The Pennsylvania State University

The Graduate School

School of Forest Resources

MULTI-SCALE LIDAR-BASED APPROACHES TO CHARACTERIZING STREAM NETWORKS, SURFACE ROUGHNESS AND LANDFORMS OF FOREST WATERSHEDS

A Dissertation in

Forest Science

by

Kristen M. Brubaker

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2011

The dissertation of Kristen Brubaker was reviewed and approved* by the following:

Elizabeth Boyer
Associate Professor of Water Resources
Co-Dissertation Advisor
Co-Chair of Committee

Wayne Myers
Professor Emeritus of Forest Biometrics
Co-Dissertation Advisor
Co-Chair of Committee

Eric Zenner
Associate Professor of Silviculture

Patrick Drohan
Assistant Professor of Pedology

Douglas Miller
Associate Professor of Geography

Michael Messina
Director, School of Forest Resources

*Signatures are on file in the Graduate School

ABSTRACT

The overall objective of this study is to utilize high resolution lidar-derived digital elevation models (DEMs) to improve classification and understanding of forested watersheds. Since geographic information systems technology became broadly used in natural resource fields in the 1980s, scientists have used digital elevation models to study aspects of forested ecosystems including the delineation of drainage networks, geomorphic modeling, and ecological classification for forest management and ecosystem management. With recently available lidar elevation data, we have improved our ability to "see" features on the landscape by orders of magnitude. Existing methodologies for assessing geomorphometry and hydrologic network delineation across the landscape may not suffice for all tasks. By taking a multi-scale, multidisciplinary approach, we can improve our understanding of headwater ecosystems and how to assess and predict the relationship between terrain and vegetation. This research was performed in the Leading Ridge experimental watersheds, the site of a long-term study analyzing the impact of forest management practices on stream water quality. The Leading Ridge experimental watersheds are also located within the Susquehanna/Shale Hills Critical Zone Observatory.

In order to assess the ability of lidar-derived DEM to improve stream network modeling, the stream network for Leading Ridge watershed number one was recorded using a GPS unit during base flow conditions. The stream network was then modeled using lidar-derived 1 m, 3 m, and 10 m resolution DEMs as well as photogrammetrically-derived NED (National Elevation Dataset) DEM. All of the lidar-derived DEMs resulted in a relatively accurate stream network model, with the 3 m DEM providing the most accurate model. There was no significant difference between any of the lidar-derived modeled stream networks, but they were all significantly different from the NED DEM-derived stream network, which was much less

accurate. Topographic index (TI) was modeled using multiple DEM products and presented very different statistical distributions and spatial patterns. The distribution of TI could have an impact on hydrologic models, while the improvements in network delineation could substantially improve our knowledge of headwater streams on the landscape. This could in part impact forest management, site planning, and ecosystem modeling.

Surface roughness was calculated for Leading Ridge using several algorithms on two different lidar-derived DEMs to evaluate patterns of roughness on the watershed. Roughness metrics included standard deviation of slope, value of pittedness in cells, standard deviation of curvature, and the difference between the original DEM and a splined surface. Micro-plots and transects were surveyed to ground truth roughness metrics. Although the scale of the 1 m DEM was too coarse to assess micro-topography at the same scale as the ground survey, unique patterns were identified on different landforms and soil types. There was also substantial interaction between the roughness algorithm and the DEM creation algorithm. The results suggest that although there are many complicating factors when assessing surface roughness using a lidar-derived DEM, there is information about soils and topography that can be obtained. Also, DEMs studied here had slightly higher elevation values (about 0.3 m) on average than the field-surveyed elevations.

In order to relate topography to vegetation, curvature was chosen to model landforms based on its importance to water transport on an ecosystem. There was evidence of curvature being reflective of underlying geology and predictive for soil properties that may affect vegetation. Leading Ridge watershed was delineated into nine curvature classes using a 10 m DEM, and patterns of curvature were used to construct four recurring formations: hidden hollows, rock ridgelets, scalloped slopes, and rounded ridges. Based on a vegetation analysis of these formations, there was a difference in both vegetation community and structure based on

formation. Similar formations were calculated for a broader region of the Ridge and Valley Province and vegetation communities on formations were identified. There was association between the identified vegetation community and the delineated formation.

Overall, methodologies were developed to explore properties of forested ecosystems in the Ridge and Valley Province. Using lidar elevation data, delineation of the stream network and characterization of terrain and micro-topography were all improved, and curvature was utilized to help classify the landforms in watershed. Further research should attempt to validate these results across a broader area, as well as work to develop techniques to use together to create a multi-scale, hierarchical classification system incorporating hydrologic data, surface roughness, and landscape level terrain data.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to thank my advisors, Dr. Wayne Myers and Dr. Beth Boyer. Dr. Myers provided me continual support, and guidance, even during his retirement, while Beth provided many exciting ideas, feedback, support, and a lot of positive energy necessary for the long process of completing a dissertation. Also, I would like to thank my committee members, Dr. Patrick Drohan, Dr. Doug Miller, and Dr. Eric Zenner for their extremely valuable insights, feedback, and for providing extra support I may have needed along the way. Many others in the School of Forest Resources also contributed to my development as a scholar including Dr. Paola Ferreri, Dr. Jeri Peck, and Lysle Sherwin. I would like to thank many graduate students for their academic help, their engaging conversations, and social support, particularly Yvette Dickinson, Sarah Johnson, Robert Guyer, and Katie Gaines who helped with field work. Finally, I would like to thank my family and friends for their continued support and for always believing I would finish my dissertation.

.

Chapter 1

Introduction

Since geographic information systems technology became broadly used in natural resources fields in the 1980s, scientists have used digital elevation models (DEMs) to study aspects of forested ecosystems including the delineation of drainage networks (Tarboton 1991), geomorphometric modeling (Moore et al. 1991), and ecological classification for forest management and ecosystem management (Franklin 1995). Traditionally, DEMs have been generated using photogrammetric methods and their utility is limited by resolution, accuracy, and consistency across a landscape. With the newly available light detection and ranging, or lidar-derived elevation data, we have improved our ability to "see" features on the landscape by orders of magnitude. Thus, lidar not only offers an improvement of resolution, but opens up a whole new way of viewing landscapes. The magnitude of the upgrade with lidar is so great that we may not be able to utilize lidar-derived DEMs in the same way we have used previously available DEMs. Models and algorithms developed using 10 m and 30 m DEMs may not function the same way with lidar-derived DEMs. Due to the improvements in scale and vertical accuracy, we shouldn't just analyze lidar data with previous methodologies and algorithms. Since lidar is available at such high resolution and accuracy, it becomes possible to analyze features at a much finer scale. For the purposes of this research, multi-scale refers to working with scales from approximately 1 m or less, up to tens, hundreds or thousands of meters. Working at these scales provides additional detail and insight in forest ecosystems.

With the availability of high-quality digital elevation information from lidar (Maune, 2007) along with improvements in processing tools including GIS and other modeling techniques, DEMs have become an integral part of studying natural resources. Lidar promises to not only

improve our ability to accurately model the earth, but also allows us to see forests and other vegetation in three dimensions. Lidar works by recording the distance between a sensor, usually mounted on a fixed wing aircraft, and the ground or other surface that may reflect light such as buildings or trees. This provides three-dimensional data that can be converted into an extremely accurate DEM, often with horizontal resolutions of approximately 1 meter, and accuracies of between 10 and 30 cm both in the vertical and horizontal direction. In addition to providing accurate terrain information, lidar returns can be used to model ecosystem functions such as net primary productivity (NPP) and total biomass of a site (Lefsky et al. 1999, Kotchenova et al. 2005, Lefsky et al. 2005).

Approach and Rationale

In order to more effectively utilize the improved resolution and accuracy of lidar, a multi-scale, multidisciplinary approach was undertaken to improve our understanding of a small, forested watershed. Because of lidar's accuracy, for this research multi-scale refers to approximately the 1 m scale up to 1000s of meters. Multi-disciplinary refers to combing soils and geology, topography, hydrology and forestry data for this research. By focusing this study in one location, extensive ground truthing was possible for stream networks, micro-topography, geomorphometry, and forest community and structure to validate the various model results and synthesize relationships and interactions.

A multi-scale approach was undertaken by exploring varying accuracies and resolutions of initial DEMs, using multiple sized windows and algorithms for analysis, and classifying terrain features at different scales. The surface expressions of underlying soil, geology, and hydrology occur at multiple scales, however for this research multi-scale is restricted to finer scaled features due to advancements in resolution and accuracy due to lidar. Since terrain is fractal in nature

(Glenn et al. 2006), metrics such as slope and roughness can be expressed very differently at varying scales and what appears as slope at one scale may be roughness at another. I explored different ways of measuring both terrain and roughness in the scale range from sub-meter to 1000s of meters. The goal of the research was to delineate features that occur at multiple scales that may impact a landscape, from micro-topography on the sub-meter scale to ridges on the scale of several kilometers. Also, hydrologic features were modeled using the finest available lidar terrain data (1 m) and generalizations from the finest scales (3 m and 10 m) to represent the stream network and wetness on the landscape in different ways.

A multi-disciplinary approach was used to emphasize the broad use of lidar-derived terrain data and to maximize the use of the site to extract information. In addition to its previous use as an experimental watershed, Leading Ridge is situated within the Shaver's Creek watershed which is being used as part of the Susquehanna/Shale Hills Critical Zone Observatory (CZO), an interdisciplinary observatory toward quantitatively predicting creation, evolution, and structure of regolith as a function of the geochemical, hydrologic, biologic, and geomorphologic processes operating in a temperate forested landscape (Anderson et al. 2008). The critical zone is defined as the "external terrestrial layer extending from the outer limits of vegetation down to and including the zone of groundwater" (Brantley et al. 2006). Because processes in the critical zone are dependent on one another, it was important to explore the impacts that certain decisions made for one field would have on another. For example, what impact would using a 1m resolution terrain dataset have on the models of topographic index (TI) that will, in turn, influence hydrologic models. A central question also involved optimizing scale for the purpose of utilizing data efficiently and effectively. Although studied on the fine scale of a single, 119 ha watershed, the goal of utilizing these techniques for a broader scale as well was a consideration.

Although properties of soils and underlying geology undoubtedly have an impact on vegetation, there is often a lack of available spatial data on soils and geology at a fine scale.

Therefore, in order to truly improve our predictive and modeling techniques, properties of the DEM must be used to infer properties of the underlying soils and geology. This research works to use lidar for delineating features that make up landforms based on the surficial expression of soils and geology in the Ridge and Valley Physiographic Province.

Research Objectives

- Improve delineation of hydrologic network using 1 m lidar DEM and generalizing up to 10 m resolution. Calculate TI for multiple resolutions of lidar to compare statistical distribution and spatial patterns.
- Characterize and model micro-topography using lidar-derived DEM across a limited range of scales. Compare lidar-derived 1m DEM to surface survey.
- Create a landform classification system that relates to patterns of landform curvature that affect vegetation communities and structure.

Previous Work

For this research, lidar is utilized to characterize a watershed based on terrain, micro-topography, hydrology, and forest community metrics. As improvements have been made in modeling, computer processing, and digital elevation data, the use of DEMs in ecosystem studies has grown (Franklin 1995). Due to known relationships between topography, climate, soil formation, and vegetation, much of the early work focused on methods of classifying landforms and modeling terrain from digital terrain models including DEMs and other digital representations. Research by Zevenbergen and Thorne (1987) and Moore et al. (1991, 1993) focused on calculating metrics such as slope, aspect, and curvature from computer based data.

There have been multiple books (Wilson and Gallant 2000, Hengl and Reuter 2009) and review papers (Blaszczynski 1997, Deng 2007) on the topic of terrain analysis using digital elevation data and there are many algorithms and techniques for calculating terrain metrics (Deng 2007).

Classifying landscapes and landforms is a common goal of scientists trying to utilize terrain data for management and predicting relationships between terrain and soil/vegetation. Pennock et al. (1987) used slope and curvature to delineate landforms into seven landform elements for the purpose of predicting soils. Gessler et al. (1995) also attempted to predict soil properties using terrain data basing their model on the concept that terrain directs the movement of water and other materials. They found that compound terrain index (CTI) and plan curvature predicted soil composition very well. Dikau (1989, 1993) and Moore (1993) also did much work with terrain classification using terrain data as well as soil and other digital data sources.

More recently, efforts have focused on automated classification systems based on multi-variate classification using clustering (k-means) and fuzzy membership rules (MacMillan et al. 1998, Burrough et al. 2000, Schmidt and Hewitt 2004, Summerell et al. 2005). These studies used many different landscape metrics in the generation of their landscape classification. The methodology of Summerell et al. (2005) focused on hydrological analysis variables, while others have relied on slope, curvature, and landscape position (Schmidt and Hewitt 2004). Myers (2000) and Kong (2006) have worked at delineating Pennsylvania into landscapes based on soil, landscape shape and landscape position.

With regard to hydrologic modeling or modeling of geomorphic processes using DEMs, O'Callahan and Mark (1984) proposed an algorithm of flow routing, the D8 method, that is still widely used today and directs flow from any cell of a DEM into the one cell with the lowest elevation adjoining that cell. There has been criticism of this method because it tends to result in parallel flow paths and cannot model flow divergence. More recently, several multiple flow direction algorithms have been proposed (Moore et al. 1993, Tarboton 1997), all of which are

capable of routing flow to multiple cells and are thought to provide a more accurate model of flow accumulation than the D8 method. Wilson et al. (2007) reviewed flow routing algorithms and found that although in general multiple direction flow routing algorithms were preferred to single direction methods, there was no agreement on a preferred methodology for all purposes.

Topographic index (TI) is another commonly used metric that was first proposed for use in TOPMODEL, a catchment scale hydrologic model (Beven and Kirkby 1979). TI, also sometimes referred to as topographic wetness index (TWI) is now used in everything from hydrologic modeling to soil modeling and prediction and ecosystem studies. TI is measured by the formula $\ln(a/\tan\beta)$ (a = upslope contributing area per unit contour; $\tan\beta$ = local slope angle). Since there are multiple algorithms for a and β, as well as differences in resolution and quality of DEMs, much work has been done to characterize the preferred method of calculating TI (Quinn et al. 1995, Sørensen et al. 2006, Sørensen and Seibert 2007, Straumann and Purves 2007). Sørensen et al. (2006) specifically explored methods of TI calculation for predicting vegetation and found that multi-directional contributing area algorithms are preferred to single direction algorithms, but that different methodologies performed better for different applications. The same seems to be true for resolution, with the common theme being that resolution does make a difference in both the statistical and geographic distribution in TI values, but there is no real agreement on the ideal resolution. Many of these studies were done, however, before the widespread use of lidar and results were based on photogrammetrically-derived DEMs.

Several studies have already demonstrated that lidar can provide improvements over traditional DEMs in the field of hydrologic modeling and stream networks delineation. Colson (2006) and Murphy et al. (2008) both used lidar-derived DEMs to improve network delineation. Lane et al. (2004) proposed an adaptation to TOPMODEL to accommodate high resolution terrain data in the calculation of TI. Murphy et al. (2011) used a depth to water index in addition

to TI to improve understanding of the soil, vegetation, and drainage properties that are associated with higher TI values.

As available elevation data has been changing and improving, there have been numerous studies examining the impact of scale and resolution on terrain and hydrologic variables. Many early studies evaluating the effects of resolution were working with photogrammetrically-derived data that lacked accuracy when compared to lidar-derived data. These studies often found no real improvements with increased resolution beyond a certain point, generally around 10 m, (Kienzle 2004, Thompson et al. 2001), although they also agreed that results in all terrain modeling can vary due to scale. Some studies conducted with lidar-derived data also sometimes have produced contradictory conclusions. Zhang et al. (2006) found that 10m lidar-derived data was the most effective scale for modeling susceptibility to erosion and that significant improvement didn't occur as resolution was increased to 4m, while Zhao et al. (2010) found that a 10m lidar DEM did not perform well in hydrologic modeling when compared to a 1 m lidar DEM, and that resolution of the DEM was more important than accuracy in hydrologic modeling. Vianello et al. (2009) used lidar-derived DEMs to model slopes of a channel network and found the ideal resolution to be between 2-5 meters, with 1 m lidar overestimating slopes. Tennenbaum et al. (2006) also found that using a smoothed lidar-derived DEM more effectively modeled soil moisture than the original data, perhaps due to the over-representation of features such as tree roots in the original dataset. Although there is no clear preferred resolution for all purposes, it seems clear that for some modeling a coarser resolution may be preferred to the highest resolution data available.

Many researchers have also focused on identifying the optimal scale for delineating particular properties or explore which properties are most sensitive to changes in scale. Curvature seems to be particularly problematic. Deng et al. (2007) found that both plan and profile curvature were more sensitive to scale than slope, with all terrain attributes responding to resolution change in characteristically different ways. Lassueur et al. (2006) also found that

curvature was very scale dependent. Tarolli et al. (2010) used multiple sized moving windows for calculating curvature and found that a 10 m resolution was the most effective in most cases, although delineating different features may require different sized windows. Tarolli and Fontana (2009) used very fine scale curvature windows to try to detect channel heads, perhaps identifying a feature that may require detection at a finer scale. The overarching message from these previous studies involving scale, accuracy, DEMs, and landforms, may be that although scale and resolution are very important, there is not an ideal scale, moving window size, or resolution for all purposes.

Because of the unique perspective lidar provides, some studies have begun to analyze surface roughness and micro-topography as an additional property of terrain. Early studies have focused on using topographic roughness signatures to identify landslide areas (McKean and Roering 2003, Glenn et al. 2006, Van Den Eeckhaut et al. 2007), alluvial fan features (Frankel and Dolan 2007), describe mangrove habitat (Knight et al. 2009), and identify stream gullies (James et al. 2007). Many different methods of calculating roughness have been used, including standard deviation of slope (Frankel and Dolan 2007), the difference between regional topography and local topography (Glenn et al. 2006), and variograms/fractal methods to examine spatial properties of differences (Glenn et al. 2006, Dragut et al. 2011). Cavelli et al. (2008) calculated roughness metrics for a stream channel to identify riffle/pool sequences using lidar.

An additional complicating factor in DEM analysis is the many different methodologies available for generating DEMs from elevation data. Lidar data are generally available in much higher density point clouds than previously available data, but there are still several different types of algorithms used to generate a DEM that can have impacts on accuracy and resulting attributes. Siska and Hung (2001) evaluated several different methodologies including generating a TIN (Triangular Irregular Network) from points and converting to a DEM, kriging, Thiessen polygons, Inverse Distance Weighted (IDW), and trend. He found TIN and kriging to be the most

accurate. Spaete et al. (2011) and Davenport et al. (2004) both examined different sources of error with lidar, with Spaete et al. (2011) finding that slope and vegetation can both negatively impact lidar accuracy, while Davenport et al. (2004) finding that over time, repeated lidar readings can experience errors up to 20-40 cm.

Understanding the impacts of terrain and water movement through an ecosystem on vegetation and other ecological parameters has been a focus on ecosystem research. Although it is understood that site is an important predictor to vegetation, there has not been agreement of which particular variables may be the most important in understanding and predicting vegetation. Many classification methodologies rely on spatial geologic or soils data that may not be available at a fine enough scale to pair with lidar. Bowersox and Ward (1972) used field-measured slope position, slope, depth of A horizon, percent silt, N content, and Mg content to predict site index. Davis and Goetz (1990) predicted live oak distribution using geology, topography, and solar radiation calculated from GIS and found that geology was the most important factor. Franklin (1995) reviewed predictive vegetation mapping and provides a nice overview of groups of methodologies. Iverson et al. (1997) and Abella et al. (2003) both used geomorphic factors in addition to mapped soils data to predict and explain vegetation distribution.

Hutchinson et al. (1999) created a moisture index to predict forest and understory species using slope and aspect, flow accumulation, and curvature, all of which can be calculated using digital elevation data. Moseley et al. (2010) defines the concept of ecological site development, which is very similar to ecological classification systems (ECSs) in ecosystem management, but from a soils perspective. Duniway et al (2010) takes the ecological site development concept a step further by incorporating the role of geomorphology in site development. Reiners et al. (1999) identified the conflict between the scale available for terrain data and other spatial variables such as soil, and chose to only use terrain variables in a landscape classification of Wyoming. Although it is understood that geology and soils data are important to understanding and

predicting the distribution of vegetation, the spatial data available for soils and geology are often not at a scale that makes their utilization with lidar appropriate. With that in mind, this research attempts to improve classification using primarily lidar elevation data and derivations of that data as the input.

References

Abella, S.R., Shelburne, V.B. & MacDonald, N.W. 2003. Multifactor classification of forest landscape ecosystems of Jocassee Gorges, southern Appalachian Mountains, South Carolina. Canadian Journal of Forest Research, 33:1933–1946.

Anderson, S.P., Bales, R.C, Duffy, C.J. 2008. Critical zone observatories: building a network to advance interdisciplinary study of Earth surface processes. Mineralogical Magazine 72: 7–10.

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

Blaszczynski, J.S. 1997. Landform characterization with geographic information systems, Photogrammetric Engineering and Remote Sensing 63:183–191.

Bowersox, T.W., W.W. Ward. 1972. Prediction of oak site index in the ridge and valley region of Pennsylvania. Forest science, 18: 192-195.

Burrough, P.A., van Gaans, P.F.M. and Macmillan, R.A., 2000. High-resolution landform classification using fuzzy k-means. Fuzzy Sets and Systems 113: 37–52.

Cavalli, M., Tarolli, P., Marchi, L., Dalla Fontana, G., 2008. The effectiveness of airborne LiDAR data in the recognition of channel bed morphology. Catena 73:249–260.

Colson, T. P. 2006. Stream Network Delineation from High Resolution Digital Elevation Models. Ph.D. Dissertation, Dept. of Forestry and Environmental Resources, NC State University, Raleigh, NC.

Davenport, I. J., Holden, N., Gurney, R. J. 2004. Characterizing errors in airborne laser altimetry data to extract soil roughness. IEEE Trans. Geosc. Remote Sensing. 42:2130-2141.

Davis, F.W. and Goetz, S., 1990. Modeling vegetation pattern using digital terrain data. Landsc. Ecol. 4: 69–80.

Deng, Y. , Wilson, J. P. and Bauer, B. O. 2007. DEM resolution dependencies of terrain attributes across a landscape, International Journal of Geographical Information Science, 21:187-213

Dikau, R. 1989. The application of a digital relief model to landform analysis in geomorphology. Three dimensional applications in geographical information systems. J. Raper, ed. Taylor and Fancis: New York. 51-77.

Dikau, R., E.E. Brabb, R.K. Mark and R.J. Pike. 1993. Morphometric landform analysis of New Mexico. Advances in Geomorphometry. Pike, R.J. and R. Dikau, eds. Gebruder Borntraeger, Berlin. 109-126.

Drăguţ, L.,C. Eisank, T. Strasser, 2011. Local variance for multi-scale analysis in geomorphometry. Geomorphology, 130:162-172.

Duniway, M. C., B T. Bestelmeyer, A Tugel. 2010. Soils processes and properties that distinguish ecological sites and states. Rangelands. December 2010: 9-15.

Frankel, K.L., Dolan, J.F., 2007. Characterizing arid-region alluvial fan surface roughness with airborne laser swath mapping digital topographic data. Journal of Geophysical Research – Earth Surface 112:F02025.

Franklin, J., 1995. Predictive vegetation mapping: geographical modeling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geogr. 19: 474–499.

Gessler P.E, Moore, I.D., McKenzie, N.J., Ryan, P.J. 1995. Soil-landscape modeling and spatial prediction of soil attributes. International Journal of Geographical Information Systems 9: 421–432.

Glenn, N. F., D. R. Streuker, D. J. Chadwick, G. D. Thackray, and S. J. Dorsch. 2006. Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity. Geomorphology. 73:131– 148.

Hengl, T. and Reuter, H. I. 2009. Geomorphometry: Concepts, Software, Applications. Developments in Soil Science, Elsevier, Amsterdam, 33, 772 pp.

Hutchinson T.F., Boerner R.E.J., Iverson L.R., Sutherland S. and Kenedy Sutherland E. 1999. Landscape patterns of understory composition and richness across a moisture and nitrogen mineralization gradient in Ohio (U.S.A.) *Quercus* forests. Plant Ecology 144: 177–189.

Iverson, L.R.; Dale, M.E.; Scott, C.T.; Prasad, A. 1997. A GIS-derived integrated moisture index to predict forest composition and productivity in Ohio forests. Landscape Ecology. 12:331-348.

James, L. A., Watson, D.G., Hansen, W.F. 2007. Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. Catena 71:132–144

Kienzle, Stefan. 2004. The effect of DEM raster resolution on first order, second order, and compound terrain derivatives. Transactions in GIS. 8:83-111.

Knight J.M., Dale P.E.R., Spencer, J., Griffin, L.,. 2009. Exploring LiDAR data for mapping the micro-topography and tidal hydrodynamics of mangrove systems: an example from southeast Queensland, Australia. Estuarine, Coastal and Shelf Science 85:593–600.

Kong, N. 2006. Topographically-based landscape-scale ecological mapping in Pennsylvania. PhD. Dissertation. Intercollege Graduate Degree in Ecology. Penn State University, University Park, PA.

Kotchenova, Svetlana, Xiangdong Song, Nikolay Shabanov, Christopher Potter, Yuri Knyazikhin, Ranga Myneni. 2004. Lidar remote sensing for modeling gross primary production of deciduous forests. Remote Sensing of Environment. 92:158-172.

Lane S.N., Brookes, C.J., Kirkby, A.J. 2004. A network-index based version of TOPMODEL for use with high-resolution digital topographic data. Hydrological Processes 18:191–201.

Lassueur, T., Joost, S., Randin, C.F., 2006. Very high resolution digital elevation models: do they improve models of plant species distribution? Ecological Modelling 198:139–153.

Lefsky, M. A., Harding, D., Cohen, W. B., & Parker, G. G. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. Remote Sensing of Environment, 67, 83– 98.

Lefsky, M. A., D. P. Turner, M. Guzy, W. B. Cohen. 2005. Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. Remote Sensing of Environment. 95:549-558.

MacMillan,R.A., Pettapiece, W.W., Nolan, S.C., 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 113: 81–109.

Maune, D. F., editor. 2007. Digital elevation models and technologies and applications: the DEM user manual, 2nd Edition. The American Society for Photogrammetry and Remote Sensing. Bethesda, MD.

McKean, J., and J. Roering. 2004. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry, Geomorphology, 57, 331– 35.

Moore, I. D., Grayson, R. B., and Ladson, A. R. 1991. Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. Hydrological Processes. 5:3-30.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A. 1993: Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57: 443-52.

Moseley , K., P. L. Shaver, H. Sanchez, and B. T. Bestelmeyer. 2010. Ecological site development: a gentle introduction. Rangelands. December:16-22.

Murphy, P.N.C. J. Ogilvie, F.-R. Meng and P.A. Arp. 2008. Stream network modeling using LiDAR and photogrammetric DEMs: a comparison and field verification, Hydrological Process 22:1747–1754.

Murphy, P.N.C. J. Ogilvie, F.-R. Meng White, B., Bhatti, J. S. and Arp, A.S. 2011. Modelling and mapping topographic variations in forest soils at high resolution: A case study Ecological Modelling 222:2314-2332.

Myers, W.L. 2000. Landscape scale ecological mapping of Pennsylvania forests. Research report ER2002-2, Environmental Resources Research Institute, University Park, PA.

O'Callaghan, J.F. and Mark, D.M., 1984, The extraction of drainage networks from digital elevation models. Computer Vision, Graphics and Image Processing, 28: 323–344.

Pennock, D.J., Zebarth, B.J. and de Jong, E., 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. Geoderma, 40:297-315.

Quinn P. F., Beven K. J., Lamb R. 1995. The ln(A/tan beta) index – How to calculate it and how to use it within the TOPMODEL framework. Hydrological Processes. 9:161-182.

Reiners, W.A., E.V. Axtmann and R.C. Thurston. 1999. Delineations of landtype associations for northwest Wyoming and the buffalo resource area. Bureau of Land Management, University of Wyoming, final report.

Schmidt, J. and Hewitt, A. 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. Geoderma 121: 243–56.

Siska, P. I. Hung. 2001. Assessment of Kriging Accuracy in the GIS Environment. Annual ESRI International User Conference, San Diego, CA, 2001.

Sørensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrology and Earth System Sciences. 10:101–112.

Sørensen, 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.

Spaete, L. P., Glenn, N. F. Derryberry, D. R., Sankey, T. T. Mitchell , J. J., Hardegree , S. P. 2010. Vegetation and slope effects on accuracy of a LiDAR-derived DEM in the sagebrush steppe, Remote Sensing Letters, 2:4, 317-326.

Straumann R.K. and Purves R.S. 2007. Resolution sensitivity of a compound terrain derivative as computed from LiDAR-based elevation data. In: Fabrikant S.I. and Wachowicz M. (eds.) Lecture Notes in Geoinformation and Cartography. The European Information Society – Leading the Way with Geo-Information, Proceedings of AGILE 2007, 8.–11 May 2007, Aalborg, Denmark, 87-109.

Summerell, G. K., J. Vaze, N. K. Tuteja, R. B. Grayson, G. Beale, and T. I. Dowling. 2005. Delineating the major landforms of catchments using an objective hydrological terrain analysis method, Water Resour. Res., 41:W12416

Tarboton D. G., Bras R.L., Rodriguez-Iturbe I. 1991. On the extraction of channel networks from digital elevation data. Hydrological Processes 5: 81–100.

Tarboton, D. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research 33: 309-320.

Tarolli, P. and D. Fontana, G. 2009. Hillslope to valley transition morphology: new opportunities from high resolution DTMs, Geomorphology, 113, 47–56, 2009.

Tarolli, P., Sofia, G., and D. Fontana, G. 2011. Geomorphic features extraction from high resolution topography: landslide crowns and bank erosion, Nat. Hazards, ISSN: 0921-030X.

Tenenbaum, D.E., L.E. Band, S.T. Kenworthy, and C.L. Tague. 2006. Analysis of soil moisture patterns in forested and suburban catchments in Baltimore, Maryland, using high-resolution photogrammetric and LIDAR digital elevation datasets. Hydrol. Proc. 20:219–240.

Thompson, J A., J.C. Bell and C.A. Butler. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil–landscape modeling. Geoderma 100:67–89.

Van Den Eeckhau M., Poesen J, Verstraeten G, Vanacker V, Moeyersons J, Nyssen J, Beek LPHv, Vandekerckhove L. 2007. Use of LIDAR derived images for mapping old landslides under forest. Earth Surface Processes and Landforms 32: 754–769.

Vianello, A., M. Cavalli and P. Tarolli, 2009. LiDAR-derived slopes for headwater channel network analysis, Catena 76:97–106.

Wilson, J. P., Gallant, 2000. J.P. Wilson and J.C. Gallant, Editors. 2000. Terrain analysis: principles and applications, Wiley, New York.

Wilson, J. P., Lam, C. S., and Deng, Y. 2007. Comparison of the performance of flow-routing algorithms used in GIS-based hydrologic analysis, Hydrological Processes, 21:1026–144.

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

Zhang, J.X., J.Q. Wu, W.J. Elliot, S. Dun and K. Chang. 2006. Effects of DEM resolution on forest hydrologic and erosion prediction using WEPP. ASABE Pap. 062179. Am. Soc. of Agric. Eng., St. Joseph, MI.

Zhao Z, Benoy G, Chow TL, Rees HW, Daigle J, Meng F. 2010. Impacts of accuracy and resolution of conventional and LiDAR based DEMs on parameters used in hydrologic modeling. Water Resources Management 24: 1363–1380.

Chapter 2

Multi-scale lidar greatly improve characterization of forested headwater streams in central Pennsylvania

Abstract

Most current hydrographic data used in Geographic Information Systems (GIS) have been derived by digitizing blue line streams from USGS topographic maps or by modeling streams using traditional digital elevations models (DEMs) in GIS. Both methods produce stream models that lack detail and accuracy, particularly in headwater streams. In addition to channel network delineation, another hydrologic attribute that is of interest to hydrologists, modelers, and ecologists, is topographic index (TI) as measured by the formula $\ln(a/\tan\beta)$ (a = upslope contributing area per unit contour; $\tan\beta$ = local slope angle). This metric and its distribution is an important component to the hydrologic model TOPMODEL and other hydrologic models, but is also used extensively to represent soil moisture in fields of ecology, forestry, and soil science.

Newly available lidar data available statewide in Pennsylvania can produce DEMs with an accuracy and resolution that far exceed previously available elevation data. In this study, streams were modeled using lidar-derived DEMs of 1 m, 3 m, and 10 m resolutions using existing GIS software programs and compared to both actual streams and streams modeled using a 10 meter National Elevation Dataset (NED) DEM. Results showed that the most accurate stream locations could be modeled using a lidar-derived DEM thinned to 3m resolution or smoothed using a mean smoothing filter. Also, when a 10 m resolution lidar-derived DEM was compared to the NED 10 m resolution DEM, the streams delineated with the 10 m lidar data were significantly better than those modeled with the 10 m NED data, showing that significant improvement in

accuracy can be achieved with no increase in data storage. When topographic index was modeled with multiple resolutions of lidar-derived DEMs, the spatial and statistical distributions were both very different, with finer resolution DEMs not accurately modeling areas of high TI. Additionally, depending on the flow accumulation algorithm used, there were differences in the change in statistical resolution with response to initial DEM resolution.

Introduction

Over the past thirty years, research has been conducted in many fields including hydrology, forestry, soil science, and geomorphology incorporating Geographic Information Systems (GIS). Commonly available hydrographic datasets are often derived from blue line streams on USGS topographic maps. These maps were created at a 1:24,000 scale and have often been shown to be missing streams, particularly first order perennial and intermittent channels (Colsen 2006). The National Hydrologic Dataset (NHD) produced by USGS is derived from maps at the 1:100,000 and 1:24,000 scale and is thus not designed for analysis at finer scales. In recent years, stream maps are being created through processing of digital elevations models (DEMs) using tools such as ArcGIS and TauDEM (Tarboton et el. 1991, Beven and Moore 1993, Garbrecht and Martz 1993). Unfortunately, the same problems with resolution, scale, and accuracy that occur with vector hydrographic models also apply to the commonly available DEMs. The most commonly available elevation datasets such as the National Elevation Dataset (NED) have not been shown to be sufficiently accurate to generate maps of headwater streams (Colsen 2006, James et al. 2007, Murphy et al. 2008).

In the United States, headwater streams, including second order streams, first order streams, intermittent streams, and ephemeral streams make up at least 80 percent of the stream network and contribute many of the solutes (e.g. nitrogen, carbon) and sediment to a stream

system (Alexander et al. 2007, Meyer et al. 2007). Unfortunately, many of these un-named tributaries do not appear on any map and ones that do are often underrepresented by stream length (Colsen 2006). In the Ridge and Valley Province in Pennsylvania, many of these headwater streams occur in heavily forested ridges and are being threatened by development and energy production. Such headwater streams are providing valuable ecosystem services such as flood control, clean drinking water, groundwater recharge, removal of pollution and sediments from the stream system, and key habitats for plants and animals (Meyer et al. 2007, Peterson et al. 2001, Meyer and Wallace 2001). Headwater streams are also hotspots for biodiversity, ecosystem services, and denitrification (Meyer et al. 2007, McClain et al. 2003.) In order to protect these ecosystem services on the landscape, accurate spatial hydrographic data are needed in these headwater areas. Management objectives may require stream network delineations at multiple scales (Corwin et al. 2006), from identifying the point that flow begins to organize into ephemeral channels, to mapping of intermittent and perennial stream channels.

Leading Ridge Experimental Watersheds, the site of this study, is in central Pennsylvania and is the site of both a long term hydrological watershed research and is situated within the Susquehanna/Shale Hills Critical Zone Observatory. Critical zone science focuses on the study of the interface between the solid earth and the atmosphere, including interaction between catchments and their streams (Brantley et al. 2007, Anderson et al. 2008). Recent advances relating terrain and topography to hydrology require not only accurate stream channel maps but also accurate terrain models of the entire catchment (Beven 2006). These improved datasets are essential as scientists are striving to gain an understanding of flow paths, residence times, and distribution of water through a catchment. Hydrologic models of the critical zone are based on many geographic attributes of headwater streams and watersheds including stream length, drainage area, channel slope, and distribution of topographic index (Beven and Moore 1993, Hornberger and Boyer 1995, Qu 2004). Nitrogen transport models also rely on accurate stream

network and watershed characteristic data because factors such as reach slope, time of travel, riparian zone characteristics, and watershed area can have significant impacts on model results (Cirmo and McDonnell 1996, Boyer et al. 2006).

A new source of data that may improve stream network delineation, lidar (light detection and ranging) elevation data, are more frequently becoming available for local and regional studies. Pennsylvania, through the PAMAP program, which aims to create a digital basemap for Pennsylvania, recently acquired lidar data and compiled a lidar-derived DEM for the entire Commonwealth of Pennsylvania. This DEM has a resolution of 3.2 ft. (about 1 m) and a targeted vertical RMSE of less than 18.5 cm for some areas (DCNR PAMAP program, http://www.dcnr.state.pa.us/topogeo/pamap/lidar.aspx). This is not only a significant improvement to resolution, but also in accuracy. Previously available DEMs typically have vertical accuracies from about 2 m (NED) to 10 m (SRTM) (Maune 2007). Due to the improvement in accuracy over NED and SRTM elevation datasets, our assumptions about resolution and hydrologic modeling may need to be re-evaluated, and it should not be assumed that the highest resolution dataset will perform the best for all purposes.

In addition to channel network delineation, another hydrologic attribute that is of interest to hydrologists, modelers, and ecologists, is topographic index (TI) as measured by the formula $\ln(a/\tan\beta)$ (a = upslope contributing area per unit contour; $\tan\beta$ = local slope angle). This metric and its distribution is an important component to the hydrologic model TOPMODEL (Beven and Kirkby 1979) and other hydrologic models, but is also used extensively in ecology (Sørensen et al. 2006), soil science (Blyth et al. 2004, Guntner et al. 2004), and forestry (Bader and Ruijten 2008). The distribution of TI has been shown to be dependent on DEM resolution and grid size (Zhang and Montgomery 2004, Bruneau et al 1995, Kienzle 2004), but only limited work has been done to quantify the difference in spatial pattern in TI and its distribution caused by both improved resolution and accuracy from lidar-derived data (Straumann and Purves 2007).

In this study, we started with a 1 m lidar-derived DEM and generalized the data to create coarser-resolution data models to test the effect of resolution on accurate delineation of stream networks. Stream networks were modeled using GIS-based flow routing algorithms. Several previous studies have completed network modeling using lidar data (Colson 2006, Murphy et al. 2008) and others have studied effects of resolution on metrics including TI (Sørensen et al. 2006, Tenenbaum et al. 2006), but few studies have examined the difference between lidar-derived and photogrammetrically derived elevation datasets at similar resolutions. Although the available accuracy of lidar-derived DEMs far exceeds traditional DEMs, this new detail can also create complications. Because of the extremely high resolution and detail, roads, bridges, and ditches can appear in the DEM, causing inaccurate flow routing. By conducting network delineation at multiple resolutions and smoothing levels of the lidar-derived DEMs, it may be found that the most accurate models of stream channels are not produced by the highest resolution and most accurate original DEM. This may be because the smoothing and generalization removes some of the erroneous detail that can cause errors. Using a coarser resolution of lidar can also allow us to work with superior lidar-derived data without increasing our storage and processing demands.

From this work, we make recommendations on a methodology of using multiple scales of lidar for stream network delineation depending on the objectives. We found that 1 m, 3 m, and 10 m lidar-derived DEMs all led to a significant improvement of stream network delineation over 10 m NED. Although there was no difference in computational efficiency between the thinned 10 m lidar dataset and the NED, the lidar dataset represents the terrain much more accurately. Also, there were benefits to not using the highest resolution data in our analysis due to error caused by features such as roads, culverts, and some vegetation. From our data, we also explore the implications of using lidar data to improve our understanding of the hydrological response of a watershed through the eyes of a standard topographic index, widely used to consider how a watershed responds to incoming precipitation in terms of wetness and propensity for saturation.

Methods

Study Area

The study area, Leading Ridge Watershed One (Figure **2-1**), is a 119 ha watershed located in the Ridge and Valley Province of Pennsylvania. The elevation of Leading Ridge Watershed One ranges from 260 m at the mouth of the watershed to 512 m at the top of Leading Ridge, which forms the north-western border of the watershed. Since this watershed is in the Ridge and Valley Province, its hydrologic network tends to form a trellis pattern instead of the more common dendritic pattern. The geologic formation underlying the watersheds consists of deeply dipping strata ranging from resistant Tuscarora quartzite and sandstone at the top of the watershed to less resistant Rose Hill shale that comprises the valley area (Shultz 1999). This terrain is consistent with that which makes up most of the Ridge and Valley Province, which is generally characterized by canoe-shaped valleys and long, linear, parallel ridges formed by differential erosion. Valleys tend to be underlain by limestone, dolomite, and shale, while ridges are formed by resistant sandstone and quartzite. Ridges tend to be extremely steep and rocky and the topography is generally well-drained (Shultz 1999).

The watersheds contain a mature oak/hickory forest approximately 100 years of age. This area was used extensively for charcoal production prior to 1900, and charcoal pits can still be found scattered throughout the site (Robinson 1959). The southern section of the watershed was probably grazed until the 1930s, when farmers were paid to abandon their land. A small salvage cut occurred on the site in 2006 following a blowdown event, and the watersheds experienced significant gypsy moth mortality over the past 20 years (pers. com. Joseph Harding 2009).

This site was chosen in part because of the history of the area being used as an experimental watershed and the vast hydrologic dataset that currently exists for the watershed.

