Adaptation and Comparison of Machine Learning Methods for Geomorphological Mapping and Terrace Prediction: A Case Study of the Buffalo River, Arkansas

by

Benjamin Thomas Swan

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science in Geography

Auburn, Alabama
August 5, 2017

Keywords: GIS, Geomorphology, LiDAR, Machine Learning, Remote Sensing, Terraces

Approved by

Stephanie Shepherd, Chair, Assistant Professor of Geosciences
Robert Griffin, Assistant Professor of Atmospheric Science (The University of Alabama in Huntsville)
Luke Marzen, Professor of Geoscience

Abstract

Working with a high-resolution LiDAR dataset, this research adapts existing techniques from machine learning to the problem of predicting general and specific geomorphic features in the complex terrain of the Buffalo River in Arkansas. This process is complicated by the fact that such geomorphic features are frequently not well-bounded and lack a one-to-one relationship between physical form and geomorphic class. These issues were addressed by analyzing terrain features within a spatial context using both local and regional land surface parameters. After selecting a horizontal resolution and degree of smoothing for creating our DEM that would maintain sufficient detail to distinguish the terraces and their boundaries while avoiding excessive noise, the study area was divided into three reaches based on general lithology and morphology. Using SAGA GIS, a free and open-source software, land surface parameters were calculated for each reach and five leaners—representing four distinct inductive biases—were tested. The results for each classifier were then validated using a dataset which combined field-mapped terraces and manually-delineated landform elements. It was found that Bayesian, Random Forest, and Support Vector Machine classifiers were the most accurate, while distance-based methods struggled to achieve acceptable accuracy. Support Vector Machines produced the smoothest class boundaries and mapped landforms in a way that was subjectively closest to manual methods; however, Bayesian and Random Forest approaches were more consistently accurate.

Acknowledgements

First and foremost, I would like to acknowledge and thank my advisor, Dr. Stephanie Shepherd. She has worked tirelessly with me through the long and twisted path of this research, resolutely taking on the unenviable task of beating good science writing into me and keeping me grounded.

I would like to thank my other committee members: Dr. Robert Griffin and Dr. Luke Marzen, first for teaching me the principles of GIS and Remote Sensing and also for always being ready to offer advice and ideas.

I would like to thank Dr. Amanda Keen-Zebert and Richard Hutto for teaching me about the geology and geomorphology of the Buffalo River.

Lastly, I must thank those people who made this possible in the first place: my parents, for their steadfast support, and my wife Kelly, for all her support and more.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

LiDAR	Light Detection and Ranging
ML	Machine Learning
GIS	Geographic Information Systems
RS	Remote Sensing
DGM	Digital Geomorphological Mapping
DEM	Digital Elevation Model
LPC	LiDAR Point Cloud
AGS	Arkansas Geological Survey
USGS	United States Geological Survey
NHD	National Hydrography Dataset
SVM	Support Vector Machine
LSP	Land Surface Parameter
RF	Random Forest
MD	Mahalanobis Distance
WTA	Winner-Takes-All
NB	Normal Bayes
VDTCN	Vertical Distance to Channel Network
TWI	Topographic Wetness Index

SWI	SAGA Wetness Index
TO	Topographic Openness

1 – Introduction

Digital Geomorphological Mapping (DGM), defined here as any computer-based landform mapping which does not use field surveys as its primary source data, is a fast-growing area of research in a number of fields, including geography, geomorphology, soil science, and ecology. Remotely sensed data, particularly stereo photographs and interferometric radar images, have long been heavily employed in landform mapping and landscape interpretation. Yet, in contrast to the plethora of sophisticated partly and fully automated techniques used in land cover classification, feature recognition in digital geomorphological mapping has, until recently, largely been a manual and interpretive process (Bishop et al., 2012; Evans, 2012; Evans et al., 2009). This is rapidly changing, however, as the increasing availability of high-resolution digital elevation data, cheap computing power, and robust software packages—often free and open source—have fueled intense research interest in applying greater levels of automation to geomorphic mapping and landform identification (Bishop et al., 2012; Dehn et al., 2001; Tarolli, 2014). Many programs and techniques use if-then logic and specified thresholds to create a deterministic partitioning of landscapes into essential forms, typically in a semi-automated manner (Klingseisen et al., 2008; Stout and Belmont, 2014; van Asselen and Seijmonsbergen, 2006). There is also growing interest in adapting the type of machine learning techniques used in remote sensing to landform classification (Arrell et al., 2007; Burrough et al., 2000; Ehsani and Quiel, 2008; Matías et al., 2009).

The research presented here is a quantitative and qualitative comparison of the performance of five supervised classifiers, chosen to represent several major families of machine learners, in mapping general and specific landforms in the heterogeneous terrain of the Buffalo River, Arkansas.

Segmentation and Classification of the Landscape

The Earth's surface is mathematically continuous, yet the division and subdivision of this surface into discrete and reliably identifiable units is important in many natural sciences. For meteorologists, ecologists, and hydrologists, landforms can represent important boundary conditions (Dehn et al., 2001; Deng, 2007; Hutchinson, 2008). Landform classification is particularly central in geomorphology and geomorphometry which both focus on how surfaces shape and are shaped by gravity-driven flows of energy and mass (Minár and Evans, 2008; Passalacqua et al., 2015). In science, a good classification system is one that allows researchers to make useful generalizations from the individual instances of different classes. In geomorphology, this means that process and ontology should define landforms, not morphometry alone (Buffington and Montgomery, 2013; Evans, 2012; Minár and Evans, 2008; Wheaton et al., 2015). Incorporating process and ontology into a DGM classification presents three major challenges. First is the inverse problem of inferring process from form: landforms do not have a one-to-one relationship with the processes by which they were created. Similar processes can produce morphologically heterogenous landforms while, in turn, distinct processes can produce landforms with similar morphologies (Bishop et al., 2012; van Asselen and Seijmonsbergen, 2006). Second, division of the continuous land surface into discrete units is dependent upon semantic models of what each landform is and how

it can be delineated from other landforms (Brandli, 1996; Dehn et al., 2001; Deng, 2007; Drăguţ and Eisank, 2011). Finally, specific landforms are frequently not well-bounded, but instead have a fuzzy and uneven transition from one form to another (Bishop et al., 2012; Evans, 2012; Wheaton et al., 2015). As a consequence of this, landform classification, even when based as much as possible on quantitative standards, is inherently to some degree subjective and interpretive (Deng, 2007; Evans, 2012; MacMillan and Shary, 2009). This problem of inferring process from form is even more acute when attempting to move beyond general landforms, such as channels and passes, to specific landforms, such as terraces or drumlins, which may be the result of processes no longer active on the landscape.

Significance of Investigation of Supervised Machine Learning Techniques

There is active development of semi-automated tools and both unsupervised and supervised machine learning techniques to deal with this complex classification task. Supervised machine learning has a number of qualities that compare favorably with semi-automated tools and unsupervised methods, making it a suitable alternative for those who want to take advantage of advances in remote sensing and GIS, particularly for those without specialized knowledge of feature classification.

Unsupervised learners, such as clustering algorithms, have long been an important tool in remote sensing; however, they have several disadvantages, as explained by Campbell and Wynne, Lillesand et al, and Tso and Mather (2011; 2015; 2009). All feature classification involves connecting patterns in input data, termed data classes, to information classes, which in this case are landform classes. Unsupervised learners create classes directly from data, which means that they are easy to train, but also means that

they have difficulty dealing with complex or ill-defined classes such as specific landforms. Since the classes created by these learners are based purely upon groupings in data, they require an expert analyst to turn these data classes into information classes. This is often a time-consuming process, made especially challenging when one information class is split into many data classes or one data class corresponds to more than one information class. Attempts to apply these methods for landform classification ran into the additional problem of data classes which had no apparent geomorphic meaning at all (Arrell et al., 2007; Burrough et al., 2000; Evans et al., 2009).

A number of purpose-built semi-automated tools are available for landform classification, examples including LandSerf (Wood, 2009, 1996), LANDFORM (Klingseisen et al., 2008), and Terrain Surface Classification (Iwahashi and Pike, 2007). While these tools are usually designed to classify general landforms, some classify specific landforms, such as the TerEx toolbar for terrace delineation (Stout and Belmont, 2014). In comparison to unsupervised methods, these tools can provide results that are directly meaningful. On the other hand, they have several limitations. Since the classes output are fixed, the semantic model of landforms used in any classification is literally hard-coded, limiting the ability of users to adapt the tools to the requirements of different studies. Similarly, because the potential number and type of input data are predetermined, users cannot experiment with the use of additional land surface parameters or ancillary data. Finally, acceptable results may depend upon precise settings for parameter values, increasing the complexity for the user and limiting the ability to transfer successful methods from one setting to another (Iwahashi and Pike, 2007; Klingseisen et al., 2008; Stout and Belmont, 2014).

Supervised machine learning techniques have qualities that help them overcome these shortcomings. Unlike unsupervised learners, they learn rules not just from data but also from training samples, supervised learners are inherently adapted to leverage expert knowledge to overcome difficult classification problems (Shalev-Shwartz and Ben-David, 2014; Tso and Mather, 2009). In comparison to semi-automated tools, their ability to learn models from data means that they are not as dependent upon the user's prior knowledge about the statistical characteristics of the input data or upon the fine-tuning of numerous parameters to the particular study area, both of which make it easier to transfer a methodology from one area to another. Finally, because the results of supervised classifications can be compared to reference data created using the same standards as training data, their accuracy can be objectively validated, even when the classes used in a study were not defined using exact mathematical values.

Research Questions

While supervised machine learning is a promising direction for research in digital geomorphological mapping, the use of these methods is still in its infancy. This research seeks to expand the body of knowledge concerning the application of supervised machine learning to general and specific landform classification by addressing two questions.

(1) Direct comparison of different machine learners in geomorphic classification has been limited, especially for more advanced learning techniques. Given the same training and input data, will learners with different inductive biases produce quantitatively or qualitatively different results?

(2) Can the incorporation of spatial-contextual information, in the form of regional Land Surface Parameters, significantly improve landform classification accuracy?

Case Study—Mapping Potential Fluvial Terraces along the Buffalo River, Arkansas, and its Major Tributaries

This project contributes to continuing investigations into the geomorphology of the Buffalo National River and its principle tributaries, concentrating on the role of lithology in determining valley width and the formation and preservation of fluvial terraces. The U.S. Geological Survey (USGS), the Arkansas Geological Survey (AGS), and (Keen-Zebert et al., 2017) have field-mapped terraces along the main-stem of the Buffalo and its larger tributaries. Taking these field-mapped geomorphologic features and a high-resolution LiDAR point-cloud dataset as source information, this research explores the utility of machine learning methods in the prediction of general and specific geomorphometric features.

Study Area

The Buffalo River in Arkansas is a 5th-Order tributary of the White River characterized by a gravel-mantled bedrock channel, deeply incised meanders, and towering bluffs (Keen-Zebert et al., 2014). Most of the Buffalo's 238 km course is protected by the Buffalo National River, a unit of the US National Park System, while the remainder flows through wilderness areas within the Ozark-St. Francis National Forest; therefore, the main-stem of the Buffalo flows almost entirely through public lands, a rare distinction among American rivers (Benke and Cushing, 2005). Additionally, its status as a National Scenic River has protected it against dams or any other hydrologic alterations,

making it one of only 42 free-flowing rivers more than 200 km in length in the contiguous United States (Benke, 1990). Easy public access and a lack of impoundments to submerge any of the geomorphic features in its watershed make the Buffalo a suitable candidate for geomorphologic study. The Buffalo flows for its entire length within the Ozark Plateaus physiographic province (Figure 1), which encompasses much of southern Missouri, northern Arkansas, and a small part of eastern Oklahoma. The Ozark region is

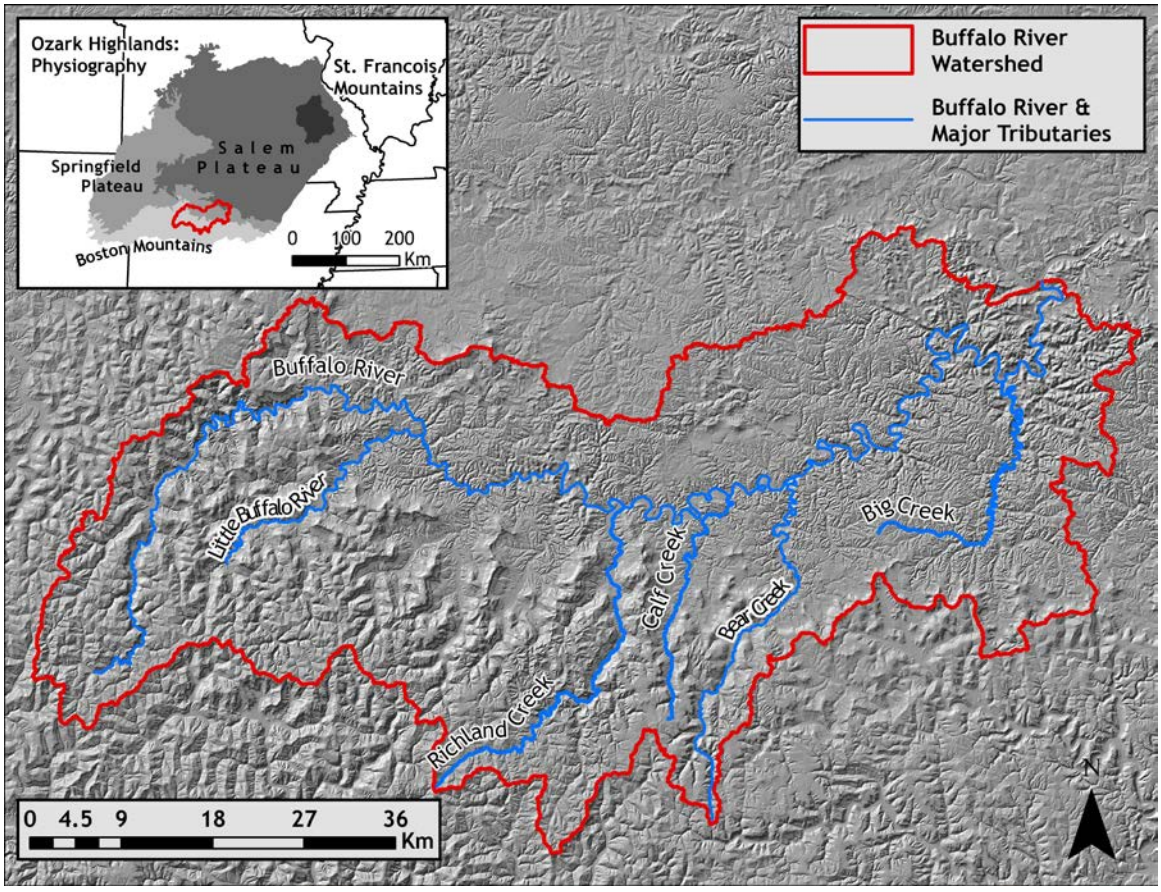

Figure 1: The Buffalo River and major tributaries shown with the physiographic provinces of the Ozark Highlands.

a broad and irregular structural dome consisting mostly of nearly flat-lying (< 3º) sedimentary strata that dip gently out and downwards from the St. Francois Mountains of east-central Missouri, an area of moderate relief underlain by pre-Cambrian igneous rocks. The River's headwaters are in the Boston Mountains, an area with shale-and-

sandstone capped ridges that reach upwards of 700 meters to be the tallest in the Ozarks. The Buffalo flows for most of its length along the Springfield Plateau, which is characterized by a karst hydrology, resulting from the cherty limestone of the Boone Formation (DiPietro, 2013; Fenneman, 1928; Keen-Zebert et al., 2017).

Terraces and the Geomorphology of Bedrock Channels

With the exception of glaciated regions, the rate and form of landscape evolution in mountainous terrain is controlled by bedrock channel streams. By controlling the lower boundary condition of hillslope erosion, they set the overall rate of denudation in uplands and are the dominant medium by which the effects of tectonic or isostatic uplift and changes in base level or climate are transmitted throughout the landscape (Hancock and Anderson, 1998; Stark, 2006; Whipple, 2004). Yet—in comparison to the current state of knowledge about alluvial streams—bedrock channel processes, such as their hydraulic geometry or the exact mechanisms of channel erosion, are still imperfectly understood (Tinkler and Wohl, 1998; Wohl, 2014). One way to gain insight into bedrock channel incision is to reconstruct their past longitudinal profiles by mapping fluvial terraces (Pazzaglia et al., 1998).

Streams may simultaneously deepen their channel by downcutting and migrate laterally across their valley. Typically, the former process occurs slowly enough that the latter widens the valley and allows floodplains to develop. When a change in boundary

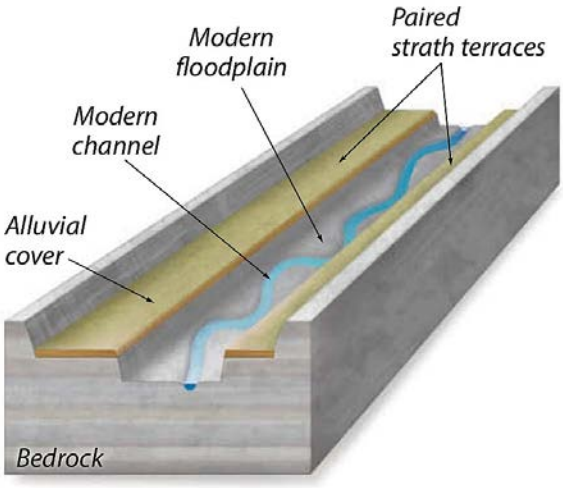

Figure 2: Strath terraces can form when a period of rapid incision results in stream eroding its channel faster than the rate of its lateral migration. Eventually, new floodplains form while the prior floodplains are left above the maximum flood level as terraces. Unlike this illustration, the vast majority of terraces along the Buffalo are unpaired. Figure from Bierman & Montgomery (2008).

condition disrupts this dynamic equilibrium, however, a stream may begin to incise its valley, leaving some portions of its previous floodplain "abandoned" far above maximum stage as terraces (Leopold et al., 1995). Terraces along the Buffalo are strath terraces. These terraces, illustrated in Figure 2, are associated primarily with bedrock channels and form when an increase in a stream's ability to cut into its bed leads to an episode of channel incision. Changes in sediment supply and/or stream power force deposition of large amounts of sediments, aggrading the stream channel and burying the original floodplain surface in a thick mantle of alluvium. Later changes in the stream's sediment carrying capacity allow it to rapidly erode through these unconsolidated sediments, cutting downwards faster than lateral migration (Bierman and Montgomery, 2014).

