Evaluation of Lidar-derived DEMs through Terrain Analysis and Field Comparison

Cody P. Gillin, Scott W. Bailey, Kevin J. McGuire, and Stephen P. Prisley

Abstract

Topographic analysis of watershed-scale soil and hydrological processes using digital elevation models (DEMs) is commonplace, but most studies have used DEMs of 10 m resolution or coarser. Availability of higher-resolution DEMs created from light detection and ranging (lidar) data is increasing but their suitability for such applications has received little critical evaluation. Two different 1 m DEMs were re-sampled to 3, 5, and 10 m resolutions and used with and without a low-pass smoothing filter to delineate catchment boundaries and calculate topographic metrics. Accuracy was assessed through comparison with field slope measurements and total station surveys. DEMs provided a good estimate of slope values when grid resolution reflected the field measurement scale. Intermediate scale DEMs were most consistent with land survey techniques in delineating catchment boundaries. Upslope accumulated area was most sensitive to grid resolution, with intermediate resolutions producing a range of UAA values useful in soil and groundwater analysis.

Introduction

Topographic analysis using digital elevation models (DEMs) has become routine in soil and hydrologic sciences, and there has been considerable assessment of the effects of grid resolution on topographic metrics. Most watershed-scale studies examined resolutions of 10 m or coarser and tended to use DEMs covering thousands of hectares. For instance, when researchers examined slope computed from DEMs of different resolutions, they observed that coarser DEMs generated lower values (e.g., Isaacson and Ripple, 1990; Jenson, 1991). Quinn *et al.* (1991) compared topographic wetness index (TWI) computed from 12.5 and 50 m DEMs and found higher values for the coarser DEM. Many other studies comparing topographic metric values computed from a range of DEMs reported lower slope, larger upslope accumulated areas (UAAs), and higher TWI values for coarser DEMs (e.g., Hancock, 2005; Saulnier *et al.*, 1997; Wolock and Price, 1994; Zhang and Montgomery, 1994). Variation in topographic metric values computed from DEMs of different resolutions is a result of discretization effects when the size of DEM grid cells is altered (which can affect the algorithm used to compute a topographic metric) and the loss of terrain detail (smoothing) that occurs through DEM coarsening (Gallant and Hutchinson, 1996).

Examination of soil and hydrologic variability of small headwater catchments may be enhanced by higher-resolution DEM data that has only recently become available through light detection and ranging (lidar) technology. Lidar-derived DEMs have been shown to be more representative of field slope measurements (Shi *et al.*, 2012) and field-determined elevations (Vaze *et al.*, 2010) than DEMs created using topographic maps. However, few studies have assessed variation in topographic metric values extracted from a range of high-resolution (10 m or less) lidar-derived DEMs. Sorensen and Seibert (2007) coarsened a 5 m lidar-derived DEM to 10, 25, and 50 m resolutions and found median TWI values increased with DEM grid cell size. Vaze *et al.* (2010) noted changes in DEM-delineated catchment boundaries across five lidar-derived DEMs as resolution decreased from 1 to 25 m.

While lidar-derived DEMs may represent field conditions better than topographic maps, their accuracy has been shown to vary depending on land cover class. For example, previous studies found elevation errors increased under forest canopy compared with open areas (Hodgson *et al.*, 2005; Reutebuch *et al.*, 2003; Su and Bork, 2006). Greater DEM elevation error associated with forest canopy may be related to a decrease in the number of lidar ground returns or off-terrain points incorrectly classified as ground (Hodgson *et al.*, 2005).

Quinn *et al.* (1991) contended that the resolution of DEMs used in hydrologic modeling must reflect topographic features vital to the hydrologic response, suggesting that resolution of early DEMs was too coarse for accurate modeling of some catchments. Two decades later, high-resolution DEMs may offer a level of topographic detail greater than that controlling surface/near surface flow pathways. For instance, Bailey *et al.* (2014) found that a 5 m DEM resulted in UAA and TWI values that were better correlated with soil horizon thickness and groundwater fluctuations than metrics calculated from a 1 m DEM. Gillin *et al.* (2014) showed that digital mapping of soils based on DEM derived topographic metrics was possible with a smoothed DEM. To mitigate landscape roughness, a DEM may be coarsened to a lower resolution through resampling or cell aggregation (e.g., Band and Moore, 1995; Sorensen and Seibert, 2007; Wu *et al.*, 2008) or smoothed through filtering (e.g., Lillesand and Kiefer, 2000; Walker and Willgoose, 1999). Filtered DEMs retain general topographic trends better than coarsened DEMs while reducing local roughness created by individual cells (Hammer *et al.*, 1995). Although filtering is a common technique for smoothing DEMs, evaluations of topographic metrics computed from filtered and unfiltered DEMs over a range of resolutions are limited (e.g., Hammer *et al.*, 1995).

This study had three principal objectives. First, we compared differences in shape and area of a catchment delineated from 1 m DEMs interpolated from lidar datasets, as well as DEMs aggregated from original 1 m resolution to coarser models (3, 5, and 10 m resolutions) and treated with low-pass smoothing

Cody P. Gillin is with Trout Unlimited, Wenatchee, WA, and formerly with the Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA.

Scott W. Bailey is with the USDA Forest Service, Northern Research Station, North Woodstock, NH (swbailey@fs.fed.us).

Kevin J. McGuire is with the Virginia Water Resources Research Center and Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA.

Stephen P. Prisley is with the Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA.

Photogrammetric Engineering & Remote Sensing
Vol. 81, No. 5, May 2015, pp. 387–396.
0099-1112/15/387–396
© 2015 American Society for Photogrammetry and Remote Sensing
doi: 10.14358/PERS.81.5.387

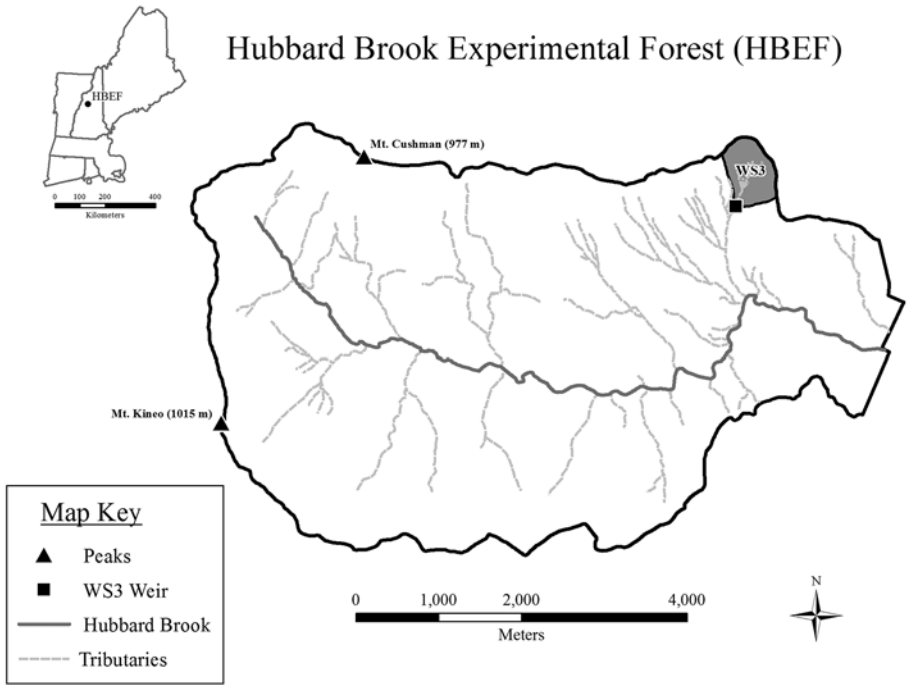

Figure 1. Map of Hubbard Brook Experimental Forest, WS3, and location within New England, United States.