The Leading Ridge Watershed Research Unit was established in Penn State's Stone Valley Experimental Forest of central Pennsylvania in 1959 as paired watersheds to study the hydrologic response of different forest practices. There are three watersheds that make up the site of 43, 104, and 123 ha in size. Each watershed received a different vegetation treatment between the years of 1967-1977. These treatments ranged from no disturbance to a clear-cut and herbicide application applied over three years to completely remove vegetation (Lynch and Corbett 1990). In addition to these treatments, Leading Ridge has a rich history as a living laboratory. In addition to historic hydrologic records covering 50 years of climatic variability, there is also a myriad of soils, climate, and water chemistry data for the watersheds (Lehman 1962, Hornbeck 1962). Watershed One, the site of this study, served as the control in these studies and did not receive any forest treatments.

In addition to being used as the Leading Ridge Watershed Research Unit, the watersheds are within the Shaver's Creek watershed which is being used as part of the Susquehanna/ Shale Hills Critical Zone Observatory (CZO), an interdisciplinary observatory aiming to quantitatively predict creation, evolution, and structure of regolith as a function of the geochemical, hydrologic, biologic, and geomorphologic processes operating in a temperate, forested landscape (Anderson et al. 2008). The Susquehanna/Shale Hills CZO is one of six watershed-scale observatories supported by the National CZO project.
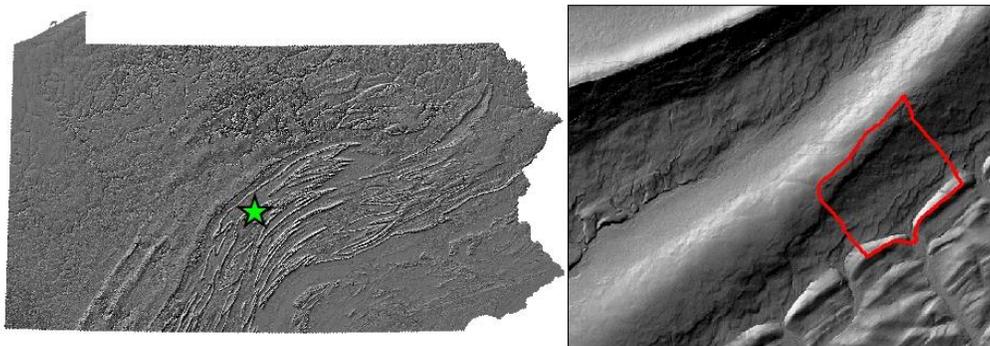
Figure **2-1**: Location of the Leading Ridge watersheds in the Ridge and Valley province of Pennsylvania.

Field Mapping of Stream Network

As a necessary first step, we needed to address the question of what the "true" stream network was in this site. By working on the ground to physically map the stream channels, we assumed for our study that this is equal to truth.

In order to measure the location of the field stream network, stream channels were mapped with a sub-meter accuracy GPS unit. Starting at the gauging station on Leading Ridge Watershed One, a reconnaissance was conducted on the stream network using a Trimble GeoXT GPS unit with differential correction during a week in January of 2009 when the temperatures did not rise above freezing to reduce the impact of snowmelt on streamflow. A total of 47 points were collected that marked the location of every visible channel convergence or divergence. All points had water flowing over them at the time of the survey. This survey continued, starting at the weir and working upstream until all channels in the watershed were marked and all points of stream convergence or divergence were recorded. Additionally, the locations of all springs were recorded. Points were post-processed using Trimble Pathfinder software to provide sub-meter accuracy.

DEM Data Sources

The heart of the study is delineating stream networks from DEMs and determining which DEM models streams most accurately. We compared ground truth streams (above) to streams delineated from DEMs: lidar- based (at multiple resolutions) and to older DEM products based on photogrammetry conducted at a scale of 1:24,000 that are today's standardly available national dataset (NED).

Lidar data were collected as part of the PAMAP lidar program in 2007 during leaf-off conditions in order to generate the most accurate bare earth model. Post spacing for the lidar return used for the DEM generation was 1.4 meters, with a vertical RMSE of 18.5 cm in open areas and 37 cm in vegetated or forested areas. This was one of the first and most accurate state-wide lidar datasets collected. As part of the project, a 3.2 ft. (approximately 1 m) resolution DEM was produced, along with 2 ft. (0.61 m) contour lines and breaklines by BAE systems. All finished products were checked for quality and accuracy (PAMAP LiDAR QAQC report 2007). For this study, the independently produced 3.2 foot resolution DEM was used as the original 1m DEM. All of these data are available publicly at the Pennsylvania Spatial Data Access (PASDA, www.pasda.psu.edu).

The lidar-generated DEM was compared to a photogrammetrically derived National Elevation Dataset (NED) 1/3 arcsecond DEM which has a horizontal resolution of approximately 10 m (Gesch 2007).

DEM Processing and Network Modeling

In order to assess the effect of multiple resolutions, accuracies, and smoothing techniques, the original lidar-derived DEM was manipulated in several ways prior to any GIS hydrologic modeling. First, the DEM was resampled using the nearest neighbor method to 3 m and 10 m resolutions. These resolutions were chosen to sample in a logarithmic manner that approximates a range of scales from the finest available resolution to traditionally available resolutions. The nearest neighbor method was used to ensure that no interpolation was taking place, but data were simply being removed from the dataset. Additionally, to test the impact of smoothing, a 3 x 3 moving window mean smoothing filter from ArcGIS spatial analyst tools was applied to the original lidar-derived DEM twice to remove localized highs and low. All

derivations of the DEMs were then filled to remove inward-draining depressions using the ArcGIS fill tool in spatial analyst. Flow direction was calculated using the flow direction tool in ArcGIS which computes flow direction using the D8 method (O'Callahan and Mark 1984), followed by flow accumulation. Flow accumulation was also calculated using the D∞ method (Tarboton 1997) in TauDEM for comparison purposes and to calculated the topographic index.

Many different methods have been proposed for modeling flow accumulation and contributing area. The simplest algorithm, D8 (O'Callaghan and Mark 1984), has been shown to model streams into straight, parallel channels that often do not accurately reflect natural stream channels. Others methods such as Rho8 (Moore et al. 1993) and D∞ (Tarboton 1997) have attempted to correct this problem by allowing flow to be directed into multiple cells. Although these methods may represent more accurate hillslope flow patterns, channels modeled using these methods tend to show dispersed flow in valleys which can be inaccurate. These algorithms are also more complicated and require greater computing power. With the improvements in elevation data provided by lidar data, we may be able to effectively use simpler algorithms, particularly in light of the increased storage and processing requirements of lidar data. With accurate elevation data, simpler algorithms such as D8 may provide an adequate model of flow accumulation, particularly when the goal is network delineation.

Stream networks were delineated by using ArcGIS Spatial Analyst raster calculator to identify cells in the D8 flow accumulation layer that had a minimum threshold of cells draining into them, which defines the drainage area of what is classified as a stream. The threshold was set at 2.5 ha, which is relatively low in order to identify all ephemeral and intermittent channels. Identifying the appropriate threshold for stream delineation is dependent on the use and purpose of the stream data and thus was outside the scope of this study. The same procedures were also performed on the NED 1/3 arcsecond DEM. Stream networks were assessed by measuring the shortest distance from each GPS point on the channel to a modeled stream channel. These

distances were analyzed using a Wilcoxon rank sum test, a non-parametric hypothesis test, to test

if there was a significant difference between the means of measured variables.

Watersheds were delineated for each of the Leading Ridge Watersheds using the weir

locations as pour points in ArcGIS and using the watershed function to identify all cells flowing

through each pour point. Watersheds were computed using all five derivations of the DEM in

order to compare results of different resolutions and accuracies of DEMs on watershed

delineation.

Topographic Index

We explored implications of improved stream networks versus poorly defined networks

on hydrological response of the watershed through the eyes of a standard topographic index,

widely used to consider how a watershed responds to incoming precipitation in terms of wetness

and propensity for saturation. For each DEM dataset, topographic index (TI) was calculated using

the formula $\ln(a/\tan\beta)$ (a = upslope contributing area per unit contour; $\tan\beta$ = local slope angle).

Contributing area was calculated using both the TauDEM D8 method (Tarboten et al. 1991) and

the TauDEM D∞ method (Tarboton1997). With the D8 method, flow accumulation was

converted to contributing area by multiplying by the cell area and dividing by cell size

(resolution). The D∞ flow accumulation method automatically calculates contributing area. For

each dataset, statistical and spatial distributions of TI were modeled and descriptive statistics

were analyzed using R (R Development Core Team, 2010). TI values were calculated that

represent the top 5% and 10%, of cells in each DEM. Percent of cells over TI values of 6 and 8

were also calculated for each DEM. Additionally, spatial distributions of TI were examined to

explore how differences in true resolution and data sources impact spatial distribution of TI.

Results

Channel Networks

All the channel networks modeled using the lidar-derived DEMs were significantly more accurate than channels modeled using the NED DEM (p =0.001). This is probably due to the ability of lidar-derived DEMs to more accurately represent small channel features on the landscape. The modeled channel networks with the smallest mean and median distance from GPS points to the channel location were from the smoothed 1 m resolution DEM and from the 3 m resolution DEM (Table **2-1**). Median distances for lidar-derived DEMs ranged from 2.35 m (3 m lidar) to 3.26 m (1 m lidar). Median distance for the NED dataset was 16.43 m. Interestingly, all of the lidar-derived DEMs that had been smoothed or generalized produced slightly more accurate channel models than the original DEM, although not significantly so. Results were analyzed using a Wilcoxon rank sum test due to the non-normal distributions of the data. There was no significant difference between the lidar-derived products, but all lidar-derived products, including the DEM thinned to 10 m performed significantly better than the NED with p values less than 0.001.

Table **2-1:** Descriptive statistics for distances between GPS locations on the channel and closest modeled channel, measured in meters. ** indicates values significantly different from the NED dataset with a p value of less than 0.001.

DEM	Mean Distance	Median Distance	SD
1 m lidar	12.4**	3.3	19.5
Smoothed 1 m lidar	8.9**	2.7	15.4
3 m lidar	8.5**	2.4	16.3
10 m lidar	12.2**	3.1	22.2
10 m NED	33.3	16.4	39.8

Boxplots of distances between each GPS point and the closest modeled channel for all five DEM derivatives are shown in Figure **2-2**. For the lidar-derived stream network models, outliers are the same stream locations, suggesting that certain stream networks were not modeled

accurately using any of the lidar-derived DEMs. Locations of modeled channels and GPS points

on those channels are shown in Figure **2-3** for each DEM derivative. Points are graduated by

color and size based on the number of standard deviations they occur from the mean. In Figure **2-3**, you can see that the same GPS'd channel locations are the three furthest outliers for all four of

the lidar-derived modeled channel networks. This is not true for the channels modeled using the

NED DEM which represents different points as outliers in the box plot.
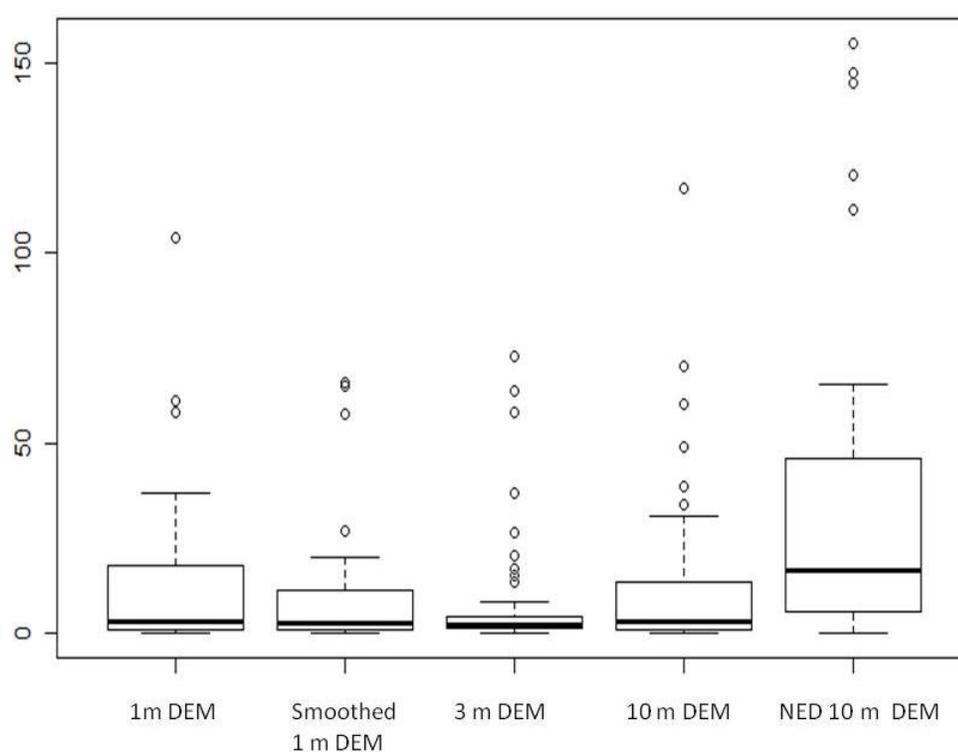
Figure **2-2:** Boxplot for distances between GPS locations on the channel and nearest modeled channel.

(a)
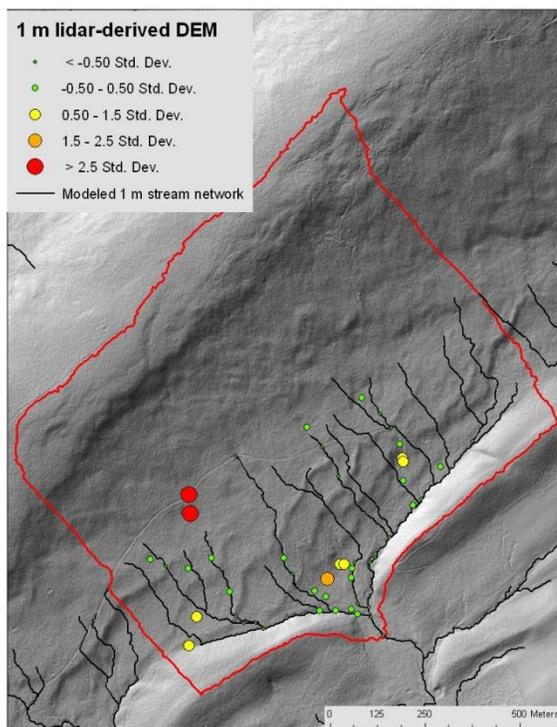
(b)
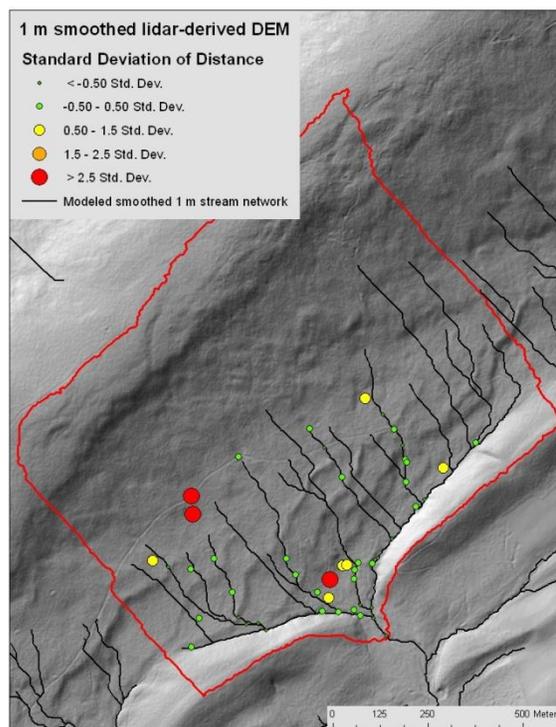
(c)
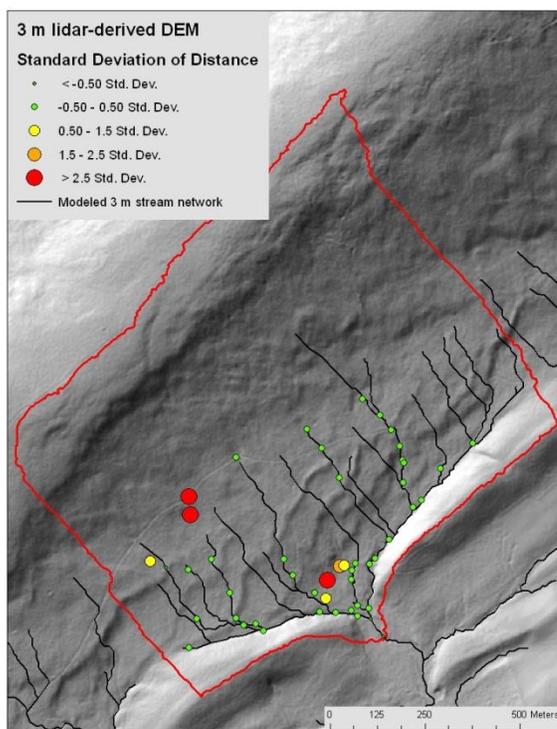
(d)
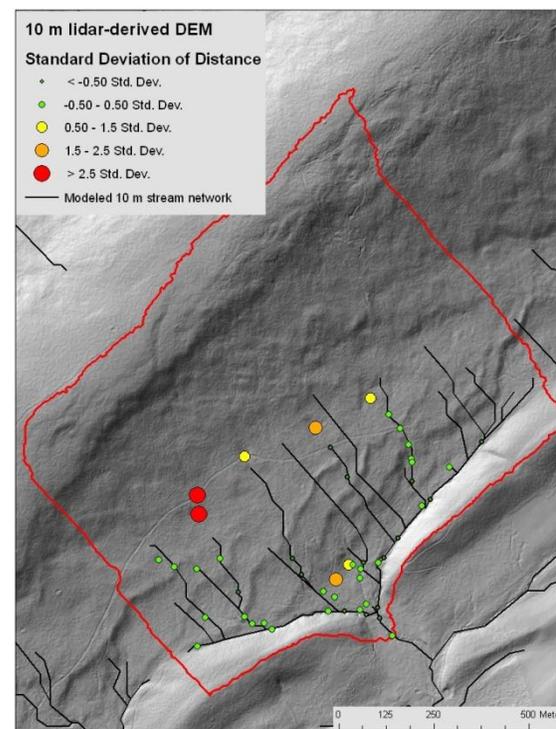
(e)
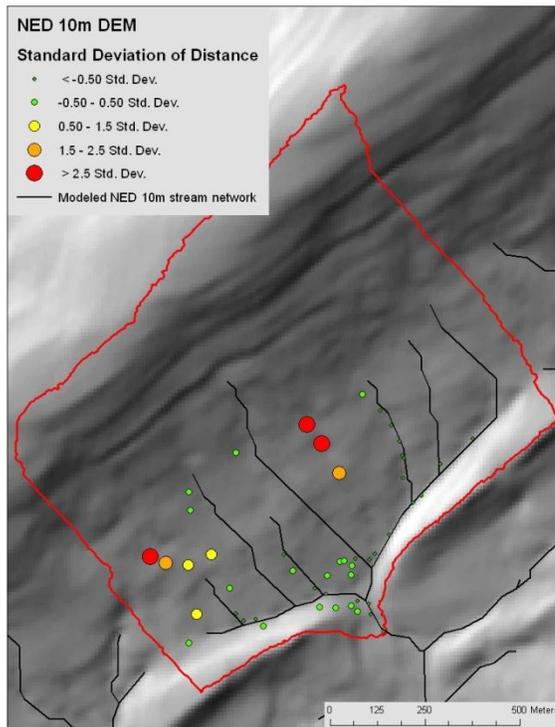
Figure **2-3:** GPS points colored by distance from closest modeled stream channel for (a) original 1 m lidar DEM, (b) 1 m lidar DEM smoothed, (c) 3 m lidar DEM, (d) 10 m lidar DEM, and (e) 10 m NED DEM. Points are graduated by number of standard deviations they fall from the mean.

Topographic Index

Topographic Index (TI) was calculated for the lidar-derived 1 m, 1 m smoothed, 3 m, and

10 m DEMs, as well the 10 m NED DEM using the D8 flow accumulation method. For the D8

modeled TIs, statistical distributions of TI were modeled and descriptive statistics were analyzed

using ArcGIS and R (Table **2-2**, Figure **2-4**). In addition to exploring the statistical distributions,

values of the top 5% and 10% of TI were calculated and mapped (Figures **2-6**, **2-7**) to represent

how the watershed would wet under different intensity hydrologic events. Additionally, the

percent of the watershed containing TI values greater than 6 and greater than 8 are shown in

Table **2-2**.

The DEMs with finer resolutions tend to model higher TI values, and a greater percentage of their area is represented by high TI values. In general, the maximum value of TI values increased as the resolution decreased, ranging from 15.45 with the NED DEM to 21.46 with the smoothed 1 m DEM. DEMs with finer resolutions also had a greater percentage of cells that have relatively high values. For example, the smoothed 1 m DEM had almost 20% of its cells with a TI value greater than 6, while the NED DEM had only 7.4% of its cells with a greater TI value than 6. Values for percent of area with TI values greater than 8 ranged from 5.71% using the 1 m lidar to 2.4% using the NED. This can have implications for hydrologists modeling the percentage of cells that would be wet in certain types of hydrologic events. The statistical distributions were also different with the NED DEM distribution having a much higher kurtosis value than the other distributions.

Table 2-2: Minimum, maximum, mean, maximum 5%, maximum 10%, percent of values over a TI of 6, and percent of values over a TI of 8, along with descriptors of the distribution of each TI calculated using the D8 method.

DEM	Min	Max	Mean	Max 5%	Max 10%	% over TI 6	% over TI 8	SD	Skew	Kurtosis
1 m lidar	-0.60	20.95	4.13	8.16	6.58	15.10%	5.71%	2.03	1.59	7.52
smoothed 1 m lidar	0.00	21.46	4.69	8.06	6.73	19.73%	5.42%	1.81	1.35	7.23
3 m lidar	-0.40	18.69	3.92	7.83	6.70	17.42%	5.09%	2.07	1.36	5.84
10 m lidar	0.33	16.48	3.83	7.40	6.26	13.04%	3.29%	1.85	1.19	5.45
10 m NED	0.63	15.45	4.23	6.58	5.79	7.40%	2.4%	1.48	1.43	8.61
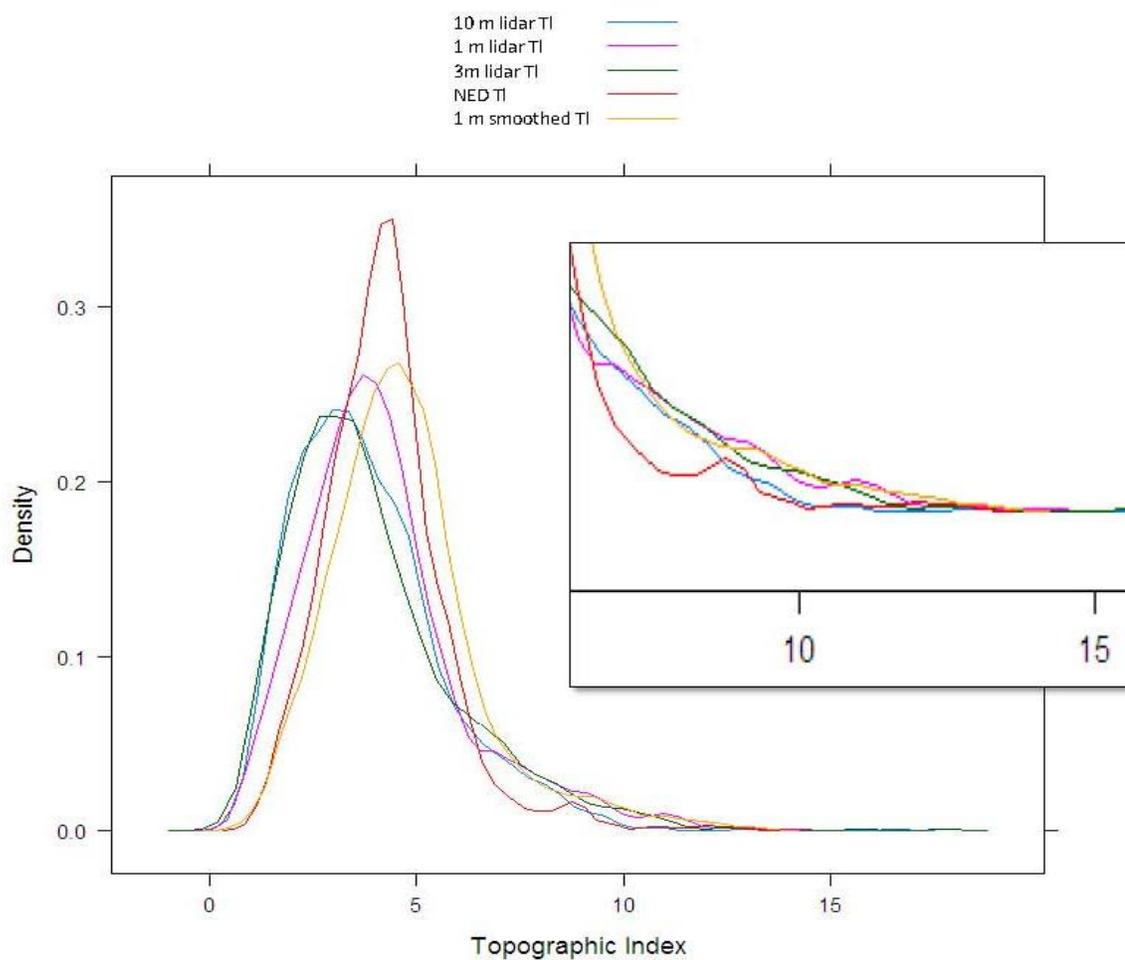
Figure **2-5:** Distributions of Topographic Index for the original 1 m lidar-derived DEM, 1 m smoothed DEM, 3 m lidar-derived DEM, 10 m lidar-derived DEM, and the NED DEM. Also included is an inset of the tail of the distribution.

(a)
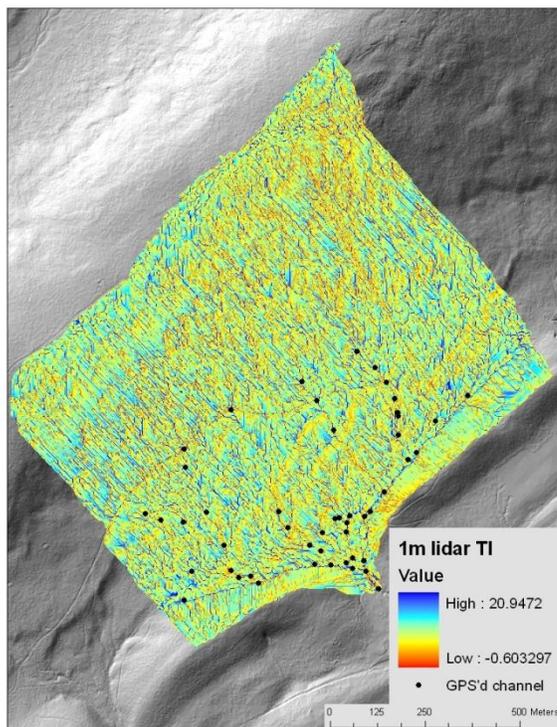
(b)
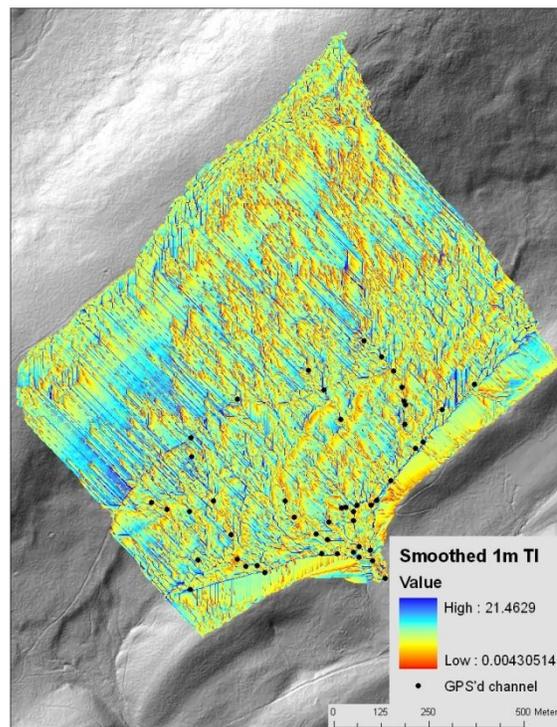
(c)
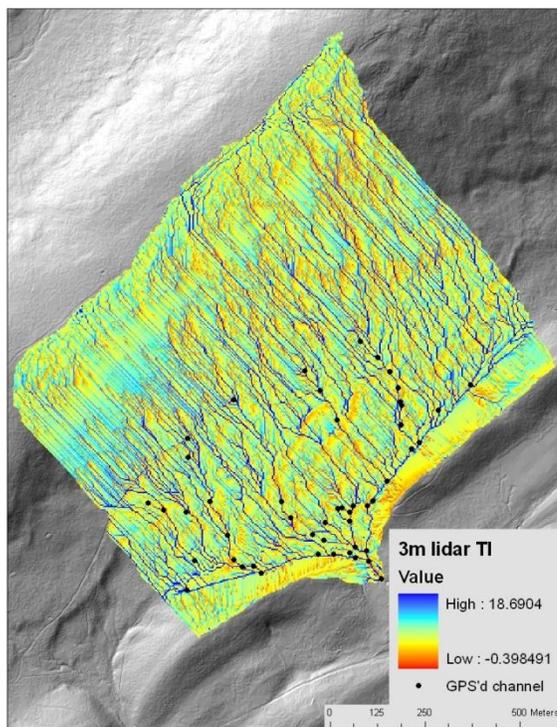
(d)
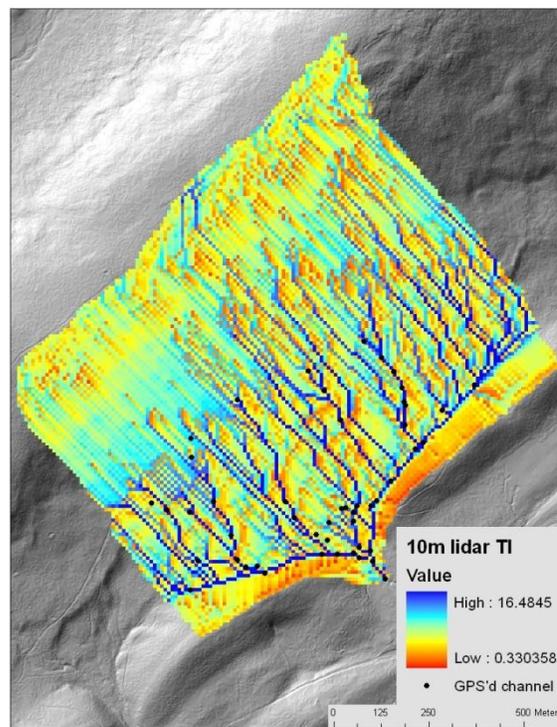
(e)
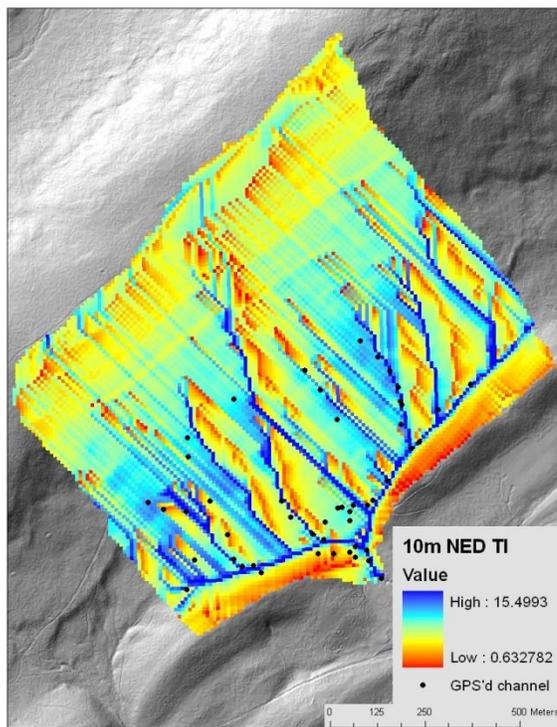
Figure **2-5:** Topographic Index values using the D8 method of flow direction for (a) original 1 m lidar DEM, (b) 1 m lidar DEM smoothed, (c) 3 m lidar DEM, (d) 10 m lidar DEM, and (e) 10 m NED DEM. Note the differences between the 10 m NED and the 10 m lidar-derived values.

(a)   *(b)*

(c)

(d)

(e)
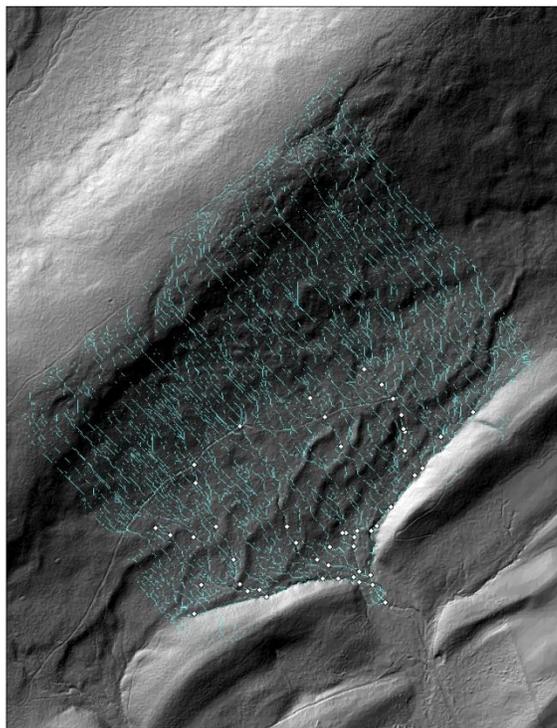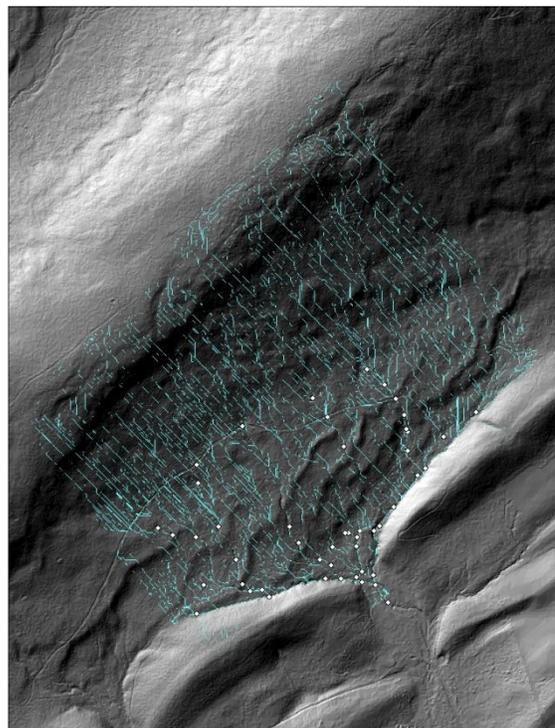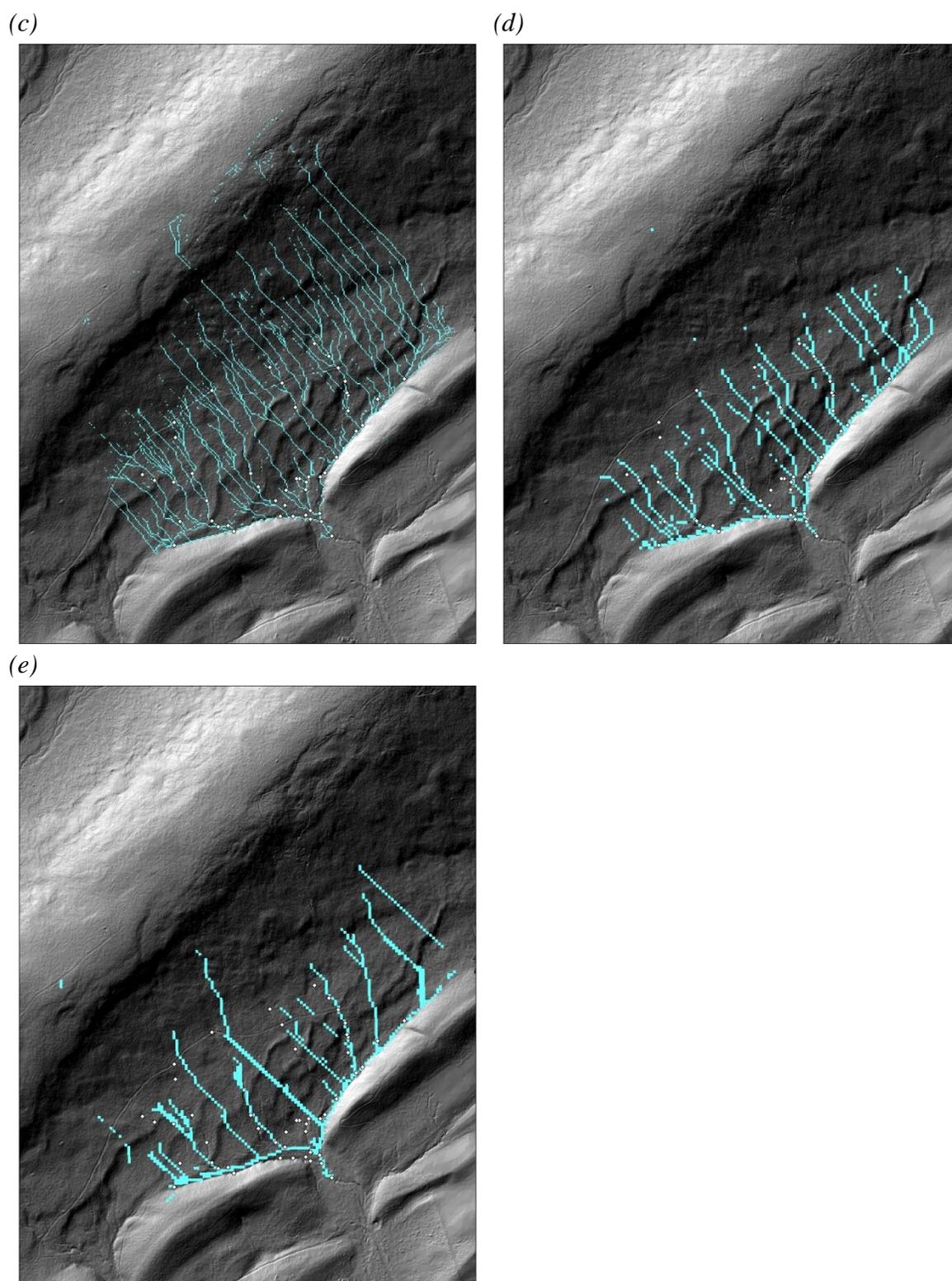
Figure **2-6:** The highest 5% of topographic index values using the D8 method of flow accululation for (a) original 1 m lidar DEM, (b) 1 m lidar DEM smoothed, (c) 3 m lidar DEM, (d) 10 m lidar DEM, and (e) 10 m NED DEM.