The geologic and geomorphic settings of the Buffalo River make it a suitable candidate for the study of the role that lithology plays in landscape evolution. Relative tectonic stability, mild isostatic uplift, and nearly-flat stratigraphy which provides little

structural control on landform development all combine to make the role of lithology more apparent in this region. Furthermore, due to Paleozoic faulting, the Buffalo cuts through repeated sequences of the Everton Formation of sandstones and dolomites and the Boone Formation of cherty limestones. This allows the effects of the lithology of a given reach to be more easily separated from effects that vary with the discharge found at different points along the river (Keen-Zebert et al., 2017). There is a consistent difference in valley width and the preservation of terraces between the two lithologies. The less resistant limestone reaches generally having wider valleys and larger terraces, whereas the more resistant sandstone and dolostone reaches had narrower valleys and smaller terraces with more of the higher (older) terraces being preserved (Keen-Zebert et al., 2017)

2 – Background

Geomorphometry

Geomorphometry, once considered a subset of geomorphology but increasingly perceived as a distinct discipline, is centered on mathematical descriptions of planetary surfaces (Pike, 1995; Pike et al., 2009). Its essence is the study of said surfaces under the influence of gravity and gravitationally-driven forces, examining how these forces shape surfaces and, in turn, those surfaces shape the forces acting on them. In this regard, geomorphometry is closely related to process geomorphology, with its similar focus on the immediate mechanisms altering landscapes and the practice of classifying landforms in an effort to better understand these mechanisms.

Modeling Land Surfaces

The mathematical descriptions of the Earth's surface used in GIS analysis are known by a number of names, including digital surface models, digital terrain models, and digital elevation models (DEM), the most common term. The key word in those terms is "model": no matter how detailed, a DEM is a simplified abstraction of the Earth's surface created with particular objectives and uses in mind. Since current models can only store a single elevation value for a given point in space, tradeoffs must be made depending on the type of analysis in question (Evans, 2012; Köthe and Bock, 2009). For instance, a civil engineer may want a bridge deck to be shown in a DEM while a hydrologist would need to remove it so that a flow routing algorithm

could properly route a stream through it. The type and quality of information in any terrain model is a function of both the source data and the way a continuous surface was created from that data. The vertical and horizontal resolution of a terrain model are primarily determined by the resolution and density of the original ground survey points or blocks. Gaps and errors in the terrain model are related to challenges of the specific terrain being surveyed, the technology used in the survey, the post-processing performed, and the method of interpolation (Passalacqua et al., 2015; Schmidt et al., 2003; Wilson, 2012).

With high vertical accuracy and horizontal point sample density, LiDAR has become very popular for high resolution terrain mapping, especially over large areas where economies of scale favor it over stereo photogrammetry, another technology capable of high resolution terrain measurement (Fowler et al., 2007; Renslow, 2012). This is evidenced by its selection by the US Geological Survey as the preferred source of elevation data for the 3D Elevation Program (Sugarbaker et al., 2014). One of the most truly revolutionary aspects of LiDAR altimetry, particularly from a user perspective, is its ability to penetrate vegetation and obtain direct ground measurements, something which is extremely difficult to do using photogrammetry, even in leaf-off conditions (Lillesand et al., 2015; McGlone, 2007; Fowler et al., 2007); furthermore, since a single pulse may be returned off of several surfaces, the same dataset can be used to produce a variety of surface models. Of additional use in geomorphology—where issues about scale are actively investigated—is the ability of the same LiDAR data to produce DEMs of varying resolutions, depending upon user needs (Evans, 2012). While initial processing of raw LiDAR data into classified point clouds requires specialized skills and programs, an

average GIS user can create custom DEMs from these point clouds using commercial, off-the-shelf software such ESRI's ArcGIS or free, open-source platforms like QGIS.

The explosion in the availability of LiDAR elevation data has helped transform the world of geomorphometry (Gessler et al., 2009). LiDAR elevation data have quickly become important in many areas of geomorphic research and, where available, LiDAR is usually the most preferred elevation measurement technology (Bishop et al., 2012; Tarolli, 2014; Wohl, 2014). The high resolution of LiDAR DEMs has facilitated precise delineation of the hillslope-to-valley transition, extraction of headwater channels in difficult, mountainous terrain, and detected the subtle topographic signature of paleochannels in floodplains (Challis, 2006; Notebaert et al., 2009; Pirotti and Tarolli, 2010; Tarolli and Dalla Fontana, 2009). It has been used to quantitatively study hillslope mass movement and produce channel slopes comparable to field survey data (Stock and Dietrich, 2003; Vianello et al., 2009). Finally, repeat LiDAR surveys have been used to monitor a variety of processes, including bank erosion and the post failure evolution of landslides (Corsini et al., 2007; Thoma et al., 2005).

Land Surface Parameters

Two general groups of land surface parameters (LSPs) are recognized based on their scope: local and regional. Local parameters are those that can be calculated for any point using only elevation data for that point and—at most—for its immediate neighborhood. Elevation and slope are prime examples. Regional parameters, on the other hand, put each unit in a larger context that usually has to do with position relative to flows of mass and energy (Deng, 2007; Olaya, 2009; Wilson, 2012). Examples are the Topographic Wetness Index (TWI), which takes upslope catchment area into account and

is used for locating channel initiation and likely habitat types, and potential incoming solar radiation. While regional LSPs have been used for landform classification (e.g. De Reu et al., 2013; Irvin et al., 1997) and there is growing interest in their use for this purpose (Olaya, 2009; Wilson, 2012), a review of recent literature shows that their use is not common practice. One of the premises of this research is that regional LSPs can potentially provide spatial context that improves automated terrace recognition.

Most local parameters are geometric, with elevation and its first two derivatives being the most important in much geomorphic analysis. Elevation, typically denoted by z, is the basis of all geometric LSPs and is the only LSP that comes close to being scale invariant (Deng, 2007; Olaya, 2009). In theory, elevation directly corresponds to the elevation of one particular location. This may be a point (vector) value, as with LiDAR, or a block value averaging the elevation over an area the size of one grid cell (stereo photogrammetry or radar). Under ideal circumstances, error in elevation will be equal to the error in the underlying measurement technology, which can be anywhere from more than a dozen meters, as with space-based IfSAR, to less than 5 cm, as with post-processed ground DGPS data.

Derivatives of elevation become more variable based upon scale, the axis being measured, and the formula used, with the propagation of underlying measurement and calculation errors increasing dramatically for higher order derivatives. The first partial derivative of elevation taken in the direction of maximum change—the gradient vector—defines the slope of a cell. The direction of this vector also gives the aspect of the cell, which is important in flow routing and calculations of potential insolation. The second derivative—the rate of change in slope—is called the curvature and it can be taken along

a number of different axes, depending upon what is being studied. By convention, positive values denote convex curvatures and negative values concave ones, with the zero-point corresponding to inflection points in calculus. Profile curvature is taken parallel to gradient and indicates acceleration or deceleration of flow. Plan curvature, on the other hand, is taken perpendicular to gradient and indicates divergence or convergence of flow. In addition to these two, there are at least fifteen more types of curvature in currently published literature, examples including unsphericity curvature (how much a surface differs from a perfect sphere) and rotor curvature (clockwise or anticlockwise turning of flow lines) (Olaya, 2009; Schmidt et al., 2003). Third-order derivatives of elevation have more recently received some use. While their meaning for landform classification and analysis is not generally clear and they are highly sensitive to any noise in the input data, they have been suggested to be useful in identifying features such as ridges and thalwegs; however, their usefulness is not widely agreed upon and they use is still experimental, as is demonstrated by the fact that no commercial or non-commercial software packages have their computation as a built-in feature (Minár et al., 2013; Olaya, 2009).

Regional LSPs attempt to describe a unit area in the broader context of hydrology and morphology beyond immediately adjacent terrain (Wilson, 2012). While they are often created with applications in hydrology, ecology, or soil mapping in mind, they are frequently employed as ancillary information to aid in landform classification. This research examined three regional LSPs for their potential to increase classification accuracy: two hydrological and one topographic. Vertical Distance to Channel Network (VDTCN) normalizes each cell according to its distance to a flow network that I defined

as any stream large enough to form a floodplain (and thus, potentially, terraces). SAGA Wetness Index (SWI), a modified version of the Topographic Wetness Index (TWI), combines local slope and upslope contributing area to predict potential soil moisture and stream initiation. Finally, Topographic Openness (TO) indicates degree of prominence or enclosure and thus may help distinguish planar areas in hollows from others.

Landforms and Landform Elements

Broadly defined, landforms are naturally occurring features that can be used to partition the Earth's surface and have typical, distinguishing attributes wherever they occur (MacMillan and Shary, 2009). Landform classification presents many challenges which are distinct from other types of feature recognition and which require innovative and synthetic methods to solve. Two fundamental concepts in remote sensing are those of information classes—which are features of interest (e.g. cornfields, water, terraces)—and spectral classes, which are clusters of spectral signatures that can be recognized and defined from the image itself. Generally speaking, unsupervised (data-driven) methods of classification involve matching information classes to algorithmically derived spectral classes, while supervised methods use examples of information classes to train an automated classifier (Lillesand, Kiefer, & Chipman, 2015). For DGM, morphometric signatures can be substituted for spectral ones. Elevation and its higher order derivatives—slope and curvature—can be analyzed in a comparable way to spectral reflectance or radiance values; however, that does not resolve all issues.

The complication in DGM comes from the fact that not all landforms have explicit, quantitative definitions of the type that could easily be adapted to algorithmic recognition. This is all the truer for specific geomorphic features such as terraces. For a

feature to be recognized in a reliable and reproducible way it must be operationally defined; furthermore, if this recognition is to be automated to any degree, this definition must be mathematical. At the same time, mathematical definitions need to retain the essential relationship to process and ontology, which is the point of DGM in the first place. In recent years, many researchers have proposed different means of mathematically defining landforms, employing variously combinations of the primary morphometric variables with different spatial relationships (Bishop et al., 2012; Evans, 2012; Minár and Evans, 2008; Wheaton et al., 2015).

Classification

Any system of landform classification must account for the fact that landforms are inherently hierarchical and scale-dependent (Bishop et al., 2012; Evans, 2012; MacMillan and Shary, 2009). A nose slope may be a component of a spur which is itself a component of a mountain. A plateau is a plain to a neighboring mountain; at the right scale, a rill is a hollow. At one level of landform classification, an entire floodplain can be classified as a single unit. At the level below this, the floodplain can be classified into several discrete units. Depending on the size and geomorphic setting of a stream, channel-floodplain interaction may produce features such as point bars, natural levees, and meander scrolls with varying degrees of development (Buffington and Montgomery, 2013; Leopold et al., 1995). Similarly, the LSPs used to quantify landforms have no single, true value, but are specific to the scale of analysis and method of calculation (Evans, 2012, 2003; I. Florinsky, 1998; MacMillan and Shary, 2009; Shary et al., 2005). Slope calculated at 90 meter intervals is highly smoothed compared what would be calculated from a high resolution LiDAR DEM (Kienzle, 2004; Wu et al., 2008).

Curvature values can be calculated at a range of radii (Wilson, 2012). The implication from these factors is that landform classification is dependent on the hierarchical level and physical scale of a given analysis (Drăguţ and Eisank, 2012; MacMillan and Shary, 2009).

Landform classification systems reflects the principle division of geomorphometry into general geomorphometry, which analyses the landscape as a continuous surface, and specific geomorphometry, which focuses on discrete landforms, such as a terrace or an alluvial fan. Of these two, general geomorphometry is more rigidly quantifiable, dealing with terrain in as purely mathematical way as possible, specific geomorphometry necessarily involves more subjectivity and interpretation of process from form in extracting discrete landforms from a continuous surface. General geomorphic forms tend to be defined according to how they affect the flow of mass and energy, such as by flow divergence or convergence. Specific geomorphic forms are usually defined by the process that formed them. Due to the difficulty of inferring process from form, efforts to use DGM for specific geomorphic mapping have met with mixed results; therefore, most landform classification systems are based more on general than specific geomorphometry and more upon how a landform currently interacts with flows of energy and mass instead of the processes which formed it. Thus, systems like those of Speight (1990) or Pennock (1987) include various forms of slope, but not landforms such as terraces.

Manually implemented systems of landform classification turn upon semantic models of salient terrain elements and incorporate, explicitly or implicitly, expert knowledge about regional, physiographic, and geomorphic process context (Argialas and Miliaresis, 1997; MacMillan and Shary, 2009). As these systems depend, to a greater or lesser

extent, upon rules and knowledge that are difficult or impossible to program, automated methods of classification are necessarily based upon directly quantifiable values. These typically are relative slope position and local curvature, but sometimes extend to different modes of mass movement, such as the distinction between alluvial and colluvial toe slopes in the system Dikau (1989) The results are classes that are typically defined by some combination of Boolean and threshold values.

Automated tools for such classifications, such as LANDFORM , implement these concepts with parameters for threshold values and scales of analysis that users can set according to their expert judgement, local knowledge, and/or experimentation (Klingseisen et al., 2008). There have also been supervised classification models—such as in ILWIS GIS—that accept a mean and standard deviation from training data to learn a classifier model. This research seeks to extend some of the work which has been done to use machine learning in a more central way, leveraging expert knowledge to derive models directly from the training data without the need for user-chosen thresholds or normal statistical analysis. Included in this is extending the range of learners used to those seldom found in the literature, such as Bayesian classifiers and SVM.

Supervised Machine Learning

Machine Learning (ML) is the subset of Artificial Intelligence (AI) that deals with the creation of algorithms that themselves create other algorithms (Kelleher et al., 2015; Shalev-Shwartz and Ben-David, 2014). Whereas traditional AI programming, such as expert systems, would create an algorithm to turn data into a desired result, ML focuses on creating an algorithm that can be given the data and the desired results and use these to create its own algorithm. To build a program that can perform classification tasks like

a human, knowledge must be obtained from experts in the field in question and incorporated that into the program. The expert systems approach works by interviewing experts and applying a top-down method of programming that knowledge into a classifier; however, expert systems are often challenged by the difficulty of obtaining sufficient information and the exponentially increasing complexity of their code. Expert knowledge is time consuming to elicit and document; furthermore, it is frequently difficult to create clear rules from expert knowledge and experts themselves may be unable to articulate rules that they implicitly use in their work (Agarwal and Tanniru, 1990; Barfield, 1986; Forsythe and Buchanan, 1989; Kemp, 1993). By learning rules from data on the basis of a training set of classified examples, supervised machine learning provides an alternative to the expert systems approach that simultaneously incorporates expert knowledge and underlying data patterns in defining class boundaries (Campbell and Wynne, 2011; Evans et al., 2009; Kelleher et al., 2015; Shalev-Shwartz and Ben-David, 2014).

Principles

Supervised machine learning algorithms accept a training dataset of instances of something labeled with their class. These instances, frequently referred to as vectors, each have a list of descriptive features which can be virtually any type of value. Surface aspect, elevation above sea level, and spectral reflectance in a Landsat band can all be descriptive features used in an ML task (Tso and Mather, 2009). The algorithm then learns a prediction model by analyzing the descriptive features of the training dataset and employing a set of assumptions—called the inductive bias—that constrain and guide the selection of model. The set of potential models for a given learner is referred to as it's

hypothesis space (Kelleher et al., 2015; Shalev-Shwartz and Ben-David, 2014). Once learned, these models are then used to predict the class of unlabeled data. In remote sensing, it is more common to use the terms classifier and classify than model and predict.

As is explained by both Kelleher et al (2015) and Shalev-Schwartz and Ben-David (2014), selection of the right learner and learner parameters for a given task is key to successful application of ML. A mistake at this stage can lead to under or overfitting the training data. A useful model is one that can create rules from the underlying patterns in its training data that will generalize well and can properly classify unseen instances. If the learner is blind to important patterns in the training data, it is said to have underfit the data and its model will be too simplistic to properly classify future instances. On the other hand, if a learner creates rules that are too specific to its training data—essentially mistaking noise and random variation in the data for real connections between descriptive features and information—then it is said to have overfit and it will not be able to generalize to unseen data. The latter is an issue that ML users need to be especially cautious about: the sheer power of learners to extract patterns from data makes at least some degree of overfitting a common occurrence, and this in turn will make a model extremely sensitive any flaws in the selection and processing of training data.

This selection process, however, is complicated by the well-known principle, commonly referred to as the No Free Lunch Theorem, that, on average, there is no best learner (Wolpert, 1996; Wolpert and Macready, 1997). A simplified version of No Free Lunch gives that, for any two ML algorithms A and B, there are as many situations in which A outperforms B as there are wherein B outperforms A. Crucially, this holds true

even if one of these algorithms is simply random guessing, meaning that there as many situations in which an algorithm may be confused by the data and perform worse than chance as there are the other way around. This means that there can be no universal learner: averaging over all possible situations, then, no algorithm performs better random guessing. The Free Lunch, of which there is none, is learning knowledge from data without prior knowledge: this prior knowledge comes in the form of an inductive bias. The inductive bias of a learner constrains the types of models it can learn and, by extension, results in class boundaries with a "shape" characteristic to the learner. A complication arises, however, in that there is no way to know *a priori* which inductive bias will perform best for a given situation. While a theoretical understanding of how different learners work can suggest which might be best suited for a specific problem, it is common practice in ML to test several algorithms before making a final choice. In light of this, five learners representing four distinct approaches were chosen, each described below.

Machine Learning Techniques

Similarity—Mahalanobis Distance & Winner-Takes-All

Similarity-based learners include some of the oldest and simplest algorithms, such as Nearest Neighbor, and are still among the most popular supervised classification methods currently employed in remote sensing. Similarity learners all depend upon creating a mathematical measure of the similarity between given instances which can be used to relate input data to training instances. These measures usually take the form of a distance calculation; therefore, these learners are often called distance-based. In order to calculate distances, each instance needs to be given spatial coordinates. To do this, a

feature space is created wherein every attribute or descriptive feature of an instance is represented as its own dimension. An example of this process applied to multispectral imagery is shown in Figure 3. While it is difficult to visualize a spatial domain that has more than three dimensions, a feature space is n-dimensional, able to use any number of dimensions that we wish to define. While image classification seldom uses more than a few dimensions (e.g. four of five bands from a Landsat scene), machine learners in other fields can and frequently do have thousands of dimensions (Kelleher et al., 2015; Shalev-Shwartz and Ben-David, 2014).

