choice of DEM grid size.

The availability of digital elevation data has been greatly increased with the development of effective spatial data acquisition tools (Huang & Chang, 2003). DEM accuracy properties including grid size could, however,
vary from source to source for an area of interest. The sensitivity of principal topographic derivatives used in hydrologic modeling to DEM resolution, however, has been systematically explored in few studies. A number of studies have shown that lower resolution DEMs under-represent slope classes (Bolstad & Stowe, 1994; Chang & Tsai, 1991; Gao, 1997; Kienzle, 2004; Thompson, Bell, & Butler, 2001; Zhang, Drake, Wainwright, & Mulligan, 1999). Zhang and Montgomery (1994) examined the effect of grid cell resolution on landscape representation and hydrologic simulations using elevation data from two small watersheds. Their results showed that increasing the grid size resulted in an increased mean topographic index because of increased contributing area and decreased slopes. Another study on DEM resolution effect on surface runoff modeling by Vieux (1993) showed that aggregation of spatial input data led to a decrease in the flow path length of the main channel. The result was believed to be attributable to the short circuiting of meanders in the original DEM. The same work also showed that the watershed area varies with increasing DEM grid size without a consistent trend.

Wilson and Gallant (2000) provided a summary on topographic properties where direct surface derivatives from DEM are categorized as primary topographic attributes. Among the primary attributes, those of most interest to hydrologists in terms of DEM resolution effect include slope, flow length, upslope contributing area, and watershed area. This work examines DEM resolution uncertainty in watershed terrain representation with the four important variables.

Flow path determination serves as the basis for extracting drainage networks and watershed delineation from DEMs in hydrologic modeling. Slope is a fundamental topographic property used in all flow path algorithms. Slope, together with aspect, determines the flow direction from which the upslope contributing area, representing the accumulated area draining into a given grid cell can be calculated. The calculation of the upslope contributing area is a cell-by-cell operation, and stream networks and watershed boundaries can be identified from its results across the DEM. Slope and upslope contributing area determine the topographic index which is popularly used in water resources applications. The index is defined as the natural logarithm of the ratio of the upslope contributing area to the ground surface slope. It has been the core element in various hydrologic models such as TOPMODEL (Beven & Kirkby, 1979) where it represents the extension of saturated areas and the spatial variation of soil moisture and groundwater levels.

Flow length is a measure of the distance along the flow path (determined by the flow direction grid) from a given cell to its drainage basin outlet. Variations with flow length may lead to significant change to the time lag between precipitation and flow peak discharge, thus resulting in a much different hydrograph. The sensitivity of watershed area to DEM resolution also deserves an examination. The rainfall coverage used to obtain rainfall input to the model must be corresponding to the watershed area determined from the applied DEM. Watershed area could also be an important factor in many hydrologic model's calibration processes as it is used in calculating unit runoff volumes. Optimized model parameters may be affected by any discrepancy between the area used in observed runoff and that for modeling prediction.

The comparative approach using different grid sizes of DEM has been commonly used for examining the effect of spatial data resolution on hydrologic models and their parameters (Saulnier, Beven, & Obled, 1997; Vieux, 1993; Wolock & Price, 1994; Wu, Li, & Huang, 2007; Zhang & Montgomery, 1994). Sensitivity profiles can be seen straightforwardly by depicting the variation of modeling results with DEM grid size. However, multiple resolutions of DEM suitable for the study are often not readily available for an area of interest. Most of the studies have used grid cell aggregation, i.e., resampling a given DEM to coarser resolutions, to explore the effect of cell size selection. In DEM resampling, the values of output cells are obtained by interpolation based on the values of input cells combined with the calculated distortion. Quite some resampling techniques exist, but only a few of them have been widely used in DEM grid cell aggregation. Multiple techniques are selected and applied in this study to see whether any significant inconsistency can be brought into modeling results by different resampling methods, ensuring that the cell aggregation approach is appropriate for the comparative study.

