MULTISCALE LANDFORM CHARACTERIZATION FOR LAND USE EVALUATION USING FUZZY SETS

by

ANDREW RAYMOND BOCK

B.S., University of Wisconsin-Stevens Point, 2004

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Master of Arts

Department of Geography

2010

This thesis entitled:
Multiscale Landform Characterization for Land Use Evaluation using Fuzzy Sets
written by Andrew Raymond Bock
has been approved for the Department of Geography

_____
Dr. Stefan Leyk, Assistant Professor of Geography

_____
Dr. Barbara Buttenfield, Professor of Geography

_____
Dr. John Pitlick, Professor of Geography

Date_____

The final copy of this thesis has been examined by the signatories, and we Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

Bock, Andrew Bock (M.A., Geography)

Multiscale Landform Characterization for Land Use Evaluation using Fuzzy Sets

Thesis directed by Assistant Professor Stefan Leyk

A multi-scale geomorphometric landform model was created through the use of fuzzy semantic import models and fuzzy overlay to measure distribution of landforms within parcels of the Conservation Reserve Program in a portion of the Delaware River Sub-basin in Northeast Kansas. Different fuzzy logic operators (intersect, algebraic mean, and fuzzy gamma) were used to test the impact of different model mechanisms on the resulting distributions of crisp and fuzzy membership values, and classification uncertainty measured by entropy values. Across scales (900 m^2 to 16,900 m^2 window sizes), only one crisp class (*drainages*) showed an optimal scale for detection area. The statistical distribution of fuzzy membership values were significantly different between the same classes derived from different overlay operators, but this had limited impact on the agreement between crisp landforms derived from the three operators at a single scale. Within each pairing of overlay operator and scale, the landform classes *backslopes* and *flats* had the highest proportional representation of classes at most entropy levels (0.95, 0.90, 0.85, and 0.75) except for locations within the single-highest entropy level (0.99), where the class proportions were more variable. This is significant, as both of these classes were dominant at different scales (*backslopes* at finer scales, and *flats* at coarser scales) within CRP Parcels. The fuzzy and multi-scale approach provides flexibility in assessing class-level uncertainty that had been previously unaddressed in the use of geomorphometric systems for mapping, modeling, and applied management applications.

Acknowledgements

First and foremost I would like to thank my parents for their continued encouragement, and support, and initiating my interest in natural sciences through numerous camping trips in Colorado and the Western United States when I was younger. Secondly, I would like to thank my advisor for pushing me to explore new areas and applications of GIScience, including previously unfamiliar and strange terrain. Thanks to Dr. Barbara Buttenfield for helping make my transition into graduate school a smooth one and providing a model for creating interest and involvement in GIS when teaching students and novices. Thanks to Dr. John Pitlick for helping me think of new ways to integrate GIS-based terrain analysis and geomorphometry into hydrologic sciences. Additionally, I'd like to thank Dr. Thomas Phillips for helping me to develop some of the algorithms used in the research.

CONTENTS

TABLES

Table

FIGURES

Figure

CHAPTER I

INTRODUCTION

1.1 Project Summary

Within landform taxonomies derived from geographic information systems, uncertainty exists as a result of three primary mechanisms. First, landforms are a scale-specific phenomenon, and the range of surface measurements used to parameterize a location, and thus the extent of landform objects and classes changes with changing scale. Secondly, because landforms are derived from data in the form of continuous fields, such as elevation, representing landforms as crisp entities assumes all important change occurs at the boundary of a specific class or object and ignores gradation and transitional states in the landscape. Finally, landforms are defined by linguistic concepts, and the semantics that define them, such as rolling, flat, or hilly or inherently vague, and can vary significantly between different paradigms.

