AWRA 2012 SPRING SPECIALTY CONFERENCE
New Orleans, Louisiana
March 26 – 28, 2012

A Proposed Workflow for Delineating Stream Networks from Lidar-Derived Digital Elevation Models to Update the National Hydrography Dataset (NHD) in the Pacific Northwest

Craig Ducey, Dan Wickwire, and Jay Stevens*

ABSTRACT: The Pacific Northwest Hydrography Framework (PNWHF) is supporting efforts to evaluate the use of lidar-derived digital elevation data to delineate hydrographic features to update the National Hydrography Dataset (NHD). This paper describes the analytical methods used to derive an initial stream network to be reviewed by NHD stewards for a 10-digit hydrologic unit (HU) in the Oregon Cascade Range.

KEY TERMS: National Hydrography Dataset (NHD), Lidar, Stream Network Delineation, Drainage Enforcement

INTRODUCTION

The goal of the National Hydrography Dataset (NHD) is to provide land managers and decision-makers with detailed representations of surface water and other hydrologic features within the United States accessible from a corporate-level GIS. When combined with other types of thematic data, users from a wide-range of disciplines access the NHD to either create cartographic reference maps or conduct complex spatial analyses of hydrologic systems. To ensure its long-term credibility in response to changing landscapes, technologies, and business needs, the Bureau of Land Management (BLM) and its Pacific Northwest Hydrography Framework (PNWHF) partners are continuously seeking new sources of baseline information and methods to improve the overall quality of the dataset's spatial geometry and attribution. Currently, attention is focused on assessing the potential of lidar-derived digital elevation data to derive more accurate and precise stream networks and hydrologic unit boundaries to correct known spatial irregularities in the current NHD and Watershed Boundary Dataset (WBD) resulting from interpretations of coarser-resolution data. Results from initial evaluations of hydrography modeled from lidar-derived digital elevation models (DEM) of bare earth surfaces appear consistent with the NHD stewardship community's requirements (Miller et al. 2004, Colson et al. 2006). Modeling complications associated with lidar's high spatial resolution, however, pose new challenges requiring innovative approaches to existing methods for deriving drainage patterns.

Recognizing the potential of lidar technology, the PNWHF sponsored the formation of a working group tasked with evaluating the use of lidar-derived DEMs as a principle data source used to update the NHD. The working group identified a successive, three-stage workflow leading to the acceptance of a revision to the national dataset: 1) preparing lidar DEMs for analysis and developing an initial stream network, 2) coordinating the review process between NHD stewards and GIS editors, and 3) migrating attribution from the original to lidar-derived stream features and metadata documentation. The project considered alternative methods in context of their impacts on limited business resources, likelihood of producing more valuable NHD products, and ability to satisfy the ultimate intended uses of the dataset. This document focuses on the PNWHF working group's experiences developing methods to accomplish the first stage of this workflow for a 10-digit hydrologic unit (HU) located in northwestern Oregon.

PILOT WATERSHED

The PNWHF selected the Little North Santiam River 10-digit HU (1709000505) as a pilot watershed for developing and evaluating a workflow for updating the NHD with hydrography modeled from lidar elevation data (Figure 1). Located on the western slopes of the Cascade Range, the watershed drains approximately 72,256 acres of heavily forested terrain into the North Santiam River near Mehama, Oregon. The area's hydrography is divided into five sub-watersheds (12-digit HUs) oriented east to west and ranging between 9,944 and 18,915 acres in size. Surface elevations range from 660 to 5,560 feet above sea level. The watershed's climate is typical of the Western Cascades physical province with precipitation occurring primarily as rain in cooler seasons and averaging between 70 and 103 inches annually (Anderson et al. 1998). Landslide

*Respectively, GIS/Remote Sensing Specialist, Hydrography Project Manager, Geographer. Bureau of Land Management Oregon State Office, 333 SW 1st Avenue, Portland, OR 97204-3421. Phone: (503) 808-6314. Email: cducey@blm.gov

activity of varying magnitude and expression is evident over 37% of the lower half of the watershed and 4% of the upper half (Sobieszczyk 2010).