Study areas

The case study areas used in this study are two small catchments located inside Back Creek Watershed in Southwest Virginia. The two catchments (signified as A and B hereafter) occupy areas of 2.83 and 4.71 km²,
respectively. Catchment A lies in a low-relief landscape close to the watershed outlet, and its surface elevation ranges from 250 to 337 m. Catchment B is situated upstream of the watershed, and its overall terrain presents distinct variability with elevation varying from 395 to 971 m. The digital elevation quadrangles at 10 m resolution containing the study areas were downloaded from the data center website of the United States Geological Survey (USGS). The catchments are delineated and extracted with the Hydrology Tools of ESRI ArcGIS Spatial Analyst (Fig. 1).

Methods

Resampling

The three most commonly used resampling techniques are applied in this study. They are nearest-neighbor interpolation, bilinear interpolation and cubic convolution (Keys, 1981; Mitchell & Netravali, 1988). The nearest-neighbor method simply assigns the value of the single closest observation to each cell. Once the location of the cell’s center on the output grid is located on the input DEM, the nearest neighbor assignment will determine the location of the closest cell center on the input grid, identify the value that is associated with the cell, and assign that value to the cell that the output cell center is associated with. The bilinear interpolation takes a weighted average of the nearest four input cells around the transformed point to determine the output cell. This method results in a smoother surface than the nearest-neighbor method. In the cubic convolution technique, the output cell value is computed by fitting a smooth surface to the nearest 16 input cells. It tends to smooth data even more than the bilinear interpolation.

The 10 m DEM of each of the two catchments is used as the base resolution, and resampled to five DEMs of 30, 60, 90, 150, and 200 m resolutions using each of the three resampling methods, respectively.
Sink filling

A common problem with drainage network delineation using DEM is the presence of sinks. A DEM sink occurs when all neighboring cells are higher than the processing cell, which has no downslope flow path to a neighbor cell. Sinks could be real components of the terrain, but are also the results of input errors or interpolation artifacts produced in DEM generation or resampling process. They cause obstacles to the calculation of flow direction, leading to inaccurate representation of flow accumulation, and thus drainage networks. Therefore the sinks are commonly removed prior to DEM processing for drainage identification. In this study, the target attributes should be derived from hydrologically sound elevation data, i.e., sink-free DEMs. There have been a number of methods developed for treating sinks in DEMs (Garbrecht & Martz, 1999). A straightforward technique is to fill the sinks by increasing the values of cells in each sink by the value of the cell with the lowest value on the sink’s boundary (Band, 1986; Jenson & Domingue, 1988; O’Callaghan & Mark, 1984). The sink filling using this approach is performed in this study on each DEM before the calculation of the topographic attributes.

Calculation of topographic attributes

The four topographic derivatives are generated with the base DEM and all the resampled DEMs using each of the three resampling techniques. Slope calculation is performed on the DEM using a \(3 \times 3\) cell neighborhood in most commonly used algorithms. The method of Horn (1981) is adopted in this study. The method is one of eight algorithms examined comparatively in a study by Jones (1998) using both synthetic and real data as test surface. The study showed that Horn’s method performed well for gradient estimation based on root-mean-square (RMS) residual error estimates. Unequal weighting coefficients are used in the method for neighboring cell elevations and they are proportional to the reciprocal of the square of the distance from the center of the window. The method is implemented in the slope tool built in the ESRI ArcGIS software which is widely applied in terrain modeling.

There are also a number of algorithms for calculating upslope contributing area. Among them, single flow direction (SFD) and multiple flow direction (MFD) methods are most popularly applied in hydrologic modeling (Arge et al., 2003). The contributing area is dependent on the flow routing method used in the calculation. D8 is a SFD algorithm which directs flow from each grid cell to one of eight nearest neighbors based on the steepest slope gradient (O’Callaghan & Mark, 1984). In MFD algorithms, water is routed from a cell to all its downslope neighbors with weighting by slopes (Freeman, 1991; Quinn, Beven, Chevallier, & Planchon, 1991). The D8 method works well for zones of convergent flow and along well-defined valleys, while the use of a MFD method seems more appropriate for overland flow analysis on hillslopes. We choose the D8 method in the calculations of upslope contributing area because it is more applicable to delineation of the drainage network for drainage areas with well-developed channels (Garbrecht & Martz, 1999).