In this thesis, a method is developed to evaluate spatial positioning of patches of conservation easement programs based on the underlying morphometric classification, which allows the estimation of relevant metrics such as erosion potential. It incorporates a multi-scale approach to derive landform classes based on differential geometry and Fuzzy Set theory in order to examine the stability and scale-dependence of the analysis, and account for the afore-mentioned uncertainties present when deriving landform taxonomies. Fuzzy sets account for inherent uncertainty in defining semantic morphometric classes by assigning membership degrees on a continuous scale. Different semantic fuzzy constructs based on terrain attributes are overlain to create fuzzy semantic landform classes. Change trajectories between the resultant classifications are evaluated through cross-tabulation matrices, similarities between fuzzy membership values

are measured through the Wilcoxon ranked-sum test, and variance and classification ambiguity are measured through Classification Entropy. Such semantic models based in geomorphometry are potential tools to evaluate conservation programs with regard to distributions of landforms, and the potential impact resulting from changes in the extent and distribution of conservation areas. This method of exploring landform patterns could become the basis for more advanced decision-making regarding enrolments in such programs.

The method is demonstrated using data from the Conservation Reserve Program, a voluntary land retirement program managed by the Farm Services Agency (F.S.A.) of the United States Department of Conservation (U.S.D.A.). Since its inception in 1986, the program has grown to more than 34 million acres nation-wide, protecting and enhancing a variety of ecological services. Over the next five years, over 21 million acres are due to expire. These potentially huge conversions of land cover change would have significant effects on natural resource health. Semantic knowledge of where conversions would occur, and subsequent potential for erosion, sedimentation, and nutrient loss would aid in forecasting landscape-scale impacts of land cover change.

Previous authors (Arrell et al., 2007) have suggested that within landform classification, high confusion or entropy indices, indicating similarity between two or more landform classes, could represent transitional areas of the landscape. Thus, transitional locations with a high degree of confusion may be understood as potentially susceptible parts of the landscapes, or areas with high misclassification potential in landform taxonomies. This research project supplements and extends previous research by focusing the fuzzy morphometric analysis on a single categorical best management practice (BMP) designation, and testing the impact of varying spatial scales on the spatial properties of the semantic landform classes.

The goals of this thesis are to investigate the following research questions:

1) How significantly do the semantic geomorphometric classes change with scale? Is one class more stable than others? How does terrain attributes impact the changes?

2) How do different fuzzy logical operators impact final crisp classifications? How similar or dissimilar are the distribution of fuzzy membership values for different operators used for fuzzy overlay?
3) Do spatial locations of high entropy represent transitional states between semantic geomorphometric classes? Are high degrees of entropy related to specific classes?

4) Based on 1-3, what are the implications for application of multi-scale geomorhpometric mapping to land use management and pedological and vegetative mapping??

1.2 Project Scope

To derive a system of simplified landform classes applicable across scales and regions from which basic mechanistic processes and physical soil properties could be inferred, a six-class system was derived (Speight, 1990; Qin et al., 2009). To do so, (1) terrain attributes were calculated using a second-order polynomial (Evans, 1972) and converted to fuzzy semantic constructs (Burrough et al., 1992), which describe the characteristics of semantic landforms, (2) fuzzy overlay was carried out to develop semantic fuzzy representations of each landform class across multiple window sizes, and (3) iteratively defuzzified landform regions were derived to examine spatial stability of resulting classification across scales for the spatial hierarchies pixel,

patch, and semantic-class. The scales examined in this study ranged from the data resolution (10 m) to the maximum window size (100 – 200 m) suggested for capturing in-hillslope variation (MacMillan and Shary, 2009) which is appropriate for the described study to be able to describe the variation in geomorphometric landform distribution within singular agricultural easement parcels whose average size is .072 km^2.

To overcome limitations in Classic Set theory, which restricts memberships to a 1 or 0, fuzzy set theory assigns the membership of a location to a range of values more suitable for dealing with uncertainty in spatial topographic data. Semantic import models based on first-order polynomials (Burrough et al., 1992; Robinson, 2003) were utilized to convert terrain attributes to fuzzy membership values on a continuous scale [0,1], in order to develop the corresponding semantic constructs for landforms. Within each model, central concepts were based on existing definitions; parameters and dispersion indices were taken from previous models and the statistical distribution of the terrain attributes over the study area to maintain a degree of generality of the approach. For each terrain attribute, multiple semantic constructs were derived to describe the relative surface shape and landscape position of each location (pixel).