(a)
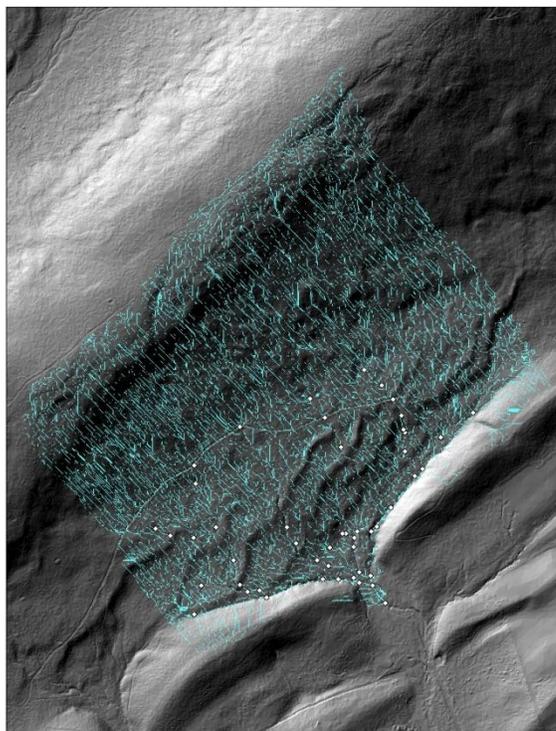
(b)
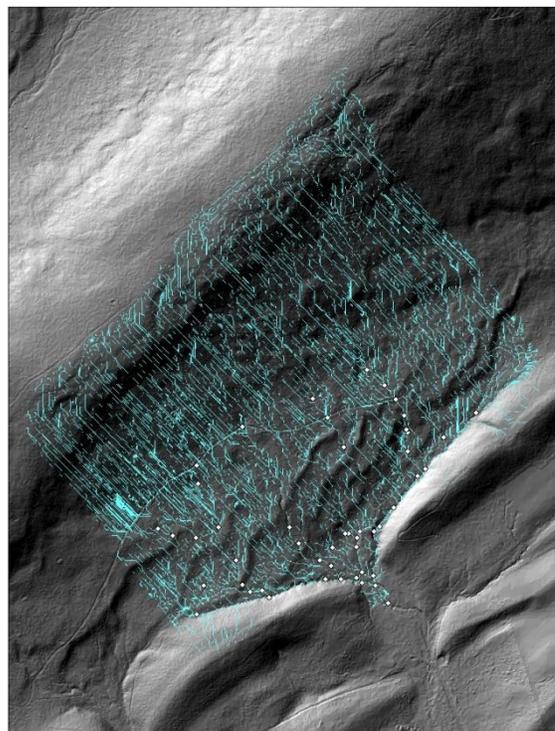
(c)
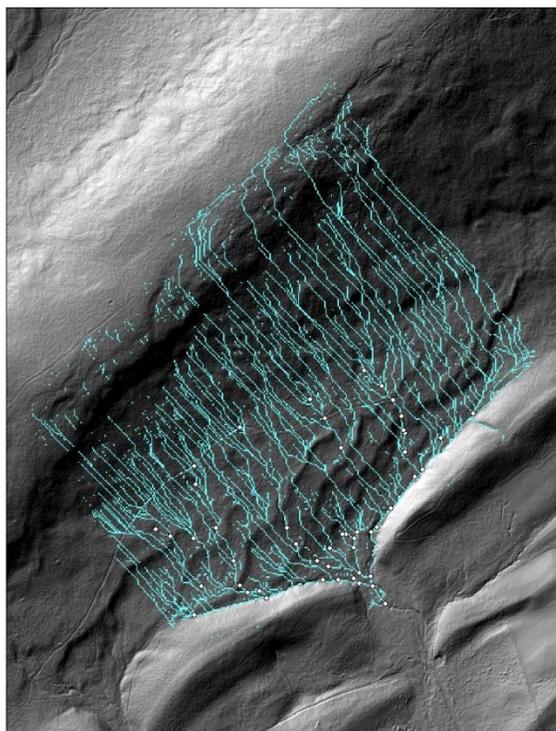
(d)
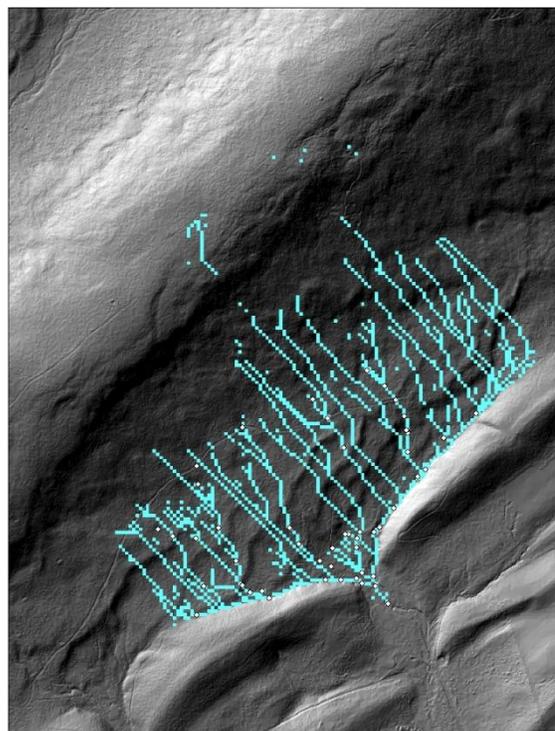
(e)
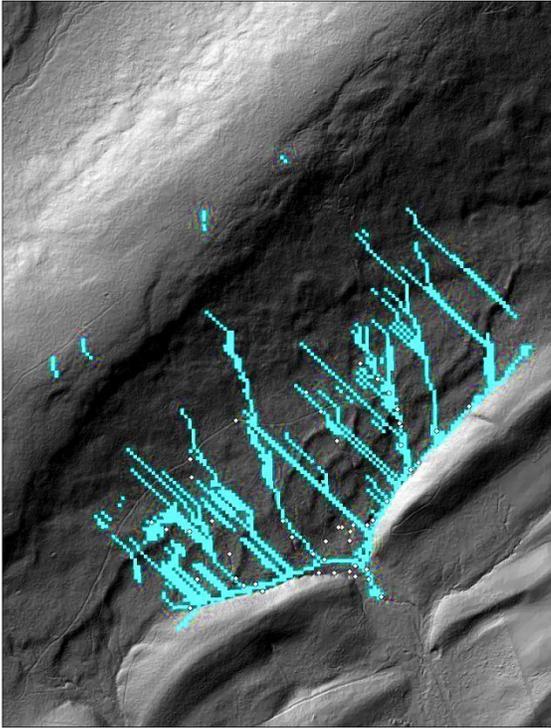
Figure **2-7:** The highest 10% of topographic index values using the D8 method of flow accumulation for (a) original 1 m lidar DEM, (b) 1 m lidar DEM smoothed, (c) 3 m lidar DEM, (d) 10 m lidar DEM, and (e) 10 m NED DEM.

Several patterns emerge when examining spatial patterns of TI (Figure **2-5**). Finer DEMs model parallel, repeating patterns of high and low TI values perpendicular to the hillslope, while coarser resolution DEMs tend to have fewer areas showing a high TI and more space in between these features. All resolutions show more areas of high TI than the number of stream channels validated with GPS on the landscape, although these areas of modeled high TI values were not field checked to determine if they were in fact wet during high volume hydrologic events. Finer resolution DEMs (1 m, 3 m) tend to model high values of TI at higher elevations in the watershed, while coarser resolutions (10 m) have highest values stopping around the mid-point of the watershed. This is noteworthy because there are no drainage features past the mid-point of the

watershed. Therefore, the finer resolution TI maps do not reflect the patterns of stream networks and surface flow on the landscape.

In order to quantify the distribution of wet cells on the watershed, a line was drawn across the watershed that connects the springs that initiate the stream network. Above this line in the watershed, no channel features or wet areas were found. For each DEM, the percentage of wet cells measured by either the top 10% or 5% of TI values that are found in the upper portion of the watershed are shown in Table **2-3**. It is clearly shown that the higher resolution DEMs contain a majority of their wet cells at both the top 5% and 10% threshold in the upper portion of the watershed. As the resolution decreases, the number of cells being represented as wet in the top of the watershed decrease. These patterns suggest that TI may be problematic when used with finer resolution data such as lidar and that patterns of high TI on the watershed may not be accurate.

Table **2-3:** Percentage of wet cells found in the upper portion of the watershed by DEM derivative.

Percent of wet cells in the upper portion of the watershed	1 m DEM	1 m smoothed DEM	3 m DEM	10 m DEM	NED DEM
Top 10% of TI Values	57.2%	58.8%	52.7%	24.9%	19.3%
Top 5% of TI Values	55.9%	56.1%	36.9%	5.6%	13.0%

TI was also calculated using D∞ flow contributing area algorithm for a 1 m lidar-derived DEM, a 3 m lidar-derived DEM, and a 10 m lidar-derived DEM for comparison purposes (Table **2-4**, Figure **2-8**, Figure **2-9**, Figure **2-10**). Overall, values were higher, with no values being identified as negative. This seems to be caused by higher values in the D∞ contributing area grid in comparison to the D8 contributing area grid. The pattern of higher D∞ TI values with relationship to resolution is reversed in comparison to the D8 method TI values, with finer resolution DEMs showing lower mean D∞ TI values (Table **2-4**). Also, the values representing the top 5% and top 10% of values are higher with the larger resolution datasets. When viewing the spatial patterns of TI on the watershed, patterns are similar to those shown by the D8 TI, with the finer resolutions showing high values across the entire watershed including the top portion,

and coarser resolutions representing more of the higher TI values in the bottom portion of the

watershed where stream networks are found (Figure **2-9**, Figure **2-10**). The patterns formed by

D∞ TI appear slightly more realistic than that delineated with the D8 method, with fewer straight,

parallel areas of alternating TI values and slightly broader areas of wet and dry regions,

particularly in the larger resolution datasets.

Table **2-4**: Minimum, maximum, mean, maximum 5%, and maximum 10% of TI values for TIs calculated using the D∞ method.

DEM	Min	Max	Mean	Max 5%	Max 10 %
1 m lidar (D∞)	1.16	24.43	6.61	10.65	9.65
3 m lidar (D∞)	2.85	23.30	7.34	10.95	10.07
10 m lidar (D∞)	4.02	22.21	7.87	11.15	10.30
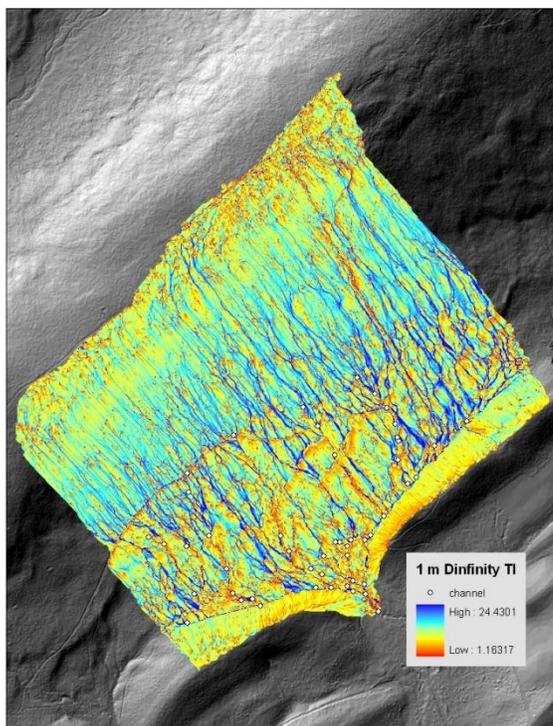
(a)   *(b)*
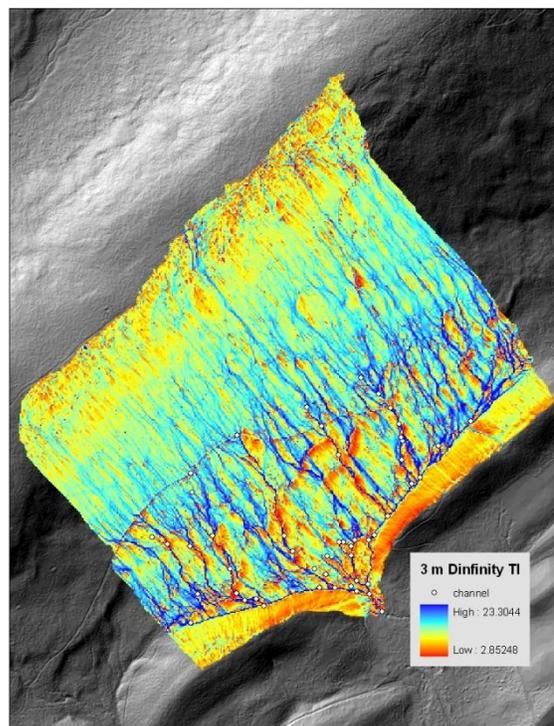
(c)
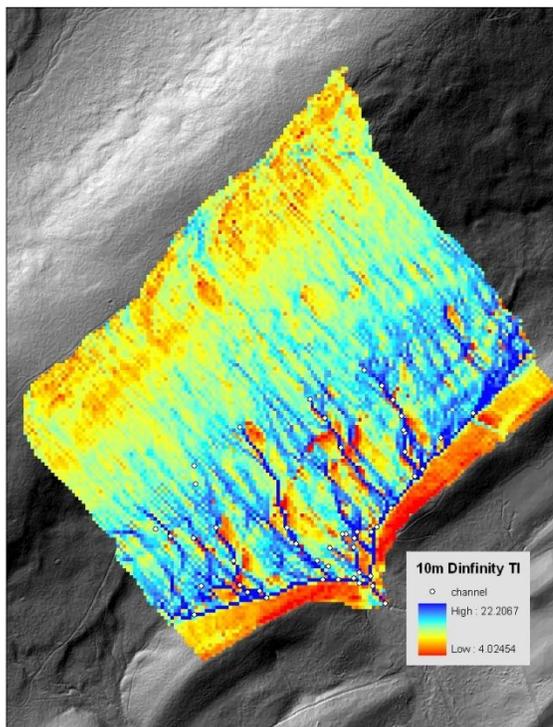
Figure **2-8**: Topographic Index values using the D∞ method of flow accumulation for (a) original 1 m lidar DEM, (b) 3 m lidar DEM, and (c) 10 m lidar DEM.

(a)
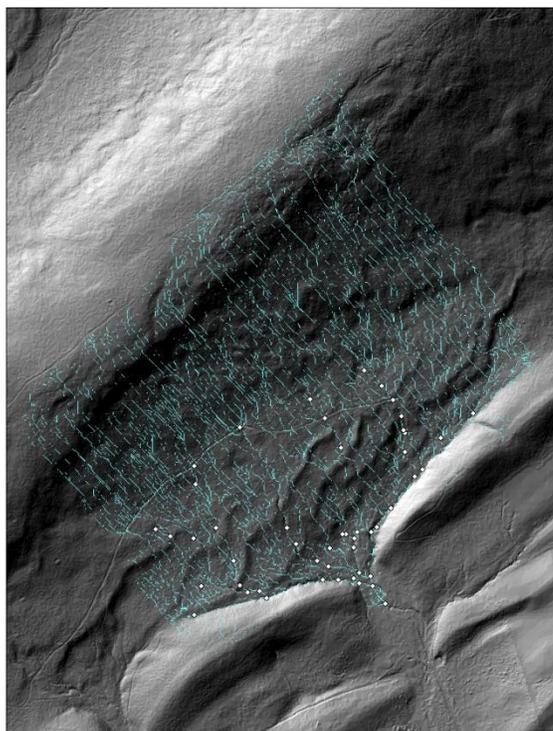
(b)
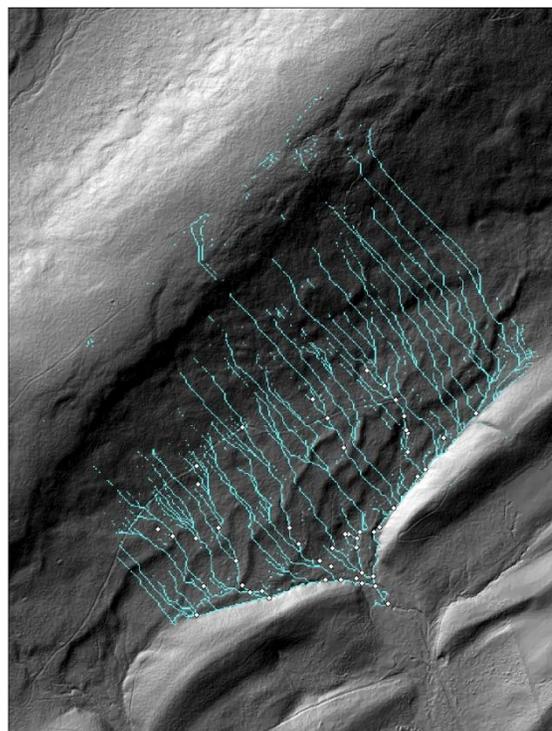
(c)
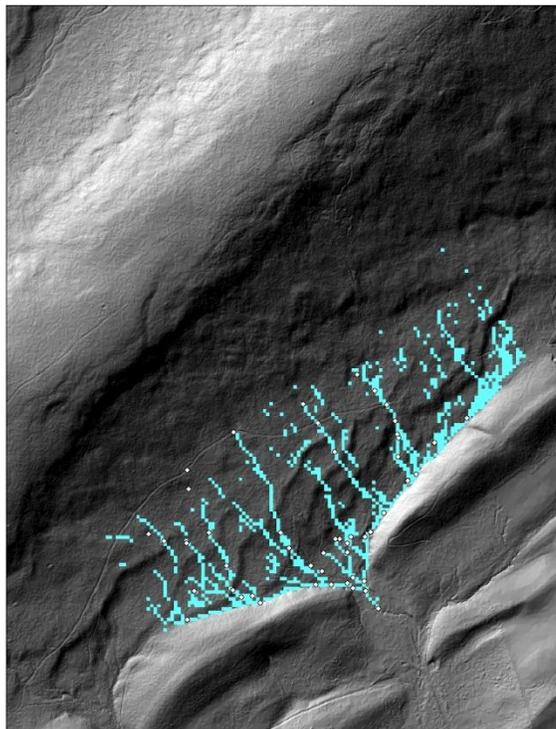
Figure **2-9**: The highest 5% of topographic index values calculated using the D∞ method of flow accumulation for (a) original 1 m lidar DEM, (b) 3 m lidar DEM, and (c) 10 m lidar DEM.
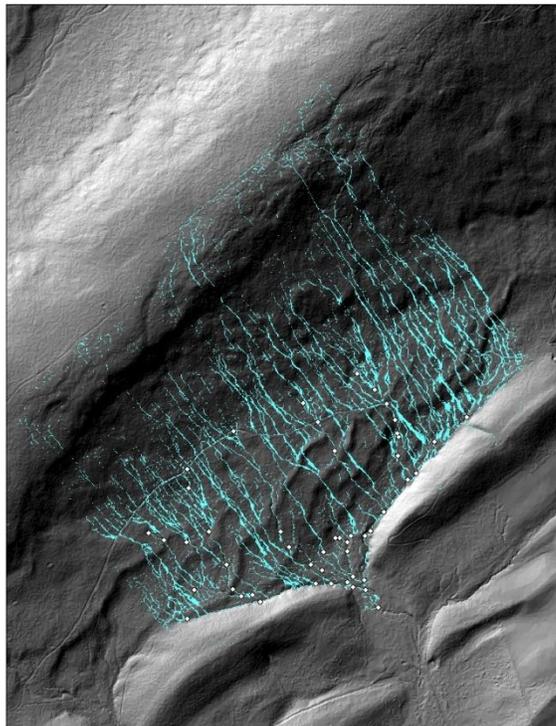
(a)   *(b)*
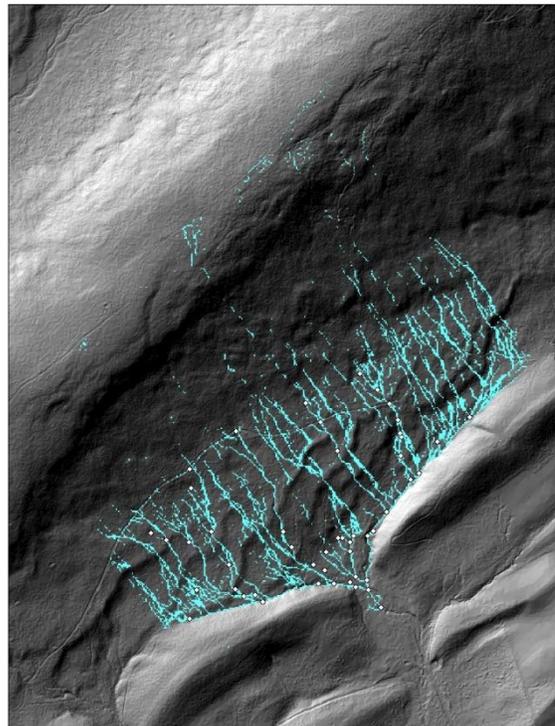
(c)
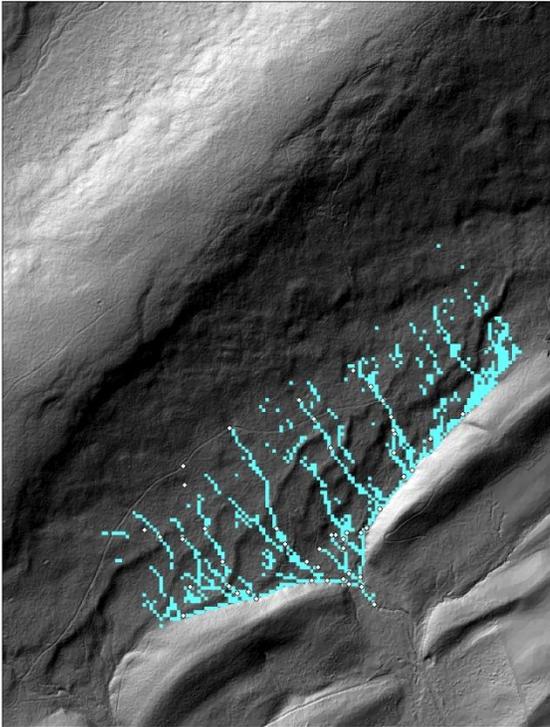
Figure **2-10**: The highest 5% of topographic index values calculated using the D∞ method of flow accumulation for (a) original 1 m lidar DEM, (b) 3 m lidar DEM, and (c) 10 m lidar DEM.

Watershed Delineations

Watersheds were delineated with the ArcGIS watershed tool using the weirs in Watersheds 1, 2, and 3 as the pour points (Table **2-5**, Figure **2-11**). Watershed areas were slightly different as measured with the different DEMs, although there was not bias with one DEM producing consistently larger or smaller watersheds (Table **2-5**). The Leading Ridge Watershed One, which showed the most variability, ranged from 121.7 hectare (ha) using the 3 m lidar-derived DEM to 113.5 ha using the NED DEM. The other two watersheds were more similar in size as modeled by all DEMs.

By examining the delineations (Figure **2-11**), you can see that watershed boundaries were slightly shifted on the hillside depending on which DEM was used for their delineation. There

was more similarity among delineations done with different resolutions of lidar than between the

10 m lidar and 10 m NED datasets. Watershed one delineated using the NED dataset is between

5.5 and 7 ha smaller than the watershed delineated with the lidar datasets. On the boundary

between watershed one and watershed two, you can see the range of boundaries, with the NED

dataset modeling the watershed boundary farthest to the east, and the 3 m lidar dataset modeling

the boundary farthest to the west. The differences among the lidar datasets is probably caused in

part by the effect of the road across the watershed on the flow modeling algorithms used to

delineate the watershed. The watershed boundary modeled with the 1 m DEM shows the

boundary further east than the other lidar datasets because flow is modeled to the west down

along the side of the road before it crosses the road. It is hard to know if this is in fact occurring

in the field, but it could be because there is no drainage feature such as a culvert crossing the road

in this area. On the other watersheds, the boundaries delineated with the NED dataset tend to be

consistently west of the lidar-delineated boundaries.

Table 2-5: Areas of Leading Ridge watersheds in hectares (ha) delineated using different DEM products.

DEM	Leading Ridge 1 (ha)	Leading Ridge 2 (ha)	Leading Ridge 3 (ha)
Lidar-derived 1 m	119	47.7	107
Smoothed 1 m DEM	120.1	44.5	106.2
Lidar-derived 3 m	121.7	44.1	106.7
Lidar-derived 10 m	120.6	44.8	107.1
NED 10m	113.5	49.5	108.7
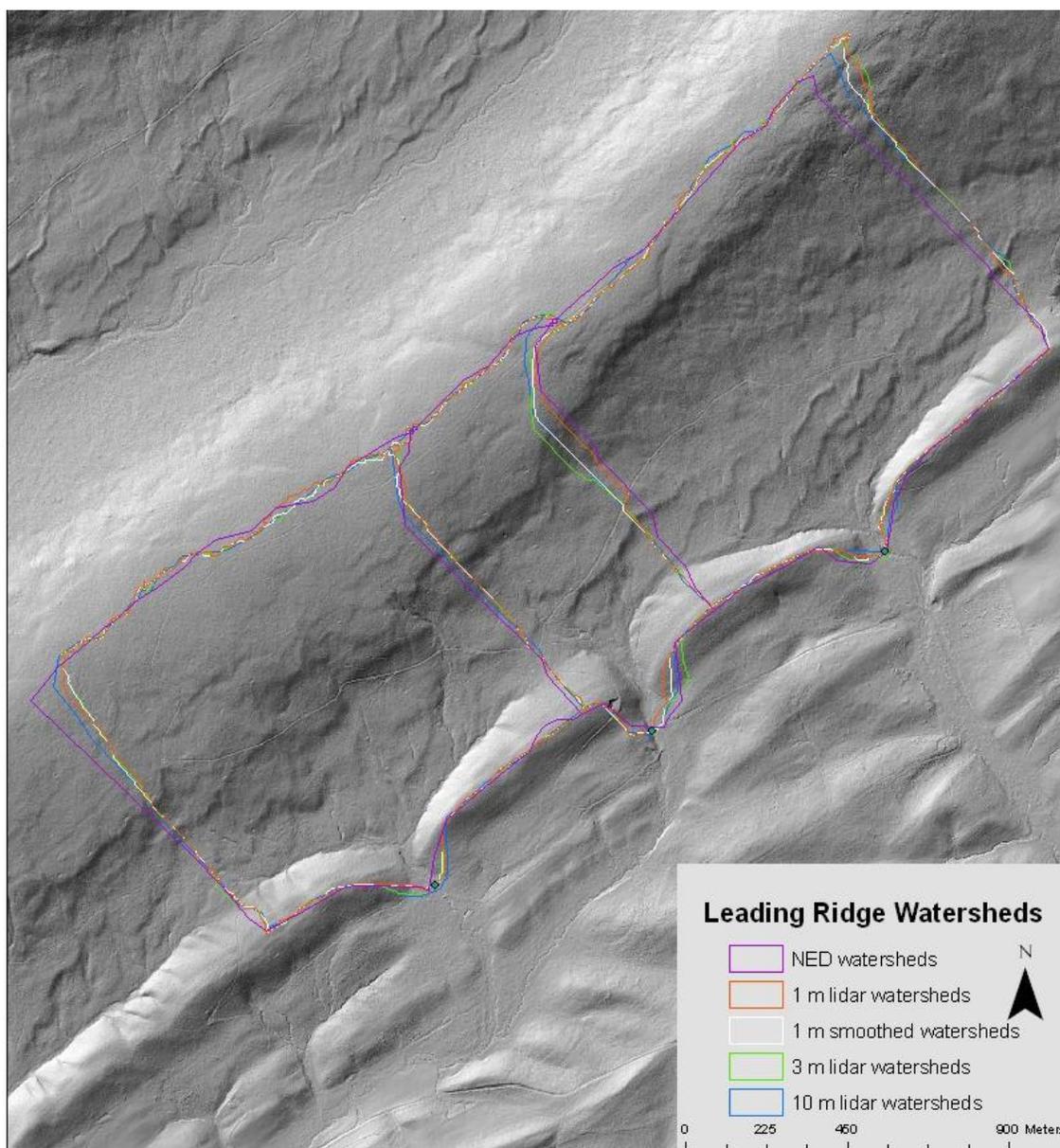
Figure **2-11:** Watersheds delineated using the NED DEM, lidar-derived 1 m DEM, smoothed 1 m lidar-derived DEM, 3 m lidar-derived DEM, and 10 m lidar-derived DEM.

Discussion

Stream Networks

Most obvious throughout the network models is the striking difference between the models produced by the lidar-derived DEM and the NED DEM. While the lidar-derived DEMs may have misrepresented a few of the side channels coming off of the hill slope, they all modeled the two branches of the main channel accurately within the resolution of the data source. This was not the case for the NED DEM which misrepresented the location of one of the main channels by over 30 meters due to errors in the original DEM (Figure **2-12**). Colson (2006) and Murphy et al. (2008) both also found improvements using lidar, but the scale of the error using the NED DEM was still notable.

a.   *b.*
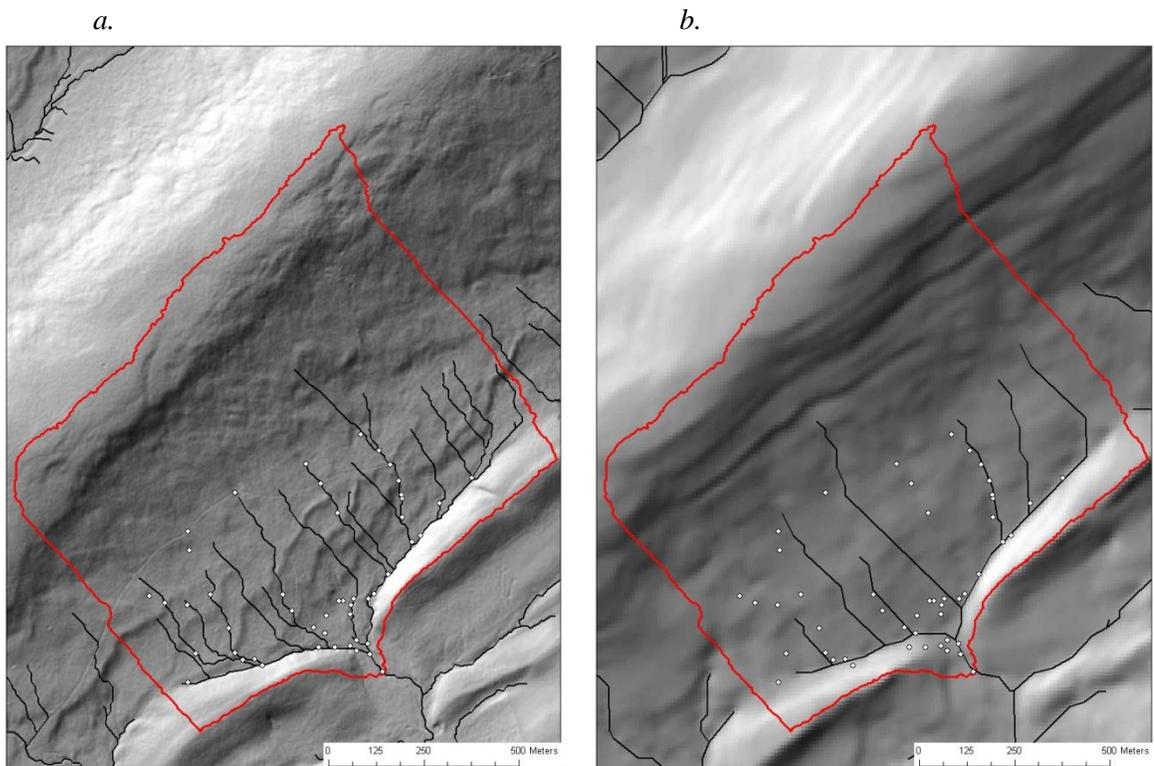
Figure **2-12:** Most accurate lidar-derived stream network (*a*) (3m DEM) compared to the NED stream network (*b*). White circles are GPS channel locations.

Despite this improvement, there were several locations in the watershed that were not accurately modeled by any of the DEMs (Figure **2-13**). These include the center of the watershed which is made up of several dispersed springs and intermittent channels, and one particular channel on the western branch of the watershed. In the case of the center of the watershed, there may be a soil or geologic feature impacting the hydrology, although this is not apparent from available soil or geologic data. It could also be caused by the road intercepting the flow from this region and directing it to the western side of the drainage. This area accounts for some of the outliers in the box plots, particularly the 3 m lidar-derived DEM since all other points were mapped so accurately. On the west side of the watershed, a channel was mapped in the field that was not modeled using any of the DEM products. This may be a result of the road through the watershed capturing water in the ditch along the road and channeling it through a culvert. While the flow would normally be dispersed through the soil, the cut from the road and ditch along its berm captured water from a broad area and channelized it under a culvert. Tenenbaum et al. (2006) found that in urban areas, lidar did not always accurately model flow due to channelized features such as this. Murphy et al. (2008) burned culverts into the DEM prior to modeling, so those issues weren't seen, although burning in culverts would not solve this modeling problem since this problem was due to an increase of field channelization and not channels being modeled in the incorrect location.
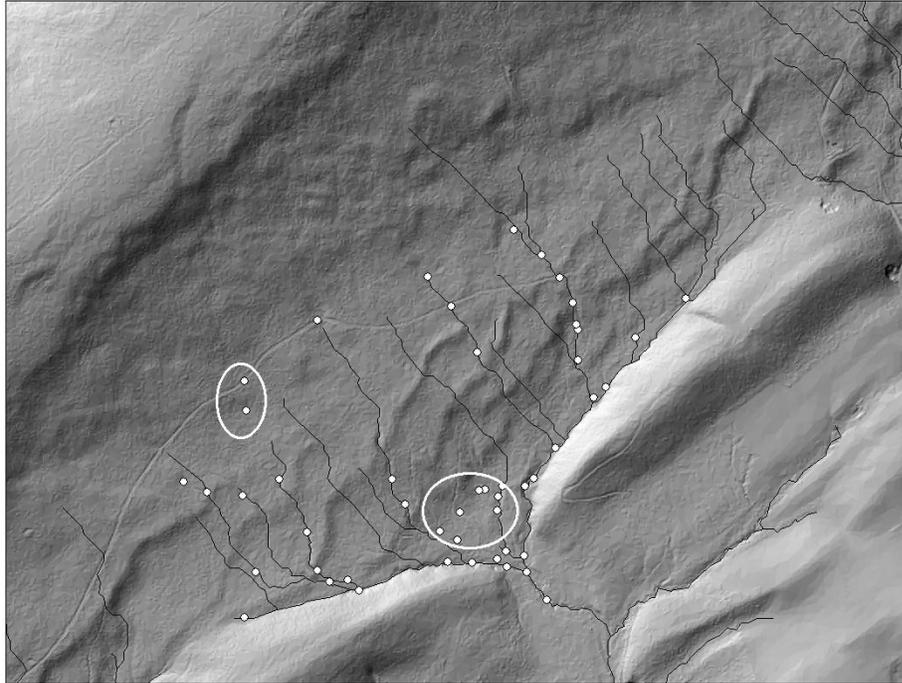
Figure **2-13**: Locations in the watershed that were not modeled accurately.

Several patterns are represented in the box plots of the distance distributions of the lidar-derived DEMs. First, all of the distributions show substantial skew with most of the GPS points lying close to the channel and a few points located far from the channel. On the boxplot of the 3 m distribution, any point located over 5 m away from the modeled channel has been identified as an outlier, highlighting how closely the modeled stream represented the actual stream. It should be noted that the median distance between the modeled stream and actual GPS points is smaller than the resolution of the dataset for the 3 m and 10 m lidar-derived DEMs. The NED modeled streams, by contrast, show large amounts of error throughout the network, with many channels being modeled tens of meters away from their actual location.

Despite its superior resolution and accuracy, the 1 m resolution lidar-derived DEM is not the most effective DEM to use for hydrologic modeling in managed watersheds. One of the channels in the eastern portion of the watershed was modeled particularly inaccurately (Figure **2-14**). Tenenbaum et al. (2006) found that errors in a 1 m lidar-derived DEM created errors in soil

moisture modeling and hypothesized that those errors may be due to inaccuracies in the initial DEM due to tree roots and other non-ground features being classified as ground. Although using the 1m smoothed or the 3 m DEM did not result in significant improvements as measured by a Wilcoxon rank sum test, both appear to result in more accurate channel models than those modeled from the original 1 m DEM. One reason may be the inability of lidar to identify features on the landscape such as small dams, culverts, and ditches. Although major features such as bridges and dams have been identified using aerial photography and burned into the lidar DEM, small dispersed features such as culverts, ditches, and water bars in small logging roads, skid trails, and old roads or trails cannot be identified. Since culvert locations are often not known in wooded, rural areas, without a way to automate culverts, smoothing and reducing resolution both effectively minimize the impact of these features on flow modeling.
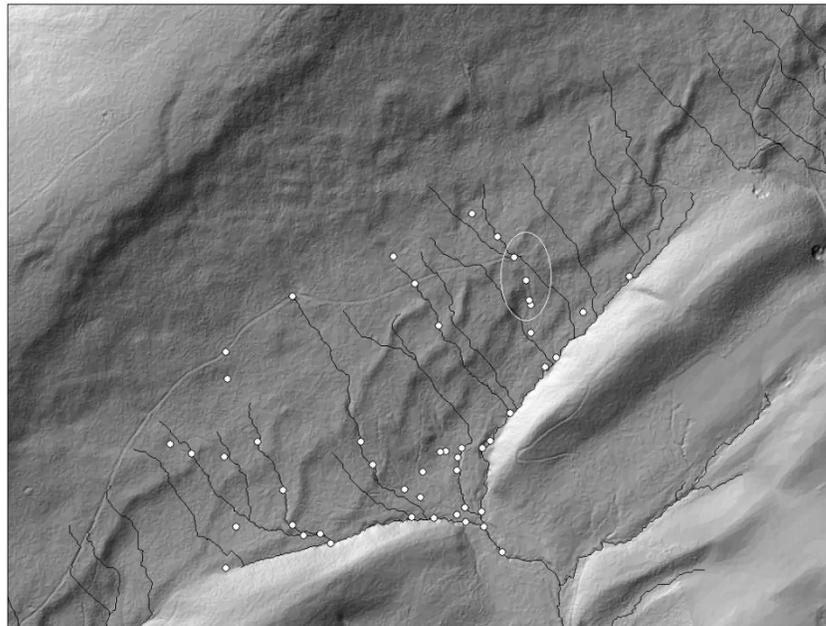
Figure **2-14**: 1m lidar-derived stream network with errors highlighted.

Another interesting result is that the lidar-derived DEM that has been thinned to 10 m resolution still performs as well as the original DEM. The mean and median distances from GPS channel to modeled channel were both slightly smaller from the 10 m lidar-derived DEM than

with the original DEM, and there was no significant difference between distances measured with any of the lidar-derived DEMs. This is contrary to the results of Zhao et al. (2010) who found that a 10 m lidar-derived DEM was outperformed at generating hydrologic modeling parameters by a higher resolution DEM. Zhang (2006) however, found that a 10 m resolution lidar-derived DEM was adequate for erosion prediction using hydrologic variables. The ability of a 10 m resolution DEM to perform adequately could be very important to managers trying to delineate stream channels on a landscape scale. The improvement achieved in going from the 1/3 arcsecond NED to the 10 m lidar-derived DEM is significant, and it is at no cost in storage and processing because both of these are the same for the 10 m lidar-derived DEM as for the 10 m NED DEM. Also, by examining the hillshades created by the different DEM products, it can be seen from Figure **2-15** that the 10 m lidar-derived DEM accurately represents smaller topographic features, whereas the NED misrepresents the finer scale topography as distorted relics of contour lines. Horizontal, parallel lines are visible in the channel network modeled using the NED DEM that are probably errors from the methodology used to generate USGS DEMs. Also, there is an error in the main stream network delineated with the NED. At one point, the major channel of the stream is modeled over 30 meters in the wrong direction. Most of the small channels coming off the hillside are also modeled in the wrong location.
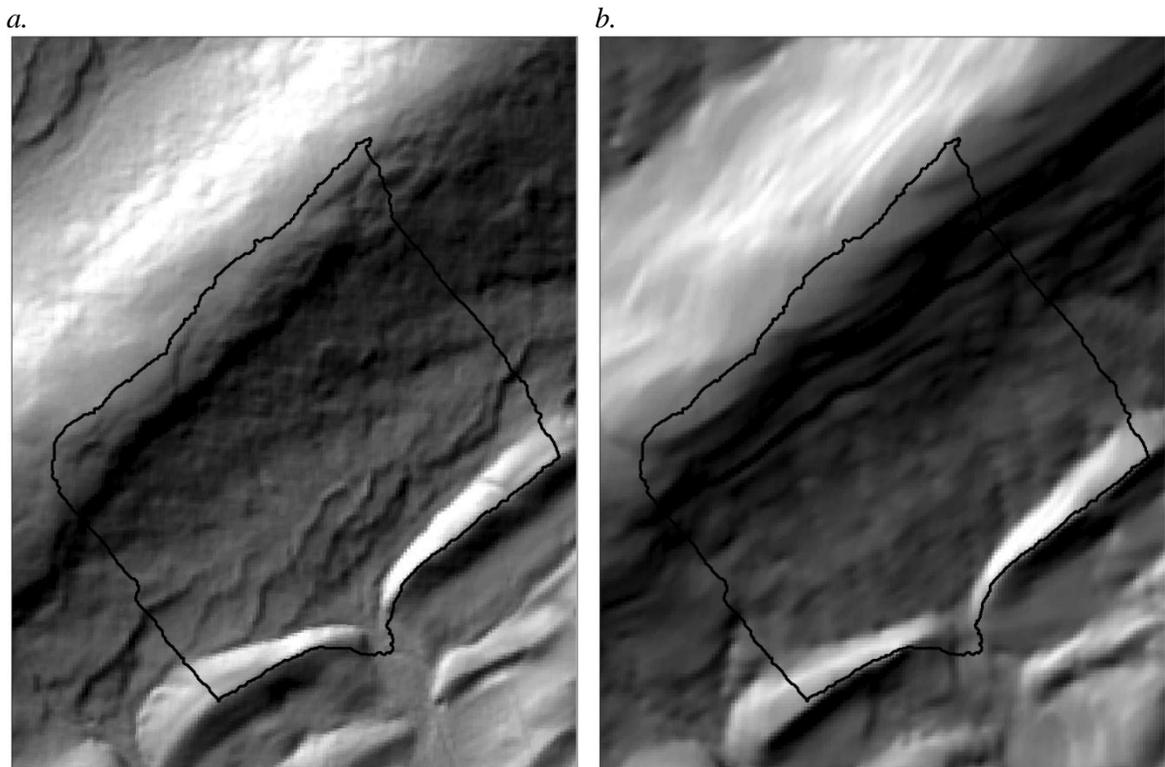
Figure **2-15:** Hillshade derived from a 10 m lidar-derived DEM (*a*) and from a 10 m NED DEM (*b*)

For the purpose of this research, any area accumulating flow of greater than 2.5 ha was classified as a stream. This relatively small threshold was chosen due to the focus of the investigation being on the ability to use lidar to improve the location accuracy of various stream networks. By choosing a higher flow accumulation threshold, fewer channel reaches would be identified. The flow accumulation threshold that a manager or modeler chooses to use depends on their goals. Despite extensive research in channel initiation (Montgomery and Dietrich 1989, Tarboton et al.1991, Dietrich et al. 1992) there tends to be uncertainty when identifying ephemeral or intermittent channels based on flow accumulation models. Due to the compounding properties of a channel network, extracting perennial channels can be accomplished fairly reliably by increasing the flow accumulation threshold to approximately 30 ha, although some intermittent and ephemeral channels will be missed at this threshold. If the goal of the user is to reduce the chance of over representing the stream network, this would be a more appropriate

threshold. If the goal of the user is to reduce the chance of missing stream reaches, a threshold closer to 2.5 ha is probably more appropriate. By decreasing the threshold, there is risk of identifying areas as streams that show no physical characteristics of a stream, but that still have water accumulating and moving through the area. This may be desirable, particularly when performing construction that may impact this flow of water such as building roads, well pads, or other construction projects.