The flow length is an estimate of the downstream distance from a cell to watershed outlet along a flow path. The flow path is determined by the D8 algorithm, as a SFD for a DEM cell is required for the estimation. The flow length distribution over a DEM is obtained using the flow length tool in ArcGIS. The distance is determined between DEM cell centers. Thus, the distance between two orthogonal cells is the grid size, and the distance between two diagonal cells is the multiplier of 1.414 and the grid size.

The computation of watershed area appears to be more straightforward. With a raster-based DEM, the watershed area is simply determined by the number of cells and cell size. The information can be easily obtained from the raster layer's attribute table in a GIS.

Results of topographic representations and discussion

Slope

The means and variances of slope decrease with the resampling of DEM resolution for both catchments (Fig. 2). This can be substantiated by the cumulative frequency distributions of watershed area against the slope value for each DEM as presented in Fig. 3. It appears that the effect is more profound at finer grid sizes,
especially for Catchment B which has a higher gradient overall. The mean slope values for both catchments drop significantly when the DEMs are resampled from the base resolution 10–30 m (Fig. 2), while the observation is much more pronounced for Catchment B. Apparently, the higher relief portion of each study area contributes a large part to the mean value decrease as presented in Fig. 3. The illustrations show that the data aggregation can cause considerable reduction of steep slope angles, although the trends of the cumulative frequency variations are the same for all the resolutions.

It can also be seen from Fig. 3 that the topography can be better preserved for a higher relief area during grid size aggregation. The slope values for Catchment B at the 200 m resolution are still distributed with a maximum of about 25° (Fig. 3b) as compared to around 55° at 10 m. On the other hand, the maximum slope value at 200 m for Catchment A is less than 6°, plunging from approximately 35° at 10 m (Fig. 3a).

It is clear, based on Figs. 2 and 3, that choice of the resampling methods has no influence on the general result of DEM resolution effect on slope distribution at the watershed scale. Minor differences exist in the slope mean values and its distributions between the three applied resampling techniques for both study areas. It can be found that, the slope values in the bilinear interpolated DEMs are, on the whole, a little smaller than those obtained from the nearest neighbor resampling. On the contrary, the slopes in the DEMs from the cubic convolution algorithm are generally greater than those of the nearest-neighbor method. The discrepancies can be well explained by comparative characteristics of the resampling techniques. Bilinear interpolation computes the output cell value using the weighted average of the nearest four input cells, and thus results in a smoother surface than the nearest-neighbor method. Cubic convolution is similar to bilinear interpolation except that the weighted average is calculated from the 16 nearest input cell centers and their values. The fact that surrounding cells are involved brings about a problem to the calculation of output values for cells close to the DEM boundary. Both bilinear interpolation and cubic convolution may lead to unreasonable interpolated
results around the DEM edge if background areas (no data) outside the given DEM are used for those close-to-border cells. However, as 16 cells nearby are used in cubic convolution as compared to four cells in bilinear interpolation, the “edge effect” can be much more significant in an output DEM from cubic convolution. The effect may produce more irregular surface close to the DEM boundary and cause higher slope values overall.

**Upslope contributing area**

A significant number of grid cells with zero flow accumulation on the DEMs of both catchments were observed when deriving distributions of the upslope contributing area with the D8 method applied. The portions of cells with the zero value on the original 10 m DEMs are 30.0% and 28.6%, respectively, for Catchment A and B. Grid cells with flow accumulation values of zero are local topographic highs, and generally correspond to peaks and ridgelines. The portion of zero value coverage increases generally for both
catchments with resampling to larger grid sizes, while fluctuation exists for most of the tests (Fig. 4). An interesting observation is that almost the same area value of zero flow accumulation for each catchment is obtained for different resampling methods when the DEM is first aggregated from 10 to 30 m. Obvious differences in the percentage can, however, be seen among the resampled DEMs at 60 m and higher grid sizes. It indicates that effect of resampling method comes up above a certain cell size (between 30 and 60 m for either catchment). Over the threshold, areas of ridgelines and hilltops on three resampled DEMs at each cell size vary to a noticeable degree due to different interpolations with the resampling methods. Also apparently, the increase of the zero flow accumulation area for Catchment A with the resampling is much more significant than that for Catchment B. The portion of zero flow accumulation area for each of the three resampled DEMs at 200 m is above 55% for Catchment A, while the percentages are all below 45% for Catchment B. This is believed to be attributable to the lower relief terrain in Catchment A.