Semantic representations of landform classes were created for each spatial scale by combining surfaces of semantic constructs in a fuzzy spatial overlay operation. Different fuzzy operators were tested, such as fuzzy intersection (MIN operator), algebraic mean, and fuzzy gamma coefficients (Robinson, 2003) to produce 6 surfaces, each representing one fuzzy semantic landform class (*crest*, *shoulderslope*, *backslope*, *footslope*, *flat*, and *drainage*).

For each spatial scale the degree of classification stability was measured by calculating the classification entropy. Low values of entropy indicate that there is one dominating class; high

values of entropy indicate there are at least two classes of similar high membership at the same location, thus this location is more ambiguous (Wood, 1996; Burrough et al., 1997). The stability of memberships of each location was tested based on the distribution (spatial and statistical) across scales.

Changes in spatial distributions of semantic landform classes (fuzzy and crisp) at the three spatial hierarchies will be examined across scales to examine the stability and scale-dependency of the final classification. The highest amount of stability and least amount of uncertainty was expected within crests and highest areas in the landscape, but the degree of flux across scales is unknown. Since CRP enrollment occurs across physiographic regions, a generic model based on land-surface parameterization could be an adaptive tool for multi-scale assessment to expand pixel-based classification and identify semantic classes in landform taxonomies that have high classification uncertainty.

CHAPTER II

BACKGROUND

2.1. The Conservation Reserve Program

2.1.1 Background

The Conservation Reserve Program (CRP) was established by the U.S. Congress through the Title XII of the Food Security Act of 1985. The program is administered by the Farm Services Agency (FSA) of the U.S. Department of Agriculture (USDA). The program was authorized during a period when the United States had excess supplies of wheat and feed grains (Secchi et al., 2009) after the expansion of agricultural exports in the 1970s skyrocketed to record levels and producers responded by bringing additional land into production (Zinn, 1997). Excessive rates of erosion, excess capacity, and insufficient markets resulted when the boom ended. The program's primary goal is to reduce soil erosion by maximizing the amount of sensitive and highly erodible cropland taken out of production. Secondary objectives are to protect the nation's long-term capability to produce food and fiber, reduce sedimentation, improve water quality, create wildlife habitat, control production of commodities, and provide financial support to agricultural producers. Benefits of the CRP have included erosion control, enhanced wildlife habitat, reduction in surplus commodities, and farm income support (Dao, 2000). Most often the program has been credited with substantial reduction in wind and water erosion of marginal croplands (Moore and Nieber, 1989).

CRP works as a voluntary program in which farmers agree to temporarily retire land from agricultural production in exchange for annual rental payments and cost-share assistance from the federal government (Ferris and Siikamaki, 2009). It is unique in that it seeks limited-duration conservation easements rather than the outright purchase of land. The initial program put aside 14.7 million acres nationally, and was eventually capped at 36.5 million acres. Under the initial guidelines, a farmer placing marginal cropland into CRP makes a commitment to retiring the land for at least ten years for approved conservation use. The program has been renewed in 1990, 1996, 2002, and 2008 with varying revisions (Ferris and Siikamaki, 2009). The 1990 bill focused on water and water quality. The 1996 re-authorization mandated stricter environmental benefits to the requirements for contract renewal or new enrollment. In 2009, CRP had a total enrollment of 34.7 million acres on 430,000 farms with a total payout of $1.8 billion (Ferris and Siikamaki, 2009). Nationwide, CRP enrollment is capped at 36,000,000 acres, with a maximum of 25% of any individual county that can be enrolled (Babcock et al,. 1995).

There have been many modifications to the initial enrollment criteria, such as increasing the threshold for years previously cropped, and the introduction of an environmental benefits system (EBI) for a more holistic natural resource management approach (Ribaudo et al., 2001). The main criteria to be eligible for enrollment must include one of the following (Ferris and Siikamaki, 2009):

- cropped in two of the previous five years (later changed to four),
- be highly erodible land,
- be located within a national or state priority area,
- be part of an expiring CRP contract,

- considered a cropped wetland,
- Be devoted to a highly beneficial environmental practice (stream buffers, etc.),

The main quantitative criterion for enrollment is that a tract of land is classified as 'highly erodible land' (HEL), based on the physiographic and pedological characteristics of the land. Highly erodible land is defined by the *Erodibility Index* (EI), developed by the NRCS to determine the inherent susceptibility of land to either sheet or rill erosion (water erosion) or wind erosion in relation to the amount of soil loss that can be tolerated for a particular soil type (Batie 1997).