Even using a conservative estimate of stream lengths, an improved methodology using a 10 m resolution DEM can increase the number of stream miles measured over either NHD datasets or blue line streams (Figure **2-16**). With this improved dataset, we can begin to protect headwaters streams and their riparian zones. The importance of the riparian zone has long been understood for protecting water quality, providing a hotspot for biodiversity, and mitigating run-off, sedimentation, and flooding (Naiman and Decamps 1997). Protection efforts have been growing, particularly in Pennsylvania with the effort to reduce nutrient inputs to the Chesapeake Bay (Baker et al. 2006). Although much progress has been made in protecting riparian zones, we can only protect what we know about, and currently almost all true first order streams are not found on digital hydrological datasets. Baker et al. (2007) found a large effect from resolution used for stream maps on buffer analysis, and recommended all buffers analysis work be done at no coarser than a 1:24,000 scale, with finer scale maps preferable.
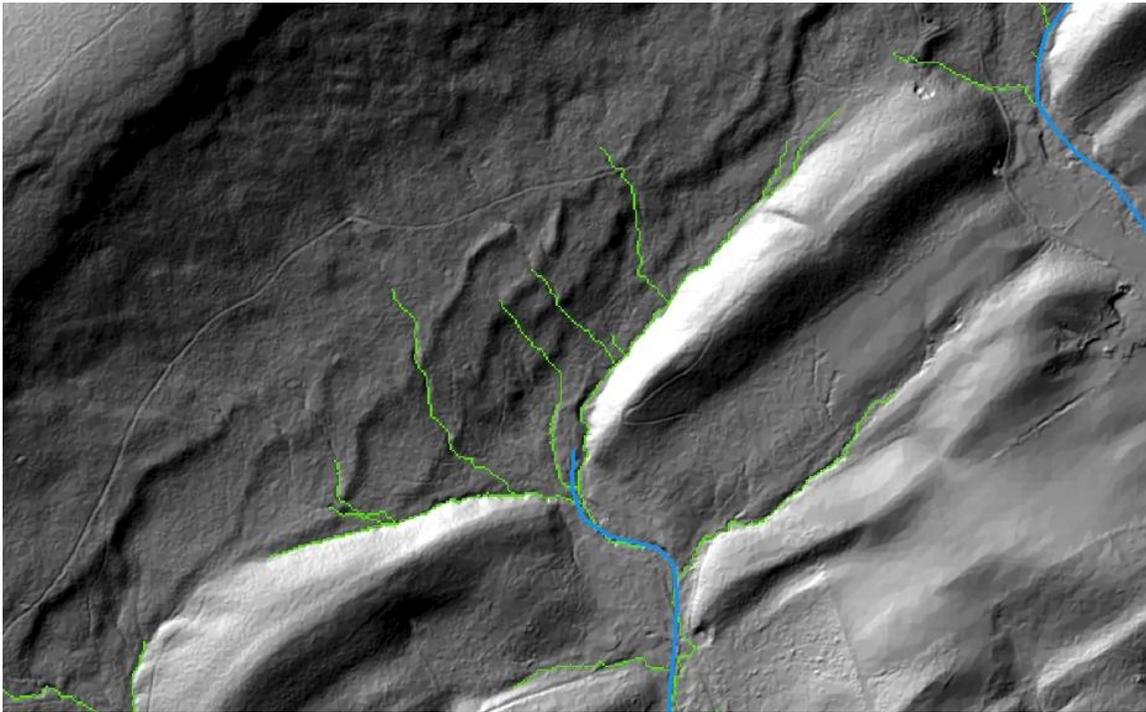
Figure **2-16:** Example of a 10 m lidar-derived stream network (green) compared to the previously available NHD stream network (blue).

Topographic Index

The D8 method was chosen to calculate flow accumulation for much of this study although some studies have suggested that a multi-directional flow algorithm may be better (Sørensen et al. 2006) in order to simplify the analysis and to use the original definition of TI that was proposed by Beven and Kirkby (1979). Also, D8 was the preferred flow accumulation model for channel modeling in this watershed. The difference between the statistical TI distribution for NED DEM and the 10m lidar-derived DEM is notable. Many studies have been conducted to explore the impact of changing resolution on TI (Zhang and Montgomery 1994, Quinn et al. 1995, Sørensen and Seibert 2007), but this demonstrates that different DEM products with the same resolution can produce very different TI distributions. Also, much of the difference in the statistical distributions is in the top 10%, which represent cells that would become water saturated

with hydrologic events. All of the statistical distributions of lidar-derived DEMs were more similar to one another than the NED derived DEM was to any of the lidar-derived DEMs.

There are also differences in spatial pattern at different resolutions. High values of topographic index are found higher up in the watershed as resolutions improve due to the dependence on contributing area. This was also noted by Sørensen and Seibert (2007). At the highest resolutions, values at the top 5% and 10% of TI can be found across the watershed, including at the highest elevations and there is no increase in the percentage of "wet" cells as you move down the watershed (Figure **2-17,** Table **2-3**). It is important to note that different resolutions of data may be providing much different information when modeling TI. There were some negative TI values which have been previously documented in TI studies evaluating lidar-derived DEMs (Straumann and Purves 2007). Overall, the results of this study suggest that TI modeled at a scale less than 10 m may not accurately depict soil moisture of vegetation patterns. Care should be taken when using TI for ecological or forest classification studies to validate the relationship between vegetation and modeled TI (Sørensen et al. 2006). Perhaps another metric such as curvature or roughness, or TI calculated using a slightly different methodology would better depict these relationships.

It is also interesting that TI calculated using a D∞ flow accumulation algorithm resulted in very different values for TI, different patterns of TI on the landscape, and different patterns in changes in TI values with changing resolution. In the literature, many different methodologies of calculating TI are often used, and this highlights the differences that can be caused in TI with changes to resolution, DEM data source, and methodology.
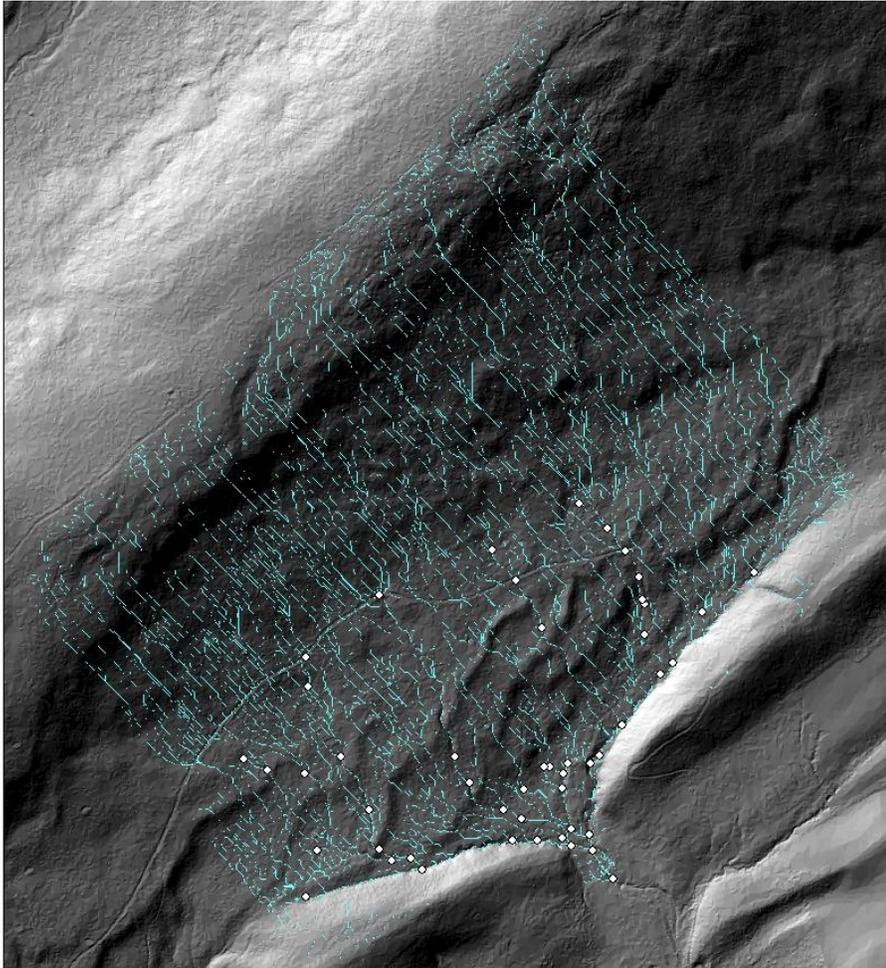
Figure **2-17**: Top 10% of TI for the 1m lidar-derived DEM. Areas of high TI extend to the top of the watershed.

There is also a notable difference between patterns formed from the lidar-derived 10 m DEM and the NED DEM. Comparing the spatial distributions of TI to the stream network models and GPS locations of springs and channels (Figure **2-18**) shows substantial discrepancy between the NED TI and channel locations. Thus, it can be misleading to use NED-derived TI to model spatial patterns of wetness in a small watershed. By using a coarser resolution of NED data (larger than 10 m), larger, dominant features may be more accurately represented, but at the 10 m resolution, the spatial locations of wet areas formed by small channels are not accurately located on a TI map.
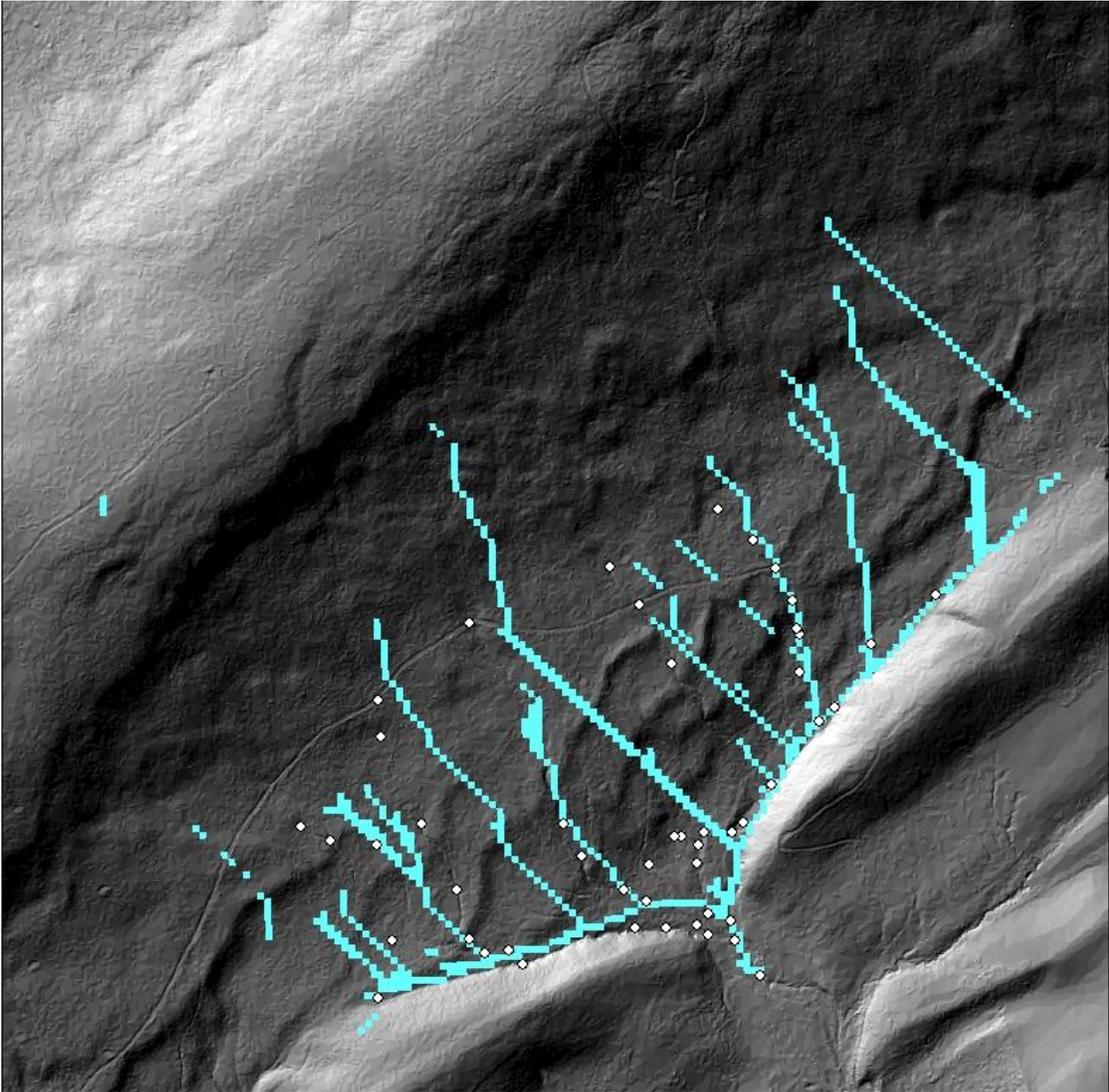
Figure **2-18:** Distribution of the top 5% of TI values as derived using the D8 method and NED DEM and stream and spring locations as mapped in the field.

Conclusions

Using lidar may improve ability to identify and protect small headwater streams that are currently missing from maps due to lack of accurate hydrographic data. For landscape scale modeling, DEMs thinned to 3 m and 10 m both perform effectively, particularly in areas where culvert and ditch placement may not be known. Headwater streams were more accurately

modeled by thinned or smoothed DEMs than by an original 1 m DEM due primarily to the presence of a forest road on the landscape. There seemed to be an advantage for using coarser resolution DEMs for modeling topographic index as well, as areas of high TI were more accurately modeled in the lower portion of the watershed by coarser resolution DEMs. Overall, the 10 m lidar-derived DEM offers a significant improvement in modeled stream channel accuracy with no increase in storage capacity, and is based on data available for the whole state of Pennsylvania.

By improving the spatial hydrographic information that managers, planners, and agencies have to work with, we can improve our ability to protect the headwater streams that play such a valuable role in ecosystems. We may also be able to improve hydrologic models and nutrient transport models to better predict hydrologic response, nutrient transport, and denitrification in headwater systems. By incorporating lidar-derived DEMs into analysis, we can significantly and consistently improve understanding of headwater channel networks and, consequently, critical zone processes, downstream responses, and ecosystem processes occurring in all stream networks.

References

Alexander, R.B., E.W. Boyer, R.A. Smith, G.E. Schwartz, R.B. Moore, 2007. The Role of Headwater Streams in Downstream Water Quality. Journal of the American Water Resources Association 43:41-59.

Anderson, S.P,, Bales, R.C, Duffy, C.J. 2008. Critical zone observatories: building a network to advance interdisciplinary study of Earth surface processes. Mineralogical Magazine 72: 7–10.

Bader, M. Y., J.J. A Ruijten. 2008. A topography-based model of forest cover at the alpine tree line in the tropical Andes. Journal of Biogeography 35:711–723.

Baker, M.E., Weller, D.E., Jordan, T.E. 2006. Improved methods for quantifying potential nutrient interception by riparian buffers. Landscape Ecol 21:1327–1345

Baker, M.E., Weller, D.E., Jordan, T.E. 2007 Effects of stream map resolution on measures of riparian buffer distribution and nutrient retention potential. Landscape Ecol 22:973–992.

Beven K.J. 1997. TOPMODEL: a critique. Hydrological Processes 11: 1069–1086.

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

Beven, K. J. and Moore, I. D. (Eds) 1993. Terrain Analysis and Distributed Modelling in Hydrology. Wiley, Chichester.

Blyth, E.M., Finch, J., Robinson, M., Rosier, P. 2004. Can soil moisture be mapped onto the terrain? Hydrology and Earth System Sciences 8: 923–930.

Boyer, E.W., Alexander, R.B. , Parton, W.J. , C.S. Li, K. Butterbach- Bahl, S.D. Donner, R.W. Skaggs, and S.J. Del Grosso, 2006. Modeling Denitrification in Terrestrial and Aquatic Ecosystems at Regional Scales. Ecological Applications 16:2123-2142.

Brantley, S. L., White, T.S., White, A.F., Sparks, D., Richter, D., Pregitzer, K., Derry, L., Chorover, J., Chadwick, O., April, R., Anderson, S., Amundson, R. 2006. Frontiers in Exploration of the Critical Zone: Report of a workshop sponsored by the National Science Foundation (NSF), October 24-26, 2005, Newark, DE, 30p.

Bruneau P., Gascuel-Odoux C., Robin P., Merot Ph., Beven K. J. 1995. Sensitivity to space and time resolution of a hydrological model using digital elevation data. Hydrological Processes 9:69-80.

Cirmo, C. P. and J. J. McDonnell, 1996. Linking the Hydrologic and Biogeochemical Controls of Nitrogen Transport in Near-Stream Zones of Temperate-Forested Catchments: A Review. Journal of Hydrology 199:88-120.

Colson, T. P. 2006. Stream Network Delineation from High Resolution Digital Elevation Models. Ph.D. Dissertation, Dept. of Forestry and Environmental Resources, NC State University, Raleigh, NC.

Corwin, D. L., Hopmans, J., and De Rooij, G. H. 2006. From field- to landscape-scale vadose zone processes: scale issues, modeling, and monitoring. Vadose Zone Journal 5: 129–139.

Dietrich, W.E., Wilson, C. J., Montgomery, D. R., McKean, J., Bauer. R. 1992. Erosion thresholds and land surface morphology. Geology. 20:675-679.

Garbrecht, J., and L. Martz. 1993. Network and subwatershed parameters extracted from digital elevation models — the Bills Creek experience, Water Resources Bulletin **29**:909–916

Gesch, D.B., 2007, The National Elevation Dataset, in Maune, D., ed., Digital Elevation Model Technologies and Applications: The DEM Users Manual, 2nd Edition: Bethesda, Maryland, American Society for Photogrammetry and Remote Sensing, p. 99-118.

Guntner, A., J. Seibert, and S. Uhlenbrook. 2004. Modeling spatial patterns of saturated areas: an evaluation of different terrain indices. Water Resources Research, 40.

Hornbeck, J. 1962. Hydrologic analysis of the three Leading Ridge experimental watersheds. M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Hornberger G.M., Boyer, E.W. 1995. Recent advances in watershed modeling. Review of Geophysics 33: 949–957.

James, L. A., Watson, D.G., Hansen, W.F. 2007. Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. Catena 71:132–144.

Kienzle, Stefan. 2004. The effect of DEM raster resolution on first order, second order, and compound terrain derivatives. Transactions in GIS. 8:83-111.

Lehman, G. S. 1962. An evaluation of the physical and hydrological properties of the soil mantle on Leading Ridge Watershed two**.** M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Lynch, J. A. and E. S. Corbett, 1990. Evaluation of Best Management Practices for Controlling Nonpoint Pollution from Silvicultural Operations. Water Resources Bulletin 26:41-52.

Maune, D. F., editor. 2007. Digital elevation models and technologies and applications: the DEM user manual, 2nd Edition. The American Society for Photogrammetry and Remote Sensing. Bethesda, MD.

McClain M.E., Boyer E.W., Dent C.L., Gergel S.E., Grimm N.B., Groffman P.M., Hart S.C., Harvey J.W., Johnston C.A., Mayorga E., McDowell W.H., Pinay G. 2003. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. Ecosystems 6:301–12.

Meyer, J.L.*,* and J.B. Wallace*,* 2001*.* Lost Linkages and Lotic Ecology: Rediscovering Small Streams*. In :* Ecology: Achievement and Challenge*,* M.C.Press*,* N.J.Huntly*,* and S.Levin (Editors). Blackwell Science, Malden, Massachusetts, pp*.* 295-317.

Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S. and Leonard, N. E. 2007. The Contribution of Headwater Streams to Biodiversity in River Networks. JAWRA Journal of the American Water Resources Association, 43: 86–103.

Montgomery, D. R., Dietrich, W. E. 1989. Source areas drainage density and channel initiation. Water Resources Research. 25:1907-1918.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A. 1993: Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57: 443-52.

Murphy, P.N.C. J. Ogilvie, F.-R. Meng and P.A. Arp. 2008. Stream network modelling using LiDAR and photogrammetric DEMs: a comparison and field verification, Hydrological Process 22:1747–1754.

Naiman RJ, Decamps H 1997. The ecology of interfaces: riparian zones. Annu Rev Ecol System 28:621–658.

O'Callaghan, J.F. and Mark, D.M., 1984. The extraction of drainage networks from digital elevation models. Computer Vision, Graphics and Image Processing, 28: 323–344.

Qu, Y. 2004. An integrated hydrologic model for multi-process simulation using semi-discrete finite volume approach, PhD dissertation, Penn State University, University Park, PA.

Quinn P. F., Beven K. J., Lamb R. 1995. The ln(A/tan beta) index – How to calculate it and how to use it within the TOPMODEL framework. Hydrological Processes. 9:161-182.

R Development Core Team 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Robinson, G. R. 1959. Tree mortality and defects on the Leading Ridge watershed number one**.** M. S. Thesis. School of Forest Resources. Penn State University, University Park, PA.

Schultz, C. H. editor. 1999. The Geology of Pennsylvania. Geologic Survey Special Publication 1.

Sørensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrology and Earth System Sciences. 10:101–112.

Sørensen, 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.

Straumann R.K. and Purves R.S. 2007. Resolution sensitivity of a compound terrain derivative as computed from LiDAR-based elevation data. In: Fabrikant S.I. and Wachowicz M. (eds.) Lecture

Notes in Geoinformation and Cartography. The European Information Society – Leading the Way with Geo-Information, Proceedings of AGILE 2007, 8.–11 May 2007, Aalborg, Denmark, 87-109.

Tarboton D. G., Bras R.L., Rodriguez-Iturbe I. 1991. On the extraction of channel networks from digital elevation data. Hydrological Processes 5: 81–100.

Tarboton, D. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research 33: 309-320.

Tenenbaum, D.E., L.E. Band, S.T. Kenworthy, C.L. Tague. 2006. Analysis of soil moisture patterns in forested and suburban catchments in Baltimore, Maryland, using high-resolution photogrammetric and LIDAR digital elevation datasets. Hydrological Processes 20:219–240.

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

Zhao Z, Benoy G, Chow TL, Rees HW, Daigle J, Meng F. 2010. Impacts of accuracy and resolution of conventional and LiDAR based DEMs on parameters used in hydrologic modeling. Water Resources Management 24: 1363–1380.

Chapter 3

Evaluating state-wide lidar data product for characterizing surface roughness in complex terrain in the forested Ridge and Valley Ecoregion of Pennsylvania

Abstract

The availability of light detection and ranging data (lidar) has resulted in a new era of landscape analysis. For example, the subsequent improvements in terrain data now make it possible to model micro topography over a large geographic area via remotely sensed techniques. In this study, two types of lidar-derived data were used: a 1 m resolution DEM available statewide in Pennsylvania and a research grade 1 m DEM generated for the Susquehanna/Shale Hills CZO. Roughness was calculated from lidar-derived DEMs using standard deviation of slope, standard deviation of curvature, a pit fill index, and as the difference between a smoothed splined surface and the original DEM. These were compared to surveyed micro-plots and transects placed in the field to obtain accurate surface models of diverse soil and terrain. Results suggest that the research grade lidar did not improve roughness modeling in comparison to the state-wide lidar and that resolution and initial point density may not be as important as the algorithm and methodology used to generate a lidar-derived DEM for roughness modeling purposes. Using lidar, patterns of roughness were identified that were associated with different landforms derived from hydro-geomorphic features such as stream channels, gullies, and depressions. Lowland areas tended to have the highest roughness values for all methods, with other areas showing conflicting but distinctive patterns of roughness.

Introduction

Lidar (light detection and ranging) is changing the way scientists view landscapes. Over the past several decades, geomorphologists, soil scientists, ecologists, foresters, and hydrologists have increasingly utilized terrain data for landscape classification (ECOMAP 1993, Cleland et al. 1997, Franklin 1995, Myers 2000), predicting forest communities (Bolstad et al. 1998), predicting soil properties (Dikau 1989, Moore et al. 1993), and understanding riparian zones and systems (McGlynn and Seibert 2003). Due to improvements in data acquisition, computing power and storage capacity, terrain data has become increasingly available at finer and finer resolutions and at broader scales, from NED and SRTM to lidar. With lidar, our ability to see features on the landscape is so improved that that we can investigate new properties of landforms remotely which results in better information for a diverse field of scientists and land managers.

Over the past ten years, the application of lidar data to address basic research and land management questions has evolved from that of a user group consisting largely of the research community to a group inclusive of land managers and practitioners. Whether using coarser state-wide data, such as Pennsylvania 1 m dataset (PAMAP) or 0.5 m research data, lidar data provide an unprecedented capability to model landscapes. However, resolution and accuracy differences result in differing capabilities to model landscape features, especially when coarser commercial grade lidar data (often > 3 m) is compared to research grade lidar (often < 3 m) typically used for site-specific, or small-watershed scale analysis. Research grade data are often collected at higher point density (which results in a more detailed model of landscape features) and potentially more accurate elevation dataset. Although lidar-derived DEMs have been shown to be extremely accurate when compared to non-lidar generated DEMs (Maune 2007), the accuracy of lidar-derived DEMs for measuring micro-topography on landscapes is debated (Kraus and Pfeifer 1998). For example, some researchers have found lidar-derived DEMs to be over-smoothed

(Kraus and Pfeifer 1998), which can minimize surface roughness and result in less topographic complexity. In contrast, others have found lidar-derived DEMs effective at identifying features such as landslides, which can have complex roughness patterns (McKean and Roering 2004). This question is confounded by abiotic terrain factors (such as slope) and biotic factors (such as evergreen vegetation and coarse woody debris) (Su and Bork 2006, Spaete et al. 2011).

Micro-topography is an important variable for modeling water movement (Dunn et al. 1991), geomorphology (Lavee et al. 1995), vegetation dynamics (Beatty 1984, Enoki 2003), and riparian communities (Naiman and Decamps 1997, Pollock et al. 1998). Presently, research on determining surface roughness has frequently focused on agricultural soils (Huang 1998, Govers et al. 2000, Kamphorst et al. 2000) or geologic features such as landslides and alluvial fan deposits (Glenn et al.2006, Frankel and Dolan 2007). Due to lidar's capability to more accurately represent micro-topography and surface roughness (Glenn et al. 2006), lidar-related research applications will likely cross over to fields such as ecology, hydrology, soil science, and forest science. Lidar data and its associated topographic variables have the potential to improve estimates of water storage and infiltration, identify wildlife habitat and surface-dependent soil properties or relief. For example, pit-and-mound topography relief is a type of micro-topography very common in natural forested ecosystems, and that is commonly driven by historical incidents of wind throw (Schaetzl et al. 1988, Rumbaitis del Rio 2006). Pit and mound topography is known to occur more frequently in areas prone to shallow rooting such as landscapes with a shallow soil depth to a water table or other restricting layer. Pit and mound topography may also result from differences in tree species, ages of stand, and other vegetation-based variables. The application of lidar data in identification of such landscapes may result in identification of some of these features which could improve our understanding of co-occurring soil and vegetation at site-specific and regional scales.

Various methods have been proposed to measure micro-topography including analysis of micro-topography fractal dimensions, (Andrle and Abrahams 1989), identifying eigenvectors parallel to micro-surfaces (McKean and Roering), and analyzing variograms associated with surfaces at multiple scales (Bai et al. 2005, Huang and Bradford 1992). Other methods include measuring the standard deviation or range of elevation over a particular scale, and calculating variability over a small scale while removing the effect of the broader scale topography (Glenn et al.2006, Frankel and Dolan 2007). Unfortunately, a consistent, preferred method of delineating surface roughness has not emerged, potentially because a one-method fits all approach is unlikely given the diverse array of needs. Thus, one of the objectives of this research is to evaluate the different roughness metrics calculated from lidar-derived DEMs.

The cost of obtaining lidar data can vary with scale, but whether a finer data resolution is always warranted is unknown. Besides differences in point cloud density across lidar data resolutions, there are also different methodologies and algorithms for generating a DEM from the lidar point cloud (Hengl and Evans 2009) which can inflate the cost. Introduced artifacts resulting in greater error during processing of the lidar-derived DEM may result in different roughness metrics at a variety of scales. While one method of generating a DEM from lidar data may work best for hydrologic modeling, another method may be more useful for assessing pit and mound topography. Additionally, cost may vary based on methodology.

I evaluated two different lidar datasets for relative accuracy using a topo-survey as a control, and roughness metrics derived from multiple methodologies to assess the effectiveness of lidar at characterizing surface roughness and microtopography. In addition, the micro-topographic signatures of several landforms within the site are modeled. Surface roughness is modeled using a 1 m DEM and field-surveyed features (at the scale of 10-100 cm): surface roughness and micro-topography are used interchangeably.

Methods

Study Area

This study took place in a 119 ha watershed located in the Ridge and Valley Province of Pennsylvania (Figure **3-1**) known as Leading Ridge Watershed One. The elevation of the watershed ranges from 260 m at the mouth of the watershed to 512 m at the top of its northwestern border. Since this watershed is in the Ridge and Valley Province, its hydrologic network tends to form a trellis pattern instead of the more common dendritic pattern typical of sandstone and shale bedrock. The geologic formation underlying the watersheds consists of deeply dipping strata ranging from resistant Tuscarora quartzite and sandstone at the top of the watershed to less resistant Rose Hill shale that comprises the valley area (Schultz 1999) (Figure **3-2**). This terrain is consistent with that which makes up most side slopes and ridges of the Ridge and Valley Province, which is generally characterized by canoe-shaped valleys and long, linear, parallel ridges formed by differential erosion (Schultz 1999). Ridges tend to be extremely steep and rocky and the topography is generally well-drained (Schultz 1999). This landscape has also been largely influenced by peri-glacial processes of the late-Pleisticene (Ciolkosz et al. 1986).
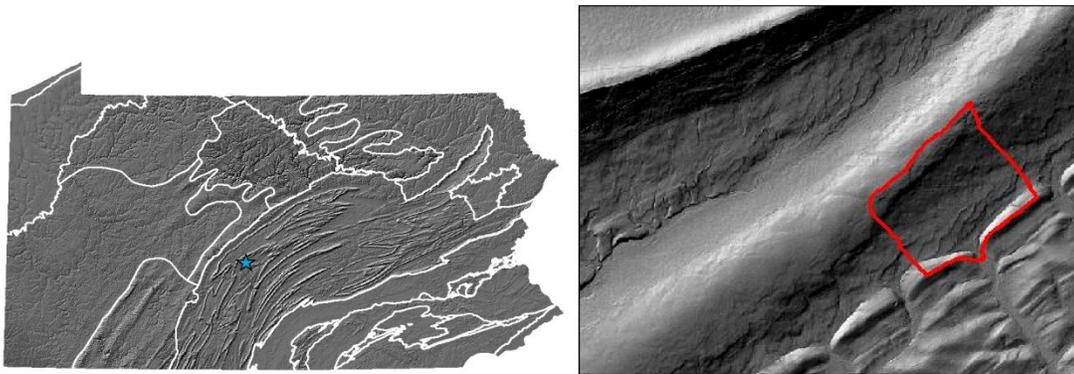
Figure **3-1**: Location of the study site in the Ridge and Valley province of Pennsylvania.

a)

b)
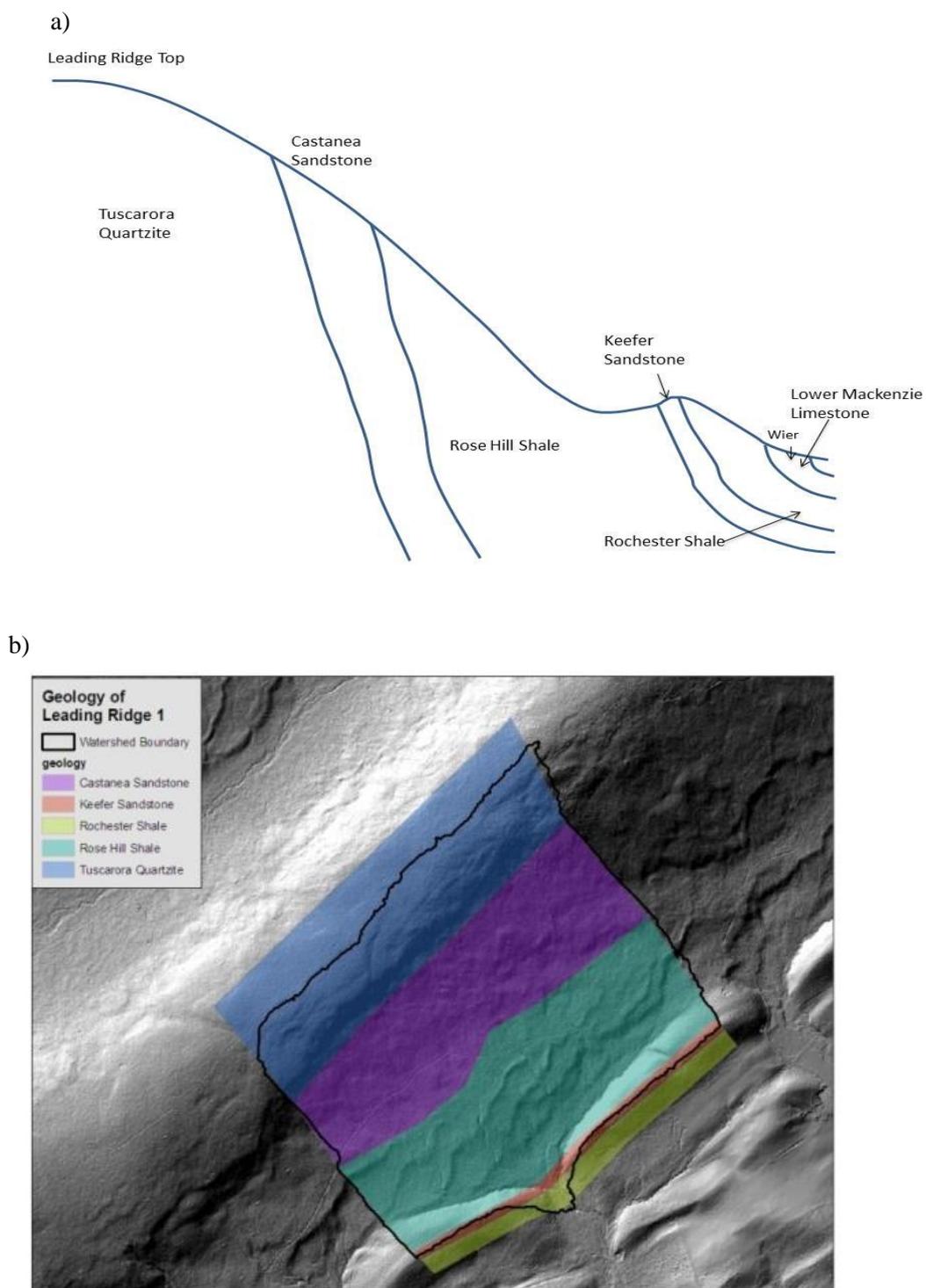
Figure **3-2** a) Schematic showing the approximate arrangement of geology in the Leading Ridge Watershed 1 (Adapted from Shields 1966) and b) map of the watershed showing approximate locations of contacts (adapted from Shields 1966).

The watersheds contain a mature oak/hickory forest approximately 100 years of age. This area has been utilized extensively for charcoal production prior to 1900, and charcoal pits can still be found scattered throughout the site (Robinson 1959). The southern section of the watershed was probably grazed until the 1930s, when farmers were paid to abandon their land. A small salvage cut was conducted in 2006 due to a blowdown event that occurred on the site, and the watersheds experienced moderate gypsy moth mortality over the past 20 years (pers, com. Joseph Harding 2009).

This site was chosen in part because of its history of being used as an experimental watershed and the vast dataset that currently exists for the site. The Leading Ridge watershed research areas were established in Penn State's Stone Valley Experimental Forest of central Pennsylvania in 1959 as paired watersheds to study the hydrologic response of different forest practices (Lynch and Corbett 1990). This watershed is also within the larger Shaver's Creek watershed which is being used as part of the Susquehanna/ Shale Hills Critical Zone Observatory (CZO), an interdisciplinary observatory toward quantitatively predicting creation, evolution, and structure of regolith as a function of the geochemical, hydrologic, biologic, and geomorphologic processes operating in a temperate, forested landscape (Brantley et al. 2007, Anderson et al. 2008).

The dominant parent materials across the watershed are sedimentary rocks formed during the Silurian period of 410-440 million years ago (Schultz 1999). Their structure is steeply dipping and outcrops are found in near vertical position (Shields 1966) (Figure **3-2**). The three rock types that underlay over 90% of the watershed are the Tuscarora quartzite, Castanea sandstone, and Rose Hill shale. Keefer sandstone is also important because although thin, it supports the shale hill formation along the lower front of the watershed. The top of the watershed, which is found on the crest of Leading Ridge, is underlain by Tuscarora quartzite, a very resistant sandstone-quartzite which is the dominant ridge forming rock in the Ridge and Valley Province of

Pennsylvania. Slightly downslope, the Castanea sandstone underlies the upper portion of the

slope area in the watershed. This sandstone is slightly softer than the Tuscarora and is comprised

primarily of greyish red sandstone and siltstone. The Rose Hill shale makes up the lower slope of

the watershed, the valley bottom, and most of the shale hill area in the front of the watershed. The

Rose Hill shale has thin limestone layers inter-bedded which contributes to the relatively higher

pH of the shale hill areas when compared to adjacent forested ridge soils (Lehman 1966). Soils

developed across the watershed tend to follow the trends of the topography and underlying

lithology (Figure **3-3**). An analysis of SSURGO (NRCS 2011) shows that the most common soil

series in the upper half of the watershed are the Hazelton-Dekalb association (sandstone

colluvian); Laidig (sandstone colluvium), which is found on the lower part of Leading Ridge;

Buchanan (sandstone/shale colluvium), which is found in the valley bottom on the southwest

portion of the watershed; Andover (sandstone/shale colluvium), which is found in the valley

bottom on the southeast portion of the watershed; and Berks-Weikert association (shale residuum,

local colluvium), which is found on the shale hills in the front of the watershed (NRCS 2011).
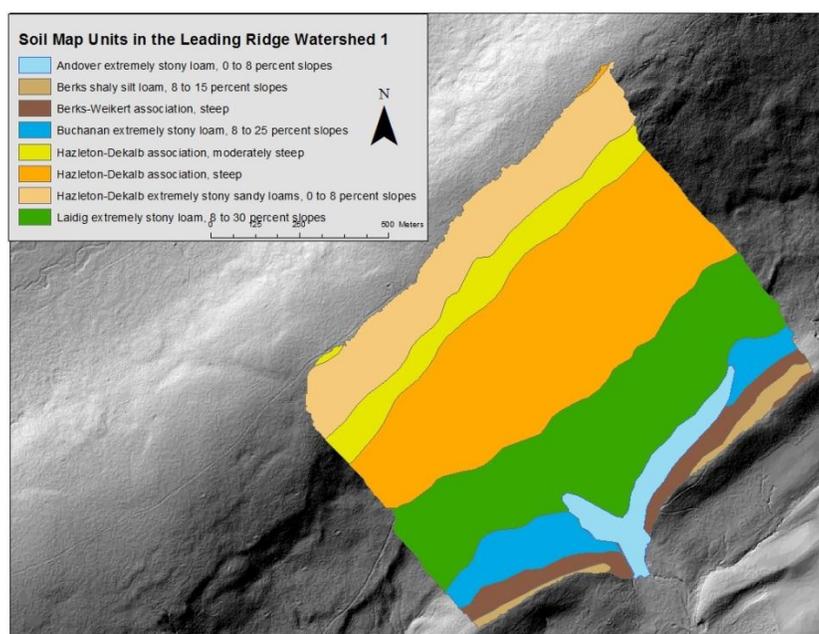
Figure **3-3:** SSURGO soil polygons for the Leading Ridge watershed.

DEM Data Sources

There were two sources of lidar-derived DEM data utilized for this project. The first lidar
data set was collected in 2007 during leaf-off conditions as part of the PAMAP lidar program.
Post spacing for the lidar returns used to generate the 1 m DEM was 1.4 meters, with a target
vertical RMSE of 18.5 cm in open areas and 37 cm in vegetated or forested areas. Points were
first classified as either ground or non-ground points, with ground points being thinned down to
create a TIN that fit the final specifications by an independent vendor, BAE Systems. Using
proprietary methods, a 3.2 foot (about 1 m) resolution DEM was produced using the TIN. All
finished products were checked for quality and accuracy (PAMAP LiDAR QAQC report 2007).

The second data set, CZO lidar, was collected in the winter of 2010-2011 by the National
Center for Airborne Laser Mapping (NCALM). Initial lidar point density was approximately 10
points per m^2, with bare earth point spacing of approximately 4 points per m^2. Bare earth points
were isolated using Terrascan, and then converted to a DEM using Golden Software's Surfer 8
Kriging algorithm using a linear variogram model with a nugget variance of 0.15 m and a search
radius of 25 m or 40 m. Complete specifications can be found in the 2010 NCALM project report
(NCALM 2010). Final resolution for the DEM used for this study was 1 m.