Table 1. Lidar Data Collection Methodologies for Datasets Acquired from the National Center for Airborne Laser Mapping (NCALM) and the White Mountain National Forest (WMNF); Information Provided by Post-project Reports and Personal Communication with Representatives of NCALM and Photo Science, Inc.

Data Collection/Processing Method	NCALM	WMNF
Area of survey	Hubbard Brook Valley (about 42 sq. km)	western White Mountain National Forest (about 484 sq. km)
Approximate altitude above sea level	1600 meters	1800 meters
Swath width	400 meters	550 meters
Overlap	50%	30%
Classified file type/size	LAS, 1000 m × 500 m tiles	LAS 1.2, 2,000m × 2,000m tiles
Ground Return Density over WS3 only (ppsm = points per square meter)	3.27 ppsm	1.16 ppsm
Vertical RMSE	0.0720 meters	0.124 meters
Ground control stations	3 (Hubbard Brook Valley only)	37 (entire WMNF)

filters, to determine what resolution/filter combination best reflected a field-surveyed catchment boundary. Second, we evaluated the accuracy of each lidar dataset and ability of lidar-derived DEMs to characterize topography and terrain features through comparison with field-determined slope measurements and total station ground surveys. Finally, we examined variation in topographic metrics computed from DEMs of varying resolution to determine the effects of grid cell size over a range of high-resolution (10 m or less) DEMs. Overall, this study aimed to provide guidance for researchers utilizing lidar-derived DEMs in watershed-scale soil and hydrologic analyses.

Methods

Study Location

The Hubbard Brook Experimental Forest (HBEF) (Figure 1), located in the White Mountains of New Hampshire (43°56'N, 71°45'W), is maintained by the United States Forest Service (USFS), Northern Research Station and is part of the National Science Foundation Long-Term Ecological Research (NSF LTER) network. Watershed Three (WS3), the hydrologic reference catchment, is underlain by mica schist of the Silurian Rangeley formation (Barton *et al.*, 1997) and partially covered by glacial till of varying thickness. Soils are predominantly Spodosols of sandy loam texture developed in glacial parent materials (Likens, 2013). Elevation ranges from 527 m to 732 m. The western side of the catchment is characterized by steep spurs flanking intermittent and perennial streams, while the eastern side contains less well-developed drainage channels and areas of subtler topography. Bedrock outcrops are most common near the catchment boundary. The catchment is dominated by second-growth northern hardwood forest including sugar maple (*Acer saccharum*), American beech (*Fagus grandifolia*), and yellow birch (*Betula alleghaniensis*) with shallow-to-bedrock areas dominated by red spruce (*Picea rubens*) and balsam fir (*Abies balsamea*) interspersed with mountain white birch (*Betula cordifolia*). Understory vegetation is comprised mainly of patches of hobblebush (*Viburnum lantanoides*), a woody shrub, with scattered herbaceous plants. The forest was partially harvested from 1870 to 1920, damaged by a hurricane in 1938, and is not currently aggrading (Likens, 2013; Siccama *et al.*, 2007).

Lidar Data Collection and DEM Interpolation

Two lidar datasets were evaluated. The first was obtained by the National Center for Airborne Laser Mapping (NCALM) in November 2009 as part of their Seed Proposal program, and the second by Photo Science, Inc. in April 2012 for the United States Forest Service White Mountain National Forest (WMNF). Both datasets were collected during leaf-off and snow-free conditions using an Optech GEMINI Airborne Laser Terrain Mapper and used to

interpolate a 1 m raster bare earth DEM from returns classified as terrain points by each respective organization. A comparison of lidar data collection methodologies is in Table 1.

Digital Terrain Analyses and GPS Data Collection

All digital terrain analyses were conducted using System for Automated Geoscientific Analyses (SAGA, version 2.1.0) and ArcGIS© (ArcMap, version 10.1) software. GPS data were collected using a Trimble Geo XT 2005 GPS unit equipped with a Trimble Hurricane Antenna and differentially corrected using Trimble Pathfinder software and Continuously Operating Reference Station (CORS) data from the National Geodetic Survey to obtain approximately 1 m precision of horizontal positions.

DEM Aggregation, Filtering, and Sink Filling

Both the NCALM and WMNF 1 m DEMs were aggregated to coarser resolutions of 3, 5, and 10 m using mean cell aggregation to create eight DEMs. Mean cell aggregation was achieved by computing the mean value for a designated cell neighborhood, then creating a single cell of the original neighborhood size and applying the mean neighborhood value. For example, to create a 3 m DEM from a 1 m DEM, the neighborhood size is nine cells (center cell plus eight adjacent cells). A 5 m DEM is created using a 25 cell neighborhood, and a 10 m DEM is created using a 100 cell neighborhood. Then, a second version of each DEM was created by treating each DEM with a simple low-pass smoothing filter using a mean filtering technique for a total of 16 DEMs. Mean low-pass filtering computes the average elevation value in a 3 × 3 cell neighborhood moving window and applies that value to the cell at the neighborhood center. Unlike cell aggregation, low-pass filtration does not change the size of the grid cells. Both cell aggregation and low-pass filtration are common methods of DEM smoothing, but a comparison of the effects of the two techniques on topographic metrics and catchment delineation applied in a soil and hydrological context is absent in the literature.

Finally, we applied a sink-filling algorithm developed by Wang and Liu (2006) to each DEM resolution/filter combination, which is common in hydrologic applications that require the derivation of flow direction and cell accumulation grids.

Watershed Boundary Delineation and Contour Line Generation

DEM-delineated catchment boundaries were established for each DEM resolution/filter combination using a differentially corrected GPS point collected at a weir defining the watershed outlet. The single flow direction algorithm (Jenson and Domingue, 1988) was used for flow direction during delineation. Each watershed polygon was buffered to a distance of 20 m to mitigate edge effects during topographic metric computation. Contour lines with a 3 m contour interval were generated using the native 1 m DEM for each lidar dataset. Finally, each DEM was clipped to the corresponding buffered watershed boundary polygon. Watershed boundaries delineated from each DEM resolution/filter combination were assessed for differences in shape and area. DEM-derived watershed boundaries and areas were compared with a manually delineated boundary measured by compass and chain survey when HBEF experimental watersheds were first established in the 1950s. The boundary has been maintained and marked since establishment and was checked for consistency with the original survey by walking it with a Trimble Geo XT 2005 GPS unit in 2011. The field-surveyed boundary was used a point of reference to compare with the DEM-derived watershed boundaries. Bearings and distances from the field WS3 survey were used to create a boundary shapefile with the weir GPS point used for georeferencing.