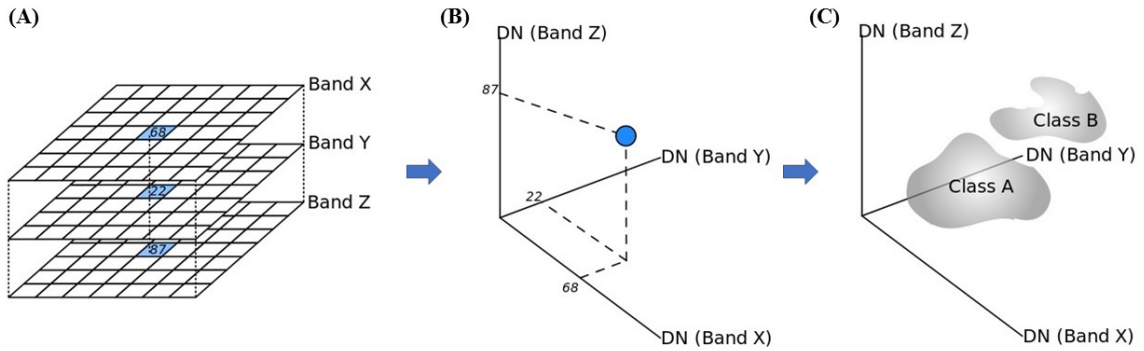

Figure 3: A generalized process for representing multispectral image data in a multidimensional feature space. For each pixel used as an instance, or vector, in the input data, the digital numbers from each spectral band (A) are used as coordinate values in a feature space (B). With all the training instances mapped in feature space, decision regions (C) are defined which then can be used to classify unseen instances. Adapted from Beck (2012).

It would be straightforward to assume that including more descriptive features in a classification is beneficial, so long as those descriptors provide additional information; however, past a given point, adding more descriptive features to a classification can have the effect of actually decreasing the classifier's accuracy, even if these extra descriptive features are meaningfully related to the classes (Bishop, 2006; Van Niel et al., 2005). This is a result of the nature of hyperspace and its effects on sample density. In order to

meaningfully relate an area of hyperspace to a class, there needs to be a reasonable density of features in that area; however, with each descriptive feature that is added to a feature space, a whole new dimension is added and thus more "space" within which the samples are spread out, decreasing the effective sample density. In a sufficiently high-dimension space, even a tightly clustered normal distribution becomes spread to its farthest tails with all instances looking equally similar—or dissimilar—to each other.

This counterintuitive trade-off between sample density and number of descriptive features, wherein adding more data about each training instance can actually decrease classification accuracy, is referred to as the curse of dimensionality and it can be a fundamental problem for many ML applications. Typical remote sensing applications for ML have both pluses and minuses when it comes to dealing with the curse of dimensionality. On the one hand, image classification usually involves only a few descriptive features, keeping the number of dimensions far lower than many ML applications in other fields. Additionally, classes in remote sensing tend to be fairly well clustered, even if they are not normally distributed. On the other hand, the complexity and cost of creating training data means that image classifiers have to work with very small input datasets (Campbell and Wynne, 2011; Mountrakis et al., 2011). Because sample density decreases exponentially with additional dimensions ($density = k^{(\frac{1}{m})}$), even a small number of added dimensions could considerably impact a classifier's accuracy (Kelleher et al., 2015). All of this means that it is paramount to carefully select training data to minimize noise and to choose and prepare descriptive feature data to achieve maximum information content with a minimum of dimensions.

There are several measures of similarity used with distance-based learners. The simplest, Nearest Neighbor, is simply the Euclidean Distance between an unclassified input point and a training sample point. This analysis used two techniques: Mahalanobis Distance (MD), an alternative to Euclidean Distance, and Winner-Takes-All (WTA), an ensemble approach. The Mahalanobis Distance is a straight-line distance measurement like Euclidean Distance; however, the coordinate space is created based upon the spread and co-variance of input data. For each measurement, an orthonormal coordinate system is created with its origin at the point of interest and its primary axis parallel to the angle of greatest variance. The units of both axes are then scaled to normalize the variance between them. The effect of this re-definition of distance is to compensate for datasets that exhibit a strong covariance that would otherwise confuse the real relationship between the variance of individual instances and the overall dataset variance if Euclidean Distance were to be used. In fact, in the absence of such covariance, a Mahalanobis Distance measurement is essentially equivalent to the Euclidean Distance (De Maesschalck et al., 2000; Kelleher et al., 2015). WTA uses all available methods of similarity measurement (in this case, Binary Encoding, Parallelepiped, Minimum Distance, Maximum Likelihood, and Spectral Angle Mapper) simultaneously, assigning the decision boundary in feature space based upon the maximum activation among all measures.

Probability—Bayesian Learners

In the 18th century, the statistician and Protestant minister Thomas Bayes proposed a theorem for updating beliefs in the light of new evidence. Formalized and

popularized in the 19th century by French scientist Pierre-Simon Laplace, it can be given mathematically as:

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

Or: The probability of Cause A, given that we observe Effect B, is equal to the Probability of Effect B, given Cause A, multiplied by the Prior Probability of Cause A, all divided by the Unconditional Probability of Effect B (Kelleher et al., 2015).

A common way of explaining this by example involves the probability of having a rare disease given a positive test with that test having a 99% accuracy. Using typical frequentist statistics, the probability of having the disease, given positive test, $\Pr(Disease|Positive)$ is 99%. However, in Bayesian statistics, the prior probability is factored in to the analysis. If this disease has an occurrence of 1:10,000 in the general populace, than this value $\Pr(Disease)$, is multiplied by the likelihood of observing a positive test, given the cause of having the disease $\Pr(Positive|Disease)$ all divided by the unconditional probability of a positive test, absent the disease $\Pr(Positive)$. This resulting value:

$$\Pr(Disease|Positive) = \frac{(0.99) * (0.0001)}{0.01} = 0.0099 = 0.99\%$$

is quite a bit smaller—and less panic inducing—than what was previously believed. That is because this new evidence was used to update the prior belief in the possibility of having the disease, which was very small. This updated probability is called the posterior probability.

There are three important advantages of using Bayesian (here Normal Bayes, NB) inference in classification of remote sensing or geomorphological data. First, Bayesian statistics do not depend upon an assumption of a normal distribution of data. Second, Bayesian methods are particularly suited to solving inverse problems, such as the inference of process from form in geomorphology. From their mathematical formulation alone, Bayesian statistics can make forward and inverse inference with equal ease (Kelleher et al., 2015). Finally, since these methods natively deal in probabilities, they are good at dealing with the degrees of uncertainty about class boundaries that make feature recognition in general, and landform classification specifically, problematic (Campbell and Wynne, 2011).

Error-based—Support Vector Machines

As first formally described by Boser, Guyon, & Vapnik (1992), Support Vector Machines (SVMs) are a group of supervised, non-parametric statistical learners. SVMs work by attempting to find a linear separating hyperplane defining a decision boundary between two or more classes in a high-dimensional space. As with other n-dimensional feature space classifiers, individual (labelled) data points are called vectors. The vectors closest to the separating hyperplane are called support vectors and they are the only ones which actually define the hyperplane. Support vectors are like swing state voters in a tight election: they are the only vectors whose vote "counts". Thus, if the support vectors are altered, deleted, or replaced, the hyperplane will be altered as well. A separating hyperplane is considered optimal when the distance from the support vectors to the plane, called the margin, is maximized, as shown in Figure 4a. For cases in which classes cannot

be separated by a straight line, SVMs can learn a smooth, curved boundary by remapping the raw data into a higher dimensional space, as in Figure 4b (Tso and Mather, 2009).

Since their introduction, SVMs have grown extremely popular due to their robust performance in many applications (Kelleher et al., 2015). While it has taken comparatively longer for SVMs to gain traction in the remote sensing community, they

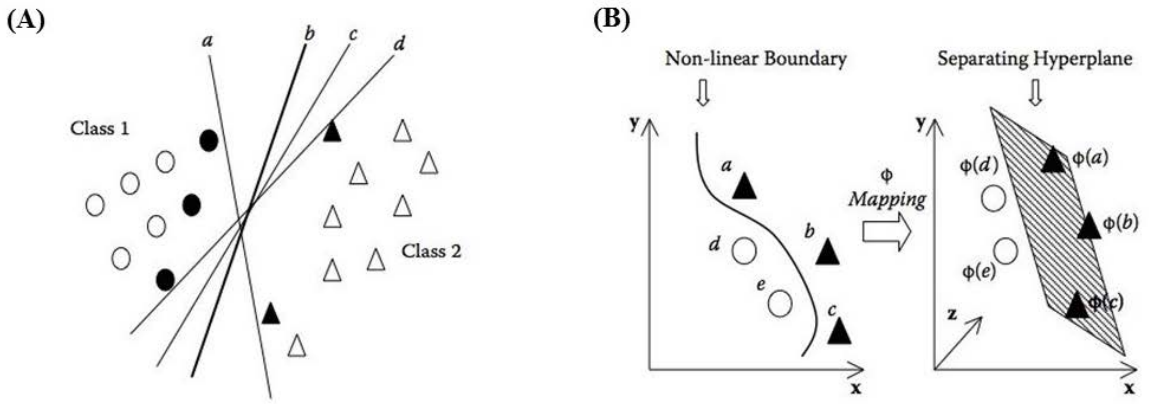

Figure 4: Margin maximization and the use of high dimensional space are important principles of SVMs. (A) shows many potential separating hyperplanes. While all of these hyperplanes would separate the training instances of the two classes in the same way, line *b* is optimal because it maximizes the distance between the decision boundary and the support vectors. As depicted in (B), SVMs can learn complex, non-linear class boundaries by mapping the input data to a higher dimensional space. Adapted from Tso and Mather (2009).

are rapidly gaining popularity due to their many strengths for dealing with feature classification issues (G. M. Foody and Mathur, 2004; Mountrakis et al., 2011). First, as a non-parametric method, SVMs make no assumptions about the distribution of their input data (Wilson, 2008). RS data are frequently not normally distributed; this does not matter with SVMs. Second, SVMs have been proven to generalize well even with a small amount of training data, data that is often costly and time consuming to produce in remote sensing (Giles M. Foody and Mathur, 2004; Mountrakis et al., 2011; Wilson, 2008). Finally, SVMs do not suffer from the curse of dimensionality, nearly as much as

most other classifiers. SVMs are routinely used with thousands of dimensions with little issue.

Ensemble—Random Forest

Ensemble methods in machine learning combine the results of a number of different learners to make a single classifier. These might be distinct types of learners or a large number of the same type of learner; Random Forest (RF) is an example of the latter. Random Forests are based upon Decision Trees, a class of information-based learners which work by using a series of queries to progressively split instances according to values of a descriptive feature until a final classification is made. Decision Trees have the advantages of being well-suited for handling the inclusion of irrelevant descriptive features—they simply don't get used in the model—and relatively little affected by the curse of dimensionality; however, due to the nature of their recursive splitting of instances, they are vulnerable to noise and random variance in their training data and thus prone to overfitting, resulting in relatively low accuracy (Kelleher et al., 2015; Svetnik et al., 2003).

Random Forest (RF) is one of a few methods developed as a way to deal with these issues (Breiman, 2001). Instead of inducing a single classifier, RF uses a large number of Decision Tree learners, each trained using a different random subset of the training data. The subsets are created by sampling with replacement, ensuring that each tree is grown from the same data distribution. The resulting trees vote for the most popular class during the classification stage. RF has been shown numerous times to be a very accurate classifier, comparing well with advanced methods such as SVMs while, at the same time, being considerably easier to use (Cracknell and Reading, 2014; Pal, 2005;

Rodriguez-Galiano et al., 2012). Despite the fact that RF grows a large number of trees, it is actually computationally more efficient than growing an equivalent number of individual Decision Trees and compares favorably with other learners (Rodriguez-Galiano et al., 2012). In situations with a very large number of descriptive features, RF has even been shown to operate more quickly than growing a single Decision Tree (Svetnik et al., 2003). Finally, RF has a number of other attractive qualities, including the ability to provide estimates of the importance of different input variables and produce an internal, unbiased estimate of generalization accuracy, reducing the need to cross-validate or test on withheld data (Cracknell and Reading, 2014; Rodriguez-Galiano et al., 2012).

Prior Work

Intense research into DGM in recent years has produced a number of different toolsets and methods. Some have been developed for primarily small, regional scale mapping (Drăguţ and Eisank, 2012; Jasiewicz et al., 2014) while other have concentrated on the mapping of general landforms, such as ridges and hollows, at a large scale (Jasiewicz and Stepinski, 2013; Sofia et al., 2014). Several researchers have worked on semi-automated methods for terrace mapping; however, three major issues have resulted in limited success.

Demoulin, et al. (2007) used a method based upon DEM segmentation and calculation of neighborhood statistics to map terraces in the Vesdre River valley of eastern Belgium. They elected to use a relatively broad operational definition which identified any area of relatively low slope bounded by higher slope segments as possible a terrace. The authors used a deterministic segmentation approach analyzing valley terrain. The main river was divided into 500m reaches on the basis of the overall length

of the river and prior knowledge of the size of terrace remnants. Then, valley segments were defined by assigning each pixel to a river segment according to the steepest path. For each of these segments, they created a smoothed bi-variate plot of percent slope versus altitude relative to the river altitude at the mouth of the reach. They considered any local minima with a slope ≤ 13% and a displacement of at least 0.7% relative to its neighboring maxima to represent terraces in the valley segment.

In terms of degree of automation, this method required a great deal of user input to optimize its parameters. Some parameters reflected information about the river and its history that was endogenous to their elevation dataset, including the mean meander length and terrace size, as well as the approximate valley altitude prior to terrace-forming incision. Other parameter values were chosen by empirical observation or trial and error. The authors of the paper identify this as a possible source of error or weakness in their method. Using the best performing combination of slope and displacement thresholds, the authors found that their method of DEM analysis correctly identified 78% of the known terraces in their validation set, compared to the 90% identified in field surveys.

Stout & Belmont (2014) developed a set of Python-based geoprocessing tools by for use in ESRI's ArcGIS software, dubbing it the TerEx toolbox. TerEx offers a semi-automated method for delineating terraces as vector features and calculating their basic attributes, such as surface area and elevation above a local stream channel segment. It represents a compromise between methods requiring an advanced understanding of information science and those more likely to be usable by non-specialists, an issue that is still being addressed in DGM research (Bishop et al., 2012). The tools are intended to work with a great deal of user input, accepting threshold values for a search area

(maximum valley width), user-defined scale of analysis (moving window), maximum allowable local relief within the scale of analysis, and minimum surface area for a valid feature. The tool also depends upon a user running multiple iterations and editing the outputs to deal with issues such as mis-classification of smooth upland surfaces as terraces and the erroneous classification of a single terrace as multiple terrace surfaces due to slight elevation variations in the DEM.

Testing their toolset on a 32 km section of the Le Sueur River (Minnesota), the authors found a tight correlation between manually mapped and TerEx derived terraces, with error normally distributed across spatial scales. They felt that the error from using their tool was similar to the variance between two different manual mappers; however, additional tests of the toolbox on different rivers were not as promising. Used along 370 km of the Minnesota River, TerEx was not able to delineate both floodplains and terraces in a single run, instead requiring a number of time consuming runs to produce an acceptable result. For the small, upland stream Gulden Gulch (Colorado), TerEx was unable to identify small terrace remnants and incorrectly linked terraces with upstream meanders, producing inaccurate height measurements. Stout & Belmont concluded that—in this instance—the considerable number of edits and iterations needed to produce a final map took more time than would have been used to manually map the terraces from the same data. Additional trials identified other factors that could increase the need for manual editing, including the presence scroll bars or paleo-channels within the flood plain, DEM noise or errors in densely-vegetated areas, and very-low stream gradients.

The TerEx toolset was also used in previous research into mapping Buffalo River terraces. The investigators found that it was necessary to aggregate the terraces delineated

in a large number of runs in order to create a usable map. A large amount of manual editing was required, both to aggregate together many true terraces that were fragmented by the tool and to delete large areas of flat upland mis-classified as terraces. Additionally, the parameters had to be carefully modified for geomorphically different river reaches (Bush and Shepherd, 2016).

Hopkins & Snyder (2016) tested the performance TerEx and two other terrace delineation methods, validating their results against visually mapped terraces in the Sheepscot River watershed (Maine). The authors rated each method according to a five-category rubric: the total time required for mapping, whether or not prior knowledge of the landscape was needed, the degree of manual editing required, the ability to produce continuous features across whole terrace surfaces, and the accuracy in total terrace area mapped. Besides the TerEx toolbox previously discussed, they employed a method introduced by Michael Rahnis (Walter et al., 2007) and the LandSerf feature classification tool. The Rahnis method maps the extent of terrace surfaces by interpolating a surface based on points manually placed by the user on a known terrace. The feature classification method, based on the LandSerf program, begins with a fully automated classification of the DEM into Channel, Ridge, Planar, and Pass surfaces. They then used elevation thresholds from an interpolated water surface to select appropriate Planar surface pixels as terraces. Hopkins & Snyder appreciated the rapid and objective results of the feature classification (LandSerf) method; however, they concluded that the TerEx toolbox was—despite its heavy dependence on user editing the best choice for accuracy across a variety of settings. By contrast, they found the Rahnis method was mostly suited for mapping single terraces in steep river valleys.

Finally, van Asselen and Seijmonsbergen (2006) used a top-down, GEOBIA method to map a sub-Alpine area in western Austria. The authors defined eight landform classes, including fluvial terraces, and created a training dataset based on a pre-existing, manually-created geomorphic map of the area. In their multistep method, the authors delineated training areas for each of the category landforms, then calculated their zonal statistics to produce parameters for fuzzy membership functions. They then employed these functions, along with a vector flow network, in a multi-level segmentation and classification within Trimble's eCognition software. They created flow paths for this network using the standard hydrology toolbox in ArcGIS, with the threshold contributing area values for channel-initiation being determined by experimenting with different values and comparing the results to existing drainage maps. The authors also derived the homogeneity criteria that were used for segmentation at the two different scale levels through trial-and-error experimentation. The classification of terraces was reasonably accurate: 69% of actual terraces were correctly identified as such, with most of the rest being mistaken for alluvial fans.

Taken together, these methods show several persistent limitations. First, as is shown in Demoulin, et al (2007) and research employing the TerEx toolbox (Bush and Shepherd, 2016; Hopkins and Snyder, 2016; Stout and Belmont, 2014), acceptable results are typically achieved only with considerable amounts of time consuming experimentation with parameters and manual editing of outputs, both of which limit the ability of these semi-automated methods to reduce the time needed for geomorphic mapping. Second, detailed prior knowledge of the specific landforms of the study area needs to be employed, such as the average width of known terraces (Demoulin et al.,

2007), the location and extent of terrace treads (Hopkins and Snyder, 2016), or delineation of landforms multiple landforms by manual interpretation (van Asselen and Seijmonsbergen, 2006). Finally, there is often considerable difficulty in transferring methods from one watershed or geomorphic setting to another, resulting in high variance in accuracy (Hopkins and Snyder, 2016) or necessity of many trial-and-error adjustments to achieve accuracy even within a single watershed (Bush and Shepherd, 2016; Demoulin et al., 2007; Stout and Belmont, 2014).