Excluding the zero values, we look at the variations of the upslope contributing areas for the cells at different DEM resolutions by depicting their cumulative frequency distributions under each resampling scheme (Fig. 5). The results clearly show that an increase of DEM grid size shifts the contributing areas towards greater values. This is consistent for all the DEMs resampled using the three techniques for both catchments. We can also see from the illustrations that the effect of data aggregation on the contributing area is stronger at finer DEM resolutions. No significant difference can be found among the distributions from the different resampling techniques, and thus there is no obvious effect by an individual resampling method that can be identified.

The sensitivities of both zero and non-zero flow accumulation areas to the DEM data aggregation should be considered in examining the overall effect of DEM resolution on upslope contributing area. The weighted mean value of the upslope contributing areas for all the cells at 10 m resolution is 15,153 m², and the value becomes 47,075 m² for the 30 m DEM resampled using the nearest-neighbor method. Fig. 6 shows the variations with grid size for all the resampling schemes. The increase of the mean of upslope contributing area presents a nearly linear relationship with grid size except for very coarser grid resolutions (>150 m) for the cubic convolution algorithm for Catchment A. The exception is consistent with the result of the cumulative frequency, and must be due to the edge effect occurring with the use of cubic convolution resampling.

It is intuitive that the data aggregation has the remarkable influence since the upslope contributing area is calculated on the cell basis and the cell coverage increases as the square of the grid size. The result we have obtained above is consistent with that presented by Zhang and Montgomery (1994) which examined the effect of DEM resolution on the portrayal of the land surface and results of hydrologic simulations with elevation data from two small watersheds. DEMs at the resolutions of 2, 4, 10, 30 and 90 m were used in the modeling for the two watersheds. Their results showed that larger grid size is in favor of larger contributing areas, and the effect is significant for TOPMODEL, a topography-based hydrologic model.
Flow length

The patterns of flow length changing with DEM grid size are also depicted using comparative cumulative frequency distributions and mean values (Figs. 7 and 8). Surprisingly, there is no definite correlation that can be found between flow length and DEM grid size from the cumulative frequency distributions of all the cases for both catchments (Fig. 7). The resampling schemes appear to have a profound impact on the flow length computations. For Catchment A, the flow length distributions at different DEM resolutions for the DEM by the nearest neighbor-resampling vary within a relatively small range. The variation is greater for the distributions from bilinear interpolation, but still no trend can be identified. The distributions from the cubic convolution have a substantial deviation. A unique observation here is that the flow length distributions keep shifting to lower values when resampling from 60 to 200 m, successively. For Catchment B, the cumulative distributions for the different resolutions of DEMs from the nearest neighbor and bilinear interpolation are mixed with minor variation. The distributions for the DEMs obtained from the cubic convolution vary a lot,

Fig. 5. Cumulative frequency distributions of upslope contributing area: (a) Catchment A, (b) Catchment B.
but no correlation with grid size can be observed. These results are corroborated by the presentations in Fig. 8 which shows variations of mean flow length values for all cells in the DEM with grid size. A common observation for all the three resampling cases for each catchment is that, when aggregating the DEM from 10 to 30 m, the flow length values shift a little higher with almost the same increase of the mean value. The curves oscillate mostly for further increase of DEM grid size.

The length of the longest flow path is also calculated for each case, and the results are given in Fig. 9. The variation patterns for the longest flow length are nearly the same as those for mean flow length. These results do not support the assertion from the published work by Vieux (1993) which investigated effects of data aggregation on surface runoff modeling. The original 30 m resolution DEM of a watershed was resampled to three more DEMs of 90, 150, and 210 m resolution, respectively, using the nearest-neighbor method. The study found that meanders in the original DEM get short circuited as grid size increases with the aggregation, causing length reduction in flow paths.

Flow length is calculated based on the distance between center points of the cells along a flow path. From the perspective of geometry, effect of DEM grid size on the length of a flow path depends mostly on the shape of the flow path. DEM resampling to a greater cell size may lead to a reduction of flow length due to short circuit in a certain part of the entire flow path, but may also cause a longer route due to stretched distances between enlarged cells in another section of the path. The overall impact of a grid size increase throughout the
path determines the change of the flow path length. Therefore, increasing the grid size does not necessarily result in a shorter path as evidenced by the cases in this study.