Modeling

Surface roughness was assessed using five methods. The focal statistics tool in ArcGIS
Spatial Analyst was used to calculate the standard deviation of elevation within a 5 m by 5 m
moving window, which resulted in a first approximation of roughness. A 5 m by 5 m moving
window was chosen because features defined as surface roughness for this study are on the order
or 1-3 meters in size. Larger moving windows may have identified larger objects such as streams,

roads, trails, and major slope breaks which were not the subject of this study. Because of the high correlation between slope and standard deviation of elevation (Figure **3-4**), this method was not further analyzed and additional metrics were calculated that attempted to remove the effect of slope on the calculation. Standard deviation of slope was calculated using focal statistics over a 5 m by 5 m moving window of a slope layer measured in percent slope using ArcGIS. A 5 m by 5 m window was again chosen in order to emphasize the scale of microtopography targeted for this analysis. Similarly, a second roughness metric was calculated using the standard deviation of curvature using a 5 m moving window. Curvature was calculated for the lidar-derived 1 m DEMs using the curvature tool in ArcGIS that measures a combination of both plan and profile curvature using the method of Zeverbergen and Thorne (1987). Calculating curvature at this scale is representing the micro-scale curvature of the surface, not dominant surface features that would be traditionally identified as curvature.
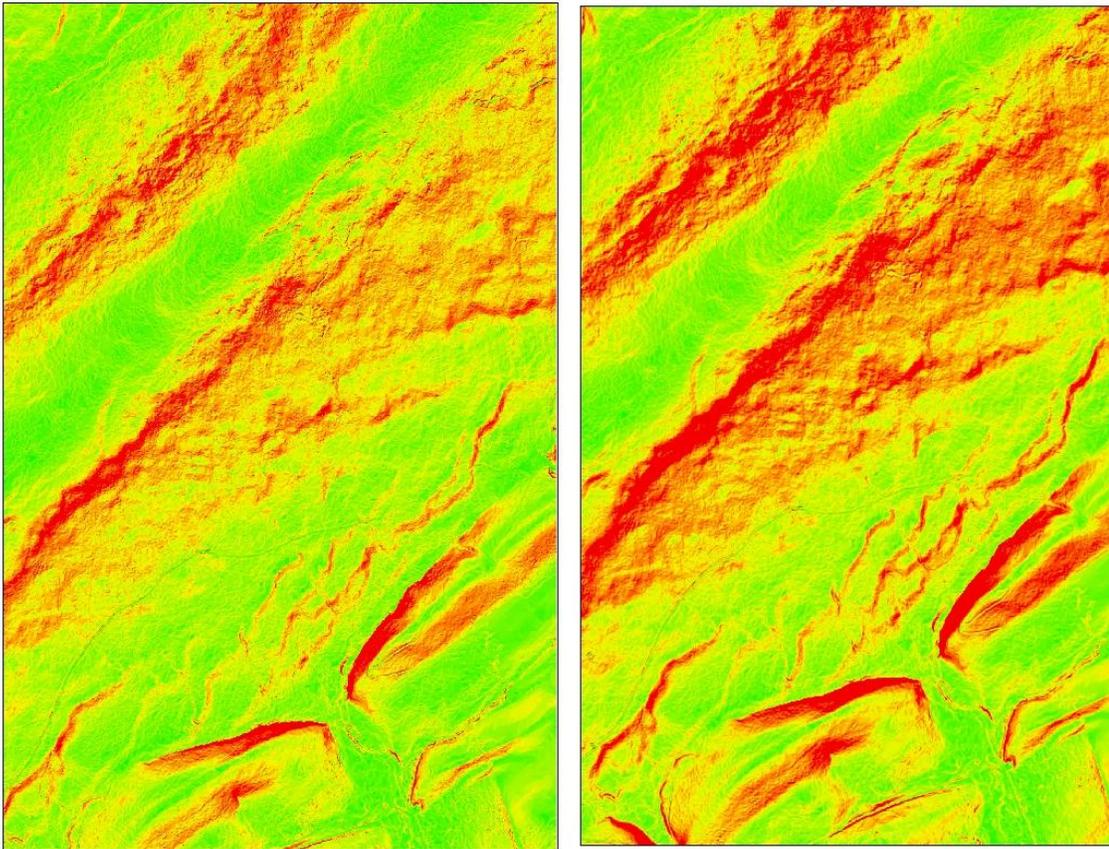
Figure **3-4**: Slope (left) and the standard deviation of elevation (right) for the PaMAP lidar 1 m DEM. Note the similarities.

A third method for measuring roughness was to isolate the micro-scale variation of the DEM from the broader scale topography of the site. This was done by generating a new, smoothed surface created by using a thin plated spline on a thinned 10 m DEM. First, the lidar-derived 1 DEM was thinned to 10 m resolution using the nearest neighbor method in ArcGIS. Using ArcGIS to fit a regularized spline with a weight of 0 in ArcGIS, the 10 m DEM was interpolated back to a 1 m DEM. The difference between the lidar-derived 1 m DEM and the resampled/splined DEM was calculated to show localized differences from the broader scale topography.

An additional metric of microtopography was the pit fill metric, which measured the difference between a hydrologically corrected (pits filled DEM) and the original DEM. In order

to calculate this layer, a filled DEM was created using the Fill tool in ArcGIS spatial analyst

extension with the lidar-derived 1 m DEMs. The original DEM was then subtracted from the

filled DEM, and values were summed over a 10 m area using the block statistics tool to improve

visualization of data.

Field Verification

Roughness metrics were validated in the field with a total station used to survey transects

and micro-plots throughout the watershed (Figure **3-5**). This methodology was intended to be

purposive and was not designed with concern for being replicable. In ArcGIS, four transects were

located perpendicular to major breaks in soil as measured by SSURGO and landform type. Each

transect was designed to be approximately 100 m in length, with one end being in one soil map

unit and landform, and the other end in a different soil map unit and landform. Two of the

transects were located at the boundary between Berks/Weikert and Buchanan soil, one transect

was located at the boundary between Buchanan and Laidig, and one was located at the boundary

between Laidig and Hazleton-Dekalb. These are the four dominant soil types found in the

watershed. These proposed transects were located in the field to ensure that transects spanned soil

types: 1 m soil cores were taken to verify transects spanned multiple series. Micro-plots

representative of the soil type/landform they represented were identified in the field. Two micro-

plots were measured on each transect for a total of 8 micro-plots. Micro-plots were surveyed in

order to create a TIN-based DEM at as high of a resolution as possible to compare to the DEMs.

This was done in order to compare the DEM to a model of the ground surface as derived from a

topographic survey.
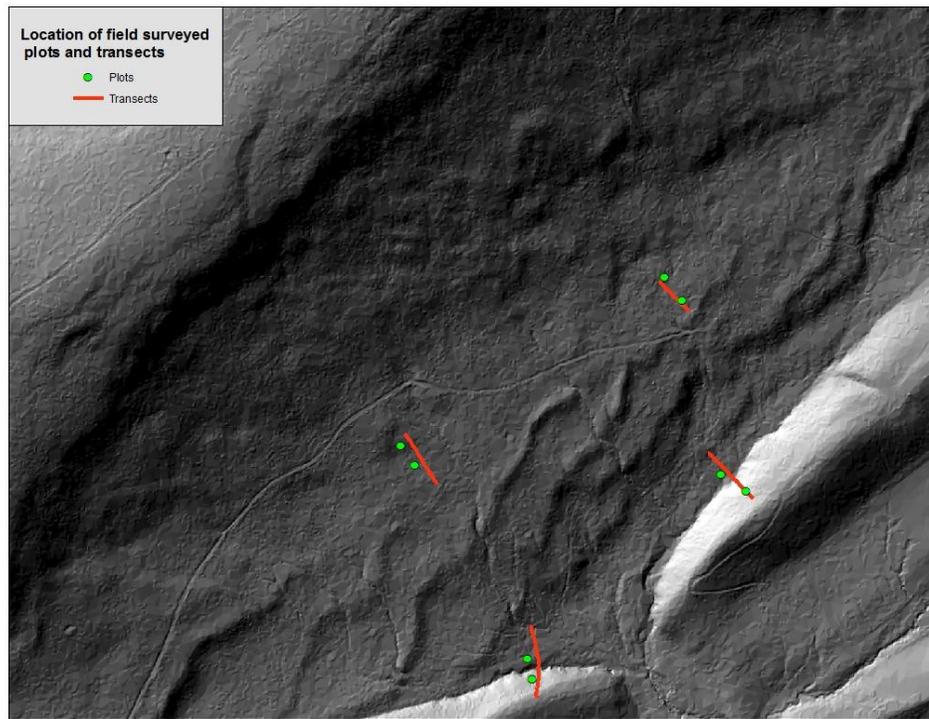
Figure **3-5**: Location of micro-plots and transects on the watershed.

For each transect, elevations were measured at every slope break in order to record the surface in two dimensions. Transects were continued until it was clear that a boundary in soil or landform was reached based on field observations of soils, seeps, springs, and a change in understory vegetation, with at least 20 meters of the transect continuing into the second soil or landform. Micro-plots (10 m by 10 m squares) were placed in locations representative of the surface in the localized area and were selected to be within survey distance of the corresponding transects. On two of the micro-plots, the slope was estimated to be nearly 100% so in these areas, the side of the plot perpendicular to the slope was measured to be 14 meters ground distance in order to create a plot that was approximately 10 m by 10 m in planimetric view. Within these micro-plots, rocks were measured as surface if they were assessed to be embedded into the ground. If they were loose, they were not considered to be part of the surface. The average rock

size was approximately 0.3 m. A topographic survey was conducted on micro-plots to record all areas of slope change in order to record as accurate of a depiction of the surface as possible.

In order to analyze field data, point elevations were converted to a triangular irregular network (TIN) using ArcGIS. A TIN model was chosen because points were located in the field to represent actual topography, with all changes in slope recorded. The TIN was converted to a raster based elevation model using a 0.3 m cell size to approximately match the scale of features that were captured in the survey. For each 0.3 m resolution plot the same roughness metrics were calculated that were used for the modeling analysis including standard deviation of slope and curvature, values of pitted cells, and difference between local elevation and a splined surface. It should be noted that since the methodology of the field survey set out to record actual topography, care was taken to record lowest and highest elevations over an area. This could be contrasted with a digital elevation model which is based on a regular gridded pattern. Therefore it would be expected that the ground surveyed plots display much greater roughness than the DEM plots.

Root Mean Square Difference

Root mean square difference (RMSD) between the DEM-modeled elevations and the surveyed elevations was calculated for the CZO lidar dataset and the PAMAP lidar dataset using the survey points as control points. There were over 700 points that were surveyed from benchmarks. These points are the same as those used in field verification. Each point was converted to a 0.3 m raster cell using the mean value of points to assign a raster value. RMSD of the lidar-derived DEMs was calculated using the formula RMSD = sqrt((1/N)*sum((x-x')^2) where N equals the number of cells/points, x is the surveyed value, and x' is the lidar DEM value.

Results

Field-based Modeling

Although the lidar-derived DEMs from the PAMAP program and the CZO lidar both are at a 1 m resolution, many differences exist between the two datasets based on their initial point density and subsequent processing techniques. The PAMAP lidar was converted to a DEM using a process based on creating a triangular irregular network (TIN) from points classified as ground points using a proprietary algorithm. This TIN was then converted to a DEM. Conversely, the CZO lidar was converted to a DEM using a kriging technique. Figure **3-6** presents shaded relief maps created from each datasets. The difference between the kriged and TIN-based DEMs is expressed in the faceted appearance of the shaded relief map of the PAMAP lidar dataset when compared to the CZO generated dataset.

In order to test and compare the accuracy of the lidar datasets, RMSD was calculated using the field survey points as control points in the watershed. RMSD for the CZO lidar was slightly larger than the RMSD for the PAMAP lidar with a value of 0.417 m and 0.410 m respectively (Table **3-1**).
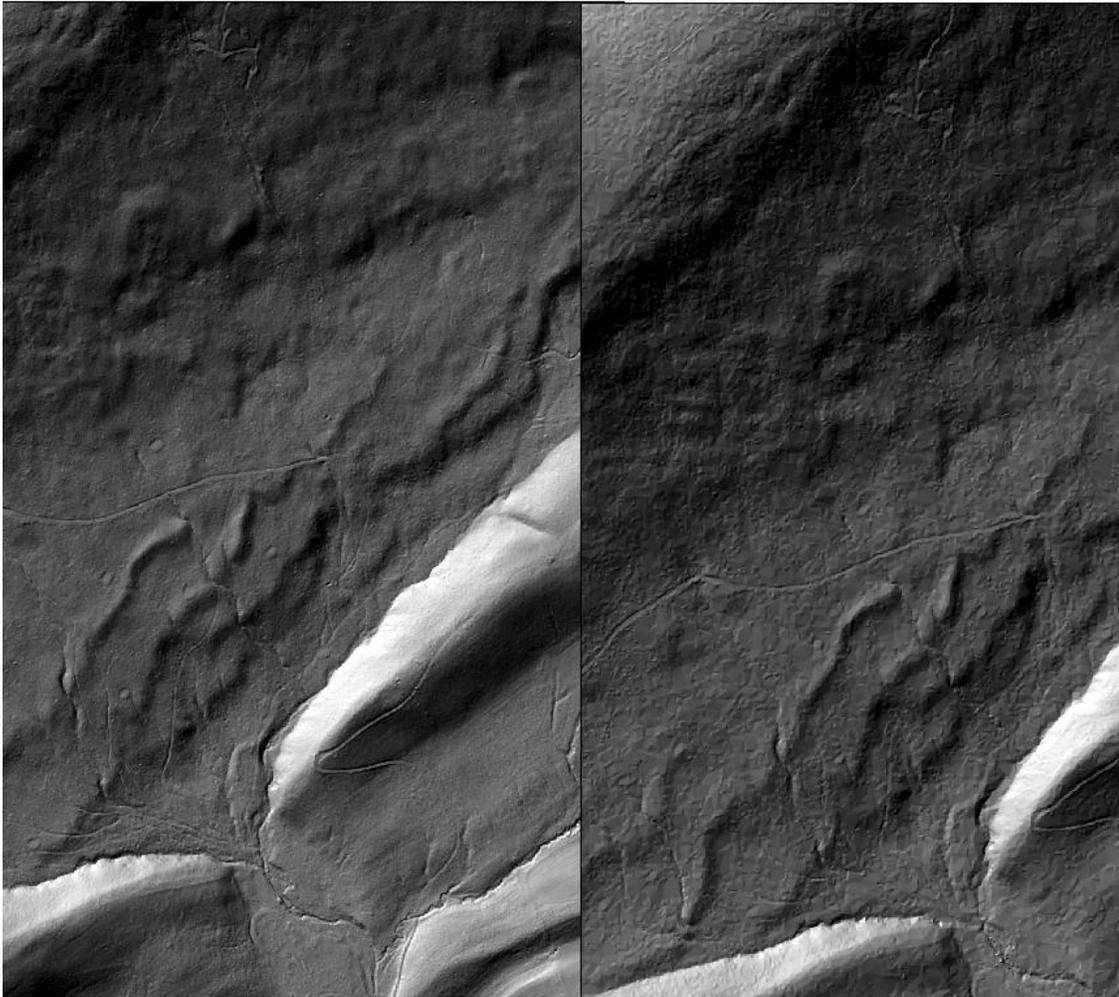
Figure **3-6**: Shaded relief maps generated from CZO lidar (left) and PAMAP lidar (right). The TIN-based algorithm was used for the PAMAP lidar and is visible in the shaded relief map.

Table **3-1**: RMSD for PAMAP and CZO lidar

Lidar Data source	CZO Lidar	PAMAP Lidar
RMSD	0.417	0.410

Errors ranged within about 1.5 meters for both datasets (Table **3-2**). Figure **3-7** shows an example of the error pattern for the PAMAP lidar dataset and CZO lidar dataset. Note the large differences found on the steep hillslope from both datasets. This was in an extremely steep area with slopes approaching 100%, which may account in part for the error. Also notice that the

region directly along the stream channel has a higher error rate for both DEM datasets. These are both areas where expected error rates would be high due to rapidly changing terrain. Although there were some differences between the PAMAP and CZO lidar, in general the errors from both datasets tended to follow the same patterns. The mean difference between surveyed and lidar-derived DEM elevations was about -0.3 meters which means the surveyed elevations were on average about a third of a meter lower than the DEM elevations.

Table **3-2**: Difference between surveyed points and the CZO and PAMAP lidar datasets

Lidar Dataset	Min difference	Max difference	Mean Difference	SD
CZO	-1.686	0.579	-0.311	0. 277
PAMAP	-1.473	0.595	-0.317	0.259
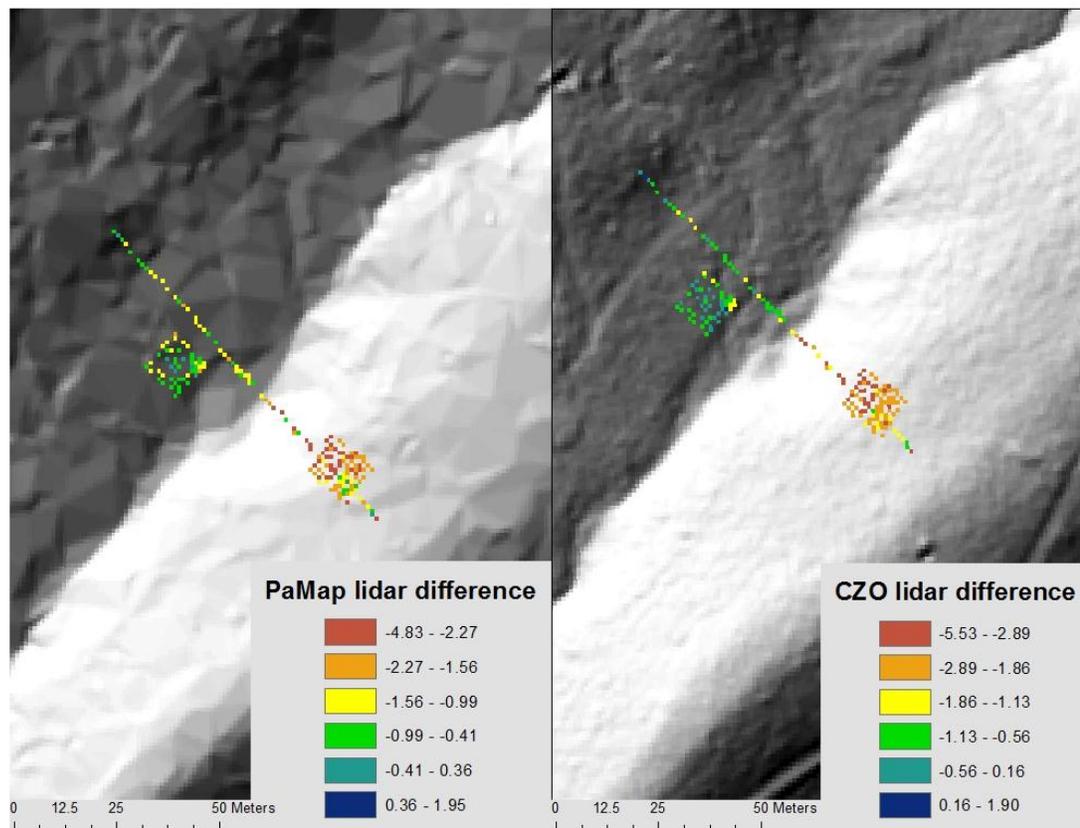
Figure **3-7**: Close up view of two plots and transect. Values represent the difference between surveyed elevations and lidar-derived elevations. Images are overlain over shaded relief maps of corresponding DEMs. Values in the key are in feet to match the original units of the DEMs.

Slope was calculated and compared for surveyed transects, PAMAP DEM, and CZO DEMs (Figure **3-8**). Although there are some localized differences (over distances of 1-2 meters), the slopes were similar between data sets. As an example Figure **3-8** shows a close-up of slopes delineated using the PAMAP lidar, the CZO lidar, and the surveyed points of one transect that range in slope from 0-90%. Transects were placed perpendicular to the steepest slope so slopes calculated from surveyed points would approximately reflect the steepest slope measurements derived from the DEMs using the ArcGIS slope algorithm.
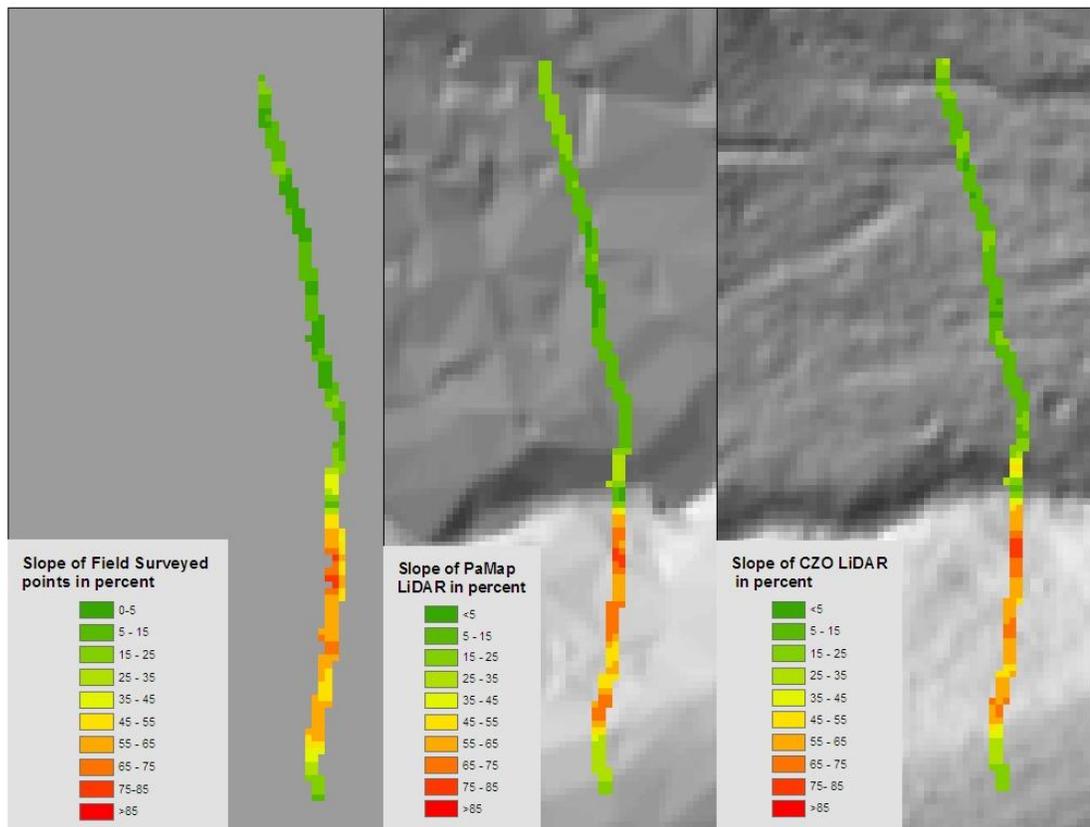
Figure **3-8**: Slope calculated along a transect located at the junction between the steep shale hill and the valley bottom. This is the transect on the southwest section of the watershed.

Roughness Modeling

Due to differences in processing algorithms between the PAMAP and CZO lidar-derived DEMs, the algorithms used to calculate the roughness metrics generated roughness maps for each lidar product that appeared very different at a fine scale, although there were consistencies among methods on the broader scale. In Figure **3-9**, the pit fill metric is shown for the PAMAP and CZO DEMs. Although the CZO lidar had much higher rates of closed depression occurrences than the PAMAP DEM, both datasets presented similar patterns, with higher values of depressions occurring on both tops of ridges and along valley bottoms. Many of these depressions tend to be located along stream areas in both datasets, and there seems to be a clear line across the middle of the slope where the prevalence of these depressions increases. There also appears to be a relationship between closed depressions and slope, with areas with higher slope having lower values of the pit fill metric. Unfortunately, the road across the middle of the watershed also produces many cells that show up as high values with this metric, making it difficult to accurately assess the natural roughness of the watershed.
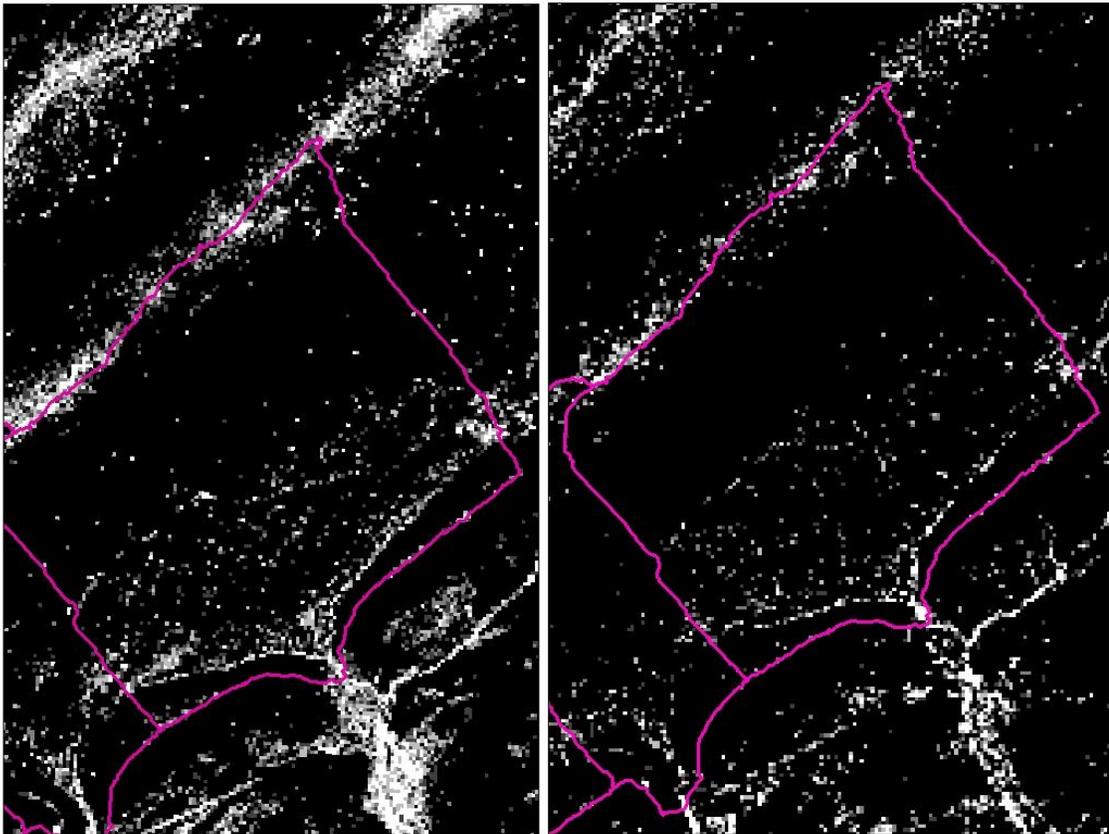
Figure **3-9**: Pit fill metric on 10 x 10 meter blocks from CZO lidar (left) and PAMAP lidar (right).

  The second roughness metric analyzed was standard deviation of curvature over a 5

meter moving window (Figure **3-10**). The results for this metric were very interesting because of

the extremely distinctive pattern that emerged in the CZO lidar dataset. There was a very clear

striping artifact pattern approximately aligned to the dominant slope. When the actual curvature

values and shaded relief maps were analyzed at a fine scale, it became apparent that this striping

was caused by the initial DEM processing, potentially by the kriging algorithm used to generate

the DEM from the initial lidar points. Despite this striping, broad scale patterns in roughness

values were visible from this layer. The PAMAP lidar did not display these striping artifacts and

instead followed the broader roughness patterns of the CZO lidar. Areas of high standard

deviation of curvature are found along steeper rocky slopes and along linear features such as

roads, stream channels, and slope breaks. Features perpendicular to the slope are prominent,

particularly when compared to the standard deviation of slope layer that is shown in Figure **3-11**.
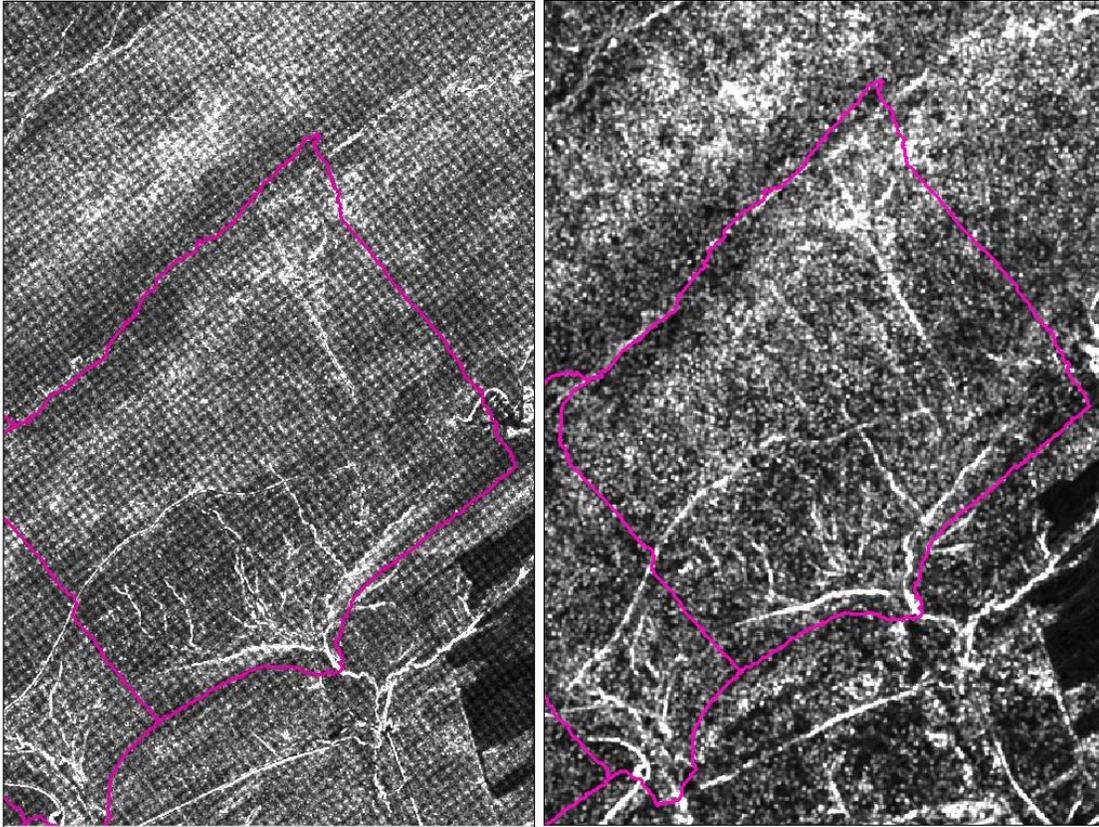
Figure **3-10:** Standard deviation of curvature values for CZO lidar (left) and PAMAP lidar (right)

Patterns in the standard deviation of slope (Figure **3-11**) are similar to the patterns found

in the standard deviation of curvature, with both methods producing very high values along linear

features such as streams and roads. Standard deviation of curvature, however, tends to highlight

features perpendicular to the dominant regional slope, while standard deviation of slope does not.

Using the SD of slope metric, the top of Leading Ridge tends to have low values, while the

steepest portion of Leading Ridge and the valley bottom both tend to have high values.

Particularly in the CZO lidar, small features such as charcoal pits and springs seem to have high

values as well.
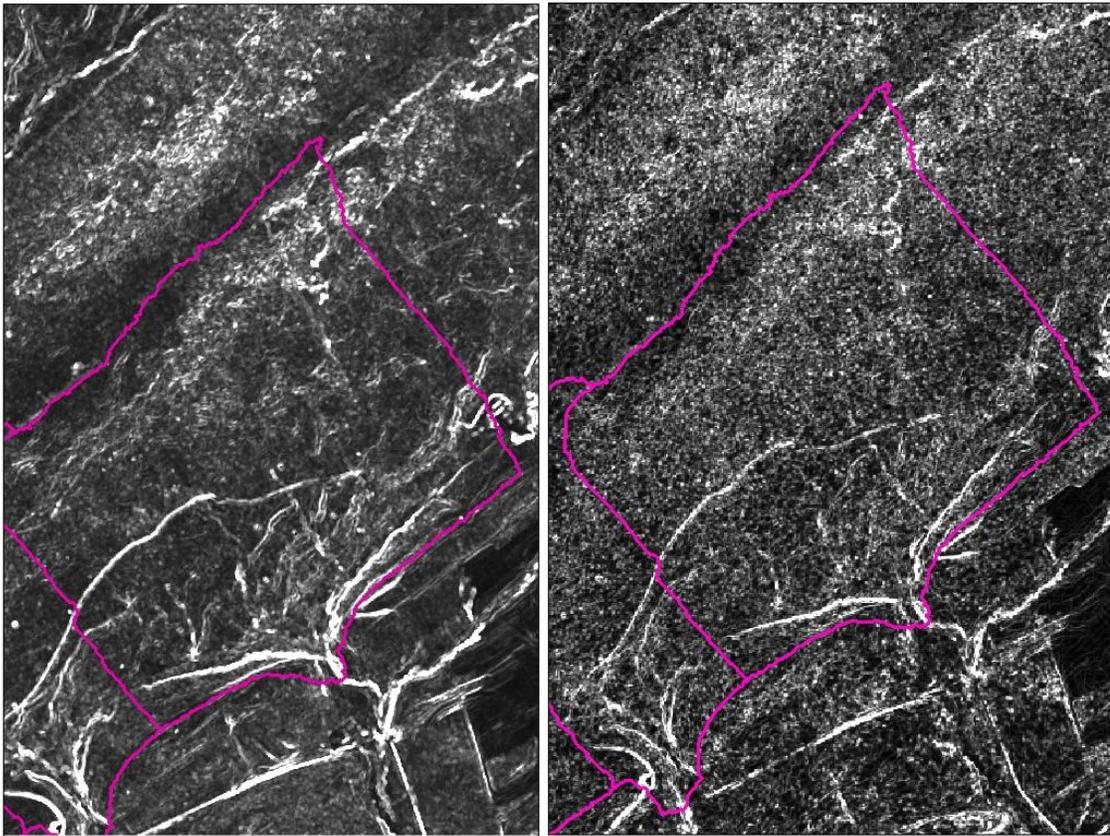
Figure **3-11**: Standard deviation of slope for CZO lidar (left) and PAMAP lidar (right)

The last roughness metric was based on the degree that the local topography differed from the regional topography and was calculated by thinning the DEM and creating a splined surface, then subtracting the original from the splined surface and measuring the absolute value to remove negative numbers (Figure **3-12**). Again, there were differences due to artifacts from algorithms from different interpolation techniques, particularly in the PAMAP lidar this time, but similar patterns were highlighted. High roughness values tended to correspond to areas of high slope. There was an area of low roughness along the top of Leading Ridge and in the mid-slope area.
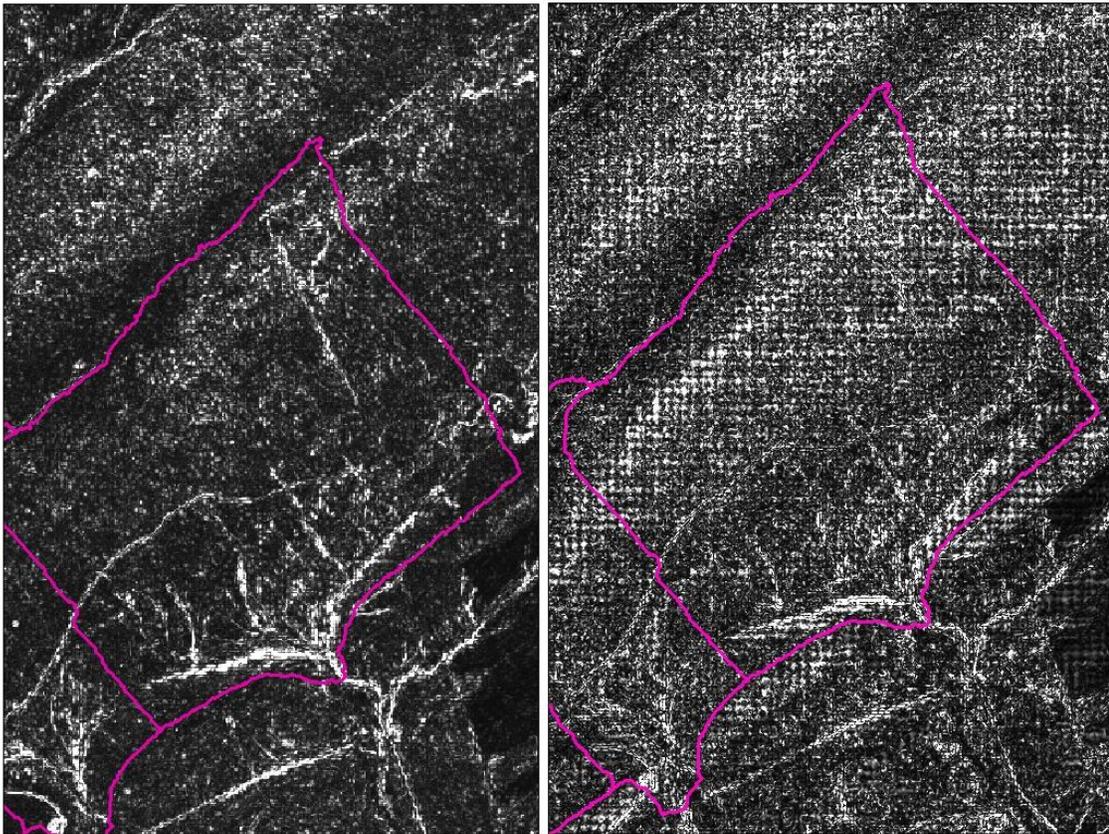
Figure **3-12**: Difference between splined surface and original DEM for CZO lidar (left) and PAMAP lidar (right).

Relating surface roughness to soil, geology, and landforms

There is a broad range of soil, geology, and topography within the Leading Ridge watersheds. Topographically the watersheds range from ridge-top settings with extensive exposures of quartzite and a dry oak heath forest type to a bottom land setting that is home to a rich, bottomland cove forest type (Hornbeck 1962). Since PAMAP and CZO lidar both exhibited similar patterns of roughness, only PAMAP lidar was used for the remainder of the analysis. Roughness metrics generated from the PAMAP lidar were analyzed by landform position to improve understanding of the surface expressions of particular soil types. Landform position was

chosen because it approximately represents both geology and soil, and soil boundaries are not available at a fine enough scale to differentiate the watershed.

The different soils found in the watershed could be expected to have different surface characteristics that could be represented as microtopography. There are five major soil series and/or associations found in the watershed including Hazelton-Dekalb extremely stony sandy loam, Laidig extremely stony loam, Andover extremely stony loam, Buchanan extremely stony loam, and Berks-Weikert shaly silt loam. Hazelton Dekalb, Laidig, and Berks-Weikert are all well-drained soils, while Buchanan and Andover are somewhat poorly drained and poorly drained, respectively. Buchanan, Andover, and Laidig are all colluvial soils or residual soils formed from colluvial material (Lynch and Corbett 1990) and all generally contain a fragipan, a layer that restricts the flow of water and root penetration (NRCS 2011). Because of these properties, areas underlain by these soils tend to have numerous springs, stream channels, seep wetlands, and other hydrologic features that may be represented as a pitted, uneven surface. There were also records of windthrow events that have recently taken place in these areas (Joseph Harding, pers. com.). In addition to these micro-topographic features, there are many areas of the watershed with rocky, talus features close to the surface but since the rocks tend to be smaller than a meter, their surface texture did not emerge using roughness metrics.

By visually analyzing the roughness maps together, it was possible to delineate features along lines of contrasting roughness patterns (Figure **3-13,** Table **3-4a**). For example, the top of the ridge was characterized by having relatively high values for pit fill metric, low values of standard deviation of slope, low values of standard deviation of curvature, and low values for the difference between splined surface and regular surface. The top of the slope exhibited an opposite pattern. The lower slope area was difficult to delineate, but was characterized by a very patchy pattern in all of the roughness metrics, along with a high value for the pit fill metric. The valley bottom contained the highest value for all of the roughness metrics, while the shale hill contained

relatively low roughness values for all of the metrics. This is probably due to the lack of

sandstone fragments present and the lack of hydrologic features on the smaller landform which

results in a smoother surface. The landforms form similar boundaries to the soils, although

Andover and Buchanan soil types are combined to form the valley bottom as there was no clear

difference between them expressed in the roughness metrics. Descriptive statistics were

calculated for each roughness metric per delineated landform and means are shown in Table **3-4b**.

a) *b)*

c) *d)*
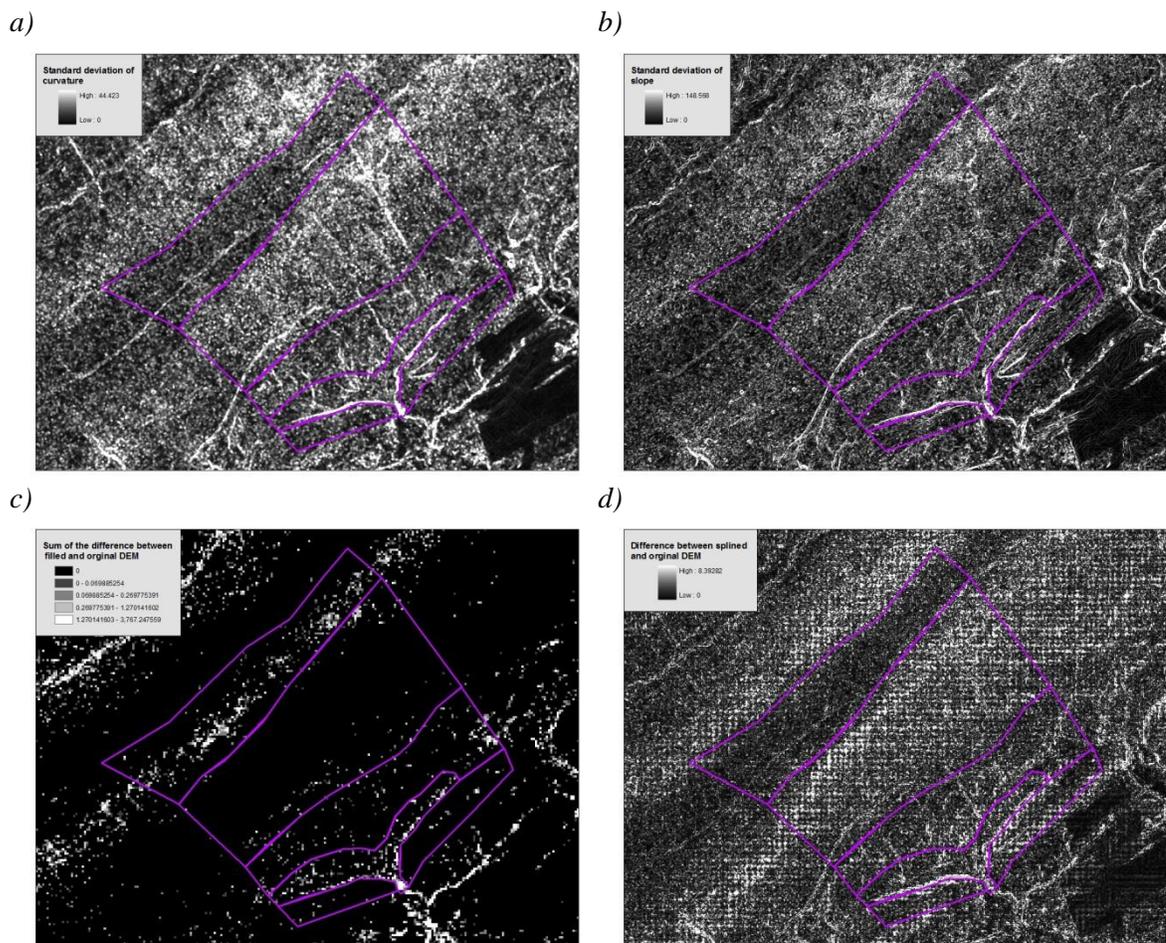
Figure **3-13**: Roughness metrics shown with the boundaries delineated using patterns of roughness. Boundary delineation was complicated by the road running through the middle of the watershed, particularly the pit fill metric.