3—Methods

Source Data

LiDAR

The key source data for this is LiDAR elevation data acquired between 2009 and 2011 which was obtained from the USGS (*2010-2011 AR_BuffaloNP - East and West Portions*, 2014) and the Arkansas GIS Office (*2009 AR_BuffaloNP - Central Portion*, 2014) as classified LiDAR Point Clouds (LPC). The nominal point spacing was 1 m, with vertical accuracy (Root Mean Squared Error) between 0.041-0.073 m and horizontal accuracy between 0.65-1 m; however, these values are not necessarily valid over the whole study area. Steep terrain and heavy vegetation, both common in the Buffalo River basin, decrease the number of pulse returns, and thus the point spacing (Fowler et al., 2007; Renslow, 2012). In turn, complex forested areas such as along the Buffalo River present many near-ground objects that are difficult to separate from the actual ground. Pulses reflected off-of things such as small boulders, fallen trees, and short shrubs can all potentially be mis-classified as bare-earth pulse returns and thus introduce noise into the elevation data, decreasing its vertical accuracy (Fowler et al., 2007; Renslow, 2012). These factors must be considered when creating DEM products from LiDAR data.

Mapped Terraces

Vector terrace data based on the field mapping of the USGS (Hudson et al., 2006; Hudson and Murray, 2004, 2003, Hudson and Turner, 2014, 2009, 2007; Turner and

Hudson, 2010), AGS (Ausbrooks et al., 2012a, 2012b, Braden et al., 2003, 2003; Chandler et al., 2011; Hutto et al., 2008), and Keen-Zebert et al (2017) were used to aid in creating training samples and was also used to perform an accuracy assessment for the Terrace/Floodplain class independent of the error matrices used to assess the overall mapping accuracy of each method.

Other Supplemental Data

The high-resolution version of the National Hydrography Dataset (USGS, 2016), a part of the USGS National Map program, was used to provide reference data on the general flow network of the Buffalo River watershed and define major channels for use in calculating the VDTCN for the study area. The NHD is based on GIS vectorization of blue lines from the original 1:24,000 USGS topographic quadrangle maps, whose hydrograph was largely created through manual interpretation of stereo-photographs and contour lines (Simley and Carswell Jr., 2009).

Data Processing Steps

The GIS workflow for this research, illustrated in Figure 5, was mostly carried out in SAGA GIS, a free and open-source software, with initial data processing being done in ArcGIS. Construction of the LiDAR-based DEM which served as the source data for all LSPs was accomplished in ArcGIS 10.4 by way of ESRI's Terrain Dataset format. LiDAR points classified as Ground or Water were converted to a Multipoint feature class and loaded into a Terrain Dataset. This Terrain Dataset was then used to interpolate a 3 meter horizontal resolution DEM using Natural Neighbors, otherwise known as the Sibson Area-Stealing algorithm. This horizontal resolution was chosen both to allow landforms of interest to be shown in the data (e.g. small tributary channels) and provide

enough sample points for each cell to reduce random noise. The interpolation used a point-thinning method that applied a height tolerance of 0.3 meters. The practical effect

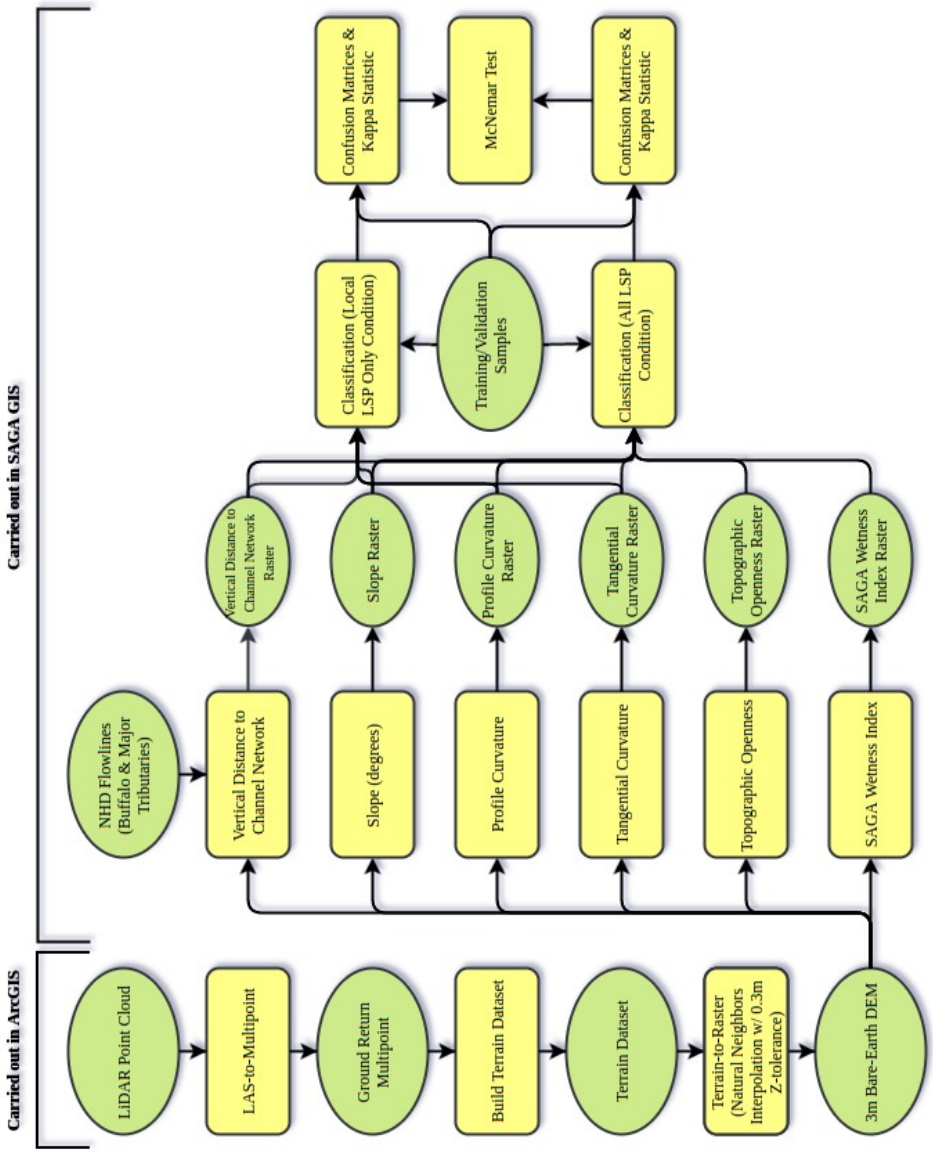

Figure 5: The GIS workflow was divided between ArcGIS, which was used to create a bare-earth DEM from the source LiDAR data, and SAGA GIS, which was used for classification and validation processes.

of this Z-tolerance based point thinning is to smooth out small-scale noise in the elevation data by filtering out points whose variance in elevation value are either the result of random noise, as given above, or micro-topography which are not relevant to the landforms used in this research.

Once the source elevation data was processed into a DEM in ArcGIS, all subsequent terrain processing and feature classification was carried out in SAGA GIS. LSPs were calculated independently for each study reach, with these LSPs being subsequently normalized to a scale of -1 – 1 in SAGA GIS before being used in classification. Normalizing different descriptive features to a common scale is an important step to prevent dimensions with high degrees of variability from overshadowing other dimensions (Lillesand et al., 2015).

Methodology for Creation of Training and Validation Data

Five landform classes were defined for use in DGM of the Buffalo River area. The number of classes chosen reflected a balance between the needs for adequately capturing the flow of the topography—particularly as it relates to gravitationally-driven flow of mass and energy—and keeping the classes general enough that landform elements would occupy relatively contiguous areas. The resulting classification scheme, detailed in Table 1 and illustrated in Figure 6, distinguishes planar uplands and lowlands (terraces and floodplains) as well as three major slope types based upon convergence or divergence of flow, but does not vertically partition hillslopes or identify individual peaks.

The Buffalo has previously been divided into four study reaches based upon the lithology of the immediate stream valley (Keen-Zebert et al., 2017); however, because this lithology is not necessarily shared by adjacent parts of the basin, this research used an alternative subdivision. The whole study area was divided into three reaches (Table 2 and Figure 7) based upon a combination of lithology, basin contributing area, and total

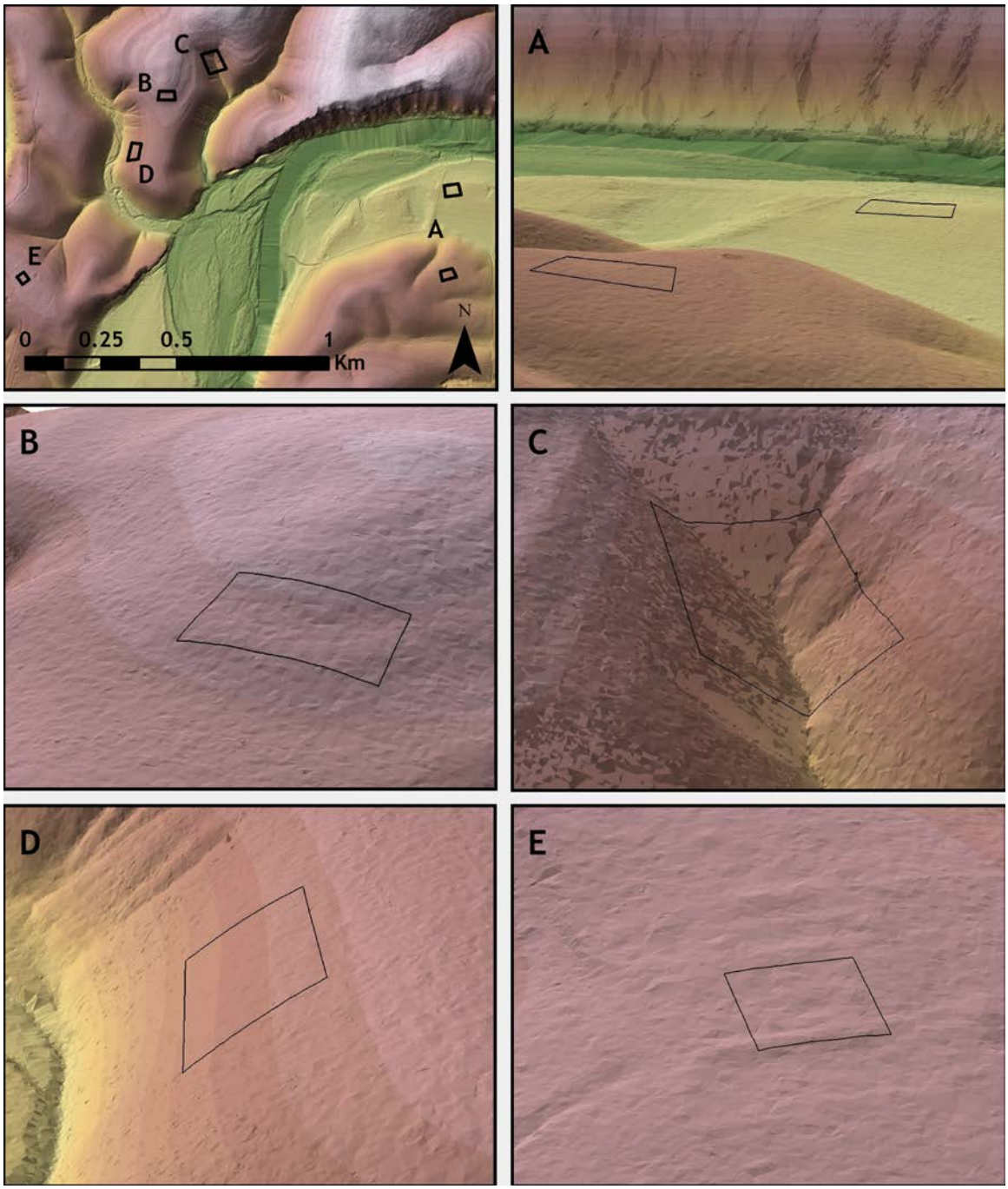

Figure 6: Representative training samples from each landform class. Terrace/Floodplain (A) samples were based upon field-mapped data, this example showing both a lower and a mid-level terrace. Divergent Hillslopes (B) are often below Planar Uplands (E) and above Convergent Hillslopes (C) and/or Planar Hillslopes (D).

relief, with the objective of making each reach as morphologically homogeneous as

possible. For each reach, visual interpretation of the terrain, as depicted by overlaying

Class Name	Description
Terrace/Floodplain	This class encompasses all planar/semi-planar fluvial landforms. These are mostly planar-to-gently sloping (usually < 4^{o}) with minimal curvature, mildly convex for both profile and longitudinal curvatures.
Divergent Hillslope (Nose Slope)	Moderate-to-steep hillslopes, convex profile, and longitudinal curvature. Areas of accelerating and divergent flow.
Convergent Hillslope (Hollow)	Moderate-to-steep hillslopes, concave profile with concave longitudinal curvature. Areas of converging flow, often some degree of channelization.
Planar Hillslope (Transportational Hillslope)	Moderate-to-steep hillslopes, may have planar-to-slightly convex profile and longitudinal curvature or moderate surface texture (close alteration between convex and concave areas) which balances out to be mostly planar in its influence on flow. Common for sides of smaller tributary valleys/hollows.
Planar Upland (Plateau)	Slopes mild-to-negligible, planar to slightly convex curvature. In the upper portions of the basin these include larger plateau segments while the middle and lower parts of the basin are predominantly narrower interfluves.

Table 1: The operational definition for each of the landform classes used in this study. These descriptions show the semantic model used to infer process from form and were created to reflect the way a human might manually map these landforms.

a partially-transparent tangential curvature raster on a multi-directional hillshade raster, was used to create a training and validation dataset.

In total, 80 training samples were created for each class in each reach, 50 being withheld for validation. The number of training samples needed for any given class is mostly a function of how many are needed to fully represent it in feature space; however, the general guideline is to have 10-30p samples, where p is the number of descriptive

features (Foody et al., 2006; Foody and Mathur, 2006; Van Niel et al., 2005). Strict adherence to this guideline has been questioned, though, with research demonstrating that accurate results can be achieved with samples as small as 2-4p (Foody et al., 2006; Foody and Mathur, 2006; Van Niel et al., 2005). Reflecting this, this research used 30 samples, or 5p, for each class, allowing for differences in the number of samples required by different classifiers (Foody and Mathur, 2006; Van Niel et al., 2005). The use of 50 validation samples for each class was chosen according to standard practice (Campbell and Wynne, 2011; Congalton and Green, 2009; Lillesand et al., 2015). There was one exception to the use of the 50 sample standard: due to the limited amount of Planar Upland areas in the Lower Buffalo reach, only 16 samples were used for validation; however, since each sample actually consisted of a cluster of cells, these 16 samples added up to 921 cells.

Reach	River km	Area km^2	Dominant Lithology along Buffalo River
Upper Buffalo	15-71	596	(17-30) Boone Formation—Limestone/Chert (30-112) Everton Formation—Sandstone/Dolostone
Middle Buffalo	71-150	196	Boone Formation—Limestone/Chert
Lower Buffalo	150-237	466	Everton Formation—Sandstone/Dolomite

Table 2: The three study reaches for this research are shown, along with the corresponding sections of the Buffalo River that they encompass. The lithology of the immediate Buffalo River valley is also given; however, this is not necessarily the dominant lithology of the study reach as a whole.

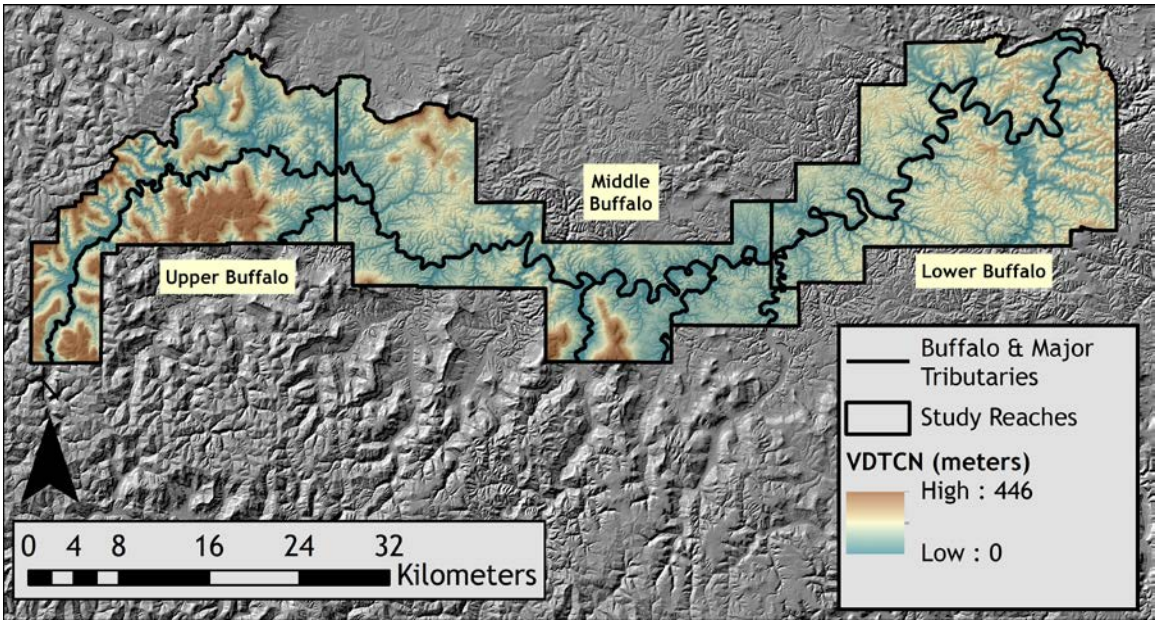

Figure 7: Three study reaches defined on the basis of general degree of relief and morphology. Vertical Distance to Channel Network (VDTCN) emphasizes relative differences in relief by normalizing elevation to the local channel level.

Input Land Surface Parameters

Three local LSPs were chosen to provide descriptive features relating each cell to local processes of mass and energy flow while an additional three regional LSPs were chosen to add descriptive features that indicate the location of each cell relative to larger scale hydrology and relief.

The algorithm used to calculate elevation derivatives was chosen to minimize the error and noise in the input data discussed previously. All derivatives were calculated using a 6 Parameter 2nd Order Polynomial function (Evans et al., 1979), which, unlike the 9 Parameter Partial Quartic function (Zevenbergen and Thorne, 1987) used in ArcGIS, is not constrained to pass through input elevation points. This results in a smoothing effect that makes the Evans method far less sensitive to noise than the

Zevenbergen method (Schmidt et al., 2003). While the addition of terms in the Zevenbergen can produce more complex curved surfaces (Schmidt et al., 2003), the real-world benefit of this is unclear; furthermore, since DEMs have inherent vertical and horizontal errors from measurement error, interpolation, and gridding, methods such as Zevenbergen may end up constraining equations to pass through elevation points which are themselves inaccurate (Evans, 1998). Several studies have found the Evans and similar methods to have greater precision and lower error than those using additional orders (Hengl and Evans, 2009; I. Florinsky, 1998; Schmidt et al., 2003; Skidmore, 1989). In addition to slope, which is the geomorphometric equivalent to velocity, profile and tangential curvatures were calculated. Profile curvature is the standard for examining flow acceleration; tangential curvature was used instead of the more commonly available planform curvature for examining flow convergence/divergence, as the latter is not recommended for use in areas with steep terrain (Olaya, 2009; Wilson, 2012).