Watershed area

The DEM-based watershed area is determined with the cell size and number of cells in GIS. The variation of calculated areas for each catchment with DEM resampling in this study is shown in Fig. 10. The two catchments possess a similar pattern of change with grid size. While the overall variation for both catchments appears to be a minor increase throughout the DEM resampling range, fluctuations are present within about ±20% and ±10% of the base areas (10 m) for Catchments A and B, respectively. The variation of the catchment area is due to the irregularity of resampled DEM boundaries, i.e., the boundaries are becoming
more and more jagged with grid size increase. The effect is in a large part dependent on the shape of the original map. The higher variation shown in Fig. 10 for Catchment A is attributable to its more irregular shape as well as its smaller area. A special case we may consider is an original DEM in a rectangle shape with each side being a multiple of all grid sizes used for the comparative study. Any expected resampling will not cause any change to the DEM coverage in the circumstances.

Vieux (1993) also examined the effect of grid size change on the watershed area obtained from DEM, and no consistent tendency is observed. The study believed that the area is varying because grid cells of different sizes cannot consistently cover the irregular shape of the watershed. The perception is corroborated by the finding on watershed area from this study. DEM aggregation may have influence on its calculated area, but the direction and extent of the effect would vary upon its base grid size, level of increase, and edge pattern of the DEM.

**Summary and concluding remarks**

Slope, upslope contributing area, flow path and watershed area are the four important topographic attributes determining watershed delineation and surface runoff profile in hydrologic modeling with raster-based GIS. Effects of elevation data resolution on the four variables are analyzed across six different DEM grid sizes for two watersheds with dissimilar overall gradient. Three commonly used resampling methods are utilized in order to obtain more objective observations on the effects of DEM resolution.
Impacts are found for all the four examined derivatives as functions of DEM resolution. Slope estimates are seen constantly decreasing with the resolution growing coarser. The reduction is more profound for higher relief areas at higher resolutions, indicating that the extent of smoothing effect would be varying with study area. Conversely, the estimates of upslope contributing area tend to increase along with the grid size increase. This is certainly true considering multiplied coverage area of each grid cell. The increase of mean contributing area is roughly in linear relationship to grid size under fair DEM resolutions (<100 m). No definite trend of bias is observed for the calculations of flow path length and watershed area while both are sensitive to the grid size change. It is suggested that the shape of each flow path plays a critical role in determining the response of its flow length to grid resolution change. Similarly, the variation of DEM-based watershed area with grid size depends on the shape of map boundary. The higher the irregularity of the boundary shape, the more unpredictable the effect on the watershed area estimate will be.

The above-examined effects on the primary attributes constitute the basis for DEM resolution uncertainty with topographic and hydrologic characterization at watershed scale. The impacts will be conveyed to secondary or compound topographic derivatives, which may be used as spatial input of a hydrologic model, thus resulting in uncertainty with prediction output. For example, TOPMODEL is a widely used topography-based model that simulates hydrologic fluxes of water through a watershed. The model is established on the assumption that the local hydraulic gradient is equal to the local surface slope, implying that all points with the same value of the topographic index will respond in a hydrologically similar way.
Both primary attributes used in the index have been examined on their sensitivities to DEM resolution in this study. Obviously, the topographic index, as a combined attribute, couples the joint sensitivities of slope and contributing area. The trend towards lower slope values and larger contributing areas with grid size increase shifts the index to higher values, leading to a varied hydrologic model output. An optimum DEM grid size (or a range) exists for a hydrologic model application with raster GIS, which can render most optimum results in prediction. The grid size would depend much on terrain complexity of the study site and topographic attributes used in the particular model.

The findings suggest that a profound understanding on the primary topographic attributes is essential to the applications of hydrologic modeling with DEMs. The understanding would ensure that hydrologic model uncertainty be rigorously assessed in terms of spatial data resolution with gaining integral insight into the association between watershed topography and hydrologic processes.

Acknowledgments

The authors would like to thank the anonymous referees for contributing insightful comments and helpful suggestions which led to a substantially improved paper.

References