Surface ownership and management responsibility in the lower three sub-watersheds is fragmented between private holdings, the State of Oregon, BLM, and US Forest Service (USFS). The USFS oversees the majority of the watershed's upper reaches as the Willamette National Forest and Opal Creek Wilderness. Approximately 380 miles of trails and roads are maintained throughout the watershed with the highest densities of highways, residential roads, active and remnant logging roads, and road right-of-ways occurring in the lower two sub-watersheds.

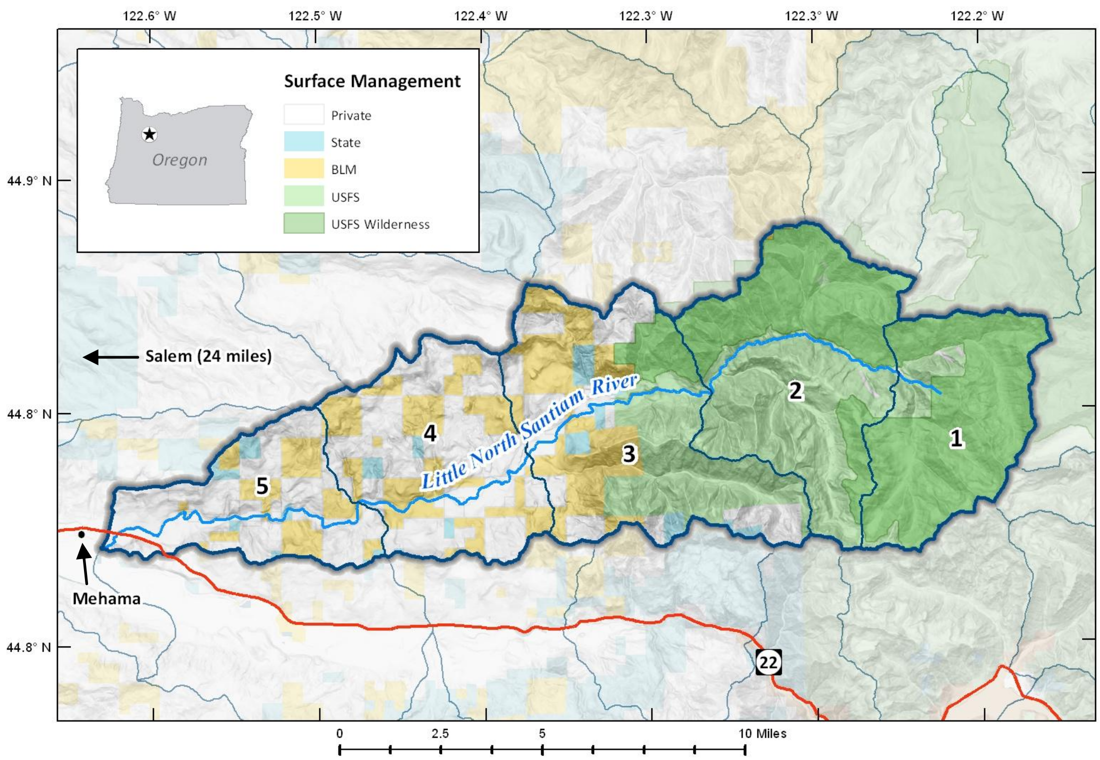

Figure 1. Surface management responsibility within the Little North Santiam River 10-digit HU and its five sub-watersheds (labeled 1 through 5 toward the watershed's pour point) is distributed between PNWHF partners and private ownership.

LIDAR ACQUISITION

Watershed Sciences, Inc. collected airborne lidar covering the study area for the Oregon Department of Geology and Mineral Industries' (DOGAMI) and Oregon Lidar Consortium between 17 September 2008 and 1 July 2009. Average pulse density for the entire acquisition is 0.73 pulses/foot^2 and average ground density is 0.14 pulses/foot^2. An absolute accuracy root mean square error (RMSE) value of 0.13 feet was determined by calculating the deviation between laser points and more than 30,000 Real Time Kinematic (RTK) ground survey points dispersed across the acquisition area (Watershed Sciences, Inc. 2009). The lidar vendor created 3-foot cell size, bare-earth DEMs from the lidar ground-classified point returns clipped to USGS 7.5 minute quad boundaries, which were visually inspected by DOGAMI for calibration and seam line offsets, tiling artifacts, and uncommonly high and low points. Following an independent vertical accuracy assessment, DOGAMI also determined the DEM tiles exceeded the maximum allowable vertical offset of 0.65 feet specified by the contract (DOGAMI 2009).