Comparison of Field and DEM Slope Measurements

We compared 75 field slope measurements with DEM-derived slope values. Percent slope was measured with a clinometer 5 m upslope and 5 m downslope from soil characterization pits and groundwater wells along the line of maximum slope. We collected and differentially corrected GPS locations for each pit/well location. GPS accuracy of approximately 1 m was sufficient for locating pits and wells within one grid cell in the finest DEM analyzed. The steeper of the upslope/downslope clinometer measurements was compared with DEM-derived percent slope values computed using the maximum slope algorithm (Travis *et al.*, 1975) from filtered and unfiltered 1, 3, 5, and 10 m resolution NCALM and WMNF DEMs. A scatterplot comparing field slope with difference between field and DEM slope was used to determine DEM resolution that best simulated field slope measurements.

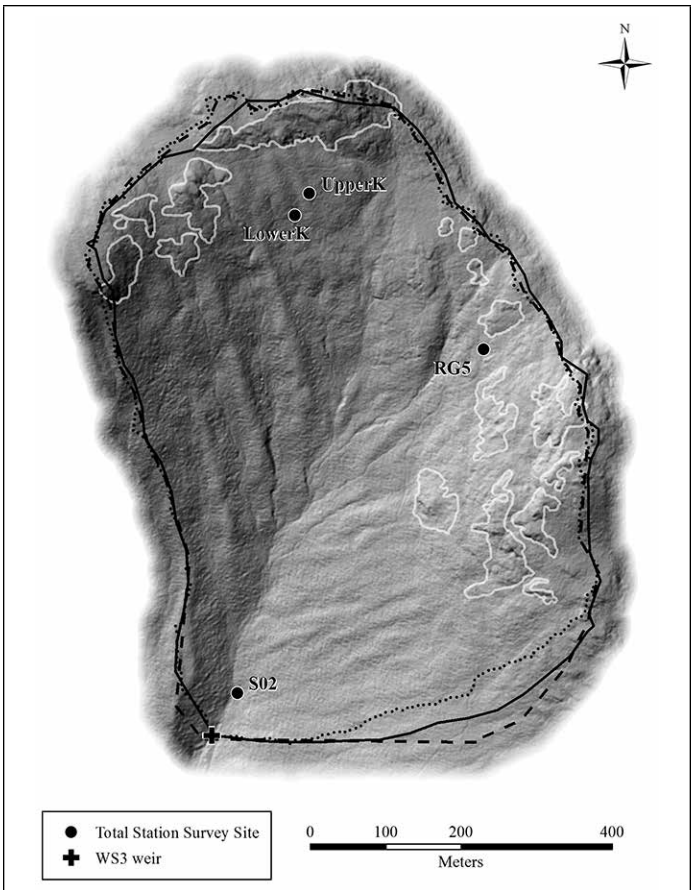

Figure 2. Catchment hillshade map with total station survey sites, catchment outlet, DEM-delineated catchment boundaries, and bedrock outcroppings (light grey polygons). Solid line represents the field-surveyed catchment boundary, the small-dashed line denotes the NCALM 1 m boundary, and the long-dashed line represents the WMNF 10 m boundary.

Total Station Ground Surveys

Terrain features (boulders, hummocky topography, and fallen tree boles) can be considered part of the ground surface, but it is not well-understood whether lidar classification methods label terrain features as ground or whether interpolation algorithms smooth these features during DEM generation. We conducted elevation ground surveys at four locations in WS3 in May and June of 2012 using a Sokia SET 610 total station to determine if the lidar-derived DEMs reflected terrain features. Survey sites incorporated diverse topography, terrain features, and vegetative cover. Three sites were located entirely under mature forest canopy (UpperK, LowerK, SO2) while a fourth site was partially located under mature forest canopy and partially in a rain gage area cleared of mature forest but with dense beech regrowth (RG5) (Figure 2). SO2 contained the greatest density of understory vegetation (primarily hobblebush) and terrain features, RG5 contained the lowest density

of understory shrubs but dense midstory beech saplings and minimal terrain features, and the K sites contained an intermediate density of vegetation and terrain features. At each site, the total station was placed directly over a point established as plot center and leveled to an XY planar accuracy of 10 seconds. Data were collected at 3 m intervals along transects established prior to each survey. Transect lengths varied depending on visibility through the understory, but generally ranged from 20 to 35 m. When a terrain feature was encountered, data points were collected adjacent to and on the feature, often resulting in greater point densities at these locations. Vertical (relative elevation), horizontal, and slope distances, as well as horizontal angle (azimuth in degrees) were recorded for each point.

Elevation values were extracted from the unfiltered/unfilled NCALM and WMNF 1 m DEMs to each survey point. Relative elevation differences between each benchmark and survey points indicated by each DEM were subtracted from the relative elevation differences between each benchmark and survey points indicated by the total station. The absolute value of the resulting values constitutes overall error. Finally, overall error was used to calculate Root Mean Squared Error (RMSE) as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}} \quad (1)$$

where e_i are the errors or differences between the elevations of the total station survey locations and their corresponding DEM elevations, and n is the number of survey points. RMSE offers a common empirical assessment of lidar-derived DEM accuracy (Raber *et al.*, 2007). We compared RMSE values between sites, as well as for canopy versus no canopy and terrain feature versus no terrain feature. A decrease in vegetative cover should correspond with lower RMSE caused by higher lidar ground return densities and fewer off-terrain features misclassified as ground. Greater RMSE for terrain features may indicate their omission from lidar-derived DEMs.

Topographic Metric Computation

DEMs were used to compute four topographic metrics commonly used in soil and hydrologic sciences: slope, planform curvature, upslope accumulated area (UAA), and topographic wetness index (TWI = $\ln(\alpha/\tan\beta)$, where α = upslope accumulated area (UAA), and tan β = local slope; (Beven and Kirkby, 1979)). Slope was calculated using the maximum slope algorithm (Travis *et al.*, 1975). Planform curvature was calculated using the equation described by Zevenbergen and Thorne (1987). UAA is a measure of the amount of area upslope of a given point on a landscape to which surface flow is attributable and was calculated using the triangular multiple flow direction algorithm (Seibert and McGlynn, 2007). TWI is a common metric used to simulate surface and shallow subsurface wetness of different points on a landscape relative to each other, particularly in hydrologic models (e.g., Niu *et al.*, 2005; Tague and Band, 2001).

To compare topographic metrics across DEM resolution/filter combinations, we generated random points for topographic metric value extraction. Point generation was constrained by a 20 m negative buffer polygon so that all points were also contained within each watershed boundary. To avoid spatial autocorrelation, we generated random points at a density of one point per 1,000 m^2 using the mean area of all DEM-delineated catchments for a total of 421 random points. Slope, planform curvature, UAA, and TWI values from each DEM resolution/filter combination were extracted to the random points. Boxplots were used to evaluate differences in topographic metric distributions across changes in DEM resolution and filtering. Additionally, UAA maps were created using the WMNF DEMs for unfiltered 1, 3, 5, and 10 m resolutions to demonstrate how DEM resolution impacts UAA values.

Results

Watershed Boundary Delineation

Watershed boundaries were most similar to the 1950s field-surveyed boundary when delineated from 3 m and 5 m resolution DEMs, regardless of whether a filter was applied (not shown). The NCALM unfiltered 1 m DEM excluded a 1 ha region (approximately 2 percent of the total watershed area) in the southeastern portion of the watershed that was included in the field-surveyed boundary, while the WMNF unfiltered 10 m DEM included a 1 ha area in this same region that was excluded from the field-surveyed boundary (Figure 2).