Regional LSPs used were VDTCN, TO, and the SWI. VDTCN was used to transform the DEMs from Geoid height to height above the nearest downstream member of the flow network. This flow network was created by rasterizing the high-resolution NHD blue lines for the Buffalo and any tributaries large enough for floodplain development. To the extent that preserved terraces reflect past river elevation, this pattern may be reflected in VDTCN. Use of this parameter also has the potential to provide more accurate measurements of terrace heights above current river level than methods such as those of Demoulin et al (2007), which assigned the elevation of all cells to that of the lowest elevation in each river segment, or of Stout and Belmont (2014), who noted the errors produced by the fact that their TerEx toolbar assigned elevations based upon

Euclidean distance instead of flowline distance. TO gives information about prominence of higher locations and valley width for lower locations, potentially helping to distinguish hollows from larger stream valleys (Anders et al., 2009; Yokoyama et al., 2002). Finally, the SWI, which is a modification of the Topographic Wetness Index intended to better represent planar areas (Bohner et al., 2002), adds information about local slope, relative slope position, and potential position in regional hydrology (Bohner et al., 2002; Böhner and Selige, 2006; Irvin et al., 1997).

Methodology for Systematic Learner Comparison

To assess the effects of adding regional LSPs, each classification was run twice. The first run was the local-only condition, using only slope, profile curvature, tangential curvature, and VDTCN. The second run was the all-LSP condition, adding TO and SWI.

Machine Learning Method	Software Implementation
Mahalanobis Distance	SAGA
Winner-Takes-All	SAGA
Random Forest	OpenCV module for SAGA
Normal Bayes	OpenCV module for SAGA
Support Vector Machine	LIBSVM for SAGA (Chang and Lin, 2011)

Table 3: The modular nature of SAGA allows for the inclusion of other software packages in addition to those designed specifically for SAGA. The module employed for each learner is specified here to give the exact version of the learning algorithm being used.

It should be noted that, while VDTCN is actually a regional LSP, it was necessary to include it in all classifier runs in order to successfully distinguish between the Terrace/Floodplain and Planar Upland classes. This resulted in a total of 30 classifications, 2 for each learner (Table 3) on each reach. Once all the classifications

were completed, they were compared both on a per-reach basis and aggregated over the whole study area.

After classification, three statistical measures were used for quantitative analysis. Classification accuracy was assessed using confusion matrices and kappa statistics, standard methods in remote sensing described in (Campbell and Wynne, 2011; Congalton and Green, 2009; Lillesand et al., 2015). A Confusion Matrix gives the overall accuracy of the classification, defined as the portion of correctly classified cells out of all classified cells compared to validation data, as well as a per class evaluation of User's and Producer's Accuracy. User's Accuracy shows the likelihood that a classification for a cell is correct by subtracting Type I errors for a given class from the total number of cells of that class which were compared to the validation data. In turn, Producer's Accuracy gives the portion of cells whose class is known from the validation data that were correctly classified by subtracting Type II errors from the total number of validation cells for a given class. The kappa statistic is another measure of classification accuracy that takes into account chance agreement between the classified and validation data. Confusion matrices and kappa were calculated for each learner and each reach in both the local-only condition and all-LSP condition. Overall accuracy for each leaner across the whole study area was also computed by joining all three of the classifications and comparing them to the whole validation set. Finally, the predicted Terrace/Floodplain areas from each learner were compared with the full dataset of field -mapped terrace and floodplain areas to calculate user and producer accuracies. In line with convention in the remote sensing literature, this research defined acceptable classification accuracy as overall classification accuracy of 0.85 or greater and/or a kappa of 0.8 or greater (Campbell and Wynne, 2011;

Congalton and Green, 2009; Evans et al., 2009). Finally, to determine if there was a significant difference in the performance of learners with different inductive biases, the McNemar Test was used for the higher-performing, all Land Surface Parameter condition. The McNemar Test is a pairwise comparison, similar to a Chi Square Test, which uses a cross-tabulation of correct and incorrect classifications on the same map area to evaluate whether the difference between two classifications are statistically significant (Campbell and Wynne, 2011). It is commonly used in comparing classifications produced by different analysts or using different methods (Campbell and Wynne, 2011; Congalton and Green, 2009; Foody, 2009, 2004).

4—Results and Discussion

Quantitative Evaluation

Comparing the overall accuracy of each learner on each reach in both the local-only and all-LSP conditions (Table 4) yields two major observations. First, the addition of regional LSPs was crucial in achieving an accurate classification, defined previously as an overall accuracy of 0.85 or greater or a kappa of 0.80 or greater. Without using regional LSPs as descriptive features, only a single classification, NB in the Upper Buffalo reach, achieved acceptable accuracy, in comparison to 12 of the 15 classifications which were acceptably accurate when using all LSPs. This result indicates both that the distribution of the chosen regional LSPs values was consistent with the semantic model of major landforms used in selection of training and validation data and was effective in increasing the distinctiveness of the target classes.

Second, the more advanced learners performed better and had greater consistency than the distance-based learners. Although MD produced the single highest classification accuracy, 92% for the Lower Buffalo, closer examination of the results suggests this may be due to change: the exception that proves the rule. All of the advanced learners were accurate in the all-LSP condition, while WTA failed to achieve acceptable accuracy in the Middle Buffalo reach and MD was only accurate for the Lower Buffalo. The spread in accuracy scores between reaches for the advanced learners was between 1 and 3%, compared to the 6% spread for WTA and 12% spread for MD. Finally, examining the

overall accuracy across all reaches (Table 5) shows that the both the difference in accuracy among the distance-based techniques and between them and the more advanced learners were greater than the difference among the advanced learners themselves.

Figures 8 and 9 illustrate the increased accuracy of the all LSP condition. Examining the

Method	*Upper Buffalo*			
	Local LSP		All LSP	
	Kappa	Accuracy	Kappa	Accuracy
MD	0.68	0.75	0.75	0.80
WTA	0.76	0.82	**0.85**	**0.89**
RF	0.75	0.81	**0.86**	**0.89**
NB	**0.81**	**0.85**	**0.85**	**0.89**
SVM	0.67	0.75	**0.88**	**0.90**
Method	*Middle Buffalo*			
	Local LSP		All LSP	
	Kappa	Accuracy	Kappa	Accuracy
MD	0.65	0.73	0.78	0.84
WTA	0.69	0.77	0.78	0.83
RF	0.71	0.78	**0.86**	**0.89**
NB	0.73	0.80	**0.85**	**0.89**
SVM	0.70	0.77	**0.84**	**0.88**
Method	*Lower Buffalo*			
	Local LSP		All LSP	
	Kappa	Accuracy	Kappa	Accuracy
MD	0.58	0.67	**0.90**	**0.92**
WTA	0.56	0.66	**0.85**	**0.89**
RF	0.69	0.76	**0.87**	**0.90**
NB	0.69	0.76	**0.89**	**0.91**
SVM	0.59	0.68	**0.88**	**0.91**

Table 4: Overall accuracy and kappa for each classifier in each reach for both the local LSP and all-LSP conditions. Overall accuracy gives the percent of correct classifications across all classes, while kappa is an additional method for evaluating overall accuracy that accounts for chance correspondence between the classified map and reference data. Values in bold meet the conventional thresholds (kappa ≥ 0.80 or overall accuracy ≥ 0.85) used to define an acceptable map accuracy.

differences in classification accuracy between the individual classes (Appendix), the greatest confusion is consistently between the different hillslope types, particularly between Planar and Convergent Hillslopes. For 19 of the 30 classifications, Planar

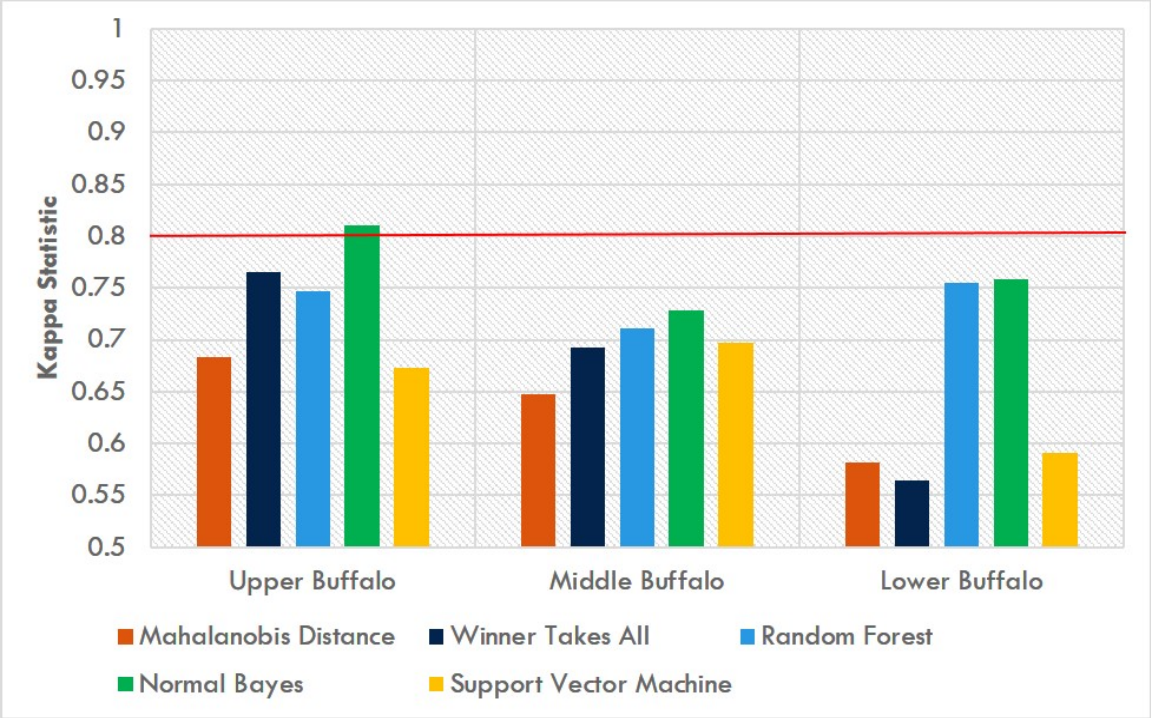

Figure 8: Kappa statistic for classifiers using local LSPs only. Red line shows the 0.8 threshold for an acceptably accurate classification. In the local-only condition, only the NB classification of the Upper Buffalo reach was accurate.

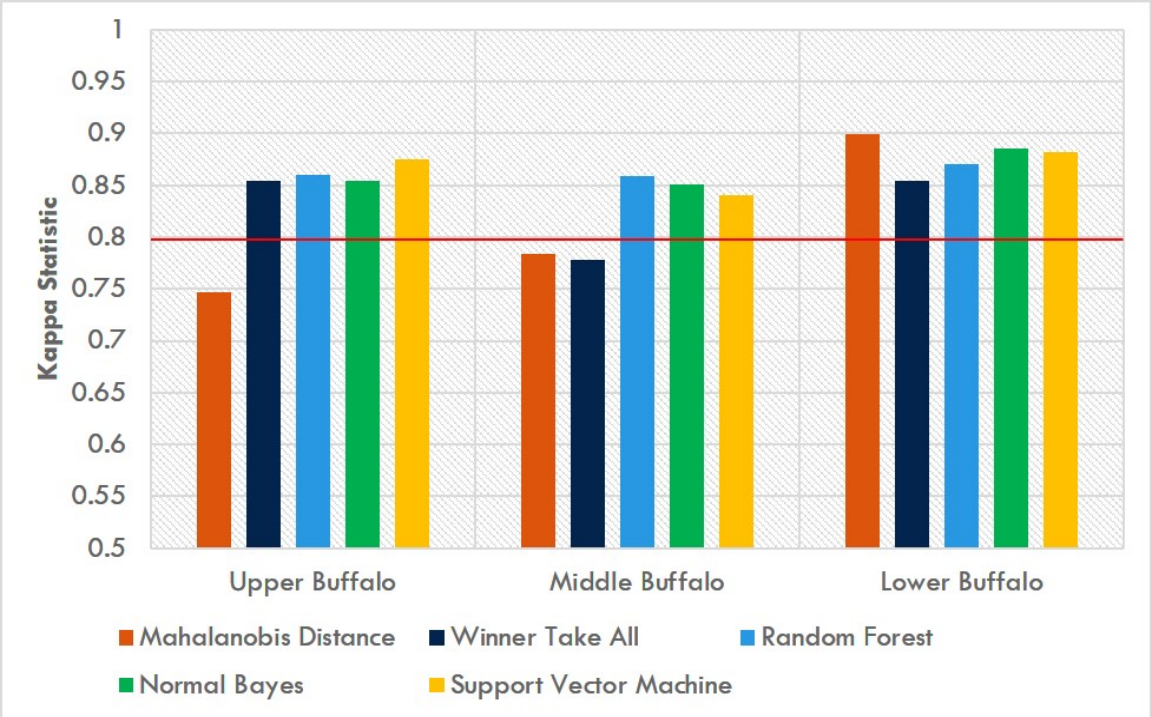

Figure 9: Kappa statistic for classifiers using all LSPs. Red line shows the 0.8 threshold for an acceptably accurate classification. Using all LSPs, 12 of the 15 classifications were accurate.

Hillslope had the lowest producer's accuracy and Convergent Hillslope the lowest user's accuracy or vice versa. Divergent and Planar Hillslopes had a similar relationship in 7 classifications, Divergent and Convergent Hillslopes in 2, with Planar Hillslope and Planar Upland and Terrace/Floodplain and Planar Upland in 1 each. The addition of regional LSPs made the largest difference with the hillslope classes, with the majority achieving acceptable accuracy in the all-LSP condition. In 13 of the 15 comparisons, Convergent and/or Planar Hillslope had the greatest increase in producer's accuracy, with Divergent Hillslope having the greatest gain in the others. Conversely, classification accuracy for the Terrace/Floodplain and Planar Upland classes changed very little with the additional LSPs and actually decreased in some instances, which might reflect floodplain or upland landforms that were not included in the classification system.

Method	Overall Accuracy (All Reaches)
Random Forest	89.3%
Support Vector Machine	89.1%
Normal Bayes	89.0%
Winner-Take-All	86.1%
Mahalanobis Distance	84.3%

Table 5: The overall accuracy of each learning technique for the whole study area. This was calculated by combining the all-LSP classifications for each of the three study reaches into one classified map and comparing this to the whole validation dataset.

Examining Table 6, which gives the statistical significance of differences between each of the learners under the higher-accuracy, all-LSP condition, shows that the greatest difference is between the distance-based and more advanced learners. While all of the

advanced learners (NB, RF, and SVM) were significantly different from (and superior to) the distance-based methods at the level of $P < 0.001$ (two-tailed), there was no significant difference (two-tailed $P < 0.05$) between the distance-based learners, nor was there any

Method	SVM v MD	SVM v RF	SVM v NB	SVM v WTA	MD v RF	MD v NB	MD v WTA	RF v NB	RF v WTA	NB v WTA
Z-Score	2.35	1.03	1.94	2.17	-2.95	-5.34	-0.68	-0.91	2.93	4.56

Table 6: Z-scores from all McNemar tests with values in bold denoting significant ($P < 0.05$) differences between the compared classifications. Positive values indicate that the first named classifier was more accurate.

significant difference between the more advanced learners. Taken together with the previously noted differences in classification accuracy, these results indicate both that the more advanced learners are better at dealing with complex classifications and that further increases in classification accuracy are more dependent on the quality of input data than the choice of learner.

Comparing the full set of predicted Terrace/Floodplain surfaces to the field mapped Terrace/Floodplain units (Table 7), all learners over-predicted total Terrace/Floodplain extent, but there were still notable differences in their classifications.

Learner	Producer's Accuracy	User's Accuracy
Mahalanobis Distance	0.49	**0.77**
Winner-Takes-All	0.70	0.57
Random Forest	0.80	0.55
Normal Bayes	0.57	0.70
Support Vector Machine	**0.85**	0.56

Table 7: Producer's accuracy and user's accuracy of the Terrace/Floodplain class when compared to the full field-mapped validation dataset. The highest producer's and user's accuracies are in bold.

Although the learners with the highest producer's accuracy, SVM and RF, also had the lowest user's accuracy, examining the relative gains and losses in these two accuracies between the different learners indicates that their high producer's accuracy is not completely a function of generally overpredicting Terrace/Floodplain surfaces. SVM correctly predicted an additional 0.05 of Terrace/Floodplain over RF at the cost of just 0.01 in user's accuracy. The difference is even greater for WTA, which had a user's accuracy 0.02 higher than SVM but a producer's accuracy 0.15 lower. The more conservative predictions of MD and NB essentially traded 0.07 between user's and producer's accuracy, respectively. The differences in how the learners delineate Terrace/Floodplain surfaces is illustrated in Figure 10, showing both young and old terraces just downstream of the confluence between the Little Buffalo and Buffalo Rivers. MD and NB (A and D) mostly or entirely miss the older terrace on the west side, while also considerably underpredicting the floodplains closer to river level. RF and SVM (C and E) come very close to predicting the whole terrace surface on the west at the cost of overpredicting floodplains and spurious terraces in the southwest quadrant. WTA partially predicts the western terrace, but fails to predict the full floodplain surface

Qualitative Evaluation

A qualitative evaluation of each learner, made by visual examination of their final classifications, provides information about how these techniques capture the semantic model of landform classes that may not be fully represented by normal statistical measures. Illustrated in Figure 11 are some of the typical differences between the maps each learner produces. Here MD (A) vastly overpredicts Convergent Hillslope and classifies Divergent Hillslope as Planar Hillslope. WTA (B) is more accurate in the

higher elevation portions of the meander bend to the southeast, yet still misclassifies the lowest hillslope tier as almost entirely Convergent Hillslope. RF (C), NB (D), and SVM (E) all perform relatively similarly; however, RF best captures the Divergent Hillslope class. The general pattern scene here of MD vastly overpredicting some classes, WTA and NB representing most classes in a fragmented manner, and RF and SVM each coming the closest to showing both accurate and consistent classifications is common across the study area.