METHODS

After considering the overall size of the Little North Santiam River watershed, as well as the spatial differences in its physical landscape and predominant land ownership, we elected to focus our analysis at the 12-digit HU level. This decision allowed us to stagger the workload between the different steps of our three-stage workflow (described above), incorporate lessons-learned as we progressed, spread out the GIS editor's and different NHD steward's workloads, and decrease the computer resource demands necessary to process high-resolution imagery. The workgroup also agreed to leverage the capabilities of ESRI ArcGIS 10 software currently used by all PNWHF partners to avoid additional costs and guarantee data, models, and results were transferable between agencies.

Image Pre-Processing

We mosaicked the 3-foot, bare-earth DEM USGS 7.5" quad tiles produced by the vendor that intersected the WBD Little North Santiam River 10-digit HU boundary using mean cell values where tiles overlapped. The mosaic was resampled to 9-foot cell sizes using cubic convolution interpolation before being divided along 250-foot buffers about each of the nested WBD 12-digit HUs. Shaded relief and slope models were created for each sub-watershed to aid visualization of the lidar data.

Deriving Drainage Patterns

The standard sequence for modeling surface drainage patterns from DEMs involves: 1) resolving depressions, 2) flow routing, 3) calculating flow accumulation values at every cell, and 4) applying a flow accumulation threshold to derive the stream network (Jenson and Domingue 1988). Creating a depressionless DEM by removing sinks is required to ensure continuous surface flow within a drainage. Flow routing using the D8 method (O'Callaghan and Mark 1984) assigns each cell a flow direction value indicating which of its immediate eight neighbors it drains into. Flow accumulation quantifies the total upslope area draining to every cell in the terrain, and can be weighted to return contribution estimates other than a cell's planimetric area. Channels are delineated by selecting a flow accumulation threshold above which surface flow is predicted to initiate. The result can be combined with the flow direction grid to determine stream order, generate a vector representation of the stream network, and delimit watershed boundaries. Despite some well-documented drawbacks over alternative flow routing and accumulation methods, these algorithms are commonly available to our partners, easily scripted, and produce comparable results (Wolock and McCabe Jr. 1995, Barber and Shortridge 2005). For this pilot project, we chose to focus our efforts on manipulating the lidar-derived DEMs to ensure the best possible flow routing and channel initiation predictions.

Drainage Enforcement

Lidar-derived, bare earth DEMs depict the elevations of the furthest pulse returns visible by the sensor interpolated across a gridded surface. As a consequence, the true path of hydrologic flow can be obscured from the sensor's view by geographic features such as bridges and road beds overlaying culverts, which act as physical barriers by forming closed depressions on their upslope sides. Depressions also result from noise introduced when sampling and gridding the source lidar point data, artifacts introduced when removing vegetation and other physical structures to create the bare earth DEM, or by naturally closed features. If closed depressions are left in place, discontinuities in the stream network occur where stream features artificially terminate at a depression's lowest elevation. Enforcing downslope drainage by either filling depressions or cutting through their borders is, therefore, an important preprocessing step in hydrologic modeling.

The most commonly used technique for generating a DEM without hydrologic impediments is to fill closed depressions by artificially raising the elevations of surrounding cells until an outlet pour point is discovered and downslope flow is achieved. Unfortunately, filling sinks does not always produce the most satisfactory results, particularly when analyzing high-resolution elevation data where microrelief between cells can cause modeled channels to deviate from their true paths. For example, the first pour point identified at depressions formed behind culverts at road-stream crossings often occurs at the elevation of relief ditches built on the upslope sides of the road bed. Rather than moving through the culvert and under the road, filling causes flow to be diverted into the relief ditch and along the road gradient (Figure 2).