Watershed areas were also compared for each DEM resolution/filter combination and with the field-surveyed boundary of 42.4 ha. Areas ranged from 41.4 ha (unfiltered NCALM 1 m) to

Table 2. Catchment Area and Mean Topographic Metrics for Each DEM Resolution/Filter Combination (LP denotes the Application of a Low-pass Filter); Mean Topographic Metric Values are Based on a Random Sampling of 10 Percent of 10 m Grid Cells that would Fit into Each Boundary; Area of the Field-surveyed Boundary is 42.3 ha

Lidar Dataset	DEM resolution/ filter	Area (ha)	Slope (%)	Planform Curvature (radians/m)	UAA (m^2)	TWI (ln(m^2))
NCALM	1m	41.4	30.6	2.8	188.2	4.3
NCALM	1mLP	42.3	29.0	0.9	161.3	4.8
NCALM	3m	41.9	28.9	0.4	571.1	6.4
NCALM	3mLP	42.1	28.5	0.1	790.5	6.6
NCALM	5m	42.1	28.8	0.4	1621.1	7.2
NCALM	5mLP	42.0	28.1	0.2	1594.1	7.3
NCALM	10m	42.2	28.4	-0.5	3286.9	8.2
NCALM	10mLP	41.5	27.4	-0.3	3410.8	8.5
WMNF	1m	42.4	30.2	2.4	173.6	4.4
WMNF	1mLP	42.3	28.8	0.9	201.8	4.9
WMNF	3m	42.3	29.3	1.4	858.8	6.4
WMNF	3mLP	41.8	28.4	0.0	937.6	6.6
WMNF	5m	42.3	28.6	-0.5	1668.7	7.2
WMNF	5mLP	42.1	28.0	0.1	1539.5	7.4
WMNF	10m	43.0	28.2	-0.2	2623.1	8.2
WMNF	10mLP	41.6	27.4	-0.6	3097.7	8.5

43.0 ha (unfiltered WMNF 10 m) (Table 2). The mean of watersheds delineated using NCALM DEMs was 42.0 ha and the mean of watersheds delineated using WMNF DEMs was 42.2 ha. In general, coarse (10 m) and fine (1 m) DEMs tended to generate watersheds containing areas least similar to the overall means, while intermediate (3 m and 5 m) DEMs tended to generate watersheds containing areas most similar to the overall means (Table 2) and the conventionally surveyed boundary.

Comparison of Field and DEM Slope Measurements

The 1 m DEMs generated slope values most different from field measurements, with the exception of the steepest slopes (Figure 3). A one-way analysis of variance indicated no statistically significant differences ($\alpha = 0.05$) between DEMs of the same resolution/filter combination derived from NCALM and WMNF lidar datasets, so for ease of visualization only NCALM scatterplot data are shown. Twenty-nine percent of the NCALM 1 m DEM slopes and 23 percent of the WMNF 1 m DEM slopes exhibited a difference from field slope greater than 10 percent. In general, difference between field and DEM slope values decreased with DEM coarsening. Filtering made no significant impact on DEM slope computation. A one-way analysis of variance comparing field slope measurements with DEM-computed slopes indicated no statistically significant differences between the means ($\alpha = 0.05$).

Total Station Ground Surveys

To determine if small scale terrain features were included when lidar returns were classified as ground points and used to interpolate a bare earth model, we compared RMSE for locations with and without terrain features. We also compared RMSE for the cleared rain gage area with RMSE for sites under mature forest canopy to determine the effects of canopy on lidar in a northern hardwood forest. In the rain gage clearing, the RMSE was 1.06 m for the NCALM DEM and 1.24 m for the WMNF DEM. Under canopy, the RMSE was 1.32 m for the NCALM DEM and 1.52 m for the WMNF DEM (Table 3). Locations on or adjacent to terrain features exhibited an RMSE of 1.49 m for the NCALM DEM and 1.66 m for the WMNF DEM, while locations without terrain features exhibited an RMSE of 1.03 m for the NCALM DEM and 1.18 m for the WMNF DEM (Table 3). RG5 had the greatest RMSE value when computed by survey site (Table 3). Greater RMSE values were observed for both datasets at locations with terrain features versus non-feature survey points. In most cases, total station measurements yielded a greater relative elevation difference than the DEM between benchmark and terrain feature survey locations.

Topographic Metric Comparison

Boxplots were used to compare the distributions of topographic metric values computed for each DEM resolution/filter combination (Figure 4). A one-way analysis of variance indicated no statistically significant differences ($\alpha = 0.05$) between DEMs of the same resolution/filter combination derived from NCALM and WMNF lidar datasets, so for ease of visualization only WMNF distributions are shown. Filtering and DEM coarsening decreased slope variance and interquartile range (IQR). Slope IQR decreased from 17 percent to 7 percent (Figure 4a). Median planform curvature did not change with filtering or cell aggregation (Figure 4b). Planform curvature variance was consistently lower for the filtered version of each DEM resolution (Levene's test statistic = 47.22, $\alpha = 0.05$). The 3 m and 5 m DEMs tended to produce planform curvature values with the narrowest distributions (Figure 4b). Median UAA values increased for both the NCALM and WMNF DEMs from 188 m^2 and 174 m^2 to 3411 m^2 and 3098 m^2, respectively. UAA maps indicated that for finer-resolution DEMs the watershed was dominated by grid cells with small UAAs (Plate 1). As DEMs were coarsened, UAA values became larger, and median TWI

Table 3. RMSE for Each Total Station Survey Site, Locations Under Mature Canopy and in the Rain Gage Clearing, and Locations With and Without Terrain Features. Relative Elevation Differences between Total Station Survey Benchmark/Survey Locations and Corresponding DEM Benchmark/Survey Locations were Used to Compute RMSE

Site ID	NCALM	WMNF
LowerK	0.69	0.73
UpperK	0.73	0.71
SO2	0.77	0.98
RG5	1.84	2.07
Canopy	1.32	1.52
Clearing	1.06	1.24
Feature	1.49	1.66
No Feature	1.03	1.18