The *EI* allows the estimation of the potential erodibility of a tract of land and is composed of two factors, the water and wind erodibility indices. The water erodibility index is comprised of factors found in the Revised Universal Soil Loss Equation (Wischmeier and Smith, 1978: National Research Council, 1986; Renard, 1996), used for calculating sheet and rill erosion from rainfall and runoff in a landscape where overland flow occurs. Less than 20% of enrollments are in regions where water erosion is greater than wind erosion, and in most cases total *EI* is measured by combining water and wind erosion indices (Ribaudo et al., 2001; Egbert, 2002). The equation for water erosion is:

$$EI_{water} = \frac{RK(LS)}{T} \qquad (2.1)$$

where *R* is the rainfall erosivity factor that accounts for energy and intensity of rainstorms (m^2 kg ha^{-1} hr^{-2}), *K* is the soil water erodibility factor (hr m^{-2}), *LS* is the dimensionless product of slope steepness and slope length factors, and *T* is the soil loss tolerance for a specific soil for high level of crop production (kg ha^{-1} hr^{-1}) (Wischmeier and Smith, 1978). *RKLS* is considered erosion potential, and can be normalized for soil type by dividing by *T*, the soil loss tolerance (Moore

and Nieber, 1989). Generally soils high in silt have greater erosion potential, correlating with higher K and T values. Wind erosion indices are computed from:

$$EI_{wind} = C\left(\frac{I}{T}\right) \qquad (2.2)$$

where *C* is the wind erosion climate factor, *I* is the soil wind erodibility factor (kg ha^{-1} hr^{-1}), and *T* is the soil loss tolerance (Wu et al., 2002). CRP enrollment uses an area-weighted *EI* threshold of eight or greater to determine whether or not land is highly erodible (Batie et al., 1997), while 'Excessively Erodible Lands' have an $EI > 15$. Erodibility indices can be computed spatially through the use of the regional climatic data and information from the Soil Survey Geographic Database (SSURGO), which contains soil attributes for each geographic entity in the database known as the map unit, a distinct polygon comprised of one to three soil types (Park and Egbert, 2005). Additionally, the original coefficients of the LS factor, which were empirically-derived from experimental plots across the United States, can be represented by multiple geospatial methods within a gridded digital elevation model (DEM) (Moore and Burch, 1986, Moore and Wilson, 1992; Mitasova, 1993; Desmet and Govers, 1996). Many of these methods primarily use surface parameters and measurements derived from a digital elevation model (DEM) as a way of better representing detachment and deposition processes, and flow convergence or divergence. Several limitations to current geospatial EI calculations exist. When using raster data, calculations are constrained by the use of the smallest window size to calculate slope and other parameters, where there is a higher magnitude of noise and errors in surface measurement. An area-weighted EI also ignores the spatial distribution of erodible lands within a parcel.

2.1.2 Conservation Reserve Program in the Northern Plains States

Currently, more than 34 million acres are enrolled within the CRP in all 50 states. Of the states participating, Kansas has the fourth highest number of active acres enrolled in CRP in the United States. The central part of the Great Plains, including Kansas, Nebraska, North and South Dakota, account for 21% of all CRP contracts in the United States (Figure 1) (Park and Egbert, 2005).

Figure 1. U.S. Department of Agriculture Farm Production Regions and percentage of Conservation Reserve Program Contracts within the Regions in 2001 (Allen, 2005).

CRP enrollments ballooned to an all-time high of around 37 million acres in 2007, and have decreased since to 33.6 million acres. One of the criticisms of CRP is that the majority of the benefits are concentrated on small percentages of the land. According to Babcock et al. (1995), 98% of total surface water quality benefits are obtained by enrolling land on less than 27% of

CRP land, 90% of wind erosion and groundwater vulnerability are concentrated on 32% of the land, and 43% of CRP lands achieve 90% of the water erosion benefits.