Table **3:4a** Qualitative interpretive table used for visual interpretation

Roughness metric	Top of Ridge	Top Slope	Lower Slope	Valley Bottom	Shale Hill
SD of curvature (1)	Low	High	Medium	High	Low
Pit fill metric (2)	High	Low	Medium	High	Low
SD of slope (3)	Low	High	Medium	High	Low
Spline (4)	Low	High	Medium	High	Low

Table **3:4b** Mean values of roughness metrics by landform delineated by roughness metrics

Roughness metric	Top of Ridge	Top Slope	Lower Slope	Valley Bottom	Shale Hill
SD of curvature (1)	2.27	3.10	2.50	3.19	2.23
Pit fill metric (2)	0.08	0.01	0.03	0.30	0.02
SD of slope (3)	2.49	3.49	2.91	4.00	2.85
Spline (4)	0.21	0.29	0.29	0.49	0.26

In order to assess the relationship between the PAMAP lidar-derived surface roughness and surveyed surface roughness, the same roughness metrics were calculated from the surveyed micro-plots. Point locations were converted to a TIN and from the TIN to a 0.3 meter DEM to create a raster-based modeled terrain of the site from a field survey. There were no plots placed on top of the ridge due to field limitations. The results of the micro-plot analysis (Table **3-5**) are very different than the results from using the original lidar-derived DEM, particularly with regard to the shale hill plots. These were the roughest plots according to three of the metrics derived from surveyed data, although appearing relatively smooth according to the lidar-derived DEM. These plots were both placed in extremely steep settings which may have affected the resulting surveyed and modeled roughness metrics.

Table **3-5** Mean values of roughness metrics calculated for each micro-plot.

Roughness metric	Upper Slope	Lower Slope	Valley Bottom	Shale Hill
SD of slope	1.31	2.15	2.29	4.72
SD of curvature	5.43	6.51	8.33	11.97
Spline	0.30	0.18	0.18	0.59
Value of Pits	0	0.00031	0.0004	0

Discussion

The PAMAP and CZO lidar-derived DEMs both have 1 m resolutions, and both have been shown to have similar RMSE, but they are still very different DEMs with contrasting surface characteristics. The PAMAP DEM has a much smoother appearance due to the reduced point density and the TIN-based interpolation method used to convert the elevation points to a DEM. The CZO lidar DEM has a fine-scale grainy appearance and appears to more accurately model the landscape due to the increased visibility of surface features and charcoal pits, although the results of the RMSD and transect slopes do not support that conclusion. Although the CZO lidar dataset is not more accurate than the less densely sampled PAMAP lidar (as measured by the RMSD), it displays topographic information and features such as trails, small channels, and charcoal pits much more clearly. This demonstrates that density and algorithm both have important impacts on DEM generation, and that different algorithms may be used on the same terrain depending on the objective. Also, density and algorithm may not be as important to depicting elevation as they may be for other purposes. In this case, there is no difference in file size between these two datasets because both lidar-derived DEMs have the same resolution. There was also no clear improvement in calculation of roughness metrics from the more densely sampled CZO lidar than from the PAMAP lidar. This may indicate that there is no clear advantage to research grade lidar for calculating roughness metrics unless accompanying resolution is also increased such as would be accomplished from a 0.5 m resolution DEM. Initial point density seemed less important than algorithm type used for both DEM generation and roughness calculation. Siska and Hung (2001) found that TIN-based DEM generation and kriging were two of the most accurate methods of generating DEMs as measured by RMSE.

By analyzing patterns of the different roughness metrics as shown in Table **3-4a**, we can start to understand what our different roughness metrics may be showing. Figure **3-14** shows

landforms with a graph of the various roughness values for each landform. For improved

visualization, roughness metrics were converted to relative roughness indices by dividing each

mean roughness by the highest mean roughness value for that metric. These values are shown in

Figure **3-14**. In Figure **3-14**, you can see that the valley bottom has the highest roughness values

for all four roughness metrics. The ridge top has relatively low roughness values except for the pit

fill metric, which was high. Notice that the pit fill metric and the spline both show much rougher

surfaces in the valley bottom than in other sites, while the SD of curvature and the SD of slope

are more similar across formations. The top of slope area has higher relative roughness values for

SD of slope and SD of curvature, while the bottom of slope and shale hills are fairly similar. The

bottom of the slope still has a relatively high value for pit fill metric, and the shale hill has almost

no filled pits.
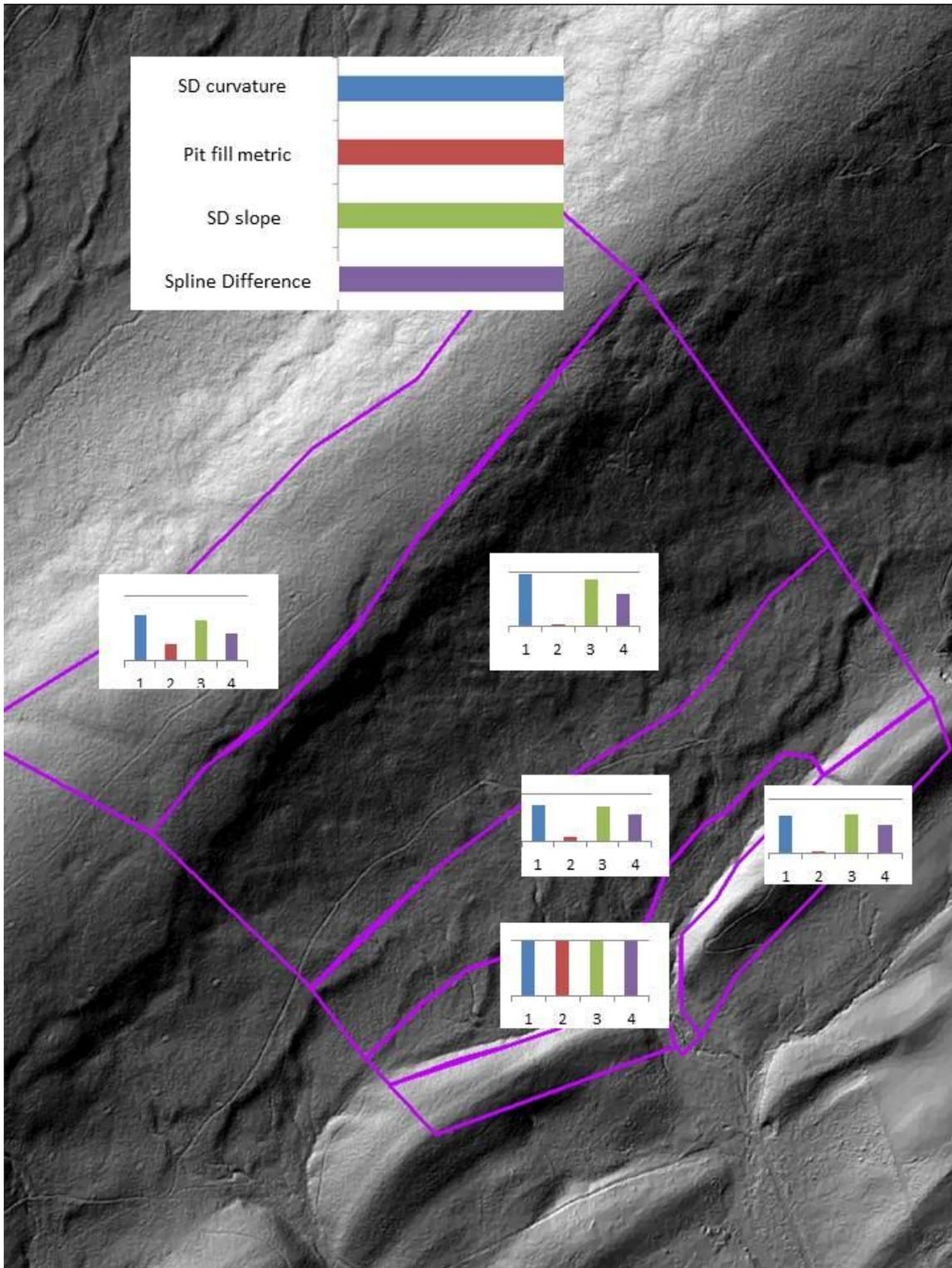
Figure **3-14:** Roughness landforms with bar graphs showing different roughness indices for each landform.

No one roughness metric stood out as being the most effective at modeling all surface roughness or micro-topography, however, using them in combination made landscape interpretation easier. There was wide variation in the results of different micro-topography algorithms as measured by high or low roughness values. There was also a wide variety of interactions between the DEM-generation algorithms and the various roughness algorithm used. Although the different methodologies created mixed effects, they all reflected and responded to similar features on the landscape.

I found that surveyed elevations were approximately 0.3 meters less than the elevations measured by both lidar datasets, with largest differences being present in areas of high slope. This agrees with Spaete et al. (2010) who found highest lidar-derived DEM error rates in areas with high vegetation cover and higher slopes. Tenenbaum et al. (2006) also suggests that tree roots may affect DEM results due to the inability of the processing algorithms to differentiate between roots and he ground. This suggests that in forested settings, particularly densely forested settings such as eastern deciduous forests, there may be error in lidar-derived DEMs caused by vegetation in the understory, coarse woody debris and roots of trees, and leaf litter, even in leaf off conditions. In this study RMSD results were relatively consistent across the watershed, but this could vary in different vegetation or landform settings. More research should be conducted to explore whether this may be due to topography or vegetation.

The top of the Leading Ridge displayed very high values in the roughness metric measuring the difference between the filled and original DEM. This may be in part because of either the rockiness of the terrain or the interaction between the lidar beam and the high density of ericaceous shrubs such as blueberry, huckleberry, and laurel. Even though the lidar was collected during leaf off conditions and blueberry and huckleberry are deciduous, the dense branching patterns of these shrubs could result in erroneous pits being represented on the landscape. Also, there are some evergreen shrubs including mountain laurel on the ridge top.

Soil type has been shown to be an important predictor of vegetation community composition and site quality on a particular site (Horsly et al. 2002, Bailey et al. 2004). Unfortunately, the accuracy of currently available spatial soil datasets is not intended to be utilized at a fine scale in forested settings. The USDA Soil Survey for the area of Leading Ridge was created at a scale of 1:24,000 (NRCS), thus the use of either lidar data set to compare to soils should also occur at this scale. I found differences in the boundaries of soil series and topographic formations when using lidar-derived DEMs which reflects the mapping scale differences between soil polygons as delineated from SSURGO data and roughness metrics. By incorporating roughness metrics into our analysis, we may be able to refine our soil mapping polygons. For example, the presence or absence of a fragipan could be expressed on the landscape as an increase in local features created from fluvial geomorphic processes such as intermittent channels, springs, and seeps (Ciolkosv et al. 1979). There could also be increased pit and mound topography caused by increased wind-throw due to more shallow rooting depth in these sites (also due to a fragipan). This could be one way to identify these features which can have important impacts on site activities such as road planning and building.

Pennsylvania's landscapes have experienced substantial land use and topographic change since European settlement over 300 years ago. This is particularly true for Leading Ridge due to the long history of management and research in the watershed (Lynch and Corbet 1990). Many of these methods for delineating roughness are particularly effective at highlighting cultural features on the landscape such as roads, skid trails, and charcoal pits. This makes it challenging to delineate natural patterns of micro-topography due to the difficulty of removing these features prior to modeling. Although creating challenges for utilizing lidar elevation data to delineate natural patterns, soils, and features, identifying these features could also be a potential use for lidar-derived roughness metrics.

It was interesting to note that despite correspondence among DEM roughness, the roughness calculated from surveyed plots did not correspond to the roughness derived from the DEMs. This may indicate that the scale of the lidar, although extremely fine, is still not sufficient for modeling surface micro-topography as would be represented using a topographic survey, or that the features represented from the topographic survey are not the same as those modeled from the DEMs. Features that are characterized as rough using these roughness metrics are larger features such as major slope breaks, stream channels, roads and trails, and other man-made features. Also, surveyed roughness values were generally higher than DEM modeled values which could be due to the fact that when surveying, the purpose of the survey was to maximize roughness by recording elevations at the extreme boundaries of features to record as accurate of a surface as possible. DEMs on the other hand, are simply capturing an elevation at a regular gridded location. Textural features such as rockiness are simply on too small of a scale to be detected and measured using a 1m DEM. Much of the modeled roughness on Leading Ridge seems to be from intermittent and ephemeral flow channels which are very dense in the lower portion of the watershed.

Conclusions

When compared to the surveyed data, CZO research grade lidar DEM and the PAMAP DEM had a RMSD of approximately 0.4 m, with surveyed elevations being on average 0.3 m lower than DEM-modeled elevations. This could be both because of the nature of the ground survey and because some points used in the generation of the DEM are actually vegetation points such as roots, the bottoms of tree boles, and dense leaf litter. There were also differences between roughness values between surveyed roughness values and roughness values generated from DEMs, not just in magnitude but in what types of features emerge as rough. This probably

indicates that roughness metrics generated from lidar are not the same as metrics assessed on a topographic survey.

Different roughness metrics have been analyzed using the 1m resolution DEM. These include standard deviation of elevation, standard deviation of slope, total value of pitted cells, and total range of elevation over different areas. No one method stood out as the most effective, but all methods highlighted different types of roughness and features. When viewed simultaneously, classes of features or landtypes were visible and associated with landforms. By incorporating roughness metrics into analyses using soils and geology, it may be possible to improve delineation of soil types and mapping units or other topographic features important in vegetation or hydrologic modeling.

References

Andrle, R. and Abrahams, A. D. 1989. Fractal techniques and the surface roughness of talus slopes, Earth Surface Processes and Landforms, 14, 197-209.

Anderson, S.P, Bales, R.C, Duffy, C.J. 2008. Critical zone observatories: building a network to advance interdisciplinary study of Earth surface processes. Mineralogical Magazine 72: 7–10.

Bai M.J., Xu D., Li Y.N., Li J.S. 2005. Evaluation of spatial and temporal variability of infiltration on a surface irrigation field. J Soil Water Conserv 19:120–123 in Chinese.

Bailey, S. W., S. B. Horsley, R. P. Long, and R. A. Hallett. 2004. Influence of edaphic factors in sugar maple nutrition and health on the Allegheny Plateau. Soil Science Society of America Journal 68:243–252.

Beatty, S.W. 1984. Influence of microtopography and canopy species on spatial patterns of forest understory plants. Ecology 65: 1406–1419.

Bolstad, P.V., Swank, W. and Vose, J., 1998. Predicting Southern Appalachian overstory vegetation with digital terrain data. Landscape Ecology 13, pp. 271–283.

Brantley, S. L., White, T.S., White, A.F., Sparks, D., Richter, D., Pregitzer, K., Derry, L., Chorover, J., Chadwick, O., April, R., Anderson, S., Amundson, R. 2006. Frontiers in Exploration of the Critical Zone: Report of a workshop sponsored by the National Science Foundation (NSF), October 24-26, 2005, Newark, DE, 30p.

Ciolkosz, E.J., Petersen, G.W. Cunningham, R.L. and Matelski, R.P., 1979. Soils developed from colluvium in the Ridge and Valley area of Pennsylvania. Soil Sci., 128: 153-162.

Ciolkosz, E.J., Cronce, R.C. and Sevon, W.D. 1986. Periglacial features in Pennsylvania. Pa. State Univ., Agron. Ser. 92, 15 pp.

Cleland, D.T., P.E. Avers, W.H. McNab , M.E. Jensen, R.G. Bailey, T. King, and W.E. Russell. 1997. National hierarchical framework of ecological units. pp.181-200. In: M.S. Boyce and A. Haney (eds.). Ecosystem management. Yale University, New Haven.

Dikau, R. 1989. The application of a digital relief model to landform analysis in geomorphology. Three dimensional applications in geographical information systems. J. Raper, ed. Taylor and Fancis: New York. 51-77.

Dunne, T., Zhang, W. and Aubry, B.F. 1991. Effects of rainfall, vegetation and microtopography on infiltration and runoff. Water Resour. Res. 27: 2271–2285.

ECOMAP. 1993. National hierarchical framework of ecological units. USDA Forest Service, Washington, D.C.

Enoki T. 2003. Micro-topography and distribution of canopy trees in a subtropical evergreen broad leaved forest in the northern part of Okinawa Island, Japan. Ecological Research 18: 103–113.

Frankel, K.L., Dolan, J.F., 2007. Characterizing arid-region alluvial fan surface roughness with airborne laser swath mapping digital topographic data. Journal of Geophysical Research – Earth Surface 112, F02025.

Franklin, J., 1995. Predictive vegetation mapping: geographical modeling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geogr. 19: 474–499.

Glenn, N. F., D. R. Streuker, D. J. Chadwick, G. D. Thackray, and S. J. Dorsch. 2006. Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity. Geomorphology. 73:131– 148.

Govers G, Takken I, Helming K. 2000. Soil roughness and overland flow. Agronomie 20: 131–146.

Hengl, T., Evans, I. S. 2009. Mathematical and digital models of the land surface. In: T. Hengl and H.I. Reuter, Editors, Geomorphometry: Concepts, Software, Applications, Elsevier, Amsterdam. pp. 31–63.

Hornbeck, J. 1962. Hydrologic analysis of the three Leading Ridge experimental watersheds. M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Horsley, S.B., R.P. Long, S.W. Bailey, R.A. Hallett, and P.M.Wargo. 2002. Health of eastern North American sugar maple forests and factors affecting decline. North. J. Appl. For. 19:34–44.

Huang, C.H., 1998. Quantification of soil microtopography and surface roughness. In: Baveye, P., Parlange, J.Y., Stewart, B. (Eds.), Fractals in Soil Science. CRC Press, Boca Raton, FL.

Huang, C. and Bradford, J.M., 1992. Applications of a laser scanner to quantify soil microtopography. Soil Science Society of America Journal 56: 14–21.

Kamphorst, E., Jetten, V., Gue´rif, J., Pitka¨nen, J., Iversen, B., Douglas, J., Paz, A. 2000. Predicting depressional storage from soil surface roughness. Soil Science Society of America Journal 64:1749–1758.

Kong, N. 2006. Topographically-based landscape-scale ecological mapping in Pennsylvania. PhD. Dissertation. Intercollege Graduate Degree in Ecology. Penn State University, University Park, PA.

Kraus K,N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing 53: 93–203.

Lavee, H., Pariente, S. 1995. Spatial variability of soil properties along a climatic gradient: guide to Judean Desert Climatalogical Gradient Excursion. IGU GERTEC Commission meeting, Israel, May 1995, pp. 16–30.

Lynch, J. A. and E. S. Corbett, 1990. Evaluation of Best Management Practices for Controlling Nonpoint Pollution from Silvicultural Operations. Water Resources Bulletin 26:41-52.

Maune, D. F., editor. 2007. Digital elevation models and technologies and applications: the DEM user manual, 2nd Edition. The American Society for Photogrammetry and Remote Sensing. Bethesda, MD.

McGlynn, B. L., and J. Seibert, Distributed assessment of contributing area and riparian buffering along stream networks, Water Resour. Res., 39: 1082.

McKean, J., and J. Roering. 2004. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry, Geomorphology, 57, 331– 351.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A. 1993: Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57: 443-52.

Myers, W.L. 2000. Landscape scale ecological mapping of Pennsylvania forests. Research report ER2002-2, Environmental Resources Research Institute, University Park, PA.

Naiman RJ, Decamps H 1997. The ecology of interfaces: riparian zones. Annual Review Ecological System 28:621–658.

NRCS. 2009. Soil Survey Geographic (SSURGO) Database. USDA Natural Resources.

Pollock, M. M., Naiman, R.J., Hanley,T.J.1998. Plant species richness in riparian wetlands — a test of biodiversity theory, Ecology 79: 94–105.

Robinson, G. R. 1959. Tree mortality and defects on the Leading Ridge watershed number one. M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Rumbaitis del Rio C.M. 2006. Changes in understory composition following catastrophic windthrow and salvage logging in a subalpine forest ecosystem. Can. J. For. Res.36:2943-2954.

Schaetzl, R.J., Burns, S.F., Johnson, D.L. and Small, T.W., 1989. Tree uprooting: review of impacts on forest ecology. Plant Ecology. 79 : 165–176.

Schultz, C. H. editor. 1999. The Geology of Pennsylvania. Geologic Survey Special Publication 1.

Shields, R R. 1966. A shallow seismic refraction study of the soil mantle and bedrock configuration of Leading Ridge Watershed Two M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Siska, P. I. Hung. 2001. Assessment of Kriging Accuracy in the GIS Environment. Annual ESRI International User Conference, San Diego, CA, 2001.

Spaete, L. P., Glenn, N. F. Derryberry, D. R., Sankey, T. T. Mitchell , J. J., Hardegree , S. P. 2010. Vegetation and slope effects on accuracy of a LiDAR-derived DEM in the sagebrush steppe, Remote Sensing Letters, 2:4, 317-326.

Su, J.G. *&* Bork, E.W. 2006. Influence of vegetation, slope and LiDAR sampling angle on DEM accuracy. Photogramm. Eng. Remote Sens. 72*:* 1265–1274.

Tenenbaum, D.E., L.E. Band, S.T. Kenworthy, and C.L. Tague. 2006. Analysis of soil moisture patterns in forested and suburban catchments in Baltimore, Maryland, using high-resolution photogrammetric and LIDAR digital elevation datasets. Hydrol. Proc. 20:219–240.

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

Chapter 4

Using lidar to improve classification of forested geomorphic landforms in the Ridge and Valley Physiographic Province of Pennsylvania

Abstract

Technology is changing the way scientists see the landscape, with lidar data providing information never before available. Scientists in many fields are looking for concise, effective methods to predict vegetation by classifying landscapes based on digital elevation model-derived terrain metrics such as slope, aspect, curvature, and topographic indices. With newly available lidar-derived elevation data, the accuracy and resolution is greatly improved but scale is so dramatically different that new approaches should be used to classify topographic features. By identifying lidar-derived patterns of curvature, major gradients impacting vegetation in a catchment including water accumulation, soil characteristics and nutrient availability can be summarized into one parsimonious metric. We applied this approach in the Ridge and Valley region of central Pennsylvania, in the broader basin encompassing the Shale Hills Critical Zone Observatory. By classifying the watershed into four dominant recurring landforms using patterns of lidar-derived curvature data, dominant vegetation communities and forest structures were successfully predicted and confirmed using multivariate statistical methods. When applied to a broader area encompassing a forested region containing several repeating ridges, the same landforms were delineated and were shown to support corresponding forest community types.

Introduction

Technology is changing the way scientists see the landscape, with new data providing accuracy never before available. Scientists in many fields have been looking for concise, effective methods to predict vegetation by classifying landscapes based on digital elevation model (DEM) derived terrain metrics such as slope, aspect, curvature and topographic indices (Dikau 1989, Moore et al. 1993, ECOMAP 1993, Cleland et al. 1997, Franklin 1995, Myers et al. 2000). Understanding landscapes and their interactions is integral to ecosystem management due to the coupled relationship between terrain and vegetation (Devlin et al. 2001.) Landscape classification can also inform managers in their goals of identifying rare or sensitive ecosystems or maintaining a mosaic of habitat types (Zenner et al. 2010). With recently available lidar-derived elevation data, the accuracy and resolution of terrain data is greatly improved but scale is much different, meaning new approaches could be utilized to classify topographic features.

Foresters and ecologists have long identified fundamental relationships between terrain and vegetation (Bowersox and Ward 1972, Pfister et al. 1977, Franklin 1995, Noss 1987). Terrain impacts physical factors such as temperature, water availability, nutrient availability, and light availability. Particularly in mountainous regions, plant communities exhibit predictable transitions as elevation or aspect change. In eastern forested ecosystems however, the relationship between elevation and vegetation communities is not so straight forward. Since multiple terrain variables are intrinsically linked and can often affect many physical factors simultaneously, it is often difficult to tease out the contributing effects of multiple variables. Bolstad et al. (1998) found no strong relationships between vegetation and slope or aspect at Coweeta Hydrologic Lab in the Southern Appalachians. Soil data, which has been used effectively for many classification systems is not described in detail for forested mountainous areas in Pennsylvania due to lack of an extensive agricultural land use history in many of these areas (Carter and Ciolkosv 1991,

Devlin et al. 2001). With improved elevation data, we may be able to improve our classification systems by identifying patterns of landforms that are influencing vegetation.

In order to effectively engage in ecosystem management, forest managers need classification systems that are detailed enough to include changes in forest community, structure, and successional trajectory (Zenner et al. 2010). In Pennsylvania, these detailed classification systems do not currently exist. The Pennsylvania Department of Natural Resources (DCNR) Bureau of Forestry manages substantial amounts of land in Pennsylvania and has adopted a landscape classification system based on the National Hierarchy of Ecological Units (ECOMAP 1993) and landscape units defined by Pennsylvania DCNR's Bureau of Topographic and Geologic Survey Landform Mapping project (Sevon 1998). These efforts focused on identifying landtype associations (LTAs) and ecological landtypes (ELTs). An ecological landtype is defined (ECOMAP 1993) as a "continuous sector of terrain that exhibits a relatively uniform influence on the landscape context" with landtype associations being composed of a combination of complementary ELTs that are spatially adjacent. Allegheny National Forest, the only national forest in Pennsylvania, defined LTAs based on general topography, geomorphic processes, surficial geology, soil families, potential natural communities, and local climates (Moriarity 1996). Their LTAs range in size from over 1000 ha to over 30,000 ha in size. The 207,603 ha comprising the Allegheny National Forest have been delineated into 160 ELTs (Moore et al. 2006).

Kong (2006) delineated LTAs and ELTs for the entire state of Pennsylvania based on a hierarchical methodology of topographic delineation that started with a system of cuplands and caplands established by Myers (2000). Cuplands are areas where the dominant processes shaping the landscape are deposition of water and sediments, where caplands are landscapes dominated by rapid runoff and erosion (Kong 2006). Regions were further subdivided using a top-down approach predominantly using a statewide 90 m DEM that resulted in over 10,000 separate LTAs

that ranged in size from 40 to 2,000 ha. ELTs were delineated on a statewide basis based on topographic and soil data. With the improvement of elevation data now available, we may be able to improve our fine scale landform delineation based on topographic data. The scale chosen for this new landform classification is finer than that of landtype association (LTA) with most landforms delineated within this context fitting within one or two LTAs. For the purpose of this research, we were interested in identifying landforms in the Ridge and Valley province of a scale that resulted in changes in vegetation community or resulted in a distinct change in forest structure.

Recently, advances in computing and GIS have led to multiple classification systems that use automated methodology to assign classes. The first group uses automated DEM analysis to create classes based on terrain features using moving windows (Dikau et al. 1989). A second landform delineation method is derived from multivariate classification systems based on combinations of terrain variables, or supervised classification where an expert assigns classes (Niemann and Howes 1991). More recently, automated classification methods using fuzzy hierarchical clustering algorithms have been used (MacMillan et al. 2000, Schmitt and Hewitt 2004). These systems can be very effective at modeling mountainous terrain where major gradients in water, temperature, and nutrient availability exist, as well as in some areas of relatively flat terrain where elevation gradients coincide with soil differences or glacial features. In forested ecosystems of eastern Pennsylvania however, where water is abundant and elevation differences across major landforms are usually not greater than 300 m, topographic causes of vegetation community variability are more difficult to classify than simply a combination of elevation and aspect. Also, landform scale and position are difficult to incorporate into automated algorithms (Schmidt and Hewitt 2004). Perhaps most importantly, expert knowledge is not considered in order to limit user bias and unpredictable replication (Kong 2006). For these reasons, an automated methodology was not used for this study.

In the Ridge and Valley Province, geomorphology is dominated by unique patterns and processes. Differential erosion augments fluvial geomorphology processes to create a terrain with strong gradients of vegetation structure and community along subtle changes in slope and landform position. The trellis drainage patterns of the province coincide with small parallel drainage networks that are driven by patterns in geology and folding. Rock formations are often on end with steep dips, particularly in ridge areas (Shultz 1999). These patterns are particularly difficult to capture through automated methodology because common topographic variables such as slope, elevation, and aspect may not have a noticeable effect on vegetation due to the overriding influence of underlying geology and soil. Almost all remaining landscape scale forested areas in the Ridge and Valley Province are found on ridges, so if forest classification is the goal, techniques must be developed to differentiate landforms that make up the larger, more dominant ridges. Since topographic structures are driven by changes in parent material, effects of terrain can be compounded by different soils and corresponding mineral contents (Ciolkosv et al. 1990). The dominant ridge and valley forming rock formations tend to occur repeatedly in the same order, therefore classification schemes developed in one region should repeat throughout the region. The study watershed displays a very classic pattern of a major ridge underlain by resistant sandstone with a minor ridge of less resistant shale in the foreground.

With the improvements in elevation data now available, we chose to use curvature and patterns of curvature to differentiate landforms. Curvature describes the rate of change of slope both down a hillslope perpendicular to contour lines (profile curvature) and along contour lines (plan curvature), with both curvatures relating to movement of water and other materials along a slope (Moore et al. 1993, McKenzie and Austin 1993). Curvature has been effectively utilized in geomorphometry and terrain analysis for many years (Wilson and Gallant 2000, Olaya 2009). Pennock et al. (1987) described a classification system based on curvature and slope used to relate soil properties to terrain. Convex areas tend to be divergent with water, nutrients, and soil

being removed from the area, while concave areas tend to be convergent areas. Since these processes are interdependent, by measuring curvature we can take advantage of one parsimonious metric that correlates with the movement of water and other materials that may affect vegetation (Moore et al. 1991). Gessler et al. (1995) found that upslope mean curvature modeled soil properties more effectively than other terrain metrics. The artifact contour errors in previously available DEMs made classifying patterns of curvature on a fine scale problematic, but by using new lidar-derived digital elevation data, curvature can be expressed much more accurately. We classified the ridge terrains into four recurring landform classes based on patterns of curvature.

A vegetation community based analysis was conducted to explore the relationship between landform and vegetation. Pennsylvania forests have been classified into vegetation communities (Fike 1999) in collaboration with the Pennsylvania DCNR Bureau of Forestry, Western Pennsylvania Conservancy, and the Nature Conservancy. Fike identified and described 23 terrestrial and 9 palustrine community types in Pennsylvania. Vegetation communities were used in collaboration with measured vegetation structure metrics to evaluate the idea that delineated landforms can inform vegetation patterns. Objectives of the study included classifying the landscape into repeating landforms that correspond to vegetation patterns and expanding that classification to include a landscape scale area of a forested region in the Ridge and Valley Province in Pennsylvania.

Methods

Study Area

This study took place in two regions exhibiting similar topography approximately 30 km apart. The first study area, Leading Ridge Watershed One, is a southwest to northeast trending 119 ha watershed located in the Ridge and Valley Province of central Pennsylvania (Figure **4-1**). The elevation of Leading Ridge Watershed One ranges from 260 m at the mouth of the watershed to 512 m at the top of Leading Ridge, which forms the north western border of the watershed. The hydrologic network of the watershed tends to form a trellis pattern. The geologic formation underlying the watersheds consists of deeply dipping strata ranging from resistant Tuscarora quartzite and sandstone at the top of the watershed to less resistant Rose Hill shale that comprises the valley area (Lehman 1962) (Figure **4-2**). This terrain is consistent with that which makes up most of the Ridge and Valley Province, which is generally characterized by canoe-shaped valleys and long, linear, parallel ridges formed by differential erosion (Schultz 1999). Valleys tend to be underlain by limestone, dolomite, and shale, while ridges are formed by resistant sandstone and quartzite. Ridges tend to be extremely steep and rocky and the topography is generally well-drained.
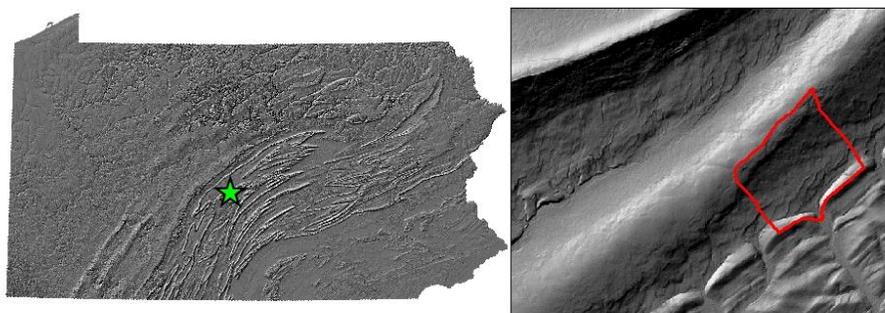
Figure **4-1**: Approximate location of the Leading Ridge Study area in the Ridge and Valley Province in Pennsylvania.
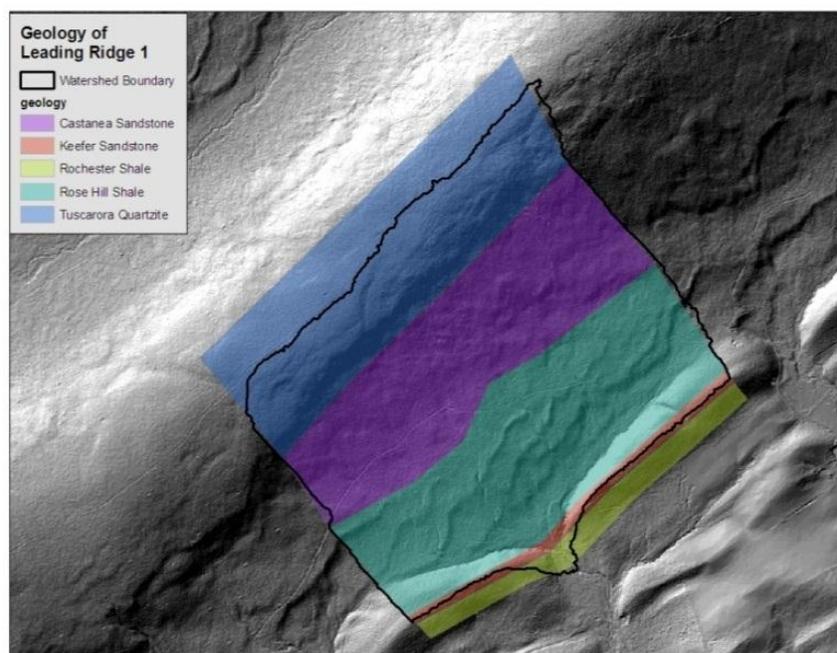
Figure **4-2**: Geology of the Leading Ridge watershed One (adapted from Shields 1966)

The watersheds contain a mature oak/hickory forest approximately 100 years of age. This area has been utilized extensively for charcoal production prior to 1900, and charcoal pits can still be found scattered throughout the site (Robinson 1959). The southern section of the watershed was probably grazed until the 1930s, when farmers were paid to vacate their land. A small salvage cut was conducted in 2006 due to a blow-down event that occurred on the site, and the watersheds experienced substantial gypsy moth mortality over the past 20 years (pers. com. Joseph Harding 2009).

This site was chosen in part because of the long rich history of being used as an experimental watershed and the vast hydrologic and other datasets that currently exist for the site. The Leading Ridge watershed research areas were established in Penn State's Stone Valley Experimental Forest of Central Pennsylvania in 1959 as paired watersheds to study the hydrologic response of different forest practices (Lynch and Corbett 1990). This watershed is also within the Shaver's Creek watershed which is being used as part of the Susquehanna/ Shale Hills

Critical Zone Observatory (CZO), an interdisciplinary observatory toward quantitatively

predicting creation, evolution, and structure of regolith as a function of the geochemical,

hydrologic, biologic, and geomorphologic processes operating in a temperate forested landscape

(Anderson et al. 2008). The Susquehanna/Shale Hills CZO is one of six watershed-scale

observatories supported by the National CZO project.

The second study area is located in the Seven Mountains region of Rothrock State Forest,

encompassing about 2600 ha. Rothrock State Forest covers an area dominated by rugged ridges in

the Ridge and Valley region of central Pennsylvania. This area was chosen to apply the results of

the Leading Ridge case study to a broader scale region of the Ridge and Valley Province.

Lidar DEM Data Sources

Lidar data was collected as part of the PAMAP lidar program in 2007 during leaf-off

conditions. Post spacing for the lidar returns used to generated the 1 m DEM was 1.4 meters, and

the resulting DEM had a target vertical RMSE of 18.5 cm in open areas and 37 cm in vegetated or

forested areas. Points were first classified as either ground or non-ground points, with ground

points being thinned down to create a TIN that fit the final specifications by an independent

vendor, BAE Systems. Using proprietary methods, a 3.2 foot resolution DEM was produced

using the TIN, along with 2 ft contour lines and break lines. All finished products were checked

for quality and accuracy (PAMAP LiDAR QAQC report 2007). For most modeling, the 1 m

DEM was thinned using the nearest neighbor method in ArcGIS 9.3 to produce a 10 m resolution

DEM.

Several studies have shown that 10 m resolution data is the appropriate resolution for

studying landforms (Kienzle 2004, Zhang and Montgomery 1994). It is important to note the

difference between traditionally available photogrammetrically-derived elevation datasets that

have been used for the majority of previously available research and newer lidar-derived datasets. Commonly used photogrammetrically derived datasets such as the National Elevation Dataset (NED) have been shown to have vertical RMSE on the order of 3 -10 m (Maune 2007), while even relatively low density lidar such as that which is available statewide in Pennsylvania has a reported RMSE of about 30 cm, even in forested settings. The difference is apparent when one views shaded relief maps of Leading Ridge with a 10 m NED dataset and a 10 m lidar-derived dataset (Figure **4-3**). In particular, in the NED shaded relief map artifact contour errors are visible on the ridge top and dominant structures on the hillslope are not shown. These errors can have strong effects on derived curvature and patterns of curvature across landforms.

a)   *b)*
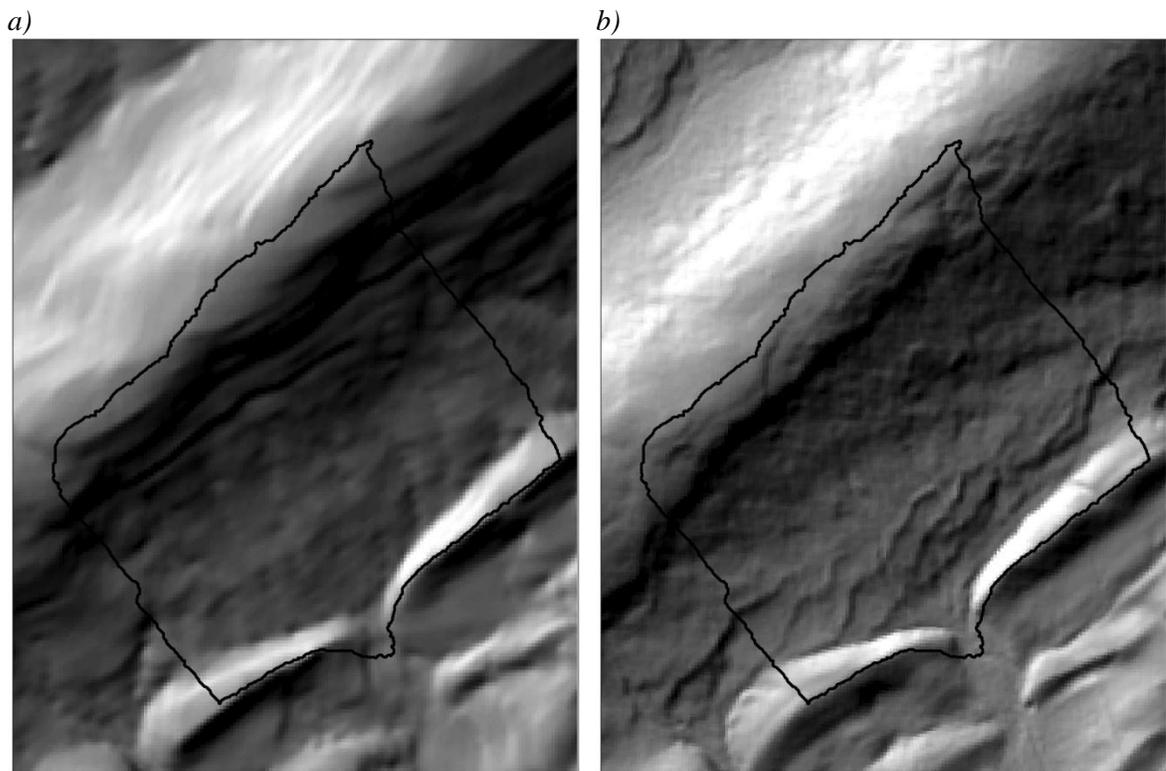
Figure **4-3**: Difference between a) photogrammetrically derived 10m hillshade and the 10m lidar-derived hillshade. Note the contour errors on top of the major ridge, along with lack of detail on the lower slope area.