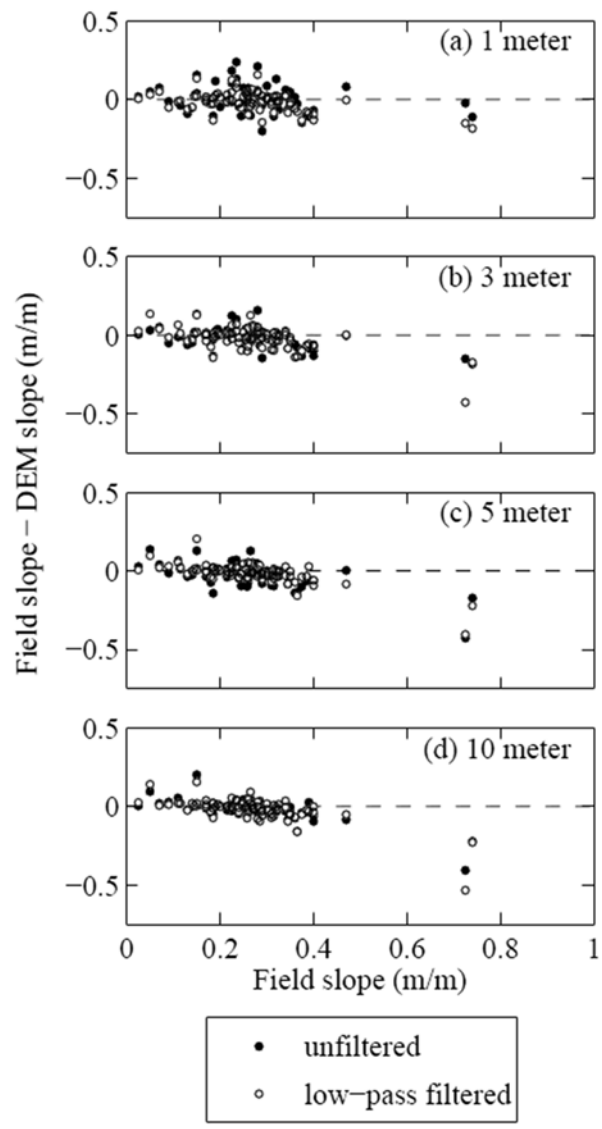

Figure 3. Field slope values measured using a clinometer for 75 locations corresponding with soil pits and shallow subsurface wells were compared with the difference between field and DEM slope measurements computed using the maximum slope algorithm for NCALM 1, 3, 5, and 10 m DEMs.

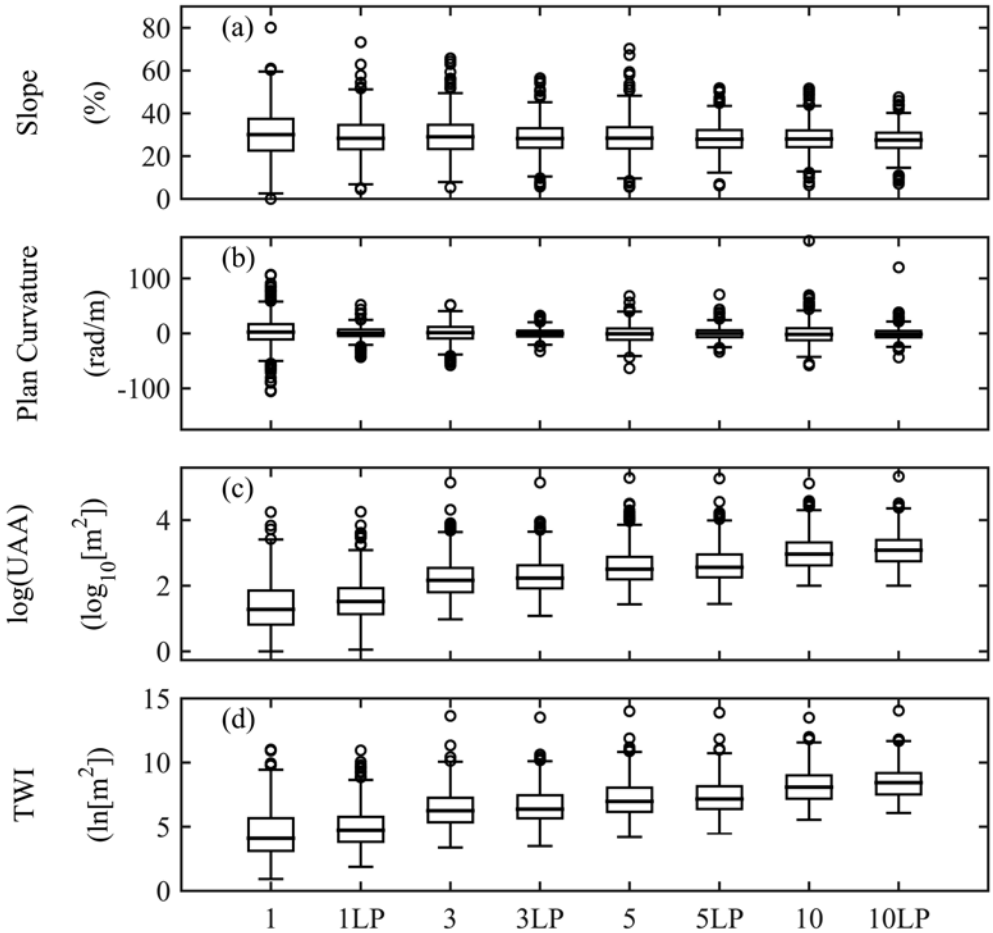

Figure 4. Box and whisker plots indicate variation in distribution of topographic metrics for WMNF: (a) slope, (b) planform curvature, (c) UAA, and (d), and TWI. Median values are represented by the thick black line, boxes represent the interquartile range, whiskers extend 1.5 times the interquartile range beyond the interquartile range box and contain approximately 99.3 percent of the data, and circles indicate points outside 99.3 percent of the data. Topographic metrics were computed for each DEM and values were extracted to 421 random points generated inside the WS3 boundary.

values also increased with filtering and coarsening from 4.1 to 8.4 when calculated using the WMNF DEMs (Figures 4c and 4d). The same overall increase in median TWI occurred for the NCALM DEMs. TWI distributions computed using the 3 m and 5 m DEMs, regardless of filtering, exhibited the greatest similarity.

Discussion

Watershed Boundary Variation with DEM Resolution and Landscape Roughness

Each DEM generated a different catchment boundary. This was especially true for the unfiltered NCALM 1 m DEM, which excluded nearly 1 ha more area than the field-surveyed boundary, and the unfiltered WMNF 10 m DEM, which included nearly 1 ha more area than the field-surveyed boundary (Table 2; Figure 2). The filtered 10 m DEMs from both datasets contained nearly 1 ha less area than the field-surveyed boundary. The area excluded or included for these four watershed boundaries was chiefly located in the southeastern portion of WS3, which is characterized by little to no channel formation and no steep spurs compared to the rest of the catchment. Variation in catchment boundary and flow accumulation across DEM resolutions is consistent with previous observations (e.g., Quinn *et al.*, 1995; Vaze *et al.*, 2010).

Such results demonstrate that care should be taken when using lidar-derived DEMs for watershed delineation, especially in regions where delineation of catchments is challenging due to subtle topography. Uncertainty in catchment area determination is a critical factor in evaluating catchment water and nutrient balances as the estimate of atmospheric precipitation inputs is dependent on this parameter. Yanai *et al.* (2014) evaluated sources of uncertainty in stream water flux in long-term catchment studies and noted that watershed area has not been critically assessed at well-known long-term catchment installations. DEM aggregation methods may not adequately preserve drainage features, which affect the delineation process. An important consideration from a hydrological perspective during DEM generation is maintaining relative elevation differences or drainage features (Ai and Li, 2010; Chen *et al.*, 2012).

Agreement between DEM and Field Slope Values

DEM-computed slope values were similar to field-measured slope values, particularly for the 5 m and 10 m DEMs (Figure 3c and 3d). In this study, field slopes were measured at a 5 m scale, which helps explain why 1 m DEMs generated slope values least similar to field measurements. Such results demonstrate that DEM resolution should reflect the desired scale of information intended for the application, e.g., operations and management decisions, erosion modeling, or soil mapping.