Within the next five years, 20 million acres of CRP contracts are scheduled to expire across the United States (Figure 2). Such a potential land conversion could have an enormous impact on a large number of environmental issues across many scales, ranging from soil loss, sedimentation, and nutrient enrichment of waterbodies, to contributing to input of the hypoxia zone of the Gulf of Mexico, and habitat preservation along the mid-continental migratory flyway.

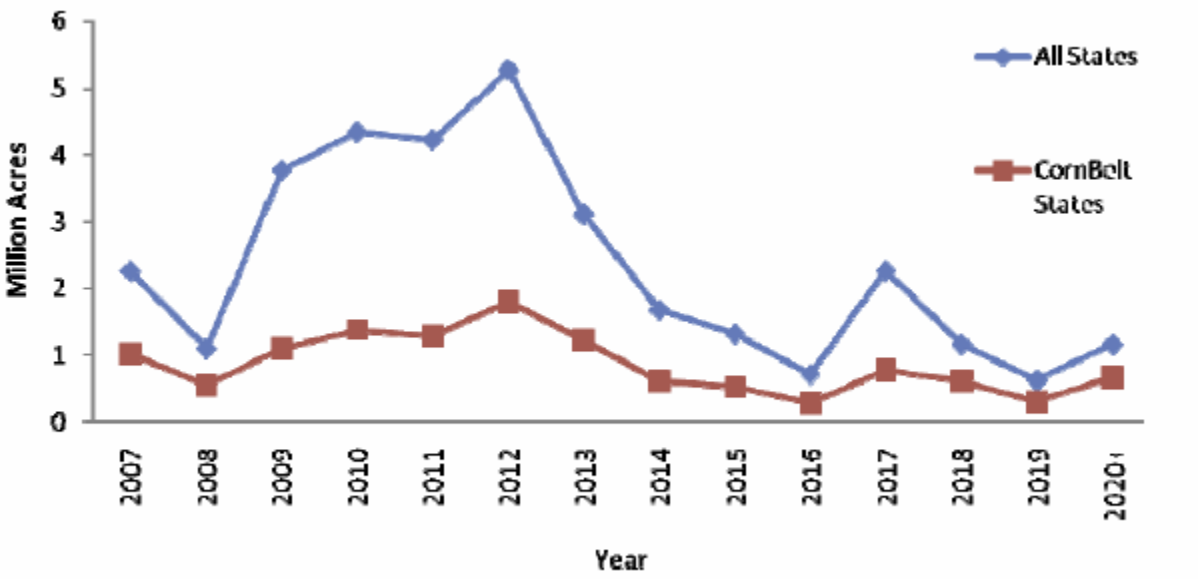


Figure 2. CRP acres expiring by year and region, including the results of 2006 re-enrollment and extension offers. (Baker and Galick, 2009)

Because of the wide variability of soil conditions across regions and even within single fields, developing a model to understand where CRP enrollments are situated in the landscape and the terrain characteristics of expiring acres and annual enrollments could be a tool for long-term planning in an uncertain future of climate change, and for adapting to the fluctuation of conservation budgets, commodity prices, and environmental goals.

2.2 Geomorphometric Terrain Analysis

Characterizing the terrain and its biological and geophysical properties of the land surface are the primary goals of geomorphometric terrain analysis. The magnitudes of many hydrological, geomorphologic, and biological processes active in the landscape are sensitive to topographic position (Moore et al., 1991). The spatial distribution of terrain attributes can therefore be used as an indirect measure of the spatial variability of these processes (Odeh, 1994). Geomorphometric measures derived from digital elevation models vary depending on the data resolution and the scale of analysis.

Land surface morphology is fundamental to mapping and many GIS applications. The field of geomorphometry is defined as "the science which treats the geometry of the landscape and attempts to describe quantitatively the form of the land surface" (Chorley et al., 1957). Identification of elementary landform units is important to the study of past and present geomorphologic processes, and spatial aspects of interaction among landforms, soil, vegetation, topoclimate, and hydrological regimes (Minar and Evans, 2008). The characterization of surface shape has evolved from the need to describe topography more exactly than terms such as 'flat' or 'rolling' (Pike 1988), to modeling how geophysical processes shape the landscape and affect ecological systems.