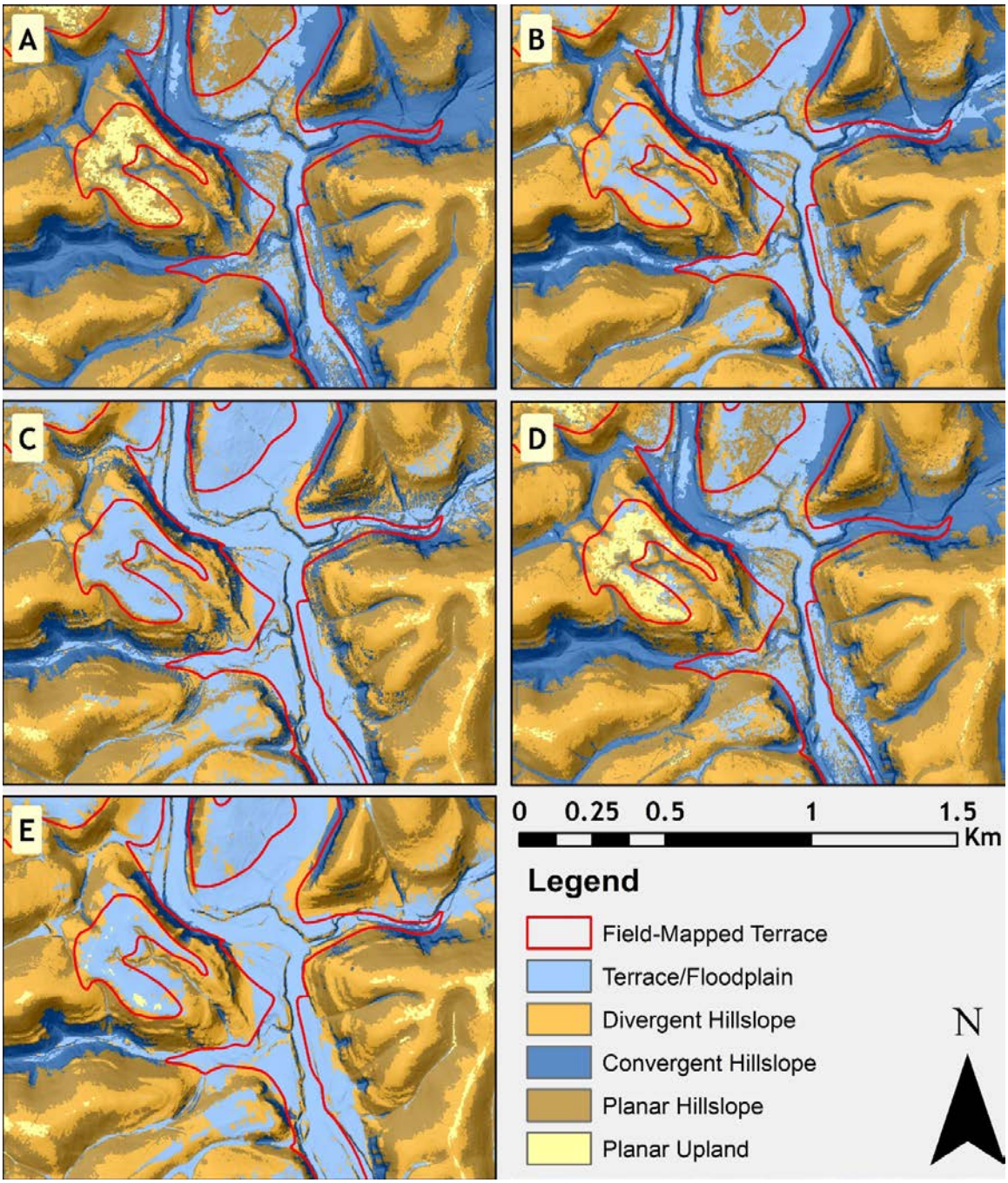

Figure 10: A comparison of the Terrace/Floodplain areas predicted by each machine learning method. Field-mapped terraces are outlined in red. The location shown is downstream of the confluence of the Little Buffalo and Buffalo Rivers. Classifications shown are: (A) Mahalanobis Distance (B) Winner-Takes-All, (C) Random Forest (D) Normal Bayes (E) Support Vector Machine

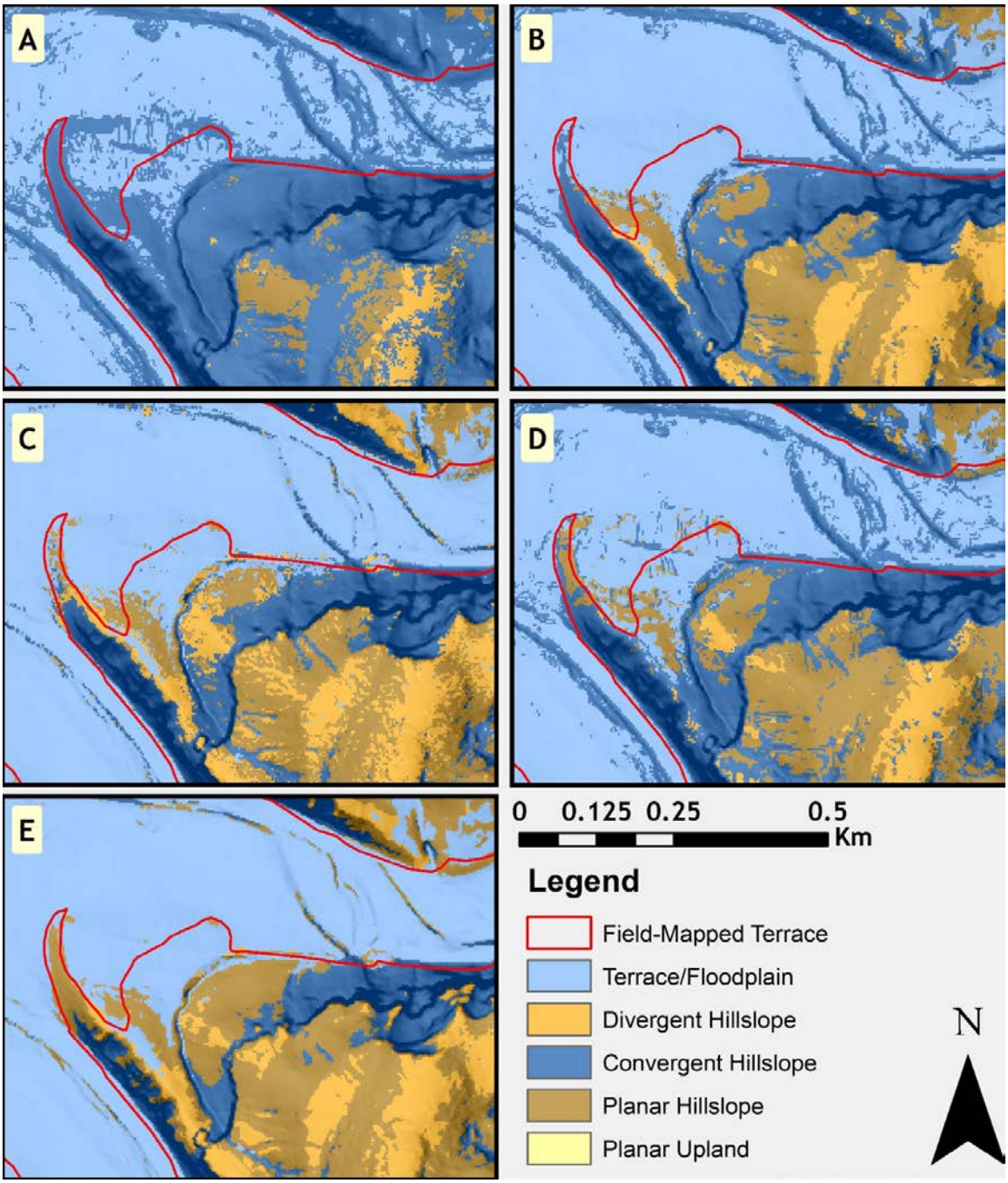

Figure 11: A qualitative comparison of how different machine learning techniques defined landform classes. Shown are: (A) Mahalanobis Distance (B) Winner-Takes-All (C) Random Forest (D) Normal Bayes (E) Support Vector Machine

Sources of Error

There are several potential sources of error in this research. First, since the study area was divided into three reaches to account for broad differences in terrain morphology, errors in the choice of these boundaries could increase the variance and overlap of landform classes. The terrace and floodplain polygons used for validation were interpolated from a limited number of GPS points using outdated Digital Raster Graphic (DRG) elevation data that has far lower vertical and horizontal resolution and far higher vertical error than the LiDAR dataset used to predict terraces. The USGS and AGS maps were based upon ~6m (20 ft) and ~12 m (40 ft), respectively, with vertical errors of up to 12 m, compared to vertical resolution of 0.01 m and error of 0.041-0.073 m (*2009 AR_BuffaloNP - Central Portion*, 2014, *2010-2011 AR_BuffaloNP - East and West Portions*, 2014; Keen-Zebert et al., 2017). Finally, the quality of the training sites selected is potentially one of the greatest sources of error, as is generally true of any supervised classification (Campbell and Wynne, 2011; Lillesand et al., 2015). Any inconsistencies or mistakes in the selection of training sites will introduce error in the model produced by the machine learning algorithm and decrease subsequent classifier accuracy (Bishop, 2006; Kelleher et al., 2015).

Conclusions and Recommendations

This research has explored how advanced supervised machine learning methods and land surface parameters derived from high-resolution elevation data can be used to classify general and specific landforms in a complex setting. Specific questions were (1) whether regional land surface parameters could improve classification accuracy by providing more information about the relative topographic and hydrologic position and

(2) whether machine learning algorithms with different inductive biases would produce results that are significantly different.

The results presented here show that the addition of regional land surface parameters resulted in a large and significant increase in classification accuracy over using local LSPs only; furthermore, not only did the accuracy improve for all classifiers across all reaches, the addition of regional LSPs was crucial in achieving acceptable classification accuracy, particularly for the three types of hillslopes, which otherwise were difficult to distinguish. Using a kappa statistic of 0.8 or higher as the definition of acceptable accuracy, only a single local LSP-only classification—NB used on the Upper Buffalo reach—exceeded the threshold for acceptability. In contrast, the addition of regional LSPs resulted in 12 of the 15 classifications achieving acceptable accuracy. Regional LSPs increased Kappa statistic scores by at least 0.064 and up to 0.317, with a mean increase of 0.162.

Three observations can be made about this increase in classification accuracy. First, it can be interpreted that the regional LSPs allowed the learners to induce better models of these landforms by adding more information about geomorphic processes. For example, while the training samples for the principle hillslope types (convergent, divergent, planar) were created by manual interpretation of their form, the semantic models for these hillslope types were based upon how they fit into the flows of mass and energy in the basin. Using only local (form-based) LSPs as descriptive features, the learners had limited success producing models which corresponded to the semantic models used to define the original classes. With the addition of regional LSPs, the models now had information about position relative to the principle channel network (VDTCN),

the accumulation of soil moisture (SWI), and the flow of terrain (TO). The subsequent increase in classification accuracy, then, may reflect the models better capturing the relationship between form and process that is at the center of geomorphology and geomorphometry. Second, this study shows that spatial context can be successfully used in supervised classification even when that classification is done on a per-pixel basis. Despite the fact that each pixel was classified independently of its neighbors and without reference to its actual spatial position, information about relative position still produced large gains in accuracy. Finally, the use of regional LSPs did not depend upon assumptions of normality, knowledge of the underlying distribution of the data, or trial-and-error experimentation with parameter or threshold values. The creation of the training samples themselves, which can be done by a user with little or no knowledge of how machine learning works, inherently sets the parameters for each class: With sufficient training data, the learner itself will build a model from the source data without the need for further human intervention (Kelleher et al., 2015; Shalev-Shwartz and Ben-David, 2014).

Looking to the validation of the Terrace/Floodplain class alone against the full dataset of field-mapped terraces and floodplains, these methods showed more limited success, as is reasonable to expect given the general difficulty of classifying specific landforms. The accuracies achieved by the learners tested in this research compare favorably to other methods employed. For instance, three of the five learners exceeded the 69% producer accuracy for Terrace/Floodplain achieved by (van Asselen and Seijmonsbergen, 2006). While the mean user accuracy was lower than that of the TerEx toolbox as used by Hopkins & Snyder (2016), this result disguises the large differences in

site-to-site accuracy seen in the TerEx results. Overall, the methods used in this study are not capable of producing an adequate finished terrace map by themselves, yet there is still potential to increase their accuracy. In particular, many of the areas misclassified as terraces were at the transition between the Everton and Boone Formations along the northern tributaries to the Buffalo which, unlike the southern tributaries, are mostly not large enough to have floodplain/terrace development. This suggests that Terrace/Floodplain classification accuracy may be increased either by the inclusion of more data about lithology or by expanding the number of classes to encompass landforms along these streams.

Comparing learners with distinct inductive biases, the results of this research show that there can be a significant difference in performance, indicating the value of considering multiple types of learners when approaching any given project. The more advanced, non-distance based learners outperformed the distance based methods by a considerable amount, ranging from Kappa statistic of 0.056 to 0.129 in the all-LSP condition, with the difference being significant at the level of $P < 0.001$. The sole exception to this was the result for MD in the Lower Buffalo reach, which had a kappa value that not only exceeded the other classifiers for that reach, but was, in fact, the highest for all classifiers across all reaches; however, given the uneven and generally mediocre performance of the MD technique, it is possible that this was mostly the result of chance.

On the other hand, the results of this research show that, in this instance, the addition of more data in the form of regional LSPs was more important than differences caused by the choice of inductive bias. First, absent the additional LSP data, the greater

performance of the more advanced methods was not enough to produce an accurate classification, with one exception noted above. Second, the mean increase in accuracy from adding regional LSPs, 0.162, was greater than the 0.128 difference between the best and worst performing classifiers. Finally, there were no statistically significant differences in performance between the more advanced classifiers, which had almost identical accuracies when summed for the whole study area. Taken together, these results indicate that, at least in this situation, classification accuracy is more limited by the types and amount of data available than by the choice of inductive bias. The nearly equal

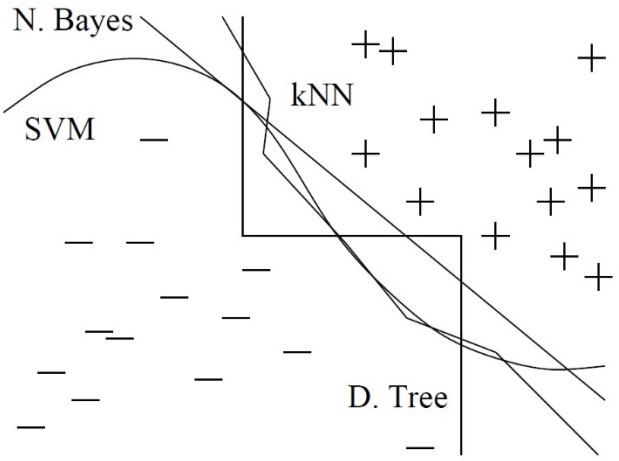

Figure 12: A schematic illustration of different decision boundaries created by different learners in a hypothetical separation of positive and negative examples of a given class. While each boundary has a shape characteristic to the inductive bias of the learner (e.g. the Decision Tree boundary is a stepped line with right-angles while the SVM boundary is a curved line), they all separate the positive and negative examples in the same way. From Domingos (2012).

performance between the top learners is not necessarily unusual in machine learning: as shown conceptually in Figure 12, learners may produce characteristically different decision boundaries, yet result in the same class separations. This also reinforces the general principle that “more data beats a cleverer algorithm” (Domingos, 2012). So, while the results of this research do support the idea that more advanced machine learning methods can produce superior classifications relative to some of the more

commonly used distance-based methods, it also reinforces the crucial role of input data, its quantity and quality, in a successful classification.

There are four major conclusions from this research that can potentially applied to future work. First, given the success of using regional LSPs in landform classification for this project, it is plausible that additional regional LSPs could further improve classification accuracy, especially for the challenging specific landform classes like terraces. Examples of LSPs that could improve classification of terraces in particular include Overland Flow Distance to Channel Network, which could be manipulated to calculate flow distance to a stream large enough to produce a floodplain, and position relative to the cut-bank and point-bar sides of a meander bend, which is useful in situations wherein most terraces are unpaired as is the case for the Buffalo River. Second, the use of image segmentation techniques, often called GeOBIA, can provide additional information about spatial context. Third, while no significant difference was found here between the higher performing classifiers, there are nonetheless ways to use these learners in a more sophisticated way. For instance, replacing NB with a Bayesian learner that accounts for conditional dependencies in the predictor variables has been shown to increase classification accuracy for geologic mapping (Porwal et al., 2006). On the other hand, the smoothness of the class boundaries created by SVMs is particularly attractive in how it honors the high degree of spatial-autocorrelation of landforms by producing contiguous landform areas without the need for further filtering or smoothing; therefore, optimization of SVMs for geomorphic mapping may be worth dealing with the admitted complexity of their use (Mountrakis et al., 2011). Finally, there is considerable room for the development of new software or software modules to capitalize upon the inherent

strengths of machine learning (Gessler et al., 2009). Utilities could help guide users through the process of creating training data, streamlining the collaboration of experts in geomorphology and machine learning. On the other side of the workflow, tools could be created to allow users to quickly refine the raw outputs of machine learners using such things as region growing algorithms and the ability to accept, reject, or modify landform boundaries. By continuing to explore more ways to describe the landscape through LSPs and capitalizing on the pattern recognition abilities of machine learners to harness expert knowledge, DGM may eventually make geomorphic mapping both faster and more accurate than manual methods.

References

2009 AR_BuffaloNP - Central Portion, 2014. . U.S. Geological Survey, Reston, VA.

2010-2011 AR_BuffaloNP - East and West Portions, 2014. . U.S. Geological Survey, Reston, VA.

Agarwal, R., Tanniru, M.R., 1990. Knowledge acquisition using structured interviewing: an empirical investigation. Journal of Management Information Systems 7, 123–140.

Anders, N.S., Seijmonsbergen, A.C., Bouten, W., others, 2009. Multi-scale and object-oriented image analysis of high-res LiDAR data for geomorphological mapping in Alpine mountains.

Argialas, D.P., Miliaresis, G.C., 1997. Landform spatial knowledge acquisition: identification, comceptualization and representation, in: Proceedings 1997 ASPRS/ASCM Annual Convention and Exposition, Technical Papers. American Society for Photogrammetry and Remote Sensing, American Congress on Surveying and Mapping Bethesda, MD, pp. 733–740.

Arrell, K.E., Fisher, P.F., Tate, N.J., Bastin, L., 2007. A fuzzy c-means classification of elevation derivatives to extract the morphometric classification of landforms in Snowdonia, Wales. Computers & Geosciences, Spatial AnalysisSpatial Analysis 33, 1366–1381. doi:10.1016/j.cageo.2007.05.005

Ausbrooks, S.M., Johnson, T.C., Nondorf, L.M., Traywick, C.L., 2012a. Geologic Map of the Cozahome Quadrangle, Marion County, Arkansas (Digital Geologic Map). Arkansas Geological Survey.