In general, depressions caused by either noise or other inaccuracies in the DEM can be satisfactorily resolved by filling. Depressions occurring behind physical features, however, are better addressed by cutting through the DEM and lowering the elevation values of cells at the depression's border to incise an outlet path oriented in the correct flow direction. Deciding which technique to apply at each of the appreciable number of depressions common within a watershed-scale, high-resolution DEM is challenging and unreasonably time-intensive. Carlson and Danner (2010) proposed a machine learning classification for isolating obstructions with bridge-like characteristics using 510 different measurements of local topographic and image properties collected at training samples manually selected throughout a DEM. Vianello et al. (2009) relied on field surveys and aerial photointerpretation to identify and embed hydrologic structures over automated techniques.

To address hydrologic barriers in our pilot watershed without the benefit of field surveys, we developed a semi-automated approach to drainage enforcement based on a modified version of the workflow presented by Poppenga et al. (2010) (Figure 2). First, all depressions are identified by filling every sink in the DEM and subtracting the original DEM from the result. Spatially adjacent cells with difference values not equal to zero represent regions around each sink where elevation values were modified to allow continuous flow. Area and depth measurements for depressions occurring behind hydrographic barriers spanned the full range of values for all depressions, and precluded us from distinguishing them using the conditional thresholds applied by Poppenga et al. (2010). Instead, we identified locations eligible for breaching by manually inspecting depressions intersecting the stream network generated after filling all sinks. Cutting through obstructions is accomplished by calculating the least accumulated cost path between the upslope depression's deepest point (source) and the position of an exterior cell with an elevation less than or equal to the depression's minimum (destination). Source points are located using zonal statistics and destination cells by incrementally buffering individual depressions outward until a qualifying cell is discovered. The destination cell's elevation is then assigned to all cells in the unfilled DEM to incise the enforcement path. Resolving one hydrographic barrier frequently reveals another obstruction downstream. Manually detecting all depressions requiring breaching, therefore, can require several iterations and extend the drainage enforcement process.

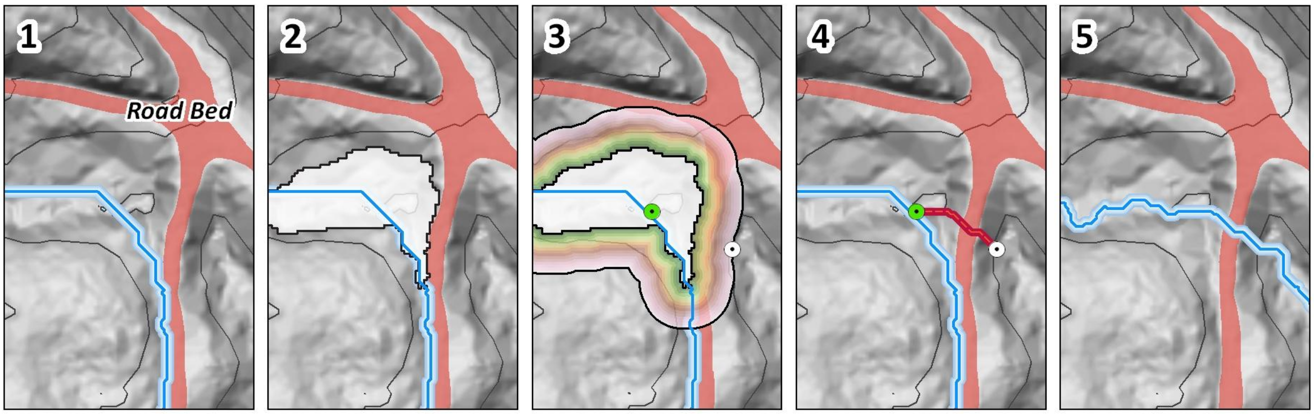

Figure 2. Obstacles impeding flow, such as road beds at stream crossings (1) typically form depressions on their upslope sides. Artificially filling depressions (white polygon in 2) can lead to erroneous flow direction predictions if the lowest elevation spill point occurs at the level of the relief ditch. To enforce drainage across the road bed, the location of the first exterior cell (white point in 3) with an elevation lower than the depression's deepest cell (green point in 3) is identified by incrementally buffering the depression outward (3). The depression is breached by assigning the exterior cell's elevation value to the least accumulated cost path connecting the two points (red line in 4) and repeating the channel delineation process (5).