The tendency for slope values to become more intermediate (decreased maximum values and increased minimum values) with DEM coarsening is consistent with previous

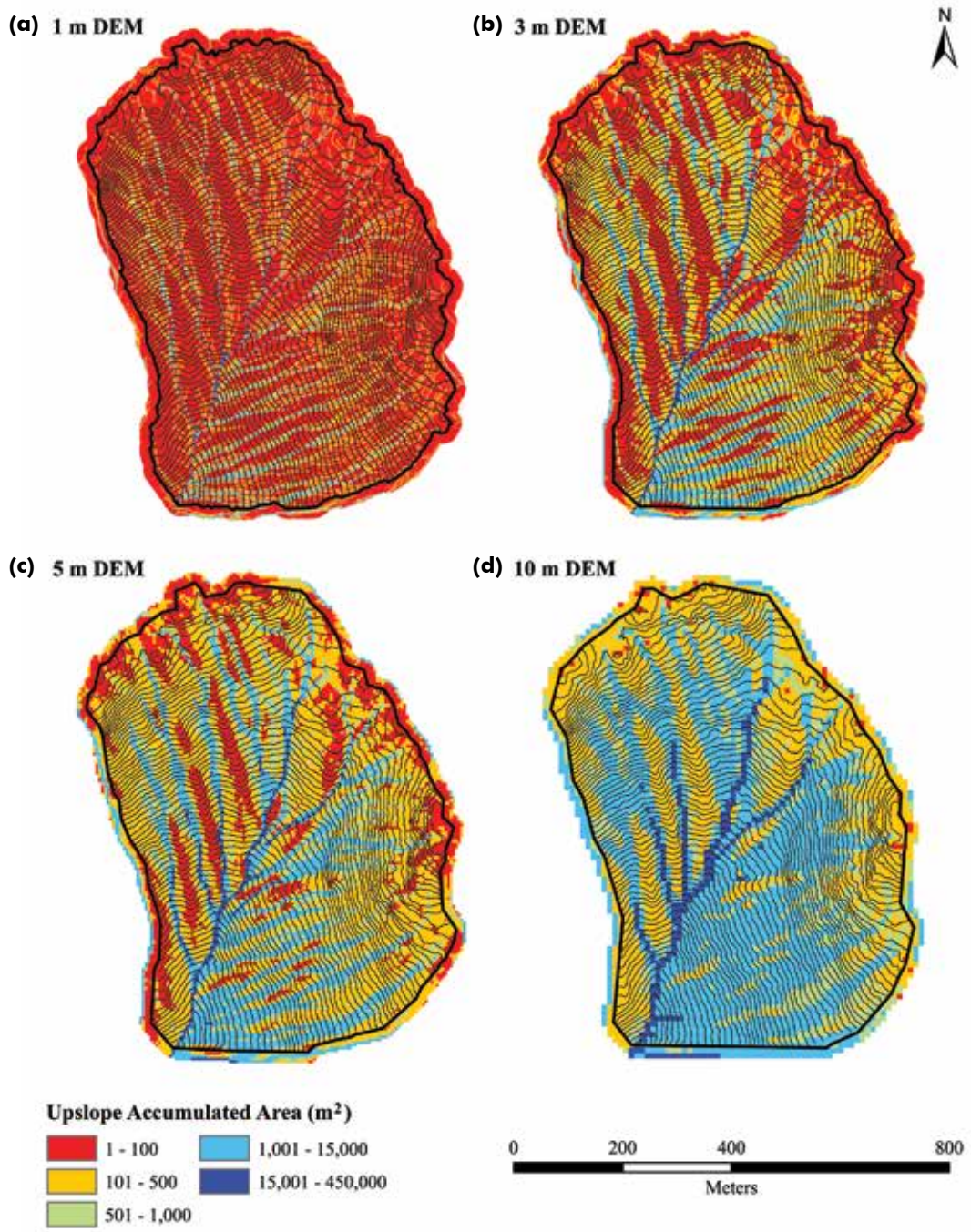

Plate 1. UAA (m^2) computed for: (a) 1 m, (b) 3 m, (c) 5 m, and (d) 10 m WMNF DEMs with 3 m contour interval. Thick black line denotes the calculated catchment boundary for each DEM.

observations (e.g., Band and Moore, 1995; Sorensen and Seibert, 2007; Vaze *et al.*, 2010). Slope values can be expected to become more intermediate when finer-resolution grid cell neighborhoods are averaged into larger cells through DEM coarsening. Our observation that the filtered version of a given DEM resolution exhibited a smaller variance than the unfiltered version for planform curvature suggests that low-pass filtering has a greater impact on local surface roughness than DEM coarsening by mean cell aggregation.

Classification and Interpolation of Lidar Returns

Total station surveys resulted in greater RMSE for locations with terrain features, such as boulders or fallen trees, than for non-feature survey points. This suggests that either the lidar pulses tended to miss terrain features or, more likely, that DEM interpolation algorithms smoothed such features. This may present challenges for researchers interested in using DEMs for studies at the hillslope scale, or to evaluate processes dependent on micro-topography. For example, hummocky terrain and large boulders may affect surface and subsurface water flowpaths, but our observations suggest that lidar interpolation algorithms mitigate roughness from these terrain features during DEM generation. Researchers interested in micro-topography may be able to utilize a DEM generated from unclassified lidar returns although our research suggested it could be difficult or impossible to distinguish boulders, hummocks, or fallen tree boles from low-lying vegetation.

RMSE was lower in forest clearings than under mature canopy for both NCALM and WMNF datasets. This result is expected, as the density of lidar ground returns used during DEM interpolation should increase as vegetative cover decreases. Higher lidar ground return density facilitates the interpolation of more accurate DEMs, and thus clearings are presumed to allow more accurate representations of the ground surface than regions with a canopy and therefore exhibit a lower RMSE. Greater RMSE for RG5 compared with other survey sites may have been caused by dense beech saplings. Young beech thickets have a higher leaf-area index (LAI) (Genet *et al.*, 2009) and perhaps a more complex branch architecture than mature stands. Limbs can cause lidar pulses to reflect in a direction other than to the sensor resulting in spurious data points known as multipath errors. Furthermore, when a pulse encounters an object and records a return, there is a lag distance of approximately three vertical meters before another return can be recorded. Hobblebush and other herbaceous understory plants found in WS3 growing to about one meter in height did not appear to interfere with lidar accuracy. However, the midstory beech thicket likely increased multipath errors and decreased the number of recorded ground returns, resulting in reduced accuracy and/or misclassification of off-terrain features as ground. Overall, our results suggest that DEM elevation error under a mature forest canopy of northern hardwoods in leaf-off conditions and in steep terrain may include as much 0.3 m to 0.5 m more error in elevation compared to open areas.

The two lidar datasets had different ground return densities in WS3; the NCALM dataset had a density of 3.27 ppsm while the WMNF dataset had a density of 1.16 ppsm (Table 1). Yet the datasets generated similar DEMs in terms of raw elevations, and the DEMs produced similar catchment boundaries, especially when coarsened to 3 m and 5 m resolutions. Furthermore, distributions of topographic metrics computed from filtered/unfiltered coarser DEMs derived from the original 1 m resolution for each dataset were similar (Table 2; Figure 4). Density of nominal post-spacing has been shown to affect DEM accuracy (e.g., Aguilar *et al.*, 2005; Hodgson *et al.*, 2004; Hu *et al.*, 2009) although several studies also indicated that in some situations lidar sampling density can be reduced by up to 50 percent with no degradation of DEM accuracy (Anderson *et al.*, 2006; Liu *et al.*, 2007). Denser post-spacing can be achieved using a higher pulse rate, lower altitude over-flight, narrower scan angle, or some combination of these variables (Raber *et al.*, 2007). The WMNF dataset contained lower nominal post-spacing and was collected from a higher altitude than the NCALM dataset. Similarity of DEM elevations, computed topographic metrics, and delineated catchment boundaries for each dataset investigated in this study suggest that post-spacing, within a modest range of variation, may not be the most limiting factor to the quality of DEMs for such watershed studies.

