Evaluating methods for automated mapping of apexes of non-linear eminences

Genevieve M. Joly*, Gaurav Sinha[a] and Wael Hassan

[a] Geography, Ohio University, Athens, United States
* gj410320@ohio.edu

Keywords: feature extraction, terrain analysis, topography, geomorphology, cartography

Introduction

The authors are collaborating with the United States Geological Survey (USGS) to extensively explore, compare, and integrate geomorphometric mapping methods for semi-automated mapping of areal extents of mesoscale landforms that people tend to recognize in the field and on topographic maps. Currently, USGS topographic maps explicitly represent landforms simplistically as point features retrieved from the Geographic Names Information System (GNIS) database. Previous work on a GEOBIA based methodology for extracting unique areal extents for landforms did not yield satisfying results (Arundel and Sinha, 2018). Now, the USGS is extensively testing machine learning methods for a wide variety of topographic mapping tasks.

In parallel, the authors are evaluating a few popular general purpose landform mapping algorithms for mesoscale landform mapping. Gruber et al. (2017) compared several such methods for soil mapping. That inspired Hassan's (2020) master's thesis which was conceived to compare Wood's (1996) geomorphometric features, Jasiewicz and Stepinski's (2013) geomorphons, and Weiss' (2001) topographic position index (TPI) methods for automated mapping of three broad landform categories: non-linear eminence (e.g., mount, hill, butte), linear eminence (e.g., ridge), and linear depressions (e.g., valley, gorge, ravine). The research discovered substantial susceptibility to parameter values and surprisingly low geometric and semantic correspondence between the methods.

Building on Hassan's work, a more systematic assessment of methods deemed suitable for mapping only non-linear eminences is underway. In this short paper, key findings are provided from an in-depth analysis of Wood's (1996) geomorphometric features and Jasiewicz and Stepinski's (2013) geomorphons for automatically identifying the apexes of visually salient non-linear eminences. If found viable, the method(s) can i) enhance peak records in topographic gazetteer databases (e.g., GNIS), and ii) help validate region growing and machine-learning algorithms for extracting areal extents of non-linear eminences.

Methodology

i) *Methods:* Wood's (1996) method constructs a local surface represented as a bivariate quadratic polynomial function for each elevation raster cell to classify each cell into one of six morphometric classes: *peak, pit, pass, ridge, channel* and *planar*. Jasiewicz and Stepinski's (2013) geomorphons are extracted based on line-of-sight analysis in eight directions around a cell, and depending on geomorphon shape, each cell is assigned one of ten classes: *flat*, *peak*, *ridge*, *shoulder*, *spur*, *slope*, *hollow*, *footslope*, *valley* and *pit*.

ii) *Parameter values:* For each method, Table 1 lists the parameters and corresponding values, and the total number of parameterized model runs needed for each study area.

iii) *Study Areas.* As in Hassan (2020) DEM rasters of 10-meter resolution were analyzed for two similar mountainous areas in the White Mountains (New Hampshire) and the Great Smoky Mountains (North Carolina and Tennessee border) and a third starkly different study area in the arid Colorado Plateau (New Mexico) to enable comparison of results between both similar and dissimilar physiographic regions.

	Morphometric Features	**Geomorphons**	
	Window size (# cells)	Outer search radius (# cells)	Inner search radius (# cells)
Neighborhood Size (# cells)	11	11	0, 5
	21	21	0, 10
	31	31	0, 10, 15
	41	41	0, 10, 15
	51	51	0, 15, 25
	61	61	0, 15, 25
Slope threshold	1, 5, 10, 15, 20	1, 5, 10, 15, 20	
Curvature	0.001, 0.0001	N/A	
Total # of runs	60	80	

Table 1: Parameter list and values for apex mapping.

iv) *Visual analysis.* Every combination of parameters for a study area yielded a separate raster map layer, which was overlaid on terrain hillshade, contour lines, and topographic basemaps. Although map overlay based visual assessment is subjective, panning across study areas is the best and only way to get a holistic perception of the quality of results. This must be first step, therefore.

v) *Confusion matrices.* Visual analysis helped narrow down the list of viable model runs, which were then compared in pairs using the familiar confusion matrix analysis. The percentage areal overlap between identical (diagonal cells) and different (off-diagonal cells) feature types was used to determine the quantitative impact of parameters on the percentage of *peak* cells.

vi) *GNIS Summit feature proximity analysis*. GNIS *Summit* features represent the apexes of a subset of topographically salient non-linear eminences. Distance from GNIS *Summit* features to the nearest extracted *peak* and *summit* polygons was measured using the ArcGIS Pro *Near* function for external semantic validation.

vii) *Workflow automation.* The *r.param.scale* and *r.geomorphon* functions in GRASS GIS (accessed via QGIS software) were used, respectively, to get outputs from the morphometric and geomorphons methods. The entire workflow from running models via QGIS to spatial analysis and creation of summary tables has been fully automated with Python scripts. Visual analysis of output rasters was done with ArcGIS Pro GIS software.