Amending Parallel Stream Segments

One undesirable artifact of delineating streams from gridded surfaces is the occurrence of parallel stream segments wherever neighboring cells have the same flow direction and both exceed the specified flow accumulation threshold. Although infrequent, their existence is viewed critically by reviewers and misrepresents the true flow path. Parallel stream segments diminish the drainage pattern's cartographic appearance, and aggravate attempts to automatically migrate attribution from the original NHD line-work. To correct instances of parallel stream segments, we developed a solution mimicking the D8 flow routing routine where output cell direction values indicate the downslope neighbor with the highest upslope contributing area rather than the angle of steepest descent (Figure 3). Using this secondary flow direction grid, the flow accumulation and stream delineation process is executed as before, effectively snapping parallel stream segments together. Resolving parallel stream segments is accomplished after the drainage enforcement efforts are completed and just prior to the final post-processing steps applied to the stream network.

Surface Area Estimation

We relied on the expert opinions of the hydrologists and NHD stewards familiar with the watershed to estimate flow accumulation threshold values and provide manual adjustments during their reviews of the proposed stream network to help inform the locations of channel initiation points and transitions in periodicity. Given the landscape's rugged terrain, we provided reviewers a more realistic estimate of the overland contributing area above each cell by weighting the flow accumulation routine with an estimate of surface area. Following the grid-based logic presented by Jenness (2004), the surface area is approximated using the elevation values of a cell and its immediately adjacent neighbors. First, triangles are formed by connecting a cell's center-point with those of its eight neighbors. The length of the triangle facets are then determined from the end-point elevations and used to estimate each triangle's area in three-dimensional space. Surface area is expressed as the sum of each triangle's area intersecting the center cell's boundary. Surface area estimates within the pilot watershed ranged from 81 ft^2 in flat areas to more than 800 ft^2 in the steepest areas (Figure 4).

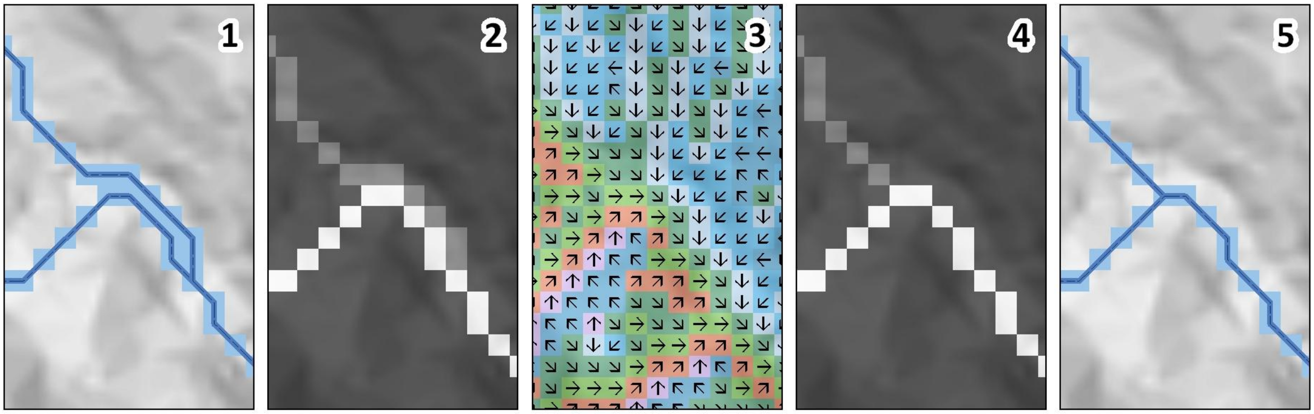

Figure 3. Parallel stream segments (1) occur where neighboring cells flow in the same direction and both exceed the specified flow accumulation threshold (2). Revisions are made by creating a secondary flow direction grid (3) indicating the downslope neighbor with the highest upslope contributing area and re-running the flow accumulation (4) and channel classification (5) routines. Image scale is 1:750.