Topographic Metric Variation with DEM Resolution

TWI distributions tended to increase with DEM coarsening and filtering, consistent with previous investigations of coarser (10 m and greater) DEMs (e.g., Quinn *et al.*, 1991; Sorensen and Seibert, 2007; Wolock and Price, 1994) but extended to a finer scale in this study. Observed increases in median UAA values with DEM coarsening (Figure 4) area also was consistent with previous observations (Zhang and Montgomery, 1994). The large differences in mean TWI values (Table 2) computed from unfiltered 1 m and filtered 10 m DEMs derived from NCALM and WMNF lidar datasets, respectively, are significant from the perspective of the hydrological modeler. This disparity is equivalent to approximately two orders of magnitude difference in simulated subsurface flow when following TOPMODEL theory (Beven and Kirkby, 1979) and using the hydraulic conductivity distribution with depth observed for WS3 (Detty and McGuire, 2010). The parameterization of TOPMODEL and other hydrologic models is often independent of topography where it is static and taken directly from available elevation data. However, small differences (less than 10 m) in DEM resolution have implications for computed topographic metric values (e.g., the mean) that affect hydrologic quantities derived from such models.

Increasing TWI values with DEM coarsening appeared to be controlled by increases in UAA values. Minimum slope values increased and mean/median slope values decreased only slightly (Figure 2; Figure 4) with DEM coarsening and therefore cannot explain the observed increases in TWI values. However, UAA boxplots (Figure 4) and UAA maps (Plate 1) demonstrated increasing values with DEM coarsening. Mean values varied over an order of magnitude from a low of 161 m^2 for the NCALM filtered 1 m DEM to a high value of 3,411 m^2 for the NCALM filtered 10 m DEM, and must therefore offset observed changes in minimum, mean, and median slope values when calculating TWI. Maximum slope value decreases, in addition to UAA increases, may at least partially explain TWI maximum value increases with DEM coarsening.

UAA values computed from the 3 m and 5 m DEMs best differentiated topographic variation seen on the catchment map with a 3 m contour interval (Plate 1). Summits and convex shoulder slopes aligned with smallest UAA values while drainages and toeslopes aligned with greatest UAA values, and backslopes aligned with moderate UAA values. In contrast, the 1 m DEMs generated small UAA values throughout much of the catchment, with much of the landscape in a variety of topographic positions having a UAA of <100 m^2 (Plate 1) This may be due to fine topographic detail and surface irregularities, which our total station survey suggests are not well represented, that interfere with the computation of UAA. On the other hand, the 10 m DEMs generated UAA values >1,000 m^2 throughout many parts of the catchment, including some areas best described as planar backslopes (Plate 1).

An example of the implications of variation in topographic metrics derived from varying DEM scale in our study area is provided by Bailey *et al.* (2014) who found that variation soil horizon sequences and thickness, and groundwater fluctuations were best correlated with TWI and UAA values derived from 3 m to 5 m DEM resolution. Five soil map units were delineated based on these hydropedological variations and occurred at hillslope positions that can be predicted by interpretation of the 3 m contour interval topographic map. Further investigation of topographic metrics calculated from DEMs at this scale is warranted as a digital soil mapping tool in this region and highlights the importance of careful consideration of DEM resolution used to compute topographic metrics, as small resolution differences can yield dramatically different results in metric values for a given point on a landscape.

Conclusions

Total station surveys suggested that small scale terrain features such as boulders and fallen trees are smoothed when a DEM is generated from lidar data, and that DEM elevation accuracy increases in the absence of vegetative cover. However, even under mature forest canopy and in rough terrain, lidar still can produce DEMs useful for soil and hydrologic analyses. The similarities we observed in watershed boundaries, topographic metrics, and RMSE values computed from each set of DEMs suggest that differences in lidar collection methods and ground return densities we studied were not sufficient to create tangible DEM accuracy differences. Methods for increasing accuracy also increase the cost of lidar acquisition and data processing/storage requirements. Our findings suggest that these costs can be reduced while generating DEMs as accurate as those developed with greater monetary and time inputs.

While finer scale DEMs generated from lidar data may be necessary for some detailed soil and hydrology studies, in our study area, the highest possible resolution DEM was not the best tool as the preponderance of low UAA values across a range of topographic positions generated by the 1 m DEM were not suitable for soil horizon and water table comparison. DEM coarsening and filtration resulted in significant changes to UAA and TWI values. Given the importance of these metrics in hydrological research and watershed management, we recommend that DEM resolution for computing UAA and TWI be carefully selected based on prior observation and expert knowledge of the scale of features controlling the hydrologic response. Blindly using the highest resolution data available may hinder the progress of the effects of topographic relationships in watershed research.

Acknowledgments

Financial support was provided by the National Science Foundation Long-Term Ecological Research (DEB 1114804), Hydrologic Sciences (EAR 1014507), and Research Experience for Undergraduate (DBI/EAR 0754678) programs. Lidar data were collected by Photo Science, Inc. for the White Mountain National Forest and as part of the National Center for Airborne Laser Mapping (NCALM) seed award program. Field work was partially conducted by Rebecca Bourgault, Margaret Burns, Erin Shoop-Volitis, J.P. Gannon, and Geoffrey Schwaner. The Hubbard Brook Experimental Forest is operated and maintained by the USDA Forest Service, Northern Research Station, Newtown Square, Pennsylvania and is part of the NSF Long-Term Ecological Research network.

References

Aguilar, F.J., F. Aguera, M.A. Aguilar, and C. Fernando, 2005. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy, *Photogrammetric Engineering & Remote Sensing*, 71(7):805-816.

Ai, T., and J. Li, 2010. A DEM generalization by minor valley branch detection and grid filling, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65:198-207.

Anderson, E.S., J. Thompson, D.A. Crouse, and R.E. Austin, 2006. Horizontal resolution and data density effects on remotely sensed lidar-based DEM, *Geoderma*, 132:406-415.

Bailey, S.W., P.A. Brousseau, K.J. McGuire, and D.S. Ross, 2014. Influence of landscape position and transient water table on soil development and carbon distribution in a steep, headwater catchment, *Geoderma*, 226-227:279-289.

Band, L.E., and I.D. Moore, 1995. Scale: Landscape attributes and geographical information systems, *Hydrological Processes*, 9:401-422.

Barton, C.C., R.H. Camerlo, and S.W. Bailey, 1997. Bedrock geologic map of Hubbard Brook Experimental Forest and maps of fractures and geology in roadcuts along Interstate-93, Grafton County, New Hampshire, US Geological Survey, *Miscellaneous Investigations Series*, Map I-2562.