2.2.1 General and Specific Geomorphmetry

There are two directions of geomorphometry, as distinguished by Evans (1972). Specific geomorphometry measures the geometry of specific types of landforms, such as drumlins,

cirques, or karst pits. General geomorphometry is the measurement and analysis of those characteristics of landforms that are applicable to any continuous surface. General geomorphometry deals with measurement of surface altitude, gradient, distance, and elevation derivatives, which allow for more variability in operational definition. This approach provides the basis for the quantitative comparison of qualitatively different landforms, and can adapt methods of surface analysis used outside geomorphology (Evans, 1972).

All measures of land surface form can be considered to be representative of the *roughness* of the surface (Mark, 1975). Roughness refers to the irregularity of a topographic surface, which can be expressed by the terms texture, grain, and relief. Geomorphologists have long sought after a single measurement to quantify the extent through which the land has been 'opened up' (dissection or aeration), especially by erosion or denudation (Evans, 1972). Earlier attempts to define such measures include periodicity and spectral analysis to quantify relief placement (Bull, 1975; Pike and Rozema, 1975; Pike, 1988) or hypsometric and slope frequency curves in drainage basin morphometry (Chorley et al., 1957).

2.2.2 Derivatives of Elevation

Any topographic surface can be represented through a $z = f(x,y)$ bivariate function, and can be analyzed through differential geometry and mathematical vector analysis (Moore et al., 1993; Jordan, 2003). Building upon Tobler (1969), Evans (1972; 1980; 1984) suggested a unifying framework for geomorphometric analysis through analysis of the zeroeth, first, and second-order derivatives of elevation, known commonly as relief, slope, and curvature. Terrain parameters should be sensitive to process as well as morphology (Wood, 1996), and these measures are

process-related and directly applicable to specific and general geomorphometry. The three derivatives provide five descriptions of surface form: altitude (zeroeth-order derivative), slope and aspect (first-order derivatives), and plan and profile curvature (second-order derivatives). Each derivative provides a map of point values producing statistical distributions that may be characterized by their mean, standard deviation, and dimensionless skewness (Evans, 1972). Correlation and factor analyses have shown that the five categories are complementary; each category measures topography in a profile that the other four do not (Evans, 1972; Mark, 1975; Pike, 1988). The five derivatives also play a role in characterizing regions, describing the surface shape and deriving areas of uniform curvature properties. The advantage of using derivatives of altitude as measures of terrain is that they can be determined for any point sampled over a surface or continuous field (Evans and Cox, 1999). A schematic diagram of the primary and secondary derivations of elevation is shown in Figure 3.

Slope is the first derivative of altitude, and is represented as a data vector with two components; gradient (slope angle) and aspect (azimuth) (Evans and Cox, 1999). Slope gradient is arguably the single-most important parameter of topographic form, since surfaces are formed completely of slopes, and slope angles control the gravitational force available for geomorphic activity (Strahler, 1952; Wood and Snell, 1960; Evans, 1972). Besides obvious uses in topoclimatic and vegetative modeling, slope aspect is used in hydrological modeling as a surrogate for flow direction and flow accumulation vectors. Statistics calculated for slope in most landscapes approach distributions similar to the Gaussian curve (Pike, 1988).