Ausbrooks, S.M., Johnson, T.C., Nondorf, L.M., Traywick, C.L., 2012b. Geologic Map of the Rea Quadrangle, Marion County, Arkansas (Digital Geologic Map). Arkansas Geological Survey.

Barfield, W., 1986. Expert-novice differences for software: Implications for problem-solving and knowledge acquisition. Behaviour & Information Technology 5, 15–29.

Beck, A., 2012. Multi Hyper-spectral Image feature space.

Benke, A.C., 1990. A Perspective on America's Vanishing Streams. Journal of the North American Benthological Society 9, 77. doi:10.2307/1467936

Benke, A.C., Cushing, C.E. (Eds.), 2005. Rivers of North America. Elsevier/Academic Press, Amsterdam ; Boston.

Bierman, P.R., Montgomery, D.R., 2014. Key concepts in geomorphology. W.H. Freeman and Company Publishers : A Macmillan Higher Education Company, New York, NY.

Bishop, C.M., 2006. Pattern recognition and machine learning, Information science and statistics. Springer, New York.

Bishop, M.P., James, L.A., Shroder, J.F., Walsh, S.J., 2012. Geospatial technologies and digital geomorphological mapping: Concepts, issues and research. Geomorphology, Geospatial Technologies and Geomorphological Mapping Proceedings of the 41st Annual Binghamton Geomorphology Symposium 137, 5–26. doi:10.1016/j.geomorph.2011.06.027

Bohner, J., Kothe, R., Conrad, O., Gross, J., Ringeler, A., Selige, T., 2002. Soil regionalisation by means of terrain analysis and process parameterisation (Research Report No. 7), Soil Classification 2001. European Soil Bureau, Luxembourg.

Böhner, J., Selige, T., 2006. Spatial prediction of soil attributes using terrain analysis and climate regionalisation. Gottinger Geographische Abhandlungen 115, 13–28.

Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers, in: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. ACM, pp. 144–152.

Braden, A., Ausbrooks, S., Mayfield, W., Clark, J., 2003. Geologic Map of the Mt. Judea Quadrangle, Newtown County, Arkansas (1:24000 scale). Arkansas Geological Survey.

Brandli, M., 1996. Hierarchical Models for the Definition and Extraction of Terrain. Geographic objects with indeterminate boundaries 2, 257.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. doi:10.1023/A:1010933404324

Buffington, J.M., Montgomery, D.R., 2013. Geomorphic classification of rivers.

Burrough, P.A., van Gaans, P.F.M., MacMillan, R.A., 2000. High-resolution landform classification using fuzzy k-means. Fuzzy Sets and Systems 113, 37–52. doi:10.1016/S0165-0114(99)00011-1

Bush, A., Shepherd, S., 2016. Automated mapping of fluvial terraces in a high relief watershed in ArcGIS using TerEx Toolbox: Buffalo River Watershed, AR.

Campbell, J.B., Wynne, R.H., 2011. Introduction to Remote Sensing, 5th ed. ed. Guilford Press, New York.

Challis, K., 2006. Airborne laser altimetry in alluviated landscapes. Archaeol. Prospect. 13, 103–127. doi:10.1002/arp.272

Chandler, A.K., Nondorf, L.M., Johnson, T.C., Traywick, C.L., 2011. Geologic Map of the Buffalo City Quadrangle, Baxter and Marion Counties, Arkansas (Digital Geologic Map). Arkansas Geological Survey.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2, 27.

Congalton, R.G., Green, K., 2009. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press, Boca Raton.

Corsini, A., Borgatti, L., Coren, F., Vellico, M., 2007. Use of multitemporal airborne lidar surveys to analyse post-failure behaviour of earth slides. Canadian Journal of Remote Sensing 33, 116–120. doi:10.5589/m07-015

Cracknell, M.J., Reading, A.M., 2014. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Computers & Geosciences 63, 22–33. doi:10.1016/j.cageo.2013.10.008

De Maesschalck, R., Jouan-Rimbaud, D., Massart, D.L., 2000. The Mahalanobis distance. Chemometrics and Intelligent Laboratory Systems 50, 1–18. doi:10.1016/S0169-7439(99)00047-7

De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., Crombé, P., 2013. Application of the topographic position index to heterogeneous landscapes. Geomorphology 186, 39–49. doi:10.1016/j.geomorph.2012.12.015

Dehn, M., Gärtner, H., Dikau, R., 2001. Principles of semantic modeling of landform structures. Computers & Geosciences, Geocomputation & Geosciences 27, 1005–1010. doi:10.1016/S0098-3004(00)00138-2

Demoulin, A., Bovy, B., Rixhon, G., Cornet, Y., 2007. An automated method to extract fluvial terraces from digital elevation models: The Vesdre valley, a case study in eastern Belgium. Geomorphology 91, 51–64. doi:10.1016/j.geomorph.2007.01.020

Deng, Y., 2007. New Trends in Digital Terrain Analysis: Landform Definition, Representation, and Classification. Progress in Physical Geography 31, 405–419. doi:10.1177/0309133307081291

Dikau, R., 1989. The Application of a Digital Relief Model to Landform Analysis in Geomorphology, in: Three Dimensional Applications in Geographical Information Systems. Taylor & Francis, pp. 51–77.

DiPietro, J.A., 2013. Landscape evolution in the United States: an introduction to the geography, geology, and natural history, 1st ed. ed. Elsevier, Amsterdam ; Boston, MA.

Domingos, P., 2012. A few useful things to know about machine learning. Communications of the ACM 55, 78–87.

Drăguţ, L., Eisank, C., 2012. Automated object-based classification of topography from SRTM data. Geomorphology 141–142, 21–33. doi:10.1016/j.geomorph.2011.12.001

Drăguţ, L., Eisank, C., 2011. Object representations at multiple scales from digital elevation models. Geomorphology 129, 183–189. doi:10.1016/j.geomorph.2011.03.003

Ehsani, A.H., Quiel, F., 2008. Geomorphometric feature analysis using morphometric parameterization and artificial neural networks. Geomorphology 99, 1–12. doi:10.1016/j.geomorph.2007.10.002

Evans, I.S., 2012. Geomorphometry and landform mapping: What is a landform? Geomorphology, Geospatial Technologies and Geomorphological Mapping Proceedings of the 41st Annual Binghamton Geomorphology Symposium 137, 94–106. doi:10.1016/j.geomorph.2010.09.029

Evans, I.S., 2003. Scale-specific landforms and aspects of the land surface. Concepts and Modelling in Geomorphology: International Perspectives. Terrapub, Tokyo 61–84.

Evans, I.S., 1998. What do terrain statistics really mean?, in: Landform Monitoring, Modelling, and Analysis, British Geomorphological Research Group Symposia Series. Wiley, Chichester ; New York, pp. 119–138.

Evans, I.S., Hengl, T., Gorsevski, P., 2009. Applications in Geomorphology, in: Geomorphometry: Concepts, Software, Applications, Developments in Soil Science. Elsevier, Amsterdam, Netherlands ; Oxford, UK ; Boston [Mass.], pp. 497–524.

Evans, I.S., Young, M., Gill, J.S., 1979. An integrated system of terrain analysis and slope mapping, final report. Univ. of Durham, Durham, NC.

Fenneman, N.M., 1928. Physiographic Divisions of the United States. Annals of the Association of American Geographers 18, 261. doi:10.2307/2560726

Foody, G.M., 2009. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. Remote Sensing of Environment 113, 1658–1663. doi:10.1016/j.rse.2009.03.014

Foody, G.M., 2004. Thematic Map Comparison. Photogrammetric Engineering & Remote Sensing 70, 627–633. doi:10.14358/PERS.70.5.627

Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sensing of Environment 103, 179–189. doi:10.1016/j.rse.2006.04.001

Foody, G. M., Mathur, A., 2004. A relative evaluation of multiclass image classification by support vector machines. IEEE Transactions on Geoscience and Remote Sensing 42, 1335–1343. doi:10.1109/TGRS.2004.827257

Foody, Giles M., Mathur, A., 2004. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. Remote Sensing of Environment 93, 107–117. doi:10.1016/j.rse.2004.06.017

Foody, G.M., Mathur, A., Sanchez-Hernandez, C., Boyd, D.S., 2006. Training set size requirements for the classification of a specific class. Remote Sensing of Environment 104, 1–14. doi:10.1016/j.rse.2006.03.004

Forsythe, D.E., Buchanan, B.G., 1989. Knowledge acquisition for Expert Systems: Some Pitfalls and Suggestions. IEEE Transactions on Systems, Man, and Cybernetics 19, 435–442. doi:10.1109/21.31050

Fowler, R.A., Samberg, A., Flood, M.J., Greaves, T.J., 2007. Topographic and Terrestrial Lidar, in: Digital Elevation Model Technologies and Applications: The DEM Users Manual. American Society for Photogrammetry and Remote Sensing, pp. 199–252.

Gessler, P., Pike, R., MacMillan, R.A., Hengl, T., Reuter, H.I., 2009. The Future of Geomorphometry, in: Geomorphometry: Concepts, Software, Applications, Developments in Soil Science. Elsevier, Amsterdam, Netherlands ; Oxford, UK ; Boston [Mass.], pp. 637–652.

Hancock, G.S., Anderson, R.S., 1998. Beyond Power: Bedrock River Incision Process and Form, in: Rivers Over Rock: Fluvial Processes in Bedrock Channels. American Geophysical Union, pp. 35–60.

Hengl, T., Evans, I.S., 2009. Mathematical and Digital Models of the Land Surface, in: Geomorphometry: Concepts, Software, and Applications, Developments in Soil

Science. Elsevier, Amsterdam, Netherlands ; Oxford, UK ; Boston [Mass.], pp. 31–63.

Hopkins, A.J., Snyder, N.P., 2016. Performance Evaluation of Three DEM-Based Fluvial Terrace Mapping Methods. Earth Surface Processes and Landforms. doi:10.1002/esp.3922

Hudson, M.R., Murray, K.E., 2004. Geologic map of the Hasty Quadrangle, Boone and Newton Counties, Arkansas (No. 2847, 1:24000 Scale), Scientific Investigations Map.

Hudson, M.R., Murray, K.E., 2003. Geologic map of the Ponca quadrangle, Newton, Boone, and Carroll Counties, Arkansas (No. 2412, 1:24000 scale), Miscellaneous Field Studies Map.

Hudson, M.R., Turner, K.J., 2014. Geologic map of the west-central Buffalo National River region, northern Arkansas (No. 3314, 1:24000 scale), Scientific Investigations Map. Reston, VA.

Hudson, M.R., Turner, K.J., 2009. Geologic map of the St. Joe quadrangle, Searcy and Marion Counties, Arkansas (No. 3074, 1:24000 scale), Scientific Investigations Map.

Hudson, M.R., Turner, K.J., 2007. Geologic Map of the Boxley Quadrangle, Newton and Madison Counties, Arkansas (No. 2991, 1:24000 scale), Scientific Investigations Map.

Hudson, M.R., Turner, K.J., Repetski, J.E., 2006. Geologic map of the Western Grove quadrangle, northwestern Arkansas (No. 2921, 1:24000 scale), Scientific Investigations Map.

Hutchinson, M.F., 2008. Adding the Z-dimension. Handbook of Geographic Information Science, Blackwell 144–168.

Hutto, R., Smart, E., Mayfield, W., 2008. Geologic Map of the Marshall Quadrangle, Searcy County, Arkansas (1:24000 scale). Arkansas Geological Survey.

I. Florinsky, 1998. Accuracy of local topographic variables derived from digital elevation models. International Journal of Geographical Information Science 47–61.

Irvin, B.J., Ventura, S.J., Slater, B.K., 1997. Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. Geoderma, Fuzzy Sets in Soil Science 77, 137–154. doi:10.1016/S0016-7061(97)00019-0

Iwahashi, J., Pike, R.J., 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. Geomorphology 86, 409–440. doi:10.1016/j.geomorph.2006.09.012

J. G. Speight, 1990. Landform, in: Australian Soil and Land Survey Field Handbook. Inkata Press, Melbourne, pp. 9–57.

Jasiewicz, J., Netzel, P., Stepinski, T.F., 2014. Landscape similarity, retrieval, and machine mapping of physiographic units. Geomorphology 221, 104–112. doi:10.1016/j.geomorph.2014.06.011

Jasiewicz, J., Stepinski, T.F., 2013. Geomorphons — a pattern recognition approach to classification and mapping of landforms. Geomorphology 182, 147–156. doi:10.1016/j.geomorph.2012.11.005

Keen-Zebert, A., Hudson, M.R., Shepherd, S.L., Thaler, E.A., 2017. The effect of lithology on valley width, terrace distribution, and bedload provenance in a tectonically stable catchment with flat-lying stratigraphy. Earth Surf. Process. Landforms. doi:10.1002/esp.4116

Keen-Zebert, A., Shepherd, S., Hudson, M.R., 2014. Quaternary Geology and Geomorphology of the Buffalo National River, Arkansas, 2014 GSA South-Central Section Meeting.

Kelleher, J.D., Mac Namee, B., D'Arcy, A., 2015. Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. The MIT Press, Cambridge, Massachusetts.

Kemp, E.A., 1993. Problems in expert systems development, in: Proceedings 1993 The First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems. Presented at the Proceedings 1993 The First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, pp. 166–167. doi:10.1109/ANNES.1993.323053

Kienzle, S., 2004. The Effect of DEM Raster Resolution on First Order, Second Order and Compound Terrain Derivatives. Transactions in GIS 8, 83–111. doi:10.1111/j.1467-9671.2004.00169.x

Klingseisen, B., Metternicht, G., Paulus, G., 2008. Geomorphometric landscape analysis using a semi-automated GIS-approach. Environmental Modelling & Software 23, 109–121. doi:10.1016/j.envsoft.2007.05.007

Köthe, R., Bock, M., 2009. Preprocessing of Digital Elevation Models–derived from Laser Scanning and Radar Interferometry–for Terrain Analysis in Geosciences. Proceedings of Geomorphometry, Zurich, Switzerland 31.

Leopold, L.B., Wolman, M.G., Miller, J.P., 1995. Fluvial processes in geomorphology. Dover Publications, New York.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2015. Remote Sensing and Image Interpretation, Seventh edition. ed. John Wiley & Sons, Inc, Hoboken, N.J.

MacMillan, R.A., Shary, P.A., 2009. Landforms and Landform Elements in Geomorphometry, in: Geomorphometry: Concepts, Software, Applications. pp. 227–254.

Matías, J.M., Ordóñez, C., Taboada, J., Rivas, T., 2009. Functional support vector machines and generalized linear models for glacier geomorphology analysis. International Journal of Computer Mathematics 86, 275–285. doi:10.1080/00207160801965305

Minár, J., Evans, I.S., 2008. Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping. Geomorphology 95, 236–259. doi:10.1016/j.geomorph.2007.06.003

Minár, J., Jenčo, M., Evans, I.S., Kadlec, M., Krcho, J., Pacina, J., Burian, L., Benová, A., 2013. Third-Order Geomorphometric Variables (derivatives): Definition, Computation and Utilization of Changes of Curvatures. International Journal of Geographical Information Science 27, 1381–1402. doi:10.1080/13658816.2013.792113

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66, 247–259. doi:10.1016/j.isprsjprs.2010.11.001

Notebaert, B., Verstraeten, G., Govers, G., Poesen, J., 2009. Qualitative and quantitative applications of LiDAR imagery in fluvial geomorphology. Earth Surf. Process. Landforms 34, 217–231. doi:10.1002/esp.1705

Olaya, V., 2009. Basic Land Surface Parameters, in: Geomorphometry: Concepts, Software, Applications. pp. 141–169.

Pal, M., 2005. Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26, 217–222. doi:10.1080/01431160412331269698

Passalacqua, P., Belmont, P., Staley, D.M., Simley, J.D., Arrowsmith, J.R., Bode, C.A., Crosby, C., DeLong, S.B., Glenn, N.F., Kelly, S.A., Lague, D., Sangireddy, H., Schaffrath, K., Tarboton, D.G., Wasklewicz, T., Wheaton, J.M., 2015. Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. Earth-Science Reviews 148, 174–193. doi:10.1016/j.earscirev.2015.05.012

Pazzaglia, F.J., Gardner, T.W., Merritts, D.J., 1998. Bedrock Fluvial Incision and Longitudinal Profile Development Over Geologic Time Scales Determined by Fluvial Terraces, in: Rivers Over Rock: Fluvial Processes in Bedrock Channels. American Geophysical Union, pp. 207–235.

Pennock, D.J., Zebarth, B.J., De Jong, E., 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. Geoderma 40, 297–315. doi:10.1016/0016-7061(87)90040-1

Pike, R.J., 1995. Geomorphometry-process, practice, and prospect. Zeitschrift fur Geomorphologie Supplementband 221–238.

Pike, R.J., Evans, I.S., Hengl, T., 2009. Geomorphometry: A Brief Guide, in: Geomorphometry: Concepts, Software, Applications, Developments in Soil Science. Elsevier, Amsterdam, Netherlands ; Oxford, UK ; Boston [Mass.], pp. 3–30.

Pirotti, F., Tarolli, P., 2010. Suitability of LiDAR point density and derived landform curvature maps for channel network extraction. Hydrol. Process. 24, 1187–1197. doi:10.1002/hyp.7582

Porwal, A., Carranza, E.J.M., Hale, M., 2006. Bayesian network classifiers for mineral potential mapping. Computers & Geosciences 32, 1–16. doi:10.1016/j.cageo.2005.03.018

Renslow, M.S. (Ed.), 2012. Manual of airborne topographic lidar. American Society for Photogrammetry and Remote Sensing, Bethesda, MD.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 67, 93–104. doi:10.1016/j.isprsjprs.2011.11.002

Schmidt, J., Evans, I.S., Brinkmann, J., 2003. Comparison of polynomial models for land surface curvature calculation. International Journal of Geographical Information Science 17, 797.

Shalev-Shwartz, S., Ben-David, S., 2014. Understanding machine learning: from theory to algorithms. Cambridge University Press, New York, NY, USA.

Shary, P.A., Sharaya, L.S., Mitusov, A.V., 2005. The problem of scale-specific and scale-free approaches in geomorphometry. Geografia Fisica e Dinamica Quaternaria 28, 81–101.

Simley, J.D., Carswell Jr., W.J., 2009. The National Map - Hydrography (No. 2009–3054), U.S. Geological Survey Fact Sheet. U.S. Geological Survey.

Skidmore, A.K., 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. International Journal of Geographical Information System 3, 323–334.