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

Chen, Y., J.P. Wilson, Q. Zhu, and Q. Zhou, 2012. Comparison of drainage-constrained methods for DEM generalization. *Computers & Geosciences*, 48:41-49.

Detty, J.M., and K.J. McGuire, 2010. Threshold changes in storm runoff generation at a till-mantled headwater catchment, *Water Resources Research*, 46:W07525, doi:10.1029/2009WR008102.

Gallant, J.C., and M.F. Hutchinson, 1996. Towards an understanding of landscape scale and structure, *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, 21-26 January, Santa Fe, New Mexico, unpaginated CD-ROM.

Genet, H., N. Breda, and E. Dufrene, 2009. Age-related variation in carbon allocation at tree and stand scales in beech (Fagus sylvatica L.) and sessile oak (Quercus petraea (Matt.) Liebl.) using a chronosequence approach, *Tree Physiology*, 30:177-192.

Gillin, C.P., S.W. Bailey, K.J. McGuire, and J.P. Gannon. 2015. Mapping of hydropedologic spatial pattens in a steep headwater catchment, *Soil Science Society of America*, doi:10.2136/sssaj2014.05.0189.

Hammer, R.D., F.J. Young, N.C. Wollenhaupt, T.L. Barney and T.W. Haithcoate, 1995. Slope class maps from soil survey and digital elevation models, *Soil Science Society of America Journal*, 59:509-519.

Hancock, G.R., 2005. The use of digital elevation models in the identification and characterization of catchments over different grid scales, *Hydrological Processes*, 19:1727-1749.

Hodgson, M.E., X. Li, and Y. Cheng, 2004. A parameterization model for transportation feature extraction. *Photogrammetric Engineering & Remote Sensing*, 70(12):1399-1404.

Hodgson, M.E., J.R. Jensen, G.T. Raber, J.A. Tullis, B.A. Davis, G. Thompson, and K. Schuckman, 2005. An evaluation of lidar-derived elevation and terrain slope in leaf-off conditions, *Photogrammetric Engineering & Remote Sensing*, 71:817-823.

Hu, P., X. Liu, and H. Hu, 2009. Accuracy assessment of digital elevation models based on approximation theory, *Photogrammetric Engineering & Remote Sensing*, 75(1):49-56.

Isaacson, D.L., and W.J. Ripple, 1990. Comparison of 7.5-minute and 1-degree digital elevation models, *Photogrammetric Engineering & Remote Sensing*, 56(12):1523-1527.

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

Jenson, S.K., 1991. Applications of hydrologic information automatically extracted from digital elevation models, *Hydrological Processes*, 5:31-44.

Likens, G.E., 2013. *Biogeochemistry of a Forested Ecosystem*, Springer, New York.

Lillesand, T.M., and R.W. Kiefer, 2000. *Remote Sensing and Image Interpretation,* John Wiley and Sons, Inc., New York.

Liu, X., Z. Zhang, J. Peterson, and S. Chandra, 2007. The effect of lidar data density on DEM accuracy, *Proceedings of the International Congress on Modelling and Simulation (MODSIM07): Land, Water and Environmental Management: Integrated Systems for Sustainability*, 10-13 December, Christchurch, New Zealand.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden, 2005. A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res., 110, D21106, doi:10.1029/2005JD006111..

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

Quinn, P., K. Beven, and R. Lamb, 1995. The ln(a/tanβ) index: how to calculate it and how to use it within the TOPMODEL framewor, *Hydrological Processes*, 9:161-182.

Raber, G.T., J.R. Jensen, M.E. Hodgson, J.A. Tullis, B.A. Davis, and J. Berglund, 2007. Impact of lidar nominal post-spacing on DEM accuracy and flood zone delineation, *Photogrammetric Engineering & Remote Sensing*, 73(6):793-804.

Reutebuch, S.E., R.J. McGaughey, H. Andersen, and W.W. Carson, 2003. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy, *Canadian Journal of Remote Sensing*, 29:527-535.

Saulnier, G.M., K. Beven, and C. Obled, 1997. Digital elevation analysis for distributed hydrological modeling: Reducing scale dependence in effective hydraulic conductivity values, *Water Resources Research*, 33:2097-2101.

Seibert, J., and B.L. McGlynn, 2007. A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models, *Water Resources Research*, 43:W04501., doi:10.1029/2006WR005128

Shi, X., L. Girod, R. Long, R. DeKett, J. Philippe, and T. Burke, 2012. A comparison of LiDAR-based DEMs and USGS-sourced DEMs in terrain analysis for knowledge-based digital soil mapping, *Geoderma*, 170:217-226.

Siccama, T.G., T.J. Fahey, C.E. Johnson, T.W. Sherry, E.G. Denny, E.B. Girdler, G.E. Likens, and P.A. Schwarz, 2007. Poplulation and biomass dynamics in a northern hardwood forest at Hubbard Brook, *Canadian Journal of Forest Resources*, 37:737-749.

Sorensen, R., and J. Seibert, 2007. Effects of DEM resolution on the calculation of topographical indices: TWI and its components, *Journal of Hydrology*, 347:79-89.

Su, J., and E. Bork, 2006. Influence of vegetation, slope, and lidar sampling angle on DEM accuracy, *Photogrammetric Engineering & Remote Sensing*, 72(11):1265–1274.

Tague, C.L., and L.E. Band, 2001. Evaluating explicit and implicit routing for watershed hydro-ecological models of forest hydrology at the small catchment scale, *Hydrological Processes*, 15:1415-1439.

Travis, M.R., G.H. Elsner, W.D. Iverson, and C.G. Johnson, 1975. VIEWIT: Computation of seen areas, slope, and aspect for land-use planning, *USDA FS General Technical Report PSW-11/1975*, Berkeley, California.

Vaze, J., J. Teng, and G. Spencer, 2010. Impact of DEM accuracy and resolution on topographic indices, *Environmental Modelling & Software*, 25:1086-1098.

Walker, J.P., and G.R. Willgoose, 1999. On the effect of digital elevation model accuracy on hydrology and geomorphology, *Water Resources Research*, 35:2259-2268.

Wang, L., and H. Liu, 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling, *International Journal of Geographical Information Science*, 20:193-213.

Wolock, D.M., and C.V. Price, 1994. Effects of digital elevation model map scale and data resolution on a topography-based watershed model, *Water Resources Research*, 30:3041-3052.

Wu, S., J. Li, and G.H. Huang, 2008. A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution, *Applied Geography*, 28:210-223.

Yanai, R.D., N. Tokuchi, J.L. Campbell, M.B. Green, E. Matsuzaki, S.N. Laseter, C.L. Brown, A.S. Bailey, P. Lyons, C.R. Levine, D.C. Buso, G.E. Likens, J.D. Knoepp, and K. Fukushima, 2015. Sources of uncertainty in estimating stream solute export from headwater catchments at three sites, *Hydrological Processes*, 29: 1793-1805.

Zevenbergen, L.W., and C.R. Thorne, 1987. Quantitative analysis of land surface topography, *Earth Surface Processes and Landforms*, 12:47-56.

Zhang, W., and D.R. Montgomery, 1994. Digital elevation model grid size, landscape presentation, and hydrologic simulation, *Water Resources Research*, 30:1019-1028.

(Received 02 February 2014; accepted 22 December 2014; final version 26 December 2014)