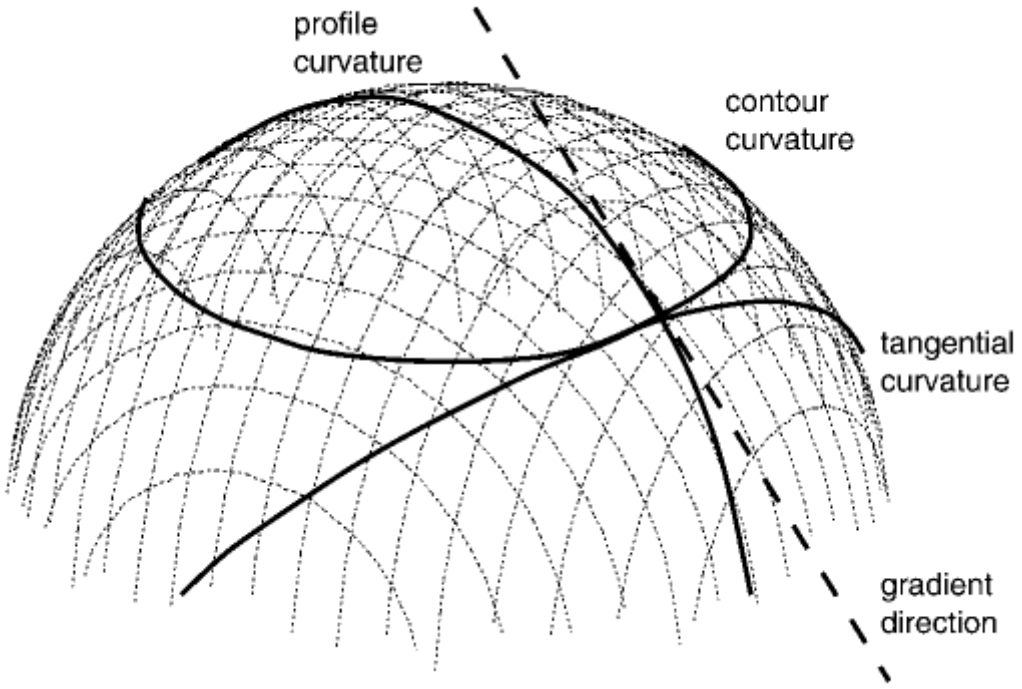


Figure 3. Measures of curvature on a surface in relation to gradient (slope). (Schmidt et al., 2003)

Slope Curvature is the second derivative of altitude. Curvature is defined as the angle subtended by three or more adjacent elevations along an elevation profile and is represented by a symmetrical 2 x 2 tensor with three independent components (Evans and Cox, 1999). The importance of profile and plan convexities rests in their control of acceleration/deceleration and convergence/divergence of near-surface flows (Evans, 1980). A signed distinction is applied to curvature by designating convex curvatures as positive and concave slopes as negative (Pike, 1988). The first two independent components, vertical curvature (rate of change down a slope line), and horizontal curvature (rate of change along a contour), can be represented by the terms profile and plan curvature (Evans and Cox, 1999). The third component of curvature is expressed by aspect change downslope, and may be too abstract for modeling most surface

processes. Convexity varies from high positive values (convexities), to high negative values (concavities) and the tails of distributions are always long with the great majority of values near zero (Evans and Cox, 1999).

Profile curvature is the derivative of the gradient vector magnitude and reflects the change in gradient along the slope line that controls the change of velocity of mass flowing down the slope curve (Jordan, 2003). At areas with convex profile curvature, flow velocity, and thus sediment detachment and entrainment increases, and at areas with concave profile curvatures, flow velocity decreases and sedimentation increases (Shary, 1995). Plan curvature measures the derivative of aspect vector magnitude and measures the angular change of the two-dimensional aspect vector. Plan Curvature reflects the influence of converging and diverging flows on a surface (Jordan, 2003). Some have advocated the use of Tangential Curvature in place of Plan Curvature because its incorporation of the slope vector (Mitasova, 1996; Wilson and Gallant, 2000). Plan and Tangential Curvature are related through the sine of the slope angle, and their computed distribution of convex and concave areas are the same, but the distribution of values for Tangential Curvature are closer to that of Profile Curvature (Jordan, 2003).

Spatial influence of the five derivatives was initially recognized in pedology by Aandahl (1948) and in geomorphology by Speight (1968) among others. Evans proofs (1972, 1980) and Young (1978) are considered the seminal work for quantitatively representing the definitions of all five derivatives. The majority of the focus in geomorphometry and geomorphology has been on slope gradient and curvature, since these are the most pertinent to physical processes such as erosion and deposition.

Surface derivates have a significant effect on surface processes. Soil water content has been related to curvature (Zaslavsky