Results

For illustration purposes, Figures 1 and 2 show the impact of parameter values on the quality of *peak* and *summit* objects for a strategically selected portion of the Great Smoky Mountains study area. For additional comparison purposes, non-linear eminences and depressions are also shown. These figures can be the reference for the discussion of the findings summarized below.

i) *Study area.* Study area choice was not found to matter much for choosing the best parameter values based on visual comparisons, suggesting parameter values can be stable across varied landscapes for mesoscale non-linear eminence mapping However, the overall quality (location and shapes) of extracted *peaks* or *summit* cells was much better for the Smoky Mountain and White Mountain mountainous areas than for the arid and flatter Colorado Plateau's isolated and smaller non-linear eminences.

ii) *Window size / search radii.* From visual analysis, window size or outer search radii is the most important parameter that affects the size and shape of extracted features. For extracting morphometric *peaks*, window sizes smaller than 300 meters (31x31) produced noisy outputs with too many peak cells, while window sizes larger than 400 meters (41x41) produced noticeably few peaks with undesirably large and inappropriately located areal patches. Similarly, for the geomorphon method, an outer search radius between 300 to 400 meters (with an inner search radius of 150 meters) was found optimal for extracting *summit* cells. However, the shape and location of many extracted objects for even the best model does not seem satisfactory.

iii) *Slope.* For morphometric *peaks*, only slope thresholds between 5°-10° should be used. Results degrade appreciably beyond that narrow range. Geomorphon *summits* are even more susceptible to slope values, with only a slope of 1° found to yield acceptable results.

iv) *Curvature*. This parameter is of interest only for morphometric peak mapping. Values of .001 produced much fewer and unacceptably large areal patches making it clear that only curvature values in the vicinity of .0001 are viable.

v) *Confusion matrix analysis*. For both methods, analysis of confusion matrices (using overall similarity, Kappa, Cramer's V, and Contingency C measures) for model runs with parameter values close to the best parameter values exhibited much higher similarity measures than other model pairs confirming the identification of optimal range of parameter values from visual analysis.

vi) *GNIS Summit proximity analysis*. GNIS *Summit* feature distances to nearest *peak* and *summit* polygon were should ideally be within 30 meters, accounting for spatial inaccuracy of GNIS feature locations. The best parameter combinations that maximized proportion of *Summit* features at shorter distances corresponded to the best models determined from visual analysis only for geomorhon *summits,* but not for morphometric *peaks*, for which a slope of 20° was found to be best. As mentioned earlier, slopes exceeding 10° produce poor quality peaks. However, even for the best models, many candidate peaks and summits emerge nearby making it quite difficult to decide on a unique corresponding polygon for any GNIS *Summit* feature.