Curvature Modeling

In order to delineate dominant landform features, curvatures were calculated for a resampled 10 m lidar-derived DEM. Plan and profile curvature were calculated using the curvature tool in ArcGIS 9.3 (Zeverbergen and Thorne 1987). These curvatures were smoothed using a 3 x 3 mean moving window to improve visual interpretation of dominant features. Both smoothed plan and profile curvatures were reclassified using the quantile method into three equal classes of curvature: concave, straight, and convex. Reclassified plan and profile curvatures were combined to create nine curvature classes based on unique combinations of plan and profile curvature (concave/concave, concave/convex, etc.). Using maps of curvature classes, distinctive landforms were delineated that captured repeating patterns of curvature. Four features were identified and named, including rounded ridges, scalloped slopes, rock ridgelets, and hidden hollows, all of which display a characteristic pattern of curvatures.

- Rounded ridge—broad profile curvature is convex or straight, undulating plan curvature, underlain by sandstone.

- Rock ridgelet—narrow, steep small ridge convex in both profile and plan curvature. Underlain by shale, soils primarily Berks-Weikert association.

- Hidden hollow—small hollow formations distinguished by being concave in both plan and profile curvature. Usually found between rounded ridges, or between rounded ridges and rock ridgelets. Often underlain by shale and Buchanan and Andover soils.

- Scalloped slope—side slopes of rounded ridges displaying undulating pattern of curvatures. Lacking clear divergent or convergent flow features.

These features were identified using the following methodology: the first features identified on a landscape were the rock ridgelets, the small convex features formed from shale. In order to delineate these features from the rest of the landscape, patches of convex/convex

curvature were identified that formed contiguous areas of at least 2 hectares that were more rounded than linear in nature. There may be isolated patches of other curvature patterns within these features, but the boundaries should be drawn along the first patch of concave/concave curvature that is reached. Each patch of rock ridgelet should be between 2 and 5 hectares and each patch should be surrounded by concave/concave curvature.

After rock ridgelets were delineated, hidden hollows were identified by isolating the linear concave/concave areas that make up headwater stream networks. These formations are narrow, long, linear connected features approximately 50 meters across, and they are composed almost entirely of concave/concave curvatures. These features are adjacent to either scalloped slopes or rock ridgelet, depending on the curvature patterns of the adjacent formation.

Scalloped slopes can be delineated based on their evenly distributed heterogeneity. All nine curvature classes are found at almost equal ratios. The lower slope end can be delineated when the curvature class becomes concave/concave over a continuous or near continuous area. The upslope of the scalloped slope can be delineated when concave/concave patches of curvature cease being found and over 95% of the area is covered by convex or straight plan curvature. It may be easier to delineate the scalloped slopes by identifying rounded ridges by their convex and straight plan curvature and putting in a boundary where the first patch of concave/concave curvature appears.

Rounded Ridges can be separated from scalloped slopes due to the lack of concave/concave curvature class. There may be isolated disconnected patches, but they are not found regularly as they are on the scalloped slope. There is generally a distinctive patch on the slope where concave/concave pockets begin forming which distinguish the top of the scalloped slope. Figure **4-4** shows a zoomed in area of Leading Ridge Watershed One to demonstrate the methodology.
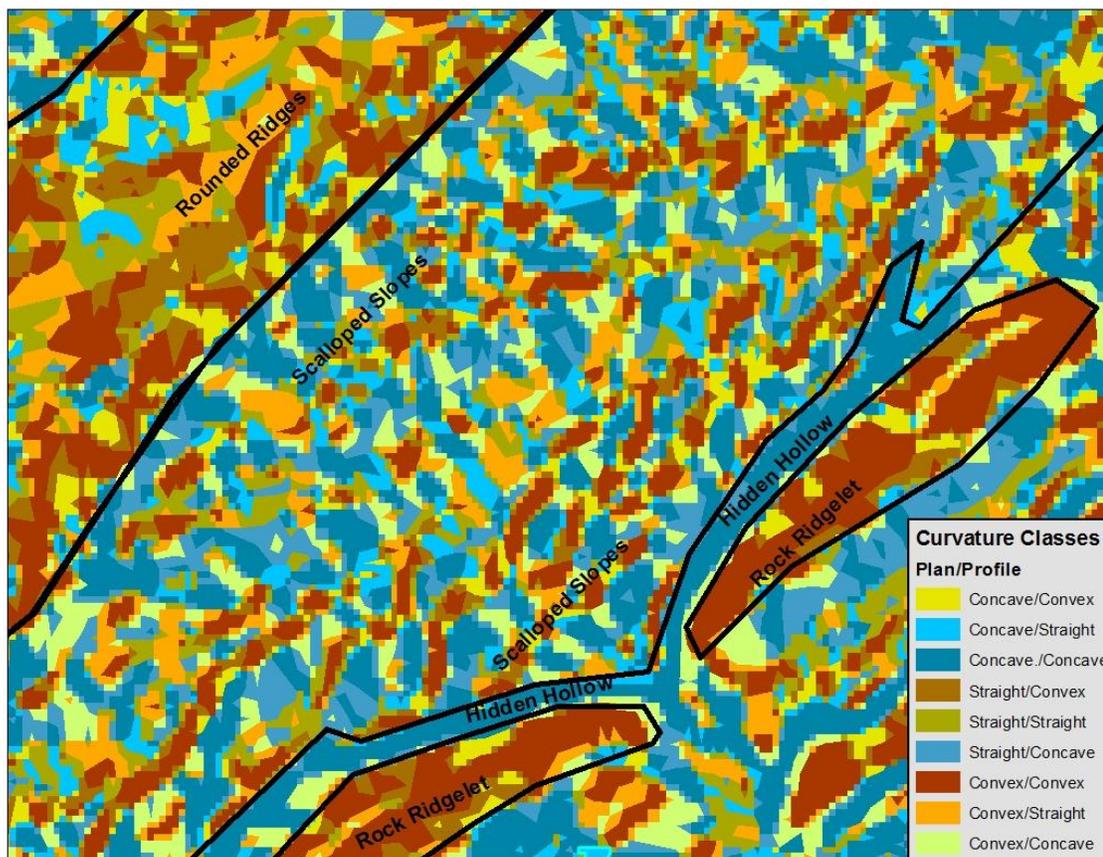
Figure **4-4:** Leading Ridge One with curvatures and landforms delineated to demonstrate the methodology for landform delineation.

Vegetation Surveys

Vegetation was surveyed during summer of 2010 at the Leading Ridge watershed one. Thirty-two 15 m radius circular plots were surveyed. In order to locate plots on the watershed that represented the range of variability in forests types, four transects were placed across the watersheds perpendicular to the topography, geology, and soil layers using GIS. The terrain was classified into four topographic formations based on patterns of curvature. Two plots were placed along the transect in each of the formations using GIS so that they were at least 30 meters apart to ensure that plots did not overlap. Two plots were placed in each of four topographic formations

for a total of 8 plots on each transect and 32 plots total. Plots were located in the field using a

Garmin GPSmap 60Cx GPS unit, and center points of plots were located at the nearest tree to the

GPS plot location.

In each plot, the following data were recorded:

1. Geographic coordinate of the plot center. Plot centers were typically a tree that was

   flagged for easy visibility for return visits.

2. Species name.The genus and species of each of the trees equal to or more than 18cm

   DBH were recorded.

3. Tree height. The "height routine" on a laser rangefinder (TruPulse 360 B, Laser

   Technology Inc.) was used for this data collection. The final tree height is the

   average of 3 data points collected.

4. Number of trees in each plot.

5. Diameter at breast height (DBH). The DBH of all trees over 18 cm (and

   corresponding species and height) were recorded.

6. Crown Class. The following categories were used:

   D Dominant: Trees with crowns extending above the general level of the crown

   cover and receiving full light from above and partly from the side.

   CD Co-dominant: Trees with crowns forming the general level of the crown

   cover and receiving full light from above, but comparatively little from the sides.

   I Intermediate: Trees shorter than those in the two classes above but with crowns

   either below or extending into the above crown cover receiving little direct light

   from above and none from the sides.

   S Suppressed: Trees with crowns entirely below the general level of the crown

   cover, receiving no direct light either from above or from the sides.

7. Leaf Area Index (LAI) Leaf area index (LAI). A LI-COR 2200 Leaf Area Meter was used with the 45 degree angle lens cap for all of the plots. Each leaf area index measurement is an average of measurements taken in the 4 cardinal directions; the value for each plot is an average of 6 different sites in each plot. The 6 different sites were measured at the following locations – the individual measuring LAI paced out these locations in each of the plots:

    a. 5 m East of plot center

    b. 5 m West of plot center

    c. 12.5 m NE of plot center

    d. 12.5 m SE of plot center

    e. 12.5 m SW of plot center

    f. 12.5 m NW of plot center

In addition to these forest species and structure metrics, a list of understory and herbaceous species mentioned in Fike's (1999) vegetation communities most likely to be found in Leading Ridge was created and presence/absence data was recorded for each species in each plot.

Raw Lidar Analysis

Using Fusion 2.90 (McGaughey), a canopy model of the forest was created by identifying the highest elevation from the raw lidar .las file within 3 meters of each cell and using that data to create a canopy surface. The bare earth DEM was subtracted from this layer to create a tree height index. In order to remove the gaps that were found in the canopy model due to the low-density leaf-off lidar product, a 3 x 3 circular moving window was moved over the surface to record the maximum height in each window with resulting modeled canopies from each high point approximately 5 m in diameter. This height index model was compared to actual field height

measurements of dominant and sub-dominant trees to relate the field measured heights to the lidar based height index.

Statistical Analysis

Vegetation metrics analyzed include total basal area, basal area of chestnut oak (*Quercus montana*), total number of trees per plot, total number of species present per plot, average height of dominant and co-dominant trees, and average leaf area index (LAI). In order to explore the spatial relationships between vegetation and landforms, maps were generated showing graduated values of each of these metrics across the watershed.

To test the hypothesis that vegetation communities on different landforms are significantly more different from one another than within landform variation, PC-ORD (McCune and Medford 2006) was used to explore the species datasets. A matrix containing presence/ absence for all tree species and selected herbaceous and understory species was used as the main matrix. A secondary matrix was created that listed on which landform each plot was located. Multi-Response Permutation Procedure (MRPP) which is a non-metric hypothesis testing procedure was used to test the significance of observed groupings of sample units (McCune and Medford 2006). MRPP uses a randomization permutation test to measure the within-group distance of sample units and compares that to randomly simulated groups of sample units to determine the significance of data groups being more similar than random groups (Peck 2010). Non-metric multi-dimensional (NMS) scaling was used to ordinate the communities and explore the gradients that may be affecting community composition (McCune and Medford 2006). A NMS stress test was run to determine that three axes should be interpreted. This ordination technique was chosen because the species dataset was not normally distributed.

An additional forest structure matrix was created that contained total basal area, basal area of chestnut oak (*Quercus montana*), total number of trees per plot, average height of dominant and co-dominant trees, total number of species, and average LAI for each plot. Since these data were all normally distributed, they were analyzed using principal components analysis in PC-ORD with a correlation matrix used as the cross produce matrix and assessed using MRPP to test the hypothesis that there is a significantly greater difference between landforms than within landforms.

Model Evaluation-Rothrock State Forest

An area of Rothrock State Forest approximately 2600 ha was used to assess the results of the curvature pattern modeling on a larger, landscape scale region. Curvature modeling was conducted on the thinned lidar-derived DEM using the same methodology as on Leading Ridge watershed one and the landscape was classified into rounded ridges, scalloped slopes, hidden hollows, and rock ridgelets using resulting patterns of curvature and previously described methodology. Sixteen plots were purposely located in areas easily accessible to roads and on public land to conduct vegetation surveys to classify the vegetation community present at each site. Four plots were located in each of the four formations. At each site, a 15 m radius plot was located approximately 50 m from the road. All trees were counted and identified by species and all understory species in Fike's classification present were recorded. Each plot was assigned a vegetation community classification based on Fike (1999).

Results

Landform Classification and Description

Patterns of curvature were used to delineate landforms because there is a curvature effect for many of the physiographic factors affecting plants including soil properties and moisture availability. The Ridge and Valley Province of Pennsylvania is distinguished by its regular repeating pattern of parallel ridges. Most of the ridges are of similar form, consisting of shorter, leading ridges paralleling the main ridge. It is important to note that for the purpose of this study, landforms are not defined based on curvature, but patterns of curvature. For example, features are not defined by belonging to one of the nine curvature classes, but on the different types of curvature classes that occur in each landform and the relationship with these classes with one another. Also, features were delineated in an interpretive manner following the methodology highlighted in the methods section by a trained interpreter. For example, the scalloped slope can be identified from the rounded ridge as one moves down the slope at the point where the profile curvature first becomes concave. Therefore, the rounded ridge can be described as having a convex or straight profile curvature and an undulating plan curvature. Scalloped slopes, in contrast, have all nine curvature classes occurring in roughly equal proportions. The features defined by the nine curvature classes are not formations, but merely the building blocks that make up the formations. When curvature classes were defined and analyzed four distinctive patterns of curvature emerged (Figure **4-5**).

a.   *b.*
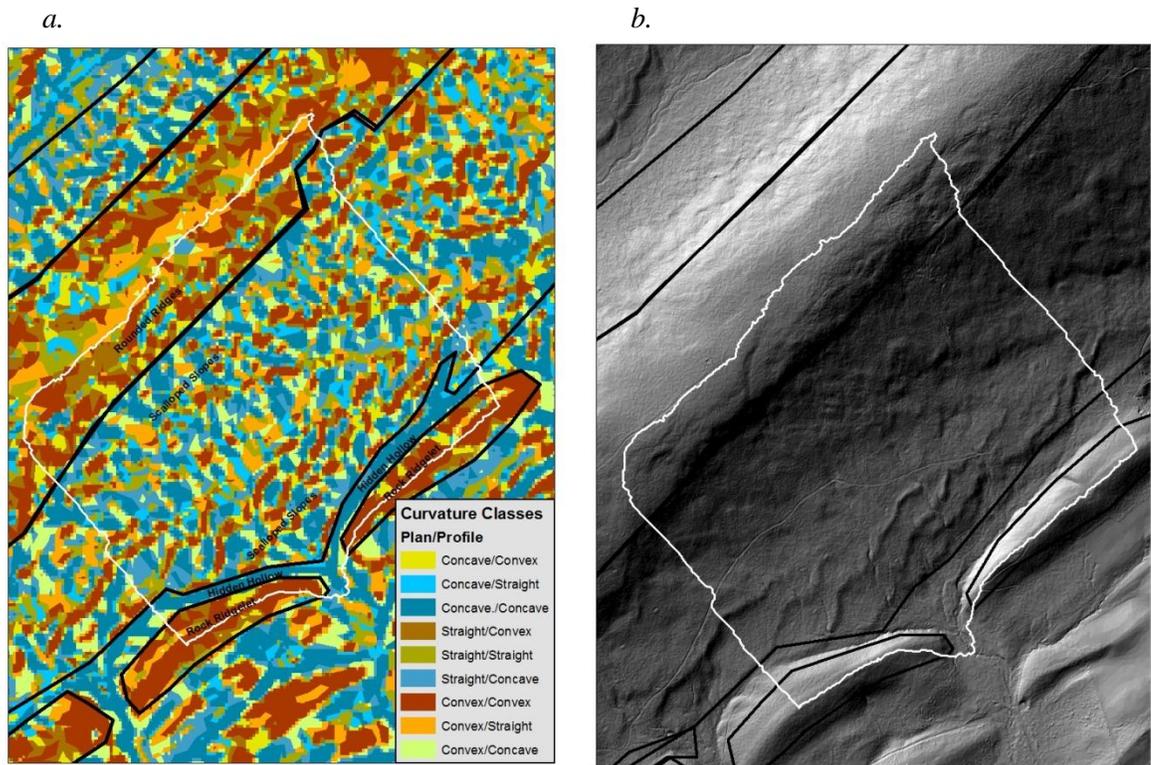
Figure **4-5:** (a) Landscape classification units with curvature classes and (b) a hillshade of the watershed showing the approximate terrain.

- Rounded ridge—broad profile curvature is primarily convex, undulating plan

  curvature, underlain by resistant Tuscarora sandstone or quartzite. These features

  make up most of the ridge tops in the physiographic province. Depending on the

  underlying geology, these ridges may be more narrow or broad, but they all show

  a characteristic pattern of broadly convex plan curvature with a mix of planar,

  concave, and convex profile curvatures as one moves along the crest. There are

  pockets of concave profile curvatures that tend to form low points or saddles

  along the ridge, but these are isolated and narrow and tend to cut across the

  whole feature in a sharp, knife-like pattern. Soils tend to be Hazelton or Dekalb.

- Rock ridgelet—These features can be identified on the landscape because they

  are convex in both plan and profile curvature. They tend to be much shorter and

narrower than rounded ridges. They are often underlain by softer sandstone or shale and can be very steep. Rock ridgelets underlain by shale such as those found in Leading Ridge have extremely shallow soils, with rock exposed in many places. They are often underlain by Berks or Berks-Weikert soils.

- Hidden hollow—small to medium sized drainage features distinguished by being concave in both plan and profile curvature. These are not the broad open valleys generally used for agricultural purposes in the Ridge and Valley, but the small forested hollows that usually contain headwaters streams. These differ from the traditional hollows of other physiographic regions in that they are often found between ridges, or between main rounded ridges and rock ridgelets. They are not cove shaped, but more linear. These formations are often transition zones in soils and geology, often being formed from softer shales. Hidden hollows often emerge through water gaps. Soils are often Buchanan or Andover.

- Scalloped slope—These features are defined by their undulating pattern of plan and profile curvature. All nine curvature classes can be found in this unit at roughly equal proportions. Areas of convex and concave plan curvature are often driven by small drainage systems affecting areas on a scale of several meters, while convex and concave profile curvature are formed by small slump-like features that are often found throughout the landscape, particularly on south facing slopes. The vegetation communities in these features tend to be very heterogeneous due to the pattern of curvature, with mixed oak and more mesic species. Soils tend to be Hazelton, Laidig, Dekalb, and are underlain by bedrock consisting primarily of sandstone and shale. The point on the bottom of the hillslope where scalloped slopes begin can be identified by the location where

profile curvature first becomes convex and on the top of the slope by the point

where profile curvature first becomes concave from the top of the rounded ridgele.

Vegetation

Chestnut oak (*Quercus montana*) was the most dominant tree in the watershed in both

tree density and basal area, followed by northern red oak (*Q. rubra*), white oak (*Q. alba*) and red

maple (*Acer rubrum*). Other trees occurring in the watershed include eastern hemlock (*Tsuga*

candedensis), eastern white pine, (*Pinus strobus*) and black oak (*Q. veluntina*) (Table **4-1**).

Height of dominant and co-dominant trees varied throughout the watershed, even among trees of

the same species. The average dominant and co-dominant tree height for each plot ranges from

15.42 m to 30.4 m. Tallest trees are found in the lower 2/3 of the watershed, while shortest trees

are found on top of both Leading Ridge and the rock ridgelet (Figure **4-6**).

Table **4-1**: Most common trees on the watershed by number of trees

Species	Number of Trees on watershed	Average DBH (cm)	Density (trees/ha)	Total BA on watershed (m^2)
Quercus montana	177	32.45	64.21	2591.36
Quercus rubra	108	39.58	39.18	1435.41
Quercus alba	54	40.52	19.59	376.07
Acer rubum	52	26.46	18.86	148.7
Tsuga canadensis	47	33.32	17.05	192.56
Quercus veluntina	45	44.26	16.32	311.53
Pinus strobus	31	33.95	11.25	87
Betula lenta	22	26.52	7.98	26.73
Nyssa sylvatica	20	22.37	7.25	15.72
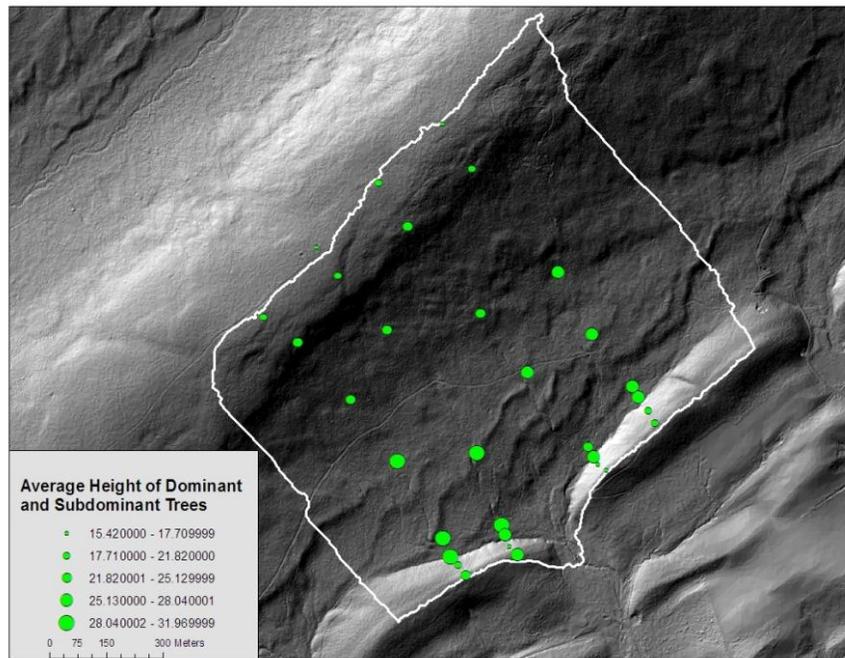
Figure **4-6**: Average height of dominant and co-dominant trees in Leading Ridge 1

Basal area (m^2/ha) ranges throughout the watershed from 10.3 to 49.7 (Figure **4-7**). Like tree height, basal area was higher closer to the bottom of the watershed, but there are some plots with a high basal area at higher elevation. Some plots on the rock ridgelet also have a relatively high basal area. The patterns of basal area of chestnut oak (Figure **4-7**) are opposite from the patterns of total basal area, with high values located on top of both the rounded ridge and the rock ridgelet and smaller values in the hidden hollow. There were only four plots that did not contain chestnut oak, all of which were located in the hidden hollow.

a)   *b)*
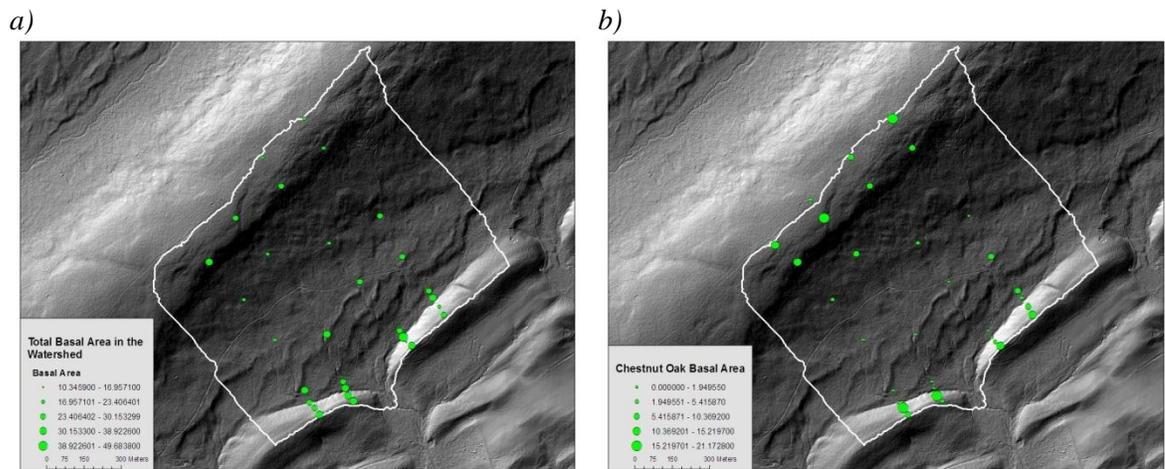
Figure **4-7**: Total basal area and *Quercus montana* basal area in Leading Ridge One

In addition to tree measurements, common herbaceous and understory vegetation were sampled for presence/absence. This was not a comprehensive survey, but included all understory species mentioned in Fike's *Terrestrial and Palustrine Plant Communities of Pennsylvania* for forest communities likely to be found in the Ridge and Valley Province. The sites with the highest diversity were found on the rock ridgelet in front of the watershed, with up to 18 species identified in one plot (Figure **4-8**). The fewest number of species identified in a plot was 6. Scalloped slope and rounded ridge plots tended to contain the fewest total number of species. Plots on top of the rounded ridge, in particular, contained a dense covering of blueberry (*vaccinium spp.)* which could have inhibited establishment of other herbaceous and understory plants. Total number of trees per plot (density) is also shown in Figure **4-8**.

a)   *b)*
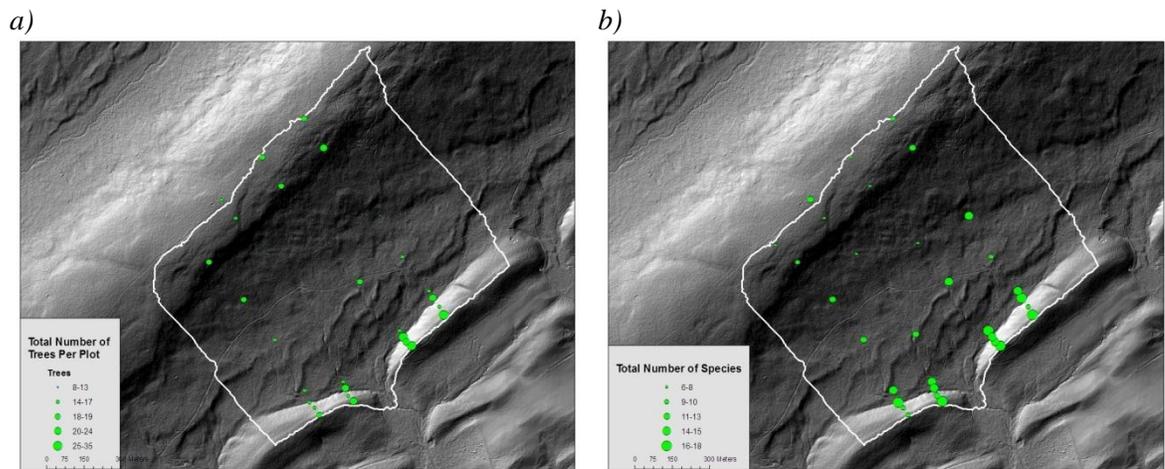
Figure **4-8**: Total number of trees per plot and total number of species at each plot for Leading Ridge One

Leaf area index (LAI) is a metric of the upper leaf surface area of vegetation divided by the total area of land. Plots on the rounded ridge had lower LAI values than other plots (Figure **4-9**), but otherwise there was heterogeneity in LAI values throughout the watershed. High values were clustered along the southwestern region of the watershed, with high values being found both on the rock ridgelet and the hidden hollow, even though tree heights were lower and site index is generally lower on the ridgelet. This may be because although trees do not grow as tall on this formation, they exhibit a stronger horizontal branching pattern and different structure that still provides a high leaf area. Also, the slope is almost 100% in this area which could introduce error into the LAI calculation due to the difference between surface area as measured along the slope and surface area measured planimetrically.
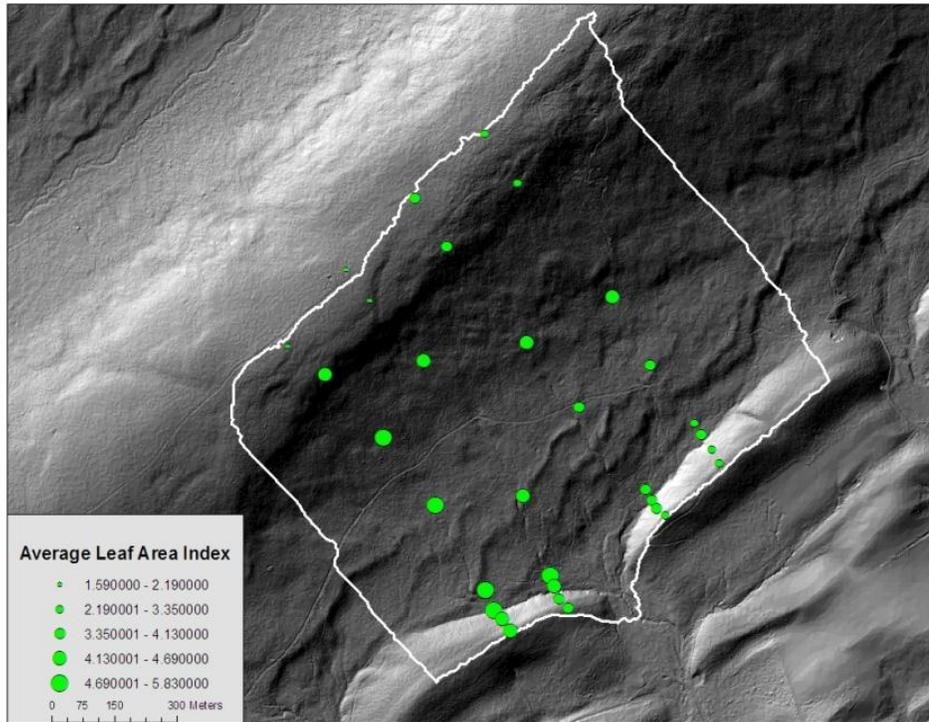
Figure **4-9**: LAI at Leading Ridge One

The results of the raw-lidar derived height index and the measured heights are shown in

Figure **4-10**. In Figure **4-11**, differences are shown between the mean field-measured heights of

dominant and co-dominant trees and the mean lidar-derived height index for each plot. In most

plots, the difference between lidar-derived heights and measured heights is less than 3 meters.

Plot 20, which shows a height difference of 9 meters, is on a plot with greater than 100% slope

that touches the base of the rock ridgelet. This error was probably caused by a dominant tree at

the bottom of the slope extending into the plot from overhead. Also worth noting is the area of

shorter trees located on the northeast strip of the watershed. This is probably due to this area

being privately owned and managed and the selective harvesting of large trees.
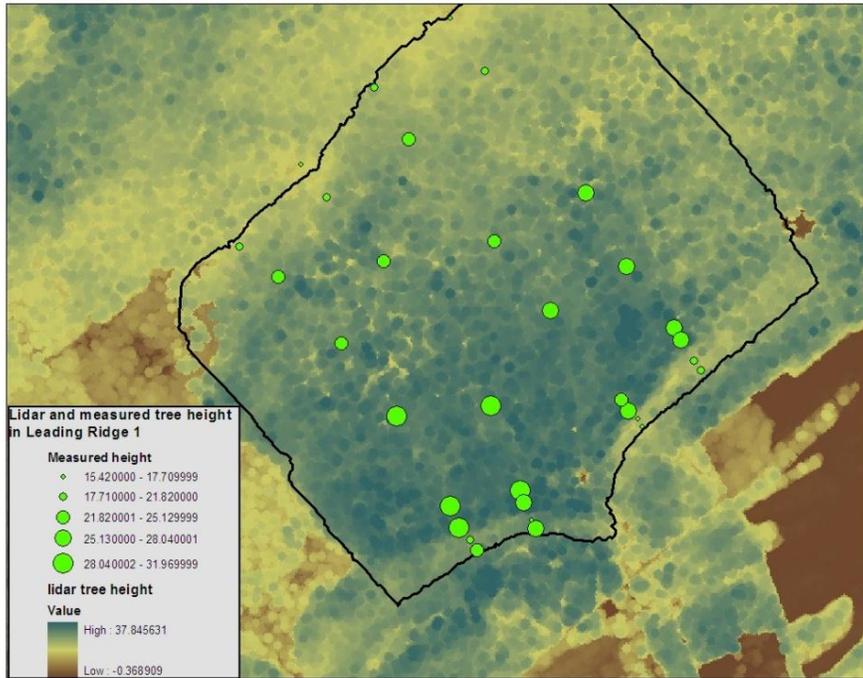
Figure **4-10**: Average tree height of dominant and co-dominant trees in each plot and the lidar-derived tree height index for Leading Ridge One.
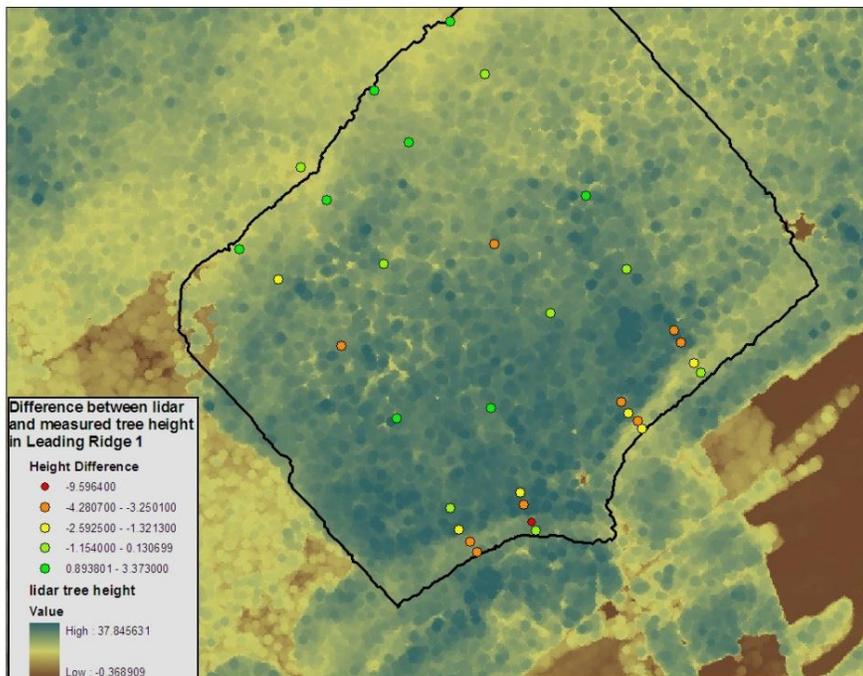
Figure **4-11**: Difference between lidar-derived height index and measured tree heights for 32 plots in Leading Ridge One. High negative values represent areas where lidar-derived height index was greater than the measured heights.

In addition to univariate spatial analysis, vegetation metrics were analyzed using multivariate methods to investigate community structure and to explore dominant gradients that may be influencing the vegetation. A set of pairwise plots was constructed in R to assess correlation across metrics, with very little correlation identified. The variables with the highest correlation coefficients were average tree height and basal area of chestnut oak which were negatively correlated with a correlation coefficient of -0.66. This is expected because chestnut oak is known to inhabit sites with a lower site index. Both metrics were kept in the matrix for analysis because it was determined that they provide different information. Two vegetation matrices were analyzed in the multivariate analysis. The first was comprised of presence and absence of herbaceous, understory, and tree species. The second matrix was comprised of vegetation structure variables. Each plot was assigned to a group based on landform occurrence (Table **4-2**).

Table **4-2**: Summary of formations and their major attributes

Group	Name	Pattern of Curvature	Geology and Soils
1	rock ridgelets	Profile curvature and plan curvature >95% convex	Underlain by shale, soils primarily Berks/Berks-Weikert.
2	hidden hollows	Profile curvature and plan curvature both >95% concave	Underlain by shale, soils primarily Andover and Buchanan
3	scalloped slopes	Profile curvature and plan curvature both undulating, similar proportions of all 9 categories	Underlain by shale and sandstone of various groups, soils of Laidig, Dekalb, or Hazelton
4	rounded ridges	Profile curvature >95% convex or flat, plan curvature more evenly undulating	Underlain primarily by Tuscarora quartzite and sandstone, soils primarily Hazelton, Dekalb, or rubble

Patterns of total basal area, number of trees, chestnut oak basal area, height, LAI, and total number of species are shown by group/formation in Table **4-3** with the standard deviation of each group in Table **4-4**. Individual metrics were not analyzed statistically for significance across groups, but metrics were analyzed collectively using MRPP.

Table **4-3**: Mean values of total basal area, total number of trees, basal area of chestnut oak, height, LAI, and total number of species summarized by group.

Group	BA	Number of Trees	Quercus montana BA	Height	LAI	Total spp
Rock ridgelet	27.58	22.13	11.58	19.86	3.57	13
Hidden hollow	33.39	19	1.816	28.11	4.40	15.38
Scalloped slopes	24.92	14.25	3.90	27.12	4.49	10.88
Rounded ridge	21.96	17.63	11.10	20.32	3.00	9.25

Table **4-4**: Standard deviation of total basal area, total number of trees, basal area of chestnut oak, height, LAI, and total number of species summarized by group.

Group	BA	Number of Trees	Quercus Montana BA	Height	LAI	Total
Rock ridgelet	7.20	6.36	6.70	3.82	0.8	3.38
Hidden hollow	8.21	5.04	3.60	2.00	0.87	1.30
Scalloped slopes	4.51	3.41	3.20	2.99	0.41	3.48
Rounded ridge	7.41	2.67	5.49	2.92	1.00	1.75

In order to test the hypothesis that vegetation structure and communities were significantly different in the four landforms, MRPP group testing was performed on a summary dataset containing total number of trees, average height of dominant and co-dominant trees, basal area, basal area of chestnut oak, LAI, and total diversity using Euclidean distance as the distance measure in PC-ORD (McCune and Medford 2006). The four groups were the landforms. The chance-corrected within-group agreement or A value was 0.20459523 with a p-value of having a smaller or equal A value of 0.0000044. MRPP group testing was also performed on a dataset containing presence absence data for tree and herbaceous species in each plot using Sørensen distance as the distance measure in PC-ORD. The chance-corrected within-group agreement or A value was 0.21286789 with a p-value of having a smaller or equal A value of 0. This verifies the

hypothesis that patterns of forest structure and forest communities are significantly more similar within landforms than between landforms.

The structure dataset was analyzed with PCA using a correlation matrix as the cross product matrix. Two axes were interpreted as significant, with axis one accounting for 45% of the variance and axis two accounting for 29% of the variance (Figure **4-12**). The plots are grouped into quadrants, with plots from rock ridgelet and rounded ridge tending to be on the upper end of the first axis, and scattered across the second axis, with the majority of plots being found close to the center on the second axis. Plots from the rounded ridge tended to be found slightly lower on the second axis than plots from the rock ridgelet. One of the plots in the rock ridgelet is a slight outlier and was located closer to plots from the hidden hollow than from other plots on the rock ridgelet or rounded ridge. This indicates that vegetation on the rounded ridge and rock ridgelet are more similar to one another than to the other formations, even though these features are very different in scale and elevation. They are also found on very different geologic and soil settings. These plots are also located furthest from one another on the landscape.

There are two dominant axes influencing vegetation structure in the watershed. Axis one is probably related to water availability because the rock ridgelet and rounded ridge plots tend to be together on one end of the axis. These plots are both located in sites with rocky soils in more convex settings. The second axis is more difficult to interpret, but may be related to nutrient availability or pH.
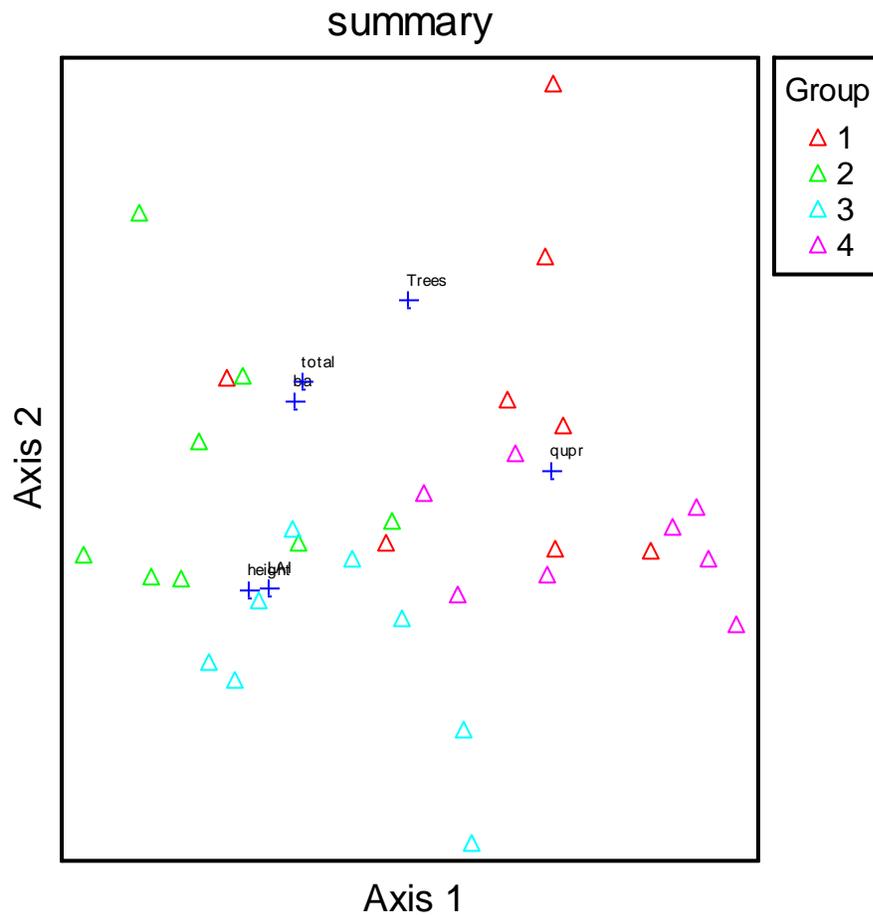
Figure **4-12**: Ordination diagram for the PCA analysis of structure dataset for Leading Ridge One.