Valley Classification and Stream Network Post-Processing

Deriving drainage patterns based solely on a predetermined flow accumulation threshold does not account for the geomorphic and environmental characteristics typically used to define a stream (Montgomery and Foufoula-Georgiou 1993). For example, different definitions of intermittent streams include characteristics such as a definable channel or evidence of annual scour and deposition (FEMAT 1993). In order to incorporate an additional level of specificity in our predictions, we required channels to initiate within areas satisfying a valley landform classification developed by Klingseisen et al. (2008) (Figure 4). Based on topographic attributes derived from the DEMs, a cell is considered in a valley if less than 40% of its neighbors within a 10-cell radius circular focal window have a lower elevation and its plan curvature is less than -0.5. Spatial discontinuities in this initial classification are addressed by recursively evaluating the relative position of a valley cell's downslope neighbors. If less than 45% of the cells within the same circular focal window surrounding a candidate cell have a lower elevation, then the candidate cell is considered in a valley. Initiation points predicted to occur above the extent of the valley classification are adjusted by masking channels with the upslope contributing area draining to the corresponding highest elevation valley cell. As a consequence of this post-processing step, channel inception points occur at variable flow accumulation values in the final, proposed stream network.

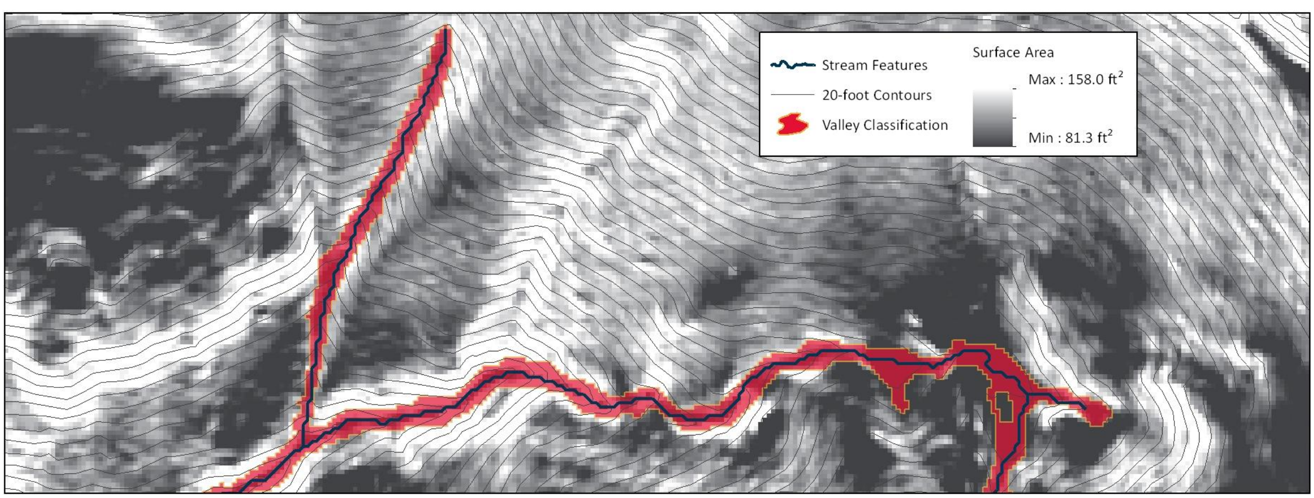

Figure 4. Surface area estimates and valley classification for a 1:4,800 area of the pilot watershed. After applying a threshold to the weighted flow accumulation to initially define channels, the stream network is post-processed to constrain channel initiation points to the valley classification.

SUMMARY AND CONCLUSIONS

The workflow developed for delineating hydrographic features in this pilot project includes a combination of automated and manual processing steps. We attempted to minimize the time required to perform each step by developing Python scripts and ESRI ArcGIS 10 geoprocessing models when possible. The cost of manually enforcing drainage at hydrographic barriers and amending parallel stream segments on a per case basis, for example, would be prohibitive without some level of automation.

Our motivation for weighting the flow accumulation routine with an estimate of surface area is to better account for terrain irregularities when modeling channel locations. Incorporating surface area effectively increases upslope contributing area values over planimetric measurements without altering the geometry of the drainage pattern. As a consequence, the surface area grid causes the density of streams to decrease by shifting initiation points downslope and reducing the occurrence of spurious channels originating on sloped, unidirectional inclines.