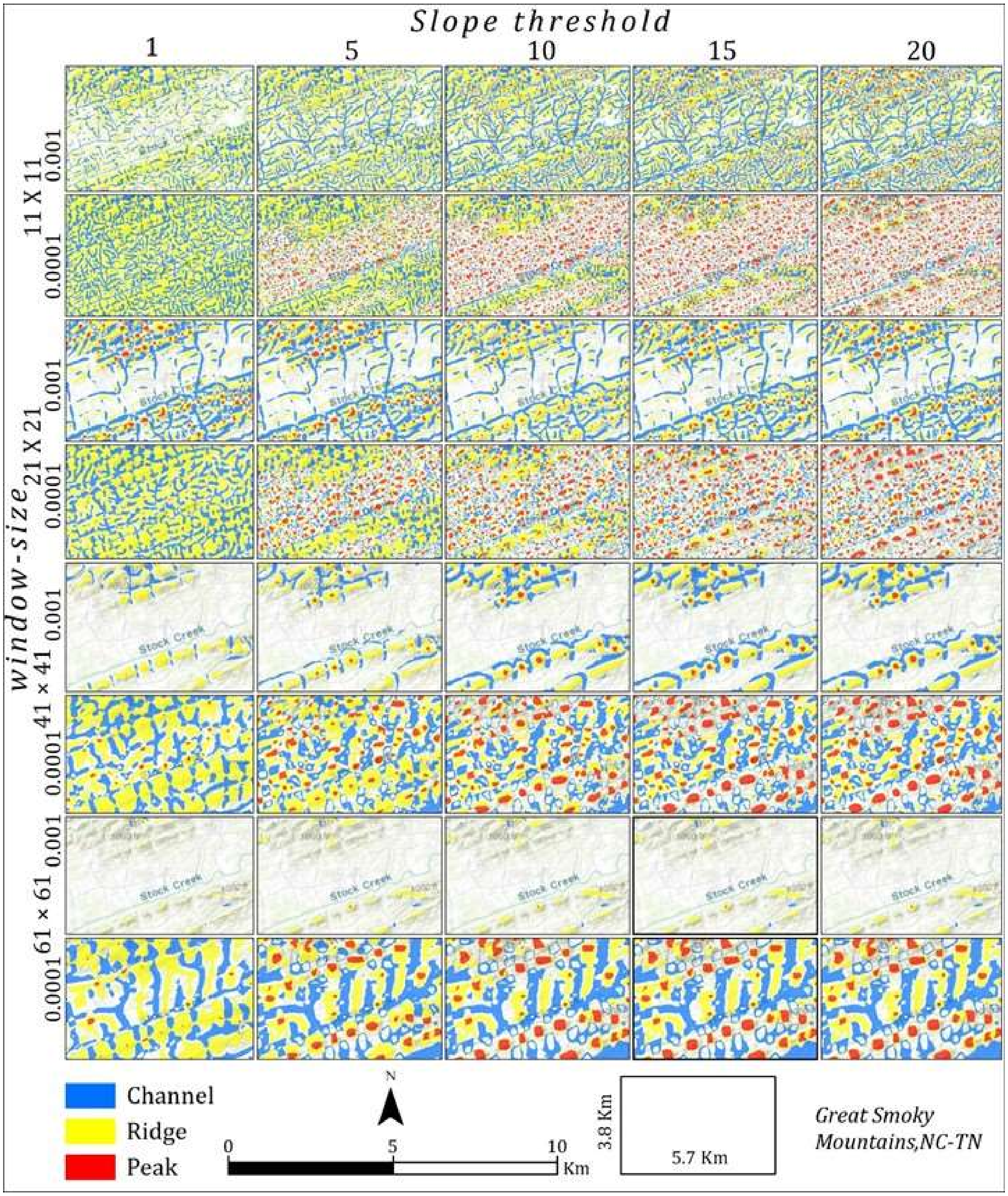

Figure 1: Mapping morphometric *features* for different parameter values for the Great Smoky Mountains study area.

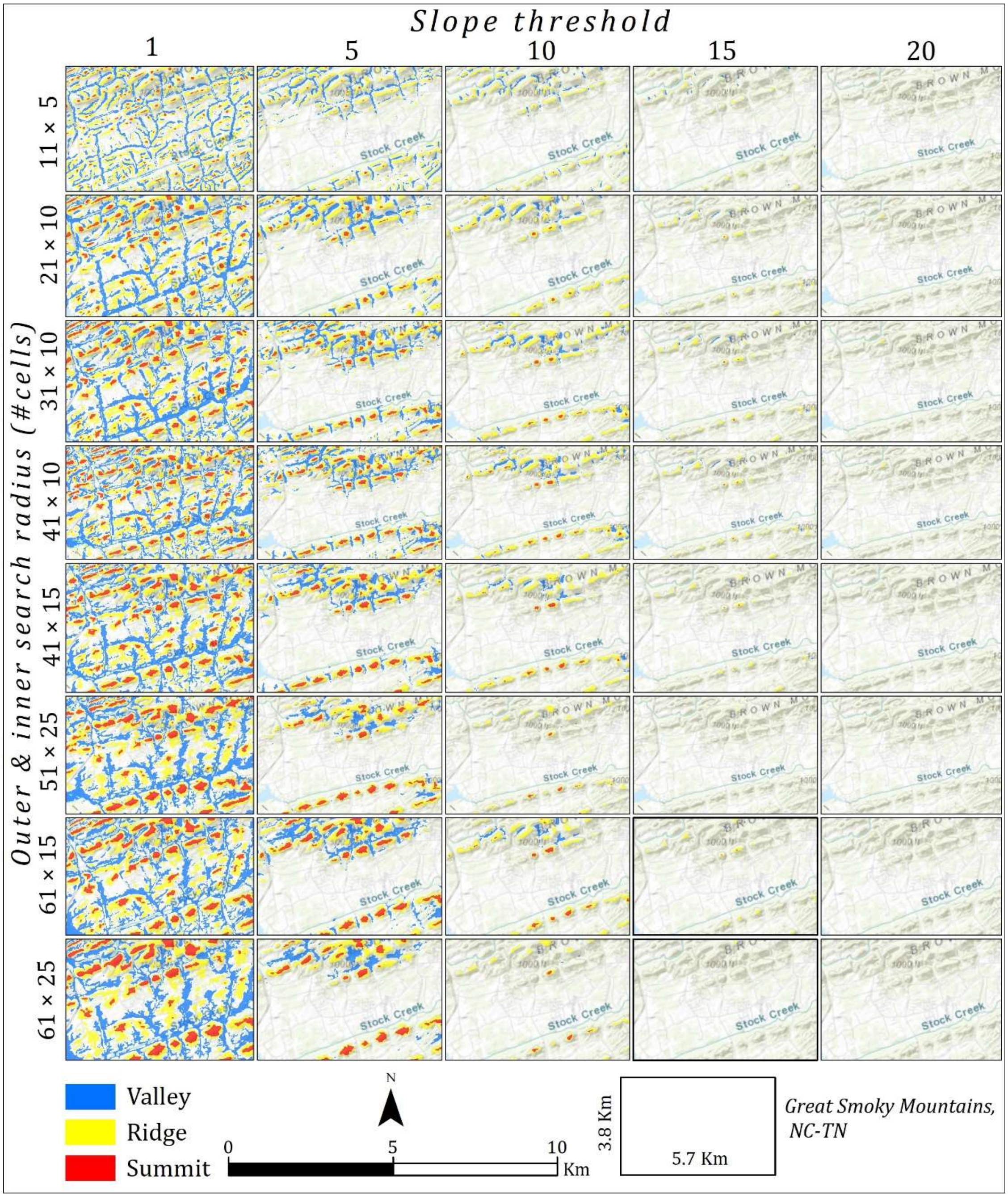

Figure 2: Mapping geomorphons based *summit* objects for different parameter values for the Great Smoky Mountains study area.

Conclusion

There is a narrow range of parameter values for which cognitively plausible morphometric features and geomorphon *summits* can be extracted as candidates for apexes of non-linear eminences. However, the shapes and locations of objects does not still match what people would identify as apex regions. Quantitative summaries can be misleading and can only be used to confirm findings from visual analysis, but not be independently used for decision making. Since the workflow is fully automated, new study areas and different parameters can be explored with ease and efficiency. While a lot of quantitative data will still be analyzed for the two methods, neither of the investigated methods is well-suited for specific geomorphometry. They are unlikely to produce objects of high quality and accuracy expected for national scale mapping of non-linear eminences. This is in contrary to what is generally assumed about these methods, and the main finding of this project. Thus, the next step will be to improve the eminence-core method (Sinha and Arundel, 2021) and test other contour-based eminence mapping methods available in the literature since they are designed to ensure reasonable shapes of extracted apexes and cores of non-linear eminences.

References

Arundel, S., & Sinha, G. (2018). Validating the use of object-based image analysis to map commonly recognized landform features in the United States. *Cartography and Geographic Information Science*, 46(5), 441-455.

Gruber, F., Baruck, J., & Geitner, C. (2017). Algorithms vs. surveyors: a comparison of automated landform delineations and surveyed topographic positions from soil mapping in an alpine environment. *Geoderma*, 308, 9-25.

Hassan, W. (2020). Comparing Geomorphometric Pattern Recognition Methods for Semi-Automated Landform Mapping. *Unpublished MS Thesis*. Department of Geography, Ohio University, Athens, OH, USA.

Jasiewicz, J., & Stepinski, T. F. (2013). Geomorphons—A pattern recognition approach to classification and mapping of landforms. *Geomorphology*, 182, 147–156.

Sinha, G. (2008). Delineation, characterization, and classification of topographic eminences. PhD Dissertation, Department of Geography, The State University of New York at Buffalo, Buffalo, NY, USA.

Sinha G., & Arundel, S. (2021). Automated extraction of areal extents for GNIS Summit features using the Eminence-Core Method. *Proceedings of the Geomorphometry 2020 Conference*. Perugia, Italy, June 2021.

Weiss, A. (2001). Topographic position and landforms analysis. Poster presentation, *Esri User Conference*, San Diego, CA, 2001.

Wood, J. (1996). The Geomorphological Characterisation of Digital Elevation Models. Unpublished PhD Thesis, Leicester University.