In order to analyze the community structure of the watershed, presence/absence data for

trees and herbaceous species was analyzed for 78 species using NMS in PC-ORD (McCune and

Medford 2006). Sørensen's distance was used with a random starting number, and three axes

were interpreted. The final stress of a 3-dimensional solution was 16.14. Patterns on the first two

axes (Figure **4-13**) were similar to the results of the PCA ordination of structure datasets, with a

pattern of plots from the rock ridgelet and rounded ridge being found together on the positive end

of the first axis, while plots from the hidden hollow and scalloped slopes are found at the lower

end of the axis. The points tend to be divided into quadrants. The pattern on axis two consists of

plots from the rock ridgelet and rounded ridge being found clustered on the negative end of the

axis, followed by plots from the scalloped slope and then the hidden hollow. Axis three (Figure **4-14**) tends to have the rounded ridge on the negative end of the axis, with the other three groups are found on the upper two thirds of the axis.
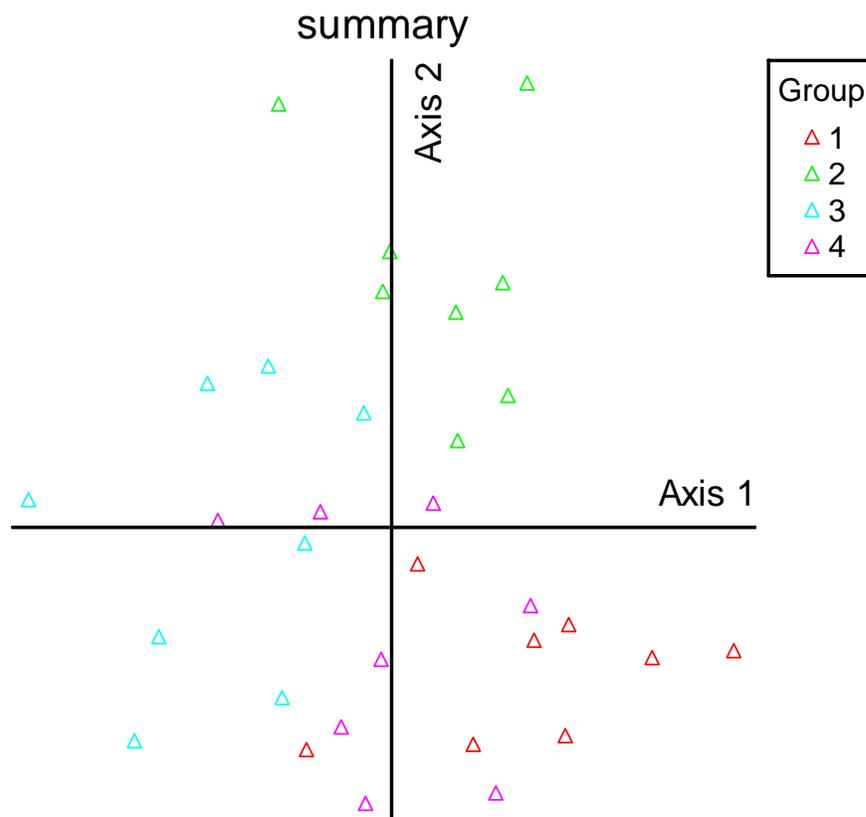
Figure **4-13**: Ordination diagram for NMS results for presence absence data for axis 1 and 2
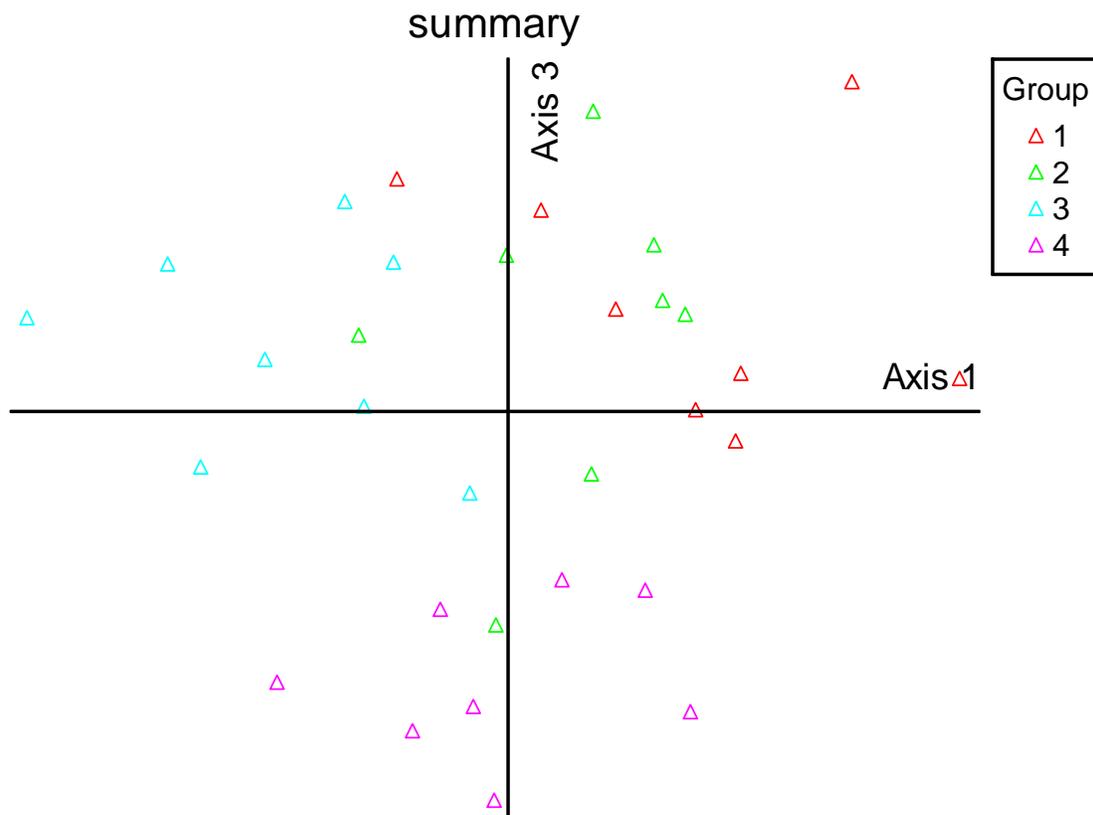
Figure **4-14**: Ordination diagram for NMS results for presence absence data for axis 2 and 3

Community Composition

In order to classify forest types to Fike's (1999) Terrestrial and Palustrine Plant Communities of Pennsylvania, community data for each formation were analyzed and each landform was assigned to a community that most accurately matched the presence absence data for species. Rounded ridges in Leading Ridge Watershed One are characterized by dry oak heath forest type, which is characterized by chestnut oak, with occasional black oak, white oak, and other trees. The shrub layer is dominated by ericaceous species including mountain laurel (*Kalmia latifolia*), huckleberry (*Gaylussacia spp.,*) and blueberry (*Vaccinium spp.*). Common

herbaceous species include teaberry *(Gaultheria procumbens)* and trailing arbutus *(Epigaea repens).* This forest type is extremely common throughout Pennsylvania.

As one moves down to the scalloped slopes, forests become more diverse with more mesic oak species sharing dominance including white oak, black oak*,* and red oak*.* Scalloped slopes in Leading Ridge One can probably be best characterized as Red oak-mixed hardwood which is dominated by northern red oak as dominant/co-dominant with black oak, white oak, hickories (*Carya spp*.), and tuliptree (*Liriodendron tulipifera*). Shrub species and herbaceous species are diverse and mixed and include maple-leaved viburnum (*Viburnum acerifolium)*, serviceberry (*Amelanchier spp*.), and mountain laurel (*Kalmia latifolia).*

Going into the hidden hollows, forests become more mesic with a much smaller proportion of oak species, particularly chestnut oak. Also, there is more conifer cover, particularly eastern white pine (*Pinus strobus)* and eastern hemlock (*Tsuga candensis)*. As conifer cover increases to greater than 25% of total, the forest is classified as belonging to the Hemlock (white pine) - red oak - mixed hardwood forest type, which is very similar to the red oak-mixed hardwood forest also found in Leading Ridge but with white pine and eastern hemlock contributing more than 25% relative cover. Herbaceous species include false Solomon's-seal (*Smilacina racemosa),* Solomon's seal (*Polygonatum biflorum.)* teaberry, Canada mayflower (*Maianthemum canadense)*, and may-apple (*Podophyllum peltatum).*

The forest type on the rock ridgelet is the most difficult to categorize because it contains primarily dry oak species such as chestnut oak but lacks the dominant heath component like the dry oak-heath forest type. It also contains a much more diverse understory than a similar forest that makes up the rounded ridge, along with a high component of eastern white pine and pitch pine (*Pinus rigida*.) The forest type it is the most similar to is the Pitch pine-mixed oak type. Shrubs include green brier (*Smilax spp*.), black huckleberry (*Gaylussacia baccata*), and low bush blueberry (*Vaccinium pallidum*) although not nearly as dominant as in the dry oak-heath type.

Herbaceous species include bracken fern (*Pteridium aquilinum*,) teaberry, and pink lady's slipper (*Cypripedium acaule*).

Rothrock State Forest

Curvature classes were calculated using the same methodology as described in the methods for a larger area that includes an area known as the Seven Mountains region of Rothrock State Forest. This area consists of multiple ridges separated by small forested valleys. As shown in Figure **4-15**, the same patterns of curvature are found in this region as were initially identified and described in the Leading Ridge One watershed. It was expected that there may be several vegetation communities found on each formation based on specific site factors and land use history.

a)
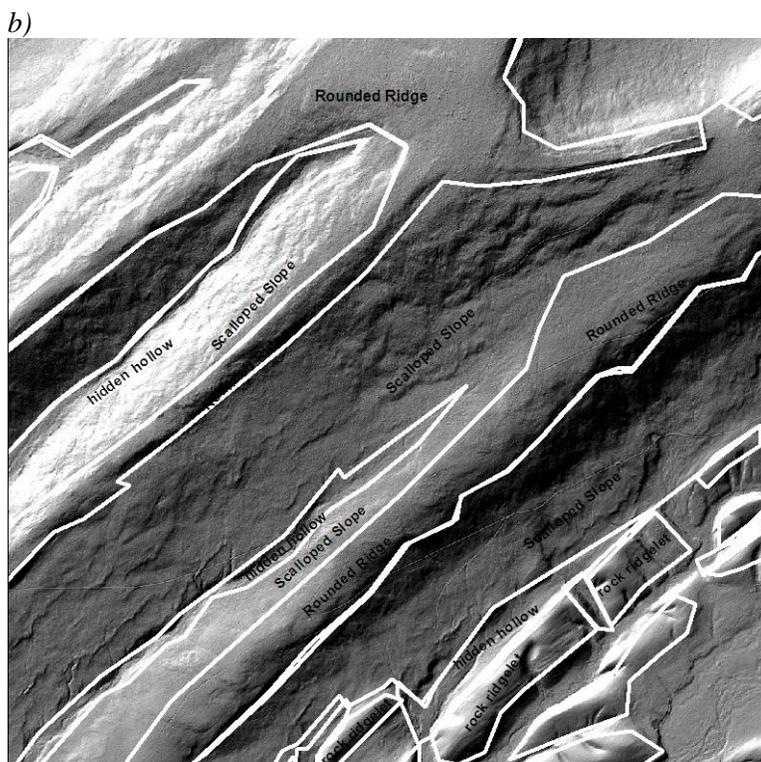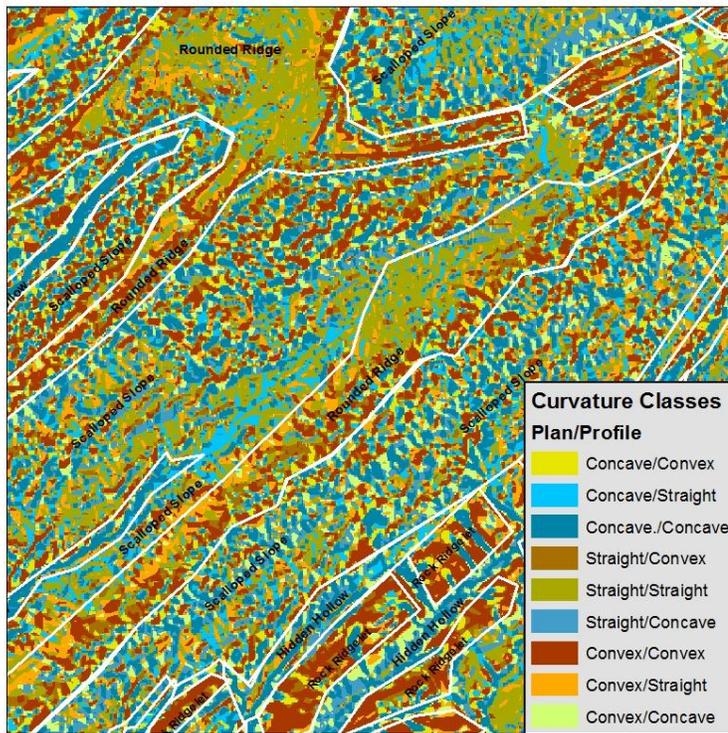
Figure **4-15:** *a*) Curvature classes in a larger region of Rothrock State Forest known as the Seven Mountains region, along with *b*) hillshade of the region to show general topography.

All 16 plots placed in Rothrock State Forest were matched to one of Fike's (1999) forest types, although in some cases this was difficult due to the small size of each plot and the confounding impact of management. Several plots that were initially planned using GIS were relocated in the field in order to avoid planted pine plantations. These sites were relocated in the field using the landform map to identify new plots. All four plots located on rounded ridges were classified as being in the dry-oak heath forest type. The four plots located on scalloped slopes were classified as being dry oak-mixed hardwood, red oak-mixed hardwood (2), or tuliptree-beech-maple. The tupliptree-beech-maple plot had been harvested within the last 30 years and did not look very much like the surrounding forests which had also recently been harvested and which had large amounts of black gum and red maple. There may have been more heterogeneity in the scalloped slope plots than the rounded ridge plots due to the changes in curvature present in this formation. Parts of the scalloped slope that are concave/concave are likely to exhibit more mesic hardwoods than the convex/convex areas that could be immediately adjacent to them.

The plots in the hidden hollows were all similar forest types: hemlock-red oak-mixed hardwood, hemlock-tuliptree-beech, tuliptree-beech-maple, and red oak-mixed hardwood. The four plots located on rock ridgelets were the most difficult to classify because of both the variation of sites on this landform and because often there was a discrepancy between the overstory and understory in these sites. For example, in one plot almost all of the overstory species were chestnut oaks but the understory was very rich and mesic, with almost no ericaceous species. The classifications were determined to be dry oak-mixed hardwood, dry white pine (hemlock) oak and pitch pine-mixed oak. They all tended to consist of dry oak species with pine of different species mixed in. None of the sites had a strong ericaceous shrub layer. There were shale fragments found throughout all of the rock ridgelets and there tended to be large amounts of moss and lichen covering the ground and tree trunks.

Discussion

In this work a method of landform classification was proposed that takes into account only one parsimonious metric based on topography instead of combining multiple terrain metrics such as topographic index (TI), elevation, aspect, and slope as many classifications do. Also, soils and geology were not directly incorporated, although it can be argued that both are inherently included due to their connection with patterns of curvature. The omission of these parameters was not because these factors were not considered important factors in vegetation community development, but that currently available data for landscape level soil and geologic information are not accurate enough or on a correct scale to delineate landforms. Reiners et al. (1999) also used an approach using only terrain data in identifying LTAs in northwest Wyoming.

In Figure **4-16**, elevation, slope, aspect, and TI are shown for Leading Ridge Watershed One. If these are compared to the landform classes in Figure **4-5**, there is little agreement between these topographic variables and the formations delineated using curvature. In the Ridge and Valley Province, rock ridgelets and rounded ridges exhibit similar vegetation structures and communities, although they are very dissimilar when viewed as a collection of their more commonly utilized topographic properties. This is in part due to the changes in geology that drive the topography. The change in water permeability from sandstone to shale influences the occurrence of the line of springs and seeps found across the center of the watershed. Additionally, soil depths above shale and sandstone tend to be different, with soil types underlain by shale being shallower than soil types underlain by sandstone (Shields 1966). There can also be differences in vegetation based on the calcium content of the parent bedrock, with more calcareous species being found over limestone and limey shale than sandstone (Fike 1999).

a)

b)

c)

d)
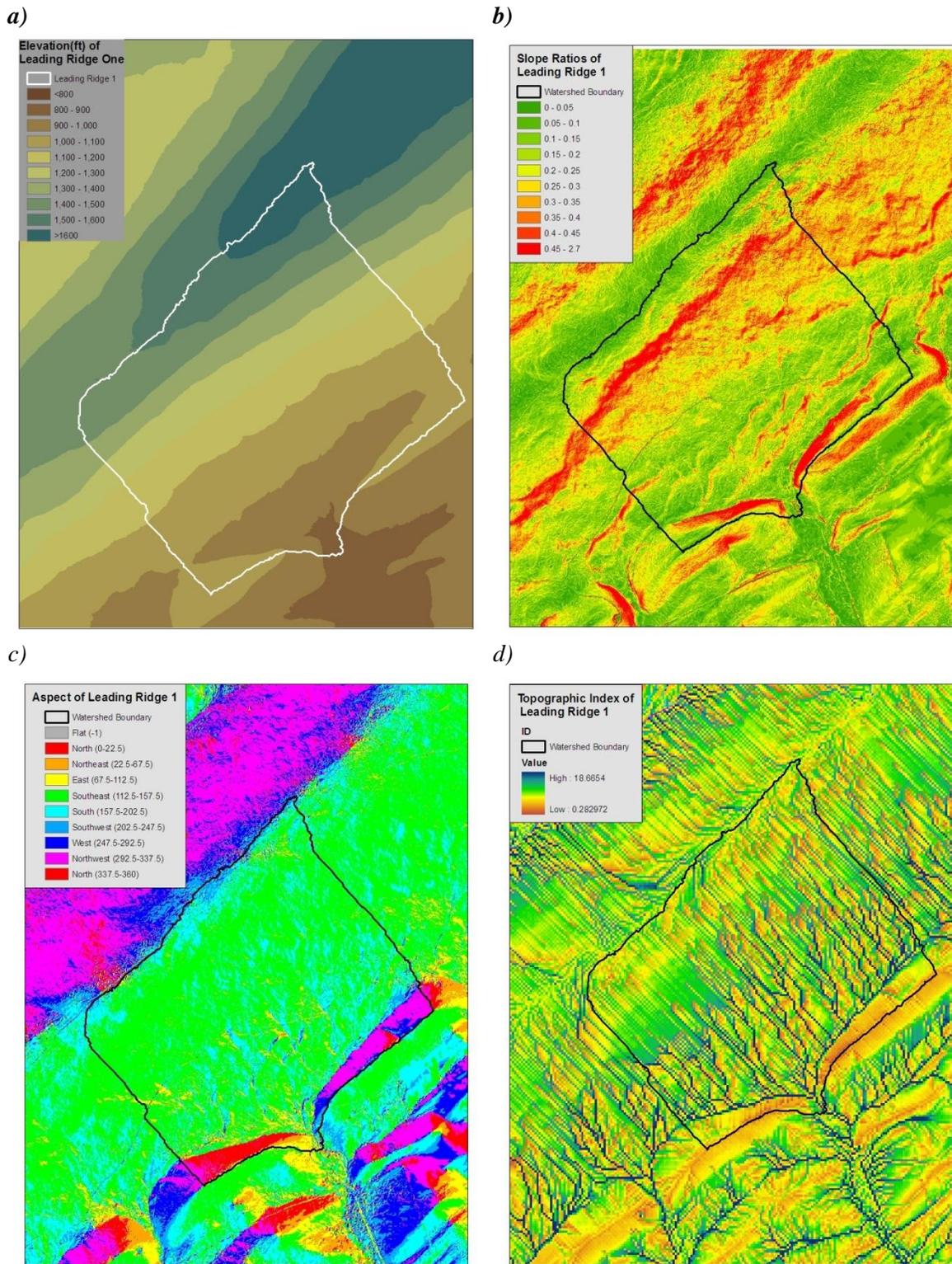
Figure **4-16**: Elevation (*a*), slope (*b*), aspect (*c*) and Topographic Index (*d*) at Leading Ridge one.

It is possible that the reason slope and TI do not effectively predict vegetation could also be related to the compounding of DEM errors when computing second and third order terrain metrics such as TI and other commonly used indices. Van Niel et al. (2004) found less DEM error with topographic position than with TI, possibly because TI is created by dividing one dataset by another. Since curvature affects water movement and transport of other materials on the landscape, by using the best available data to focus on delineating curvature, we can more accurately and easily delineate landforms and predict vegetation communities. Pennock (1987) used slope in addition to curvature for his classification system. In this case, slope could be used to refine the classification, particularly in partitioning scalloped slopes into upper and lower slopes as there is generally an obvious slope break at some point in the landform. Slope may also help differentiate the more mesic rock ridgelet sites from the more xeric locations.

Scale is another important factor that influences the design and implementation of classification systems. Previous work completed in Pennsylvania (Myers 2000, Kong 2006) has focused on landscape level classification, with Kong (2006) also including smaller ELTs that fit into larger features in a hierarchical way. With the improvement of elevation data, improvements can be made to these methodologies that efficiently utilize the refined elevation data at the appropriate scale.

In analyzing results of the ordinations, there are two dominant environmental factors affecting both vegetation structure and community, with a third factor emerging slightly in the community analysis. The most obvious patterns are found on the first axes, with plots from rock ridgelets and rounded ridges occurring on one end of the axis and plots from the other two communities occurring at the other end. This axis is probably driven by water availability, because the rock ridgelet and rounded ridge sites are both located on convex types of formations with thin or rocky soils. The primary difference between these two types of sites is that rounded ridges tend to be found on sandstone, while rock ridgelets are generally on shale, which is also

reflected by a difference of soil texture and type. Although it is difficult to identify what factors are affecting the rest of the ordination, there are similar patterns occurring in all of the ordinations, suggesting that vegetation communities and structure may be responding to the same variables.

The interaction between vegetation type and formation is not a direct one-to-one relationship, and individual landforms may be host to different vegetation communities. Part of this is due to the inherent heterogeneity involved in the delineation of the scalloped slope landform in particular. Although the small shifts in curvature are not large or dominant enough to classify as unique landforms, it would be assumed that these shifts affect the movement of water across the slope, which would impact vegetation. It is expected to find the more mesic hardwood species mixed in with more xeric oak and hickory species. This is also in part because vegetation communities can respond very strongly to land use and management histories which may be completely unrelated to landform. For example, invasive species have and continue to have a devastating impact on Pennsylvania forests, with the complete loss of the American Chestnut (*Castanea dentata*), reduction in oak due to gypsy moth (*Lymantria dispar)* defoliation, and recently, the loss of Eastern hemlock due to hemlock wooly adelgid (*Adelges tsugae*).

Despite these complications, curvature does provide a tool to help delineate the dominant terrain features that make up a landscape and affect vegetation. The results from both Leading Ridge Watershed One and Rothrock State Forest both suggest that the same types of landforms occur throughout the Ridge and Valley Province. Although patterns of curvature cannot be used to directly predict vegetation, they can be used to interpretively delineate landforms which guide community development on a site. There could also be associations between structure variables such as tree density and height among landforms, although this was not analyzed for the broader Rothrock State Forest study area.

The most difficult formation on which to predict vegetation community are the rock ridgelet sites, which contained considerable variation depending on the specific site. Additionally, these areas tended to be preferentially planted in conifer plantations and may have been preferentially used for agriculture, making an assessment of their natural vegetation more difficult. Even in the smaller Leading Ridge One Watershed, one of the rock ridgelet plots was an outlier in the initial analysis, with a much richer, more mesic composition than the other rock ridgelet sites. This problem may have been exacerbated by the lack of an appropriate vegetation community description for these unique landforms. There is a large amount of variability of vegetation communities on rock ridgelets depending on slope, location, and specific structure of the formation, but they are all distinctly different than the adjoining formations. It is possible that by further analyzing rock ridgelet sites using texture and curvature, we could improve classification into more xeric sites and more mesic sites, based on either curvature or slope patterns. Despite the variability of rock ridgelet sites, they were distinguished by high amounts of pine and chestnut oak, low amounts of ericaceous understory species and high moss and lichen cover on both trees and the ground.

Conclusions

In an age of available lidar elevation data and with the continued need for improved ecological and landscape classification, lidar can be used in new ways to improve our classification methodologies. In the Ridge and Valley Province of Pennsylvania, a method of landscape classification for forested uplands has been explored based on recurring patterns of curvature on the landscape. This method reduces the sources of topographic and site variability to one parsimonious variable. Due to the patterns of geology in the Ridge and Valley Province, this

classification system is effective at predicting the presence of potential vegetation communities. Recurring formations include small steep rock ridgelets, narrow and rich hidden hollows, diverse scalloped slopes, and dominant rounded ridges. Although each of these formations has the potential to host several vegetation communities, there are defined patterns in vegetation community and structure that occur on these landforms.

The Ridge and Valley Province offers a terrain that facilitates the use of a new classification system for improving understanding of the relationships between terrain and vegetation. It is an important ecosystem to understand because eastern forests are some of the most diverse temperate ecosystems in the world. Tree species in the eastern deciduous forests are also important for timber and wood products, with some species in demand throughout the world. This makes understanding terrain for ecosystem management essential. Some commonly used terrain metrics such as elevation, slope, aspect, and TI do not adequately describe the complex relationships between vegetation and site in this region due to the impact of soil and parent bedrock.

Lidar data offers advantages over previously available elevation data, with greatly improved accuracy and resolution. It allows us to see and understand small scale features and micro-topography that can affect vegetation communities. Using traditionally available DEMs would not represent the topography and patterns of curvature the same way that lidar data currently can do. Lidar may also allow us to integrate scale in a way that was previously impossible and in the future facilitate classifying formations as sums of their smaller, more diverse parts. By expanding on this work, this may lead to the development a multi-scale hierarchical classification for the Ridge and Valley Province.

References

Anderson, S.P*., Bales, R.C, Duffy, C.J. 2008. Critical zone observatories: building a network to advance interdisciplinary study of Earth surface processes. Mineralogical Magazine 72*: 7–10.

Bolstad, P.V., Swank, W. and Vose, J. 1998. Predicting Southern Appalachian overstory vegetation with digital terrain data. Landscape Ecology 13, pp. 271–283.

Bowersox, T.W., W.W. Ward. 1972. Prediction of oak site index in the ridge and valley region of Pennsylvania. Forest science, 18: 192-195.

Carter, B.J. and Ciolkosz, E.J., 1991. Slope gradient and aspect effects on soils developed from sandstone in Pennsylvania. Geoderma 49: 199–213.

Ciolkosz, E.J., Carter, B.J., Hoover, M.T., Cronce, R.C., Waltman, W.J. and Dobos, R.R. 1990. Genesis of soils and landscapes in the Ridge and Valley province of central Pennsylvania. Geomorphology, 3:245-261.

Cleland, D.T., P.E. Avers, W.H. McNab, M.E. Jensen, R.G. Bailey, T. King, and W.E. Russell. 1997. National hierarchical framework of ecological units. Pp.181-200. In: M.S. Boyce and A. Haney (eds.). Ecosystem management. Yale University, New Haven.

ECOMAP. 1993. National hierarchical framework of ecological units. USDA Forest Service, Washington, D.C.

Devlin, D.A., W.L. Myers, W.D. Sevon and D.M. Hoskins. 2001. Use of landtype associations and landforms in managing Pennsylvania's state forests. Landtype associations conference: Development and use in natural resources management, planning and research. University of Wisconsin Madison, Wisconsin.

Dikau, R. 1989. The application of a digital relief model to landform analysis in geomorphology. Three dimensional applications in geographical information systems. J. Raper, ed. Taylor and Fancis: New York. 51-77.

ECOMAP. 1993. National hierarchical framework of ecological units. USDA Forest Service, Washington, D.C.

Fike, J. 1999. Terrestrial and palustrine plant communities of Pennsylvania. Pennsylvania natural diversity inventory.

Franklin, J., 1995. Predictive vegetation mapping: geographical modeling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geogr. **19**: 474–499.

Gessler P.E, Moore, I.D., McKenzie, N.J., Ryan, P.J. 1995. Soil-landscape modeling and spatial prediction of soil attributes. International Journal of Geographical Information Systems 9: 421–432.

Kienzle, Stefan. 2004. The effect of DEM raster resolution on first order, second order, and compound terrain derivatives. Transactions in GIS. 8:83-111.

Kong, N. 2006. Topographically-based landscape-scale ecological mapping in Pennsylvania. PhD. Dissertation. Intercollege Graduate Degree in Ecology. Penn State University, University Park, PA.

Lehman, G. S. 1962. An evaluation of the physical and hydrological properties of the soil mantle on Leading Ridge Watershed two**. M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Lynch, J. A. and E. S. Corbett, 1990. Evaluation of Best Management Practices for Controlling Nonpoint Pollution from Silvicultural Operations. Water Resources Bulletin 26:41-52.

MacMillan,R.A., Pettapiece, W.W., Nolan, S.C., 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 113**: 81–109.

Maune, D. F., editor. 2007. Digital elevation models and technologies and applications: the DEM user manual, 2nd Edition. The American Society for Photogrammetry and Remote Sensing. Bethesda, MD.

McCune, B. and Medford, M. J., PC_ORD Multivariate Analysis of Ecological Data, 5.18 ed., Gleneden Beach, OR: MJM Software, 2006.

McGaughey. Fusion version 2.90, available at http://forsys.cfr.washington.edu/fusion/fusionlatest.html.

McKenzie, N.J. and Austin, M.P., 1993. A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. Geoderma 57: 329–355

Moore, I. D., Grayson, R. B., and Ladson, A. R. 1991. Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. Hydrological Processes. 5:3-30.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A. 1993: Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57: 443-52.

Moriarity, B. 1996. Ecological landtype mapping unit legend for the Allegheny National Forest.

Myers, W.L. 2000. Landscape scale ecological mapping of Pennsylvania forests. Research report ER2002-2, Environmental Resources Research Institute, University Park, PA.

Niemann, K.O., Howes, D.E. 1991. Applicability of digital terrain models for slope stability assessment. ITC J 1991-3 : 127–138

Noss, R. F. 1987. From plant communities to landscapes in conservation inventories: a look at The Nature Conservancy. Biological Conservation 41: 11-37.

Olaya, V. 2009. Basic land-surface parameters. In: T. Hengl and H.I. Reuter, Editors, Geomorphometry — Concepts, Software, Applications, Developments in Soil Science vol. 33, Elsevier, Amsterdam pp. 141–169.

Peck, J.E. 2010. Multivariate Analysis for Community Ecologists: Step-by-Step using PC-ORD. MjM Software Design.

Pennock, D.J., Zebarth, B.J. and de Jong, E. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. Geoderma, 40:297-315.

Pfister, R. D., Kovalchik, B. L., Arno, S.F., Presby, R. C. 1977. Forest habitat types of Montana, USDA-Forest Service Gen. Tech. Rep. INT-34, Intermountain Forest and Range Experiment Station, Ogden, UT 174 pp.

Reiners, W.A., E.V. Axtmann and R.C. Thurston. 1999. Delineations of landtype associations for northwest Wyoming and the buffalo resource area. Bureau of Land Management, University of Wyoming, final report.

Robinson, G. R. 1959. Tree mortality and defects on the Leading Ridge watershed number one. M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Schmidt, J. and Hewitt, A. 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. Geoderma 121: 243–56.

Schultz, C. H. editor. 1999. The Geology of Pennsylvania. Geologic Survey Special Publication 1.

Sevon, W. D., 1998, Landform map of Pennsylvania: Geological Society of America Abstracts with Programs, v. 30, no. 1, p. 73.

Shields, R R. 1966. A shallow seismic refraction study of the soil mantle and bedrock configuration of Leading Ridge Watershed Two M. S. Thesis. School of Forest Resources. Penn State University, University Park PA.

Van Niel, Kimberly P, Laffan, S.W., Lees, B. G. 2004 Effect of error in the DEM on environmental variables for predictive vegetation modeling Journal of Vegetation Science 15: 747-756.

Wilson, J. P., Gallant, 2000. J.P. Wilson and J.C. Gallant, Editors, Terrain analysis: principles and applications, Wiley, New York.

Zenner, E.K., J.E. Peck, K. Brubaker, B. Gamble, C. Gilbert, D. Heggenstaller, J. Hickey, K. Sitch, and R. Withington. 2010. Combining ecological classification systems and conservation filters could facilitate the integration of wildlife and forest management. Journal of Forestry 108: 296-300.

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

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

Chapter 5

Conclusions

The Leading Ridge watersheds in the Ridge and Valley Province provide an interesting terrain that facilitates the use of new characterizations that can enhance the understanding of the relationships between terrain and vegetation. Commonly used terrain metrics such as elevation, slope, aspect, and TI do not adequately describe the complex relationships between vegetation and site in this region. For example, scalloped slopes are heterogeneous due to their complex washboard like pattern of being comprised by nine curvature classes in relatively equal proportions. The heterogeneity of their forest communities can be explained by this pattern, with concave areas being found immediately adjacent to convex areas. Using lidar-derived elevation data, we can identify these landforms occurring on the scale of 1-100 m, as well as the roughness of these landforms. This can facilitate an improved understanding of forested watersheds and the interactions between water, soil, topography, and vegetation. Processes are working on multiple scales simultaneously in an interactive manner, so by using a multi-scale, multidisciplinary approach to landform modeling, landform classification and vegetation predictions can be improved.

Using lidar, we can identify and protect small, headwater streams that are currently missing in the analysis due to lack of accurate hydrographic data. Riparian areas are known to be extremely important for protecting valuable ecosystem services and are hotspots for biodiversity, denitrification, flood control, and ground water recharge. In Leading Ridge, we are able to identify many more headwaters stream reaches than by using previously available data which can facilitate the identification of important riparian habitat on the landscape and also may lead to improvements in our ecosystem services models. For the purpose of stream network modeling,

resolution of the DEM cell does not have as important of an effect as vertical accuracy, at least in resolutions up to a 10 m. There was no statistical difference between the streams modeled using the 1 m or 10 m resolution lidar-derived DEMs, but there was a statistical difference between streams modeling using a 10 m lidar-derived and 10 m NED DEMs, with the 10 m lidar-derived DEM being significantly more accurate than that modeled using the NED DEM. Using a 10 m lidar-derived DEM, we could improve our stream network delineation over both site and landscape scales, greatly improving currently available spatial hydrographic datasets for forest planning, ecosystem management, and site planning. By using a smoothed DEM or a 3 to 10 m resolution DEM, errors caused by cultural hydrologic features such as culverts, roads, and trails can be reduced in the stream network mapping. When utilizing TI for ecosystem studies or hydrologic modeling, care must be taken when evaluating the effects of resolution, accuracy, and methodology. Using a fine resolution DEM with a single direction flow routing algorithm may inaccurately model extremely wet areas as parallel lines well up a watershed, while identifying areas immediately adjacent to streams as relatively dry due to the impact of the channelized contributing area on the TI calculation.

In addition to the considerations of resolution and accuracy for DEM-based modeling, there can be a substantial error and extraneous features introduced via the algorithm used for DEM-generation. Although QAQC RMSE may be low for some types of algorithms such as one based on the generation of the triangulated irregular network (TIN), this may not mean that it effectively models conditions such as surface roughness. RMSE of research grade and lower density topographic grade lidar are very similar for the area of Leading Ridge, although the DEMs exhibit different properties and appearances. Neither DEM matches surveyed surface roughness at the sub meter scale, although textural patterns are visible using both DEMs that seem to be responding to topography, geology, and soil. Using micro-topography derived from a 1 m DEM, surface expressions of changing soil and geology can be identified and delineated

interpretively. There was not one "best" metric, methodology, or resolution that modeled roughness in all settings but patterns emerged that suggest that metrics may all be reflecting common features.

By analyzing features on a broader scale than roughness modeling, dominant landforms emerge based on patterns of curvature. Curvature may be an efficient tool for modeling landforms because it is both responsive to differential patterns of weathering and is an important predictor in soil formation and hydrologic processes. Certain geologic features weather into different patterns of curvature, with resulting curvature directing water and other substances across the ecosystem which continues to perpetuate these differences. For example, Tuscarora quartzite weathers into ridges that exhibit the curvature pattern of rounded ridges, while Rose Hill shale forms ridges that are convex in both plan and profile curvature. Vegetation communities are following these patterns of curvature, with dry oak-heath forest types being found exclusively on rounded ridges, and more mesic compositions being found on other landform types. Although there is heterogeneity across formations of the same type, particularly within formations defined in part by the heterogeneity of their curvature, these formations can serve as an important tool for classifying the landscape. Also, some of the heterogeneity could probably be predicted by coupling landform analysis with a finer scale analysis such as the meter scale roughness analysis.

Lidar-derived DEMs offer advantages over previously available elevation data, with improved accuracy and resolution. It allows us to see and understand small scale features and micro-topography that can influence vegetation communities. Modeling patterns of curvature using traditionally derived DEMs would not represent the topography and patterns of curvature the same way that lidar data can do. Lidar may allow us to integrate scale into classification in a way that was previously impossible by classifying formations in terms of sums of their smaller, more diverse parts. Future work in the multi-scale nature of landform classification should refine major groups identified here using finer scale patterns of curvature or roughness data.

Another direction for future research is to broaden scale of this work and see what kinds of patterns emerge in different geologic and topographic settings. A first step would be to apply curvature and roughness modeling to regions of adjoining physiographic provinces such as the Appalachian Plateau Province and South Mountain Province, both areas which contain similar vegetation communities. It would be interesting to see if the same curvature patterns, stream delineation methods, and roughness signatures appear in these settings and if they have a similar interaction with vegetation communities. It would also be useful to analyze the fine scale roughness patterns that emerge in different vegetation settings, such as in areas of extremely dense evergreen shrubs, early successional forests, or old growth settings in order to expand the utility of this work. A third way this research can be continued is to use the accurate terrain and hydrological network data to explore the implications these data may have for currently utilized hydrologic models.

Much of the recent work in catchment science and critical zone science is in trying to understand water movement through watersheds above the point of delineated streams. This was a very difficult area to study prior to lidar because of the difficulty and expense of field sampling for soil properties. Research is currently being done to understand water transit times in a catchment, water storage in an ecosystem, and how water moves through an ecosystem from the time that it hits the surface until the time it becomes stream run-off. This water is difficult to trace, but through lidar, we may finally be able to predict where water is going. For example, in some areas of the watershed, streams were modeled using lidar that did not exist on the surface. In other areas, springs and streams emerged at locations not modeled using surface DEMs. By focusing on these areas, we can begin to sample and improve understanding of sub-surface water movement, and the role of surficial roughness features on flow routing.

There should also be a relationship between surface roughness of terrain and water storage. By integrating roughness metrics with flow accumulation algorithms, we may begin to

understand these temporary and long-term water storage predictors and the role of these micro-scale features on the hydrologic landscape. Combining these metrics with hydrologic models and discharge measurements, we may better understand short term storage and identify areas particularly important to ecosystems services such as denitrification and flood mitigation.

Despite all of these advantages, there are challenges to utilizing lidar related to the care that must be used to integrate data of various scales and from various lidar products simultaneously. Although there is a good deal of new information that can be extracted from a lidar-derived DEM, there is also much mis-information that can result if done incorrectly or without taking suitable precautions. Lidar can be said to provide us with a new set of eyes, and in order to take advantage of the tool, we should re-evaluate all previous assumptions regarding scale, terrain analysis, and DEMs. This work endeavors to do that, but much more work can still be done to verify and apply these results to a broader scale and to get this data into the hands of practitioners and managers.

Kristen M. Brubaker

Education
The Pennsylvania State University, University Park, PA**,** 2007-2011
    Ph.D. School of Forest Resources, watershed stewardship option
Mississippi State University, Starkville, MS, 2004-2006
    M.S. in Geosciences in the Teachers in Geosciences program
The Pennsylvania State University, University Park, PA, 1999-2003
    B.S. in Secondary Education, Earth and Space Science option

College Teaching Experience

2011-2012	Visiting Postdoctoral Fellow with the Center for Sustainability Education, **Dickinson College**
2009-2011	Natural Resources GIS (FOR 496A) **Penn State University**, Instructor
2009-2010	Remote Sensing and Spatial Data Handling (FOR 455), **Penn State University,** Instructor
Spring 2009	Introduction to Geographic Information Systems (EENV 360), **Susquehanna University**, Adjunct Professor

Publications

E.K. Zenner, J.E. Peck, **K. Brubaker**, B. Gamble, C. Gilbert, D. Heggenstaller, J. Hickey, K. Sitch, and R. Withington. 2010. **Combining ecological classification systems and conservation filters could facilitate the integration of wildlife and forest management. Journal of Forestry** 108(6): 296-300.

Sherwin, L., **K. Brubaker**, J. Kozak, A Lashaway, M. May. Center for Watershed Stewardship. In press. Natural Resource Condition Assessment of Valley Forge National Historic Park. Natural Resources Report. National Park Service.

Center for Watershed Stewardship. 2009. Upper Bald Eagle Creek Watershed Assessment. Penn State University.

Conferences, Meetings and Presentations

July 2011	Lidar imagery improves classification of forest function in the Ridge and Valley physiographic province of Pennsylvania. (poster) Gordon Research Conference for Catchment Science, Bates College, ME.
May 2011	LiDAR imagery improves classification of forested landforms in the Shale Hills/Susquehanna Critical Zone Observatory region of Pennsylvania. (poster) CZO All Hands Meeting, Tuscon, AZ.
Dec 2007	Organized and taught a workshop for college lecturers on remote sensing/GIS, Jalgaon, India.
Dec 2007	Presented GIS/Remote Sensing at Jalgaon DG INT Group Meeting for Digital Governance and Hotspot GeoInformatics , December 16-23, 2007.

Grants and Awards

Jan 2010	College of Agricultural Sciences Student technology grant $2500.00. "Professional Grade GPS unit for Forest Management"
Aug 2007-May 2008	Penn State University Graduate Fellowship