We recognize the parameters chosen for valley classification are not field-validated and could be more finely tuned to better depict this geomorphic component within the watershed. Reactions from NHD stewards, however, are positive and their revisions are not limited to the valley classification's extent. In addition, the valley classification was a valuable visual aid when manually identifying drainage enforcement locations. Future research should explore the potential value of landform classifications to inform flow routing algorithms and predictions of channel initiation and periodicity.

Efforts to improve the NHD with lidar-derived bare earth DEMs have merit. Our experience shows the quality of resulting stream features exceeds that of the original sources used for the Oregon dataset, and can effectively be added to the NHD as part of Oregon's stewardship program. The lidar-based delineations represent a significant workload that should not be underestimated. The efforts required of GIS editors and NHD stewards to ensure acceptance of the dataset are considerable, and often involve qualitative interpretations of the drainage pattern. Ultimately, programmatic use of lidar-derived DEMs to update NHD will be determined by the perceived benefits to the resource management programs relative to available budgets.

ACKNOWLEDGMENTS

We are grateful to the PNWHF for supporting this pilot project. We recognize the efforts of the BLM remote sensing and GIS staff, as well as the time and expertise provided by hydrologists and other natural resource specialists at the BLM, USFS, and Oregon Department of Forestry.

REFERENCES CITED

Anderson, E.W., Borman, M.M., and K.C. Krueger, 1998. The ecological provinces of Oregon – A treatise on the basic ecological geography of the state. Oregon State University. Agricultural Experiment Station. 138 p.

Barber, C.P. and A. Shortridge, 2005. Lidar elevation data for surface hydrologic modeling: Resolution and representation issues. Cartography and Geographic Information Science 32:401-410.

Carlson, R. and A. Danner, 2010. Bridge detection in grid terrains and improved drainage enforcement. *In* Proceedings of the 18th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. San Jose, CA.

Colson, T.P., Gregory, J.D., Mitasova, H., S.A.C. Nelson, 2006. Comparison of stream extraction models using lidar DEMs. *In* Proceedings of the Geographic Information Systems and Water Resources IV – AWRA Spring Specialty Conference. Houston, TX.

DOGAMI (Oregon Department of Geology and Mineral Industries), 2009. Willamette Valley lidar project, 2009 – delivery 13 QC analysis. Oregon Department of Geology and Mineral Industries, 16 p.

FEMAT (Forest Ecosystem Management Assessment Team), 1993. Forest ecosystem management: an ecological, economic, and social assessment. Number 1993-793-071. U.S. Government Printing Office, Washington, D.C., USA. pp. V-36.

Jenness, J.S., 2004. Calculating landscape surface area from digital elevation models. Wildlife Society Bulletin 32:829-839.

Jenson, S.K. and J.O. Domingue, 1988. Extracting topographic structure from Digital Elevation Data for Geographic Information System analysis. Photogrammetric Engineering and Remote Sensing 54:1593-1600.

Klingseisen, B., Metternicht, G., and G. Paulus, 2008. Geomorphometric landscape analysis using a semi-automated GIS-approach. Environmental Modelling & Software 23:109-121.

Miller, S.N., Shrestha, S.R., and D. Semmens, 2004. Semi-automated extraction and validation of channel morphology from LIDAR and IFSAR terrain data. *In* ASPRS Annual Conference Proceedings. Denver, Colorado.

Montgomery, D.R. and E. Foufoula-Georgiou, 1993. Channel network source representation using Digital Elevation Models. Water Resources Research 29:3925-3934.

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

Poppenga, S.K., Worstell, B.B., Stoker, J.M., and S.K. Greenlee, 2010. Using selective drainage methods to extract continuous surface flow from 1-meter lidar-derived digital elevation data: US Geological Survey Scientific Investigations Report 2010-5059, 12 p.

Sobieszczyk, S., 2010. Using turbidity monitoring and lidar-derived imagery to investigate sources of suspended sediment in the Little North Santiam River Basin, Oregon. Masters Thesis, Portland State University, Portland, OR.

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

Watershed Sciences, Inc., 2009. Lidar remote sensing data collection: DOGAMI, Willamette Valley Phase I Study Area. Watershed Sciences, Inc., 39 p.

Wolock, D.M. and G.J. McCabe Jr., 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. Water Resources Research 31:1315-1324.