Sofia, G., Fontana, G.D., Tarolli, P., 2014. High-resolution topography and anthropogenic feature extraction: testing geomorphometric parameters in floodplains. Hydrol. Process. 28, 2046–2061. doi:10.1002/hyp.9727

Stark, C.P., 2006. A self-regulating model of bedrock river channel geometry. Geophys. Res. Lett. 33, L04402. doi:10.1029/2005GL023193

Stock, J., Dietrich, W.E., 2003. Valley incision by debris flows: Evidence of a topographic signature. Water Resour. Res. 39, 1089. doi:10.1029/2001WR001057

Stout, J.C., Belmont, P., 2014. TerEx Toolbox for semi-automated selection of fluvial terrace and floodplain features from lidar: AUTOMATED TERRACE SELECTION. Earth Surface Processes and Landforms 39, 569–580. doi:10.1002/esp.3464

Sugarbaker, L., Constance, E.W., Heidemann, H.K., Jason, A.L., Lucas, V., Saghy, D., Stoker, J.M., 2014. The 3D elevation program initiative: a call for action, Circular. U.S. Geological Survey, Reston, Virginia.

Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. J. Chem. Inf. Comput. Sci. 43, 1947–1958. doi:10.1021/ci034160g

Tarolli, P., 2014. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. Geomorphology 216, 295–312. doi:10.1016/j.geomorph.2014.03.008

Tarolli, P., Dalla Fontana, G., 2009. Hillslope-to-valley transition morphology: New opportunities from high resolution DTMs. Geomorphology, Understanding earth surface processes from remotely sensed digital terrain models 113, 47–56. doi:10.1016/j.geomorph.2009.02.006

Thoma, D.P., Gupta, S.C., Bauer, M.E., Kirchoff, C.E., 2005. Airborne laser scanning for riverbank erosion assessment. Remote Sensing of Environment 95, 493–501. doi:10.1016/j.rse.2005.01.012

Tinkler, K., Wohl, E., 1998. A Primer on Bedrock Channels, in: Rivers Over Rock: Fluvial Processes in Bedrock Channels, Geophysical Monograph. American Geophysical Union, Washington, D.C., pp. 1–18.

Tso, B., Mather, P.M., 2009. Classification Methods for Remotely Sensed Data, 2nd ed. ed. CRC Press, Boca Raton.

Turner, K.J., Hudson, M.R., 2010. Geologic map of the Maumee quadrangle, Searcy and Marion Counties, Arkansas (No. 3134, 1:24000 scale), Scientific Investigations Map.

USGS, 2016. USGS National Hydrography Dataset (NHD) Best Resolution for Arkansas 20160412 State or Territory FileGDB 10.1. U.S. Geological Survey, Reston Virginia.

van Asselen, S., Seijmonsbergen, A.C., 2006. Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser DTM. Geomorphology 78, 309–320. doi:10.1016/j.geomorph.2006.01.037

Van Niel, T.G., McVicar, T.R., Datt, B., 2005. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. Remote Sensing of Environment 98, 468–480. doi:10.1016/j.rse.2005.08.011

Vianello, A., Cavalli, M., Tarolli, P., 2009. LiDAR-derived slopes for headwater channel network analysis. CATENA 76, 97–106. doi:10.1016/j.catena.2008.09.012

Walter, R., Merritts, D., Rahnis, M., 2007. Estimating volume, nutrient content, and rates of stream bank erosion of legacy sediment in the Piedmont and Valley and Ridge physiographic provinces, southeastern and central PA. Report to the Pennsylvania Department of Environmental Protection, Lancaster, Pennsylvania.

Wheaton, J.M., Fryirs, K.A., Brierley, G., Bangen, S.G., Bouwes, N., O'Brien, G., 2015. Geomorphic mapping and taxonomy of fluvial landforms. Geomorphology 248, 273–295. doi:10.1016/j.geomorph.2015.07.010

Whipple, K.X., 2004. Bedrock Rivers and the Geomorphology of Active Orogens. Annual Review of Earth and Planetary Sciences 32, 151–185. doi:10.1146/annurev.earth.32.101802.120356

Wilson, J.P., 2012. Digital Terrain Modeling. Geomorphology, Geospatial Technologies and Geomorphological Mapping Proceedings of the 41st Annual Binghamton Geomorphology Symposium 137, 107–121. doi:10.1016/j.geomorph.2011.03.012

Wilson, M.D., 2008. Support Vector Machines, in: Jørgensen, S.E., Fath, B.D. (Eds.), Encyclopedia of Ecology. Academic Press, Oxford, pp. 3431–3437. doi:10.1016/B978-008045405-4.00168-3

Wohl, E., 2014. Time and the Rivers flowing: Fluvial Geomorphology since 1960. Geomorphology 216, 263–282. doi:10.1016/j.geomorph.2014.04.012

Wolpert, D.H., 1996. The Lack of A Priori Distinctions Between Learning Algorithms. Neural Computation 8, 1341–1390. doi:10.1162/neco.1996.8.7.1341

Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation 1, 67–82. doi:10.1109/4235.585893

Wood, J., 2009. Geomorphometry in LandSerf. Developments in soil science 33, 333–349.

Wood, J., 1996. The geomorphological characterisation of digital elevation models. University of Leicester.

Wu, S., Li, J., Huang, G.H., 2008. A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution. Applied Geography 28, 210–223. doi:10.1016/j.apgeog.2008.02.006

Yokoyama, R., Shirasawa, M., Pike, R.J., 2002. Visualizing topography by openness: a new application of image processing to digital elevation models. Photogrammetric Engineering and Remote Sensing 68, 257–266.

Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth surface processes and landforms 12, 47–56.

Appendix

Upper Buffalo MD Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**927**	0	0	7	0	99.3
Divergent Hillslope	36	**1775**	6	216	39	85.7
Convergent Hillslope	42	534	**1496**	1324	42	43.5
Planar Hillslope	88	683	25	**1450**	77	62.4
Planar Upland	2	7	0	0	**3854**	99.8
Producer's Accuracy	84.7	59.2	98.0	48.4	96.1	
Overall Accuracy						**75.2%**
Upper Buffalo MD All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**980**	0	0	7	0	99.3
Divergent Hillslope	6	**2141**	0	91	2	95.6
Convergent Hillslope	78	457	**1517**	1277	96	44.3
Planar Hillslope	31	385	10	**1622**	44	77.5
Planar Upland	0	16	0	0	**3870**	99.6
Producer's Accuracy	89.5	71.4	99.3	54.1	96.5	
Overall Accuracy						**80.2%**

Upper Buffalo WTA Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1071**	161	8	29	0	84.4
Divergent Hillslope	0	**2112**	3	452	57	80.5
Convergent Hillslope	12	62	**1452**	630	0	67.3
Planar Hillslope	12	591	64	**1776**	27	71.9
Planar Upland	0	73	0	110	**3928**	95.5
Producer's Accuracy	97.8	70.4	95.1	59.3	97.9	
Overall Accuracy						**81.7%**
Upper Buffalo WTA All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1063**	24	25	35	19	91.2
Divergent Hillslope	13	**2572**	1	322	43	87.2
Convergent Hillslope	12	44	**1466**	432	0	75.0
Planar Hillslope	7	324	35	**2188**	27	84.8
Planar Upland	0	35	0	20	**3923**	98.6
Producer's Accuracy	97.1	85.8	96.0	73.0	97.8	
Overall Accuracy						**88.8%**
Upper Buffalo RF Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1082**	30	32	34	0	91.9
Divergent Hillslope	0	**2140**	45	772	44	71.3
Convergent Hillslope	0	3	**1214**	298	0	80.1
Planar Hillslope	13	719	234	**1886**	111	63.7
Planar Upland	0	107	2	7	**3857**	97.1
Producer's Accuracy	98.8	71.4	79.5	62.9	96.1	
Overall Accuracy						**80.6%**

Upper Buffalo RF All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1055**	82	28	41	0	87.5
Divergent Hillslope	5	**2440**	13	297	87	85.9
Convergent Hillslope	2	2	**1443**	210	0	87.1
Planar Hillslope	12	324	43	**2440**	30	85.6
Planar Upland	21	151	0	9	**3895**	95.6
Producer's Accuracy	96.3	81.4	94.5	81.4	97.1	
Overall Accuracy						**89.3%**
Upper Buffalo Bayes Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1026**	2	0	13	0	98.6
Divergent Hillslope	23	**2584**	14	723	65	75.8
Convergent Hillslope	11	13	**1323**	305	0	80.1
Planar Hillslope	33	317	190	**1948**	30	77.4
Planar Upland	2	83	0	8	**3917**	97.7
Producer's Accuracy	93.7	86.2	86.6	65.0	97.6	
Overall Accuracy						**85.5%**
Upper Buffalo Bayes All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1039**	0	2	16	0	98.3
Divergent Hillslope	7	**2406**	1	177	6	92.6
Convergent Hillslope	28	20	**1459**	419	10	75.4
Planar Hillslope	21	481	65	**2384**	63	79.1
Planar Upland	0	92	0	1	**3933**	97.7
Producer's Accuracy	94.9	80.2	95.5	79.5	98.0	
Overall Accuracy						**88.8%**

Upper Buffalo SVM Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1084**	46	38	34	0	90.2
Divergent Hillslope	1	**2347**	80	803	59	71.3
Convergent Hillslope	0	0	**50**	0	0	100.0
Planar Hillslope	10	517	1359	**2094**	0	52.6
Planar Upland	0	89	0	66	**3953**	96.2
Producer's Accuracy	99.0	78.3	3.3	69.9	98.5	
Overall Accuracy						**75.4%**
Upper Buffalo SVM All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1083**	52	54	34	0	88.6
Divergent Hillslope	0	**2495**	4	161	6	93.6
Convergent Hillslope	0	0	**1280**	110	0	92.1
Planar Hillslope	12	315	189	**2594**	32	82.6
Planar Upland	0	137	0	98	**3974**	94.4
Producer's Accuracy	98.9	83.2	83.8	86.6	99.1	
Overall Accuracy						**90.5%**
Middle Buffalo MD Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1865**	0	0	0	0	100.0
Divergent Hillslope	603	**3478**	11	159	65	80.6
Convergent Hillslope	1	42	**1668**	709	0	68.9
Planar Hillslope	258	2871	727	**6372**	106	61.7
Planar Upland	290	113	0	95	**3394**	87.2
Producer's Accuracy	61.8	53.5	69.3	86.9	95.2	
Overall Accuracy						**73.5%**

Middle Buffalo MD All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2265**	5	0	0	0	99.8
Divergent Hillslope	81	**4419**	0	3	0	98.1
Convergent Hillslope	1	15	**1997**	395	0	82.9
Planar Hillslope	537	1931	409	**6937**	72	70.2
Planar Upland	133	134	0	0	**3493**	92.9
Producer's Accuracy	75.1	67.9	83.0	94.6	98.0	
Overall Accuracy						**83.7%**
Middle Buffalo WTA Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2594**	198	8	0	354	82.2
Divergent Hillslope	325	**4866**	41	670	813	72.5
Convergent Hillslope	0	34	**1530**	521	0	73.4
Planar Hillslope	92	1242	827	**6144**	1	74.0
Planar Upland	6	164	0	0	**2397**	93.4
Producer's Accuracy	86.0	74.8	63.6	83.8	67.2	
Overall Accuracy						**76.8%**
Middle Buffalo WTA All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2379**	12	88	0	574	77.9
Divergent Hillslope	248	**5360**	7	236	243	88.0
Convergent Hillslope	0	5	**1636**	195	0	89.1
Planar Hillslope	390	976	674	**6904**	12	77.1
Planar Upland	0	151	1	0	**2736**	94.7
Producer's Accuracy	78.9	82.4	68.0	94.1	76.7	
Overall Accuracy						**83.3%**

Middle Buffalo RF Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2744**	152	35	4	280	85.3
Divergent Hillslope	245	**5329**	135	1420	40	74.3
Convergent Hillslope	9	1	**913**	149	0	85.2
Planar Hillslope	19	721	1294	**5638**	5	73.4
Planar Upland	0	301	29	124	**3240**	87.7
Producer's Accuracy	91.0	81.9	37.9	76.9	90.9	
Overall Accuracy						**78.3%**
Middle Buffalo RF All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2931**	188	8	14	188	88.0
Divergent Hillslope	49	**5925**	22	236	53	94.3
Convergent Hillslope	0	2	**1346**	82	0	94.1
Planar Hillslope	37	274	1001	**6876**	0	84.0
Planar Upland	0	115	29	127	**3324**	92.5
Producer's Accuracy	97.1	91.1	55.9	93.7	93.2	
Overall Accuracy						**89.4%**
Middle Buffalo Bayes Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2496**	95	0	0	0	96.3
Divergent Hillslope	414	**5242**	79	1378	119	72.5
Convergent Hillslope	1	7	**1339**	220	0	85.4
Planar Hillslope	34	974	988	**5632**	4	73.8
Planar Upland	72	186	0	105	**3442**	90.5
Producer's Accuracy	82.7	80.6	55.7	76.8	96.5	
Overall Accuracy						**79.5%**

Middle Buffalo Bayes All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2556**	69	0	0	1	97.3
Divergent Hillslope	222	**5429**	0	206	31	92.2
Convergent Hillslope	1	6	**1819**	115	0	93.7
Planar Hillslope	164	804	587	**6927**	17	81.5
Planar Upland	74	196	0	87	**3516**	90.8
Producer's Accuracy	84.7	83.5	75.6	94.4	98.6	
Overall Accuracy						**88.7%**
Middle Buffalo SVM Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2552**	160	89	0	0	91.1
Divergent Hillslope	450	**4683**	453	682	37	74.3
Convergent Hillslope	0	0	**435**	63	0	87.3
Planar Hillslope	0	1082	1386	**6461**	0	72.4
Planar Upland	15	579	43	129	**3528**	82.2
Producer's Accuracy	84.6	72.0	18.1	88.1	99.0	
Overall Accuracy						**77.4%**
Middle Buffalo SVM All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**2786**	64	23	0	15	96.5
Divergent Hillslope	222	**5516**	14	140	0	93.6
Convergent Hillslope	0	0	**1199**	38	0	96.9
Planar Hillslope	9	646	1153	**7028**	2	79.5
Planar Upland	0	278	17	129	**3548**	89.3
Producer's Accuracy	92.3	84.8	49.8	95.8	99.5	
Overall Accuracy						**88.0%**

Lower Buffalo MD Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**896**	0	0	0	0	100.0
Divergent Hillslope	237	**898**	2	388	2	58.8
Convergent Hillslope	62	157	**1732**	1427	0	51.3
Planar Hillslope	0	325	58	**1221**	0	76.1
Planar Upland	92	12	0	16	**919**	88.5
Producer's Accuracy	69.6	64.5	96.7	40.0	99.8	
Overall Accuracy						**67.1%**
Lower Buffalo MD All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**984**	0	0	0	0	100.0
Divergent Hillslope	56	**1284**	0	106	0	88.8
Convergent Hillslope	117	6	**1778**	89	0	89.3
Planar Hillslope	43	98	14	**2826**	0	94.8
Planar Upland	87	4	0	31	**921**	88.3
Producer's Accuracy	76.5	92.2	99.2	92.6	100.0	
Overall Accuracy						**92.3%**
Lower Buffalo WTA Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1106**	0	0	9	0	99.2
Divergent Hillslope	149	**880**	12	300	104	60.9
Convergent Hillslope	32	92	**1496**	1238	0	52.3
Planar Hillslope	0	198	252	**1238**	0	73.3
Planar Upland	0	222	32	267	**817**	61.1
Producer's Accuracy	85.9	63.2	83.5	40.6	88.7	
Overall Accuracy						**65.8%**

Lower Buffalo WTA All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1117**	0	0	0	0	100.0
Divergent Hillslope	92	**1294**	8	229	15	79.0
Convergent Hillslope	41	23	**1658**	125	0	89.8
Planar Hillslope	6	70	126	**2640**	0	92.9
Planar Upland	31	5	0	58	**906**	90.6
Producer's Accuracy	86.8	93.0	92.5	86.5	98.4	
Overall Accuracy						**90.2%**
Lower Buffalo RF Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1135**	0	3	0	9	99.0
Divergent Hillslope	87	**1217**	29	764	0	58.0
Convergent Hillslope	38	20	**1378**	464	0	72.5
Planar Hillslope	0	147	382	**1797**	0	77.3
Planar Upland	27	8	0	27	**912**	93.6
Producer's Accuracy	88.2	87.4	76.9	58.9	99.0	
Overall Accuracy						**76.3%**
Lower Buffalo RF All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1198**	0	3	12	5	98.4
Divergent Hillslope	45	**1345**	0	553	0	69.2
Convergent Hillslope	5	2	**1736**	76	0	95.4
Planar Hillslope	9	45	53	**2405**	0	95.7
Planar Upland	30	0	0	6	**916**	96.2
Producer's Accuracy	93.1	96.6	96.9	78.8	99.5	
Overall Accuracy						**90.0%**

Lower Buffalo Bayes Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**998**	0	0	0	0	100.0
Divergent Hillslope	170	**1118**	4	794	1	53.6
Convergent Hillslope	26	6	**1244**	88	0	91.2
Planar Hillslope	0	232	544	**2124**	0	73.2
Planar Upland	93	36	0	46	**920**	84.0
Producer's Accuracy	77.5	80.3	69.4	69.6	99.9	
Overall Accuracy						**75.8%**
Lower Buffalo Bayes All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1038**	0	0	0	0	100.0
Divergent Hillslope	54	**1346**	0	344	0	77.2
Convergent Hillslope	51	1	**1755**	10	0	96.6
Planar Hillslope	49	39	37	**2638**	0	95.5
Planar Upland	95	6	0	60	**921**	85.1
Producer's Accuracy	80.7	96.7	97.9	86.4	100.0	
Overall Accuracy						**91.2%**
Lower Buffalo SVM Local LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1125**	0	11	0	0	99.0
Divergent Hillslope	111	**1134**	205	819	8	49.8
Convergent Hillslope	0	12	**430**	541	0	43.7
Planar Hillslope	0	230	1142	**1692**	0	55.2
Planar Upland	51	16	4	0	**913**	92.8
Producer's Accuracy	87.4	81.5	24.0	55.4	99.1	
Overall Accuracy						**62.7%**

Lower Buffalo SVM All LSP	Actual Class					
Predicted Class	*Terrace/Floodplain*	*Divergent Hillslope*	*Convergent Hillslope*	*Planar Hillslope*	*Planar Upland*	*User's Accuracy*
Terrace/Floodplain	**1162**	0	15	19	1	97.1
Divergent Hillslope	61	**1342**	0	451	0	72.4
Convergent Hillslope	0	0	**1690**	1	0	99.9
Planar Hillslope	0	50	80	**2558**	0	95.2
Planar Upland	64	0	7	23	**920**	90.7
Producer's Accuracy	90.3	96.4	94.3	83.8	99.9	
Overall Accuracy						**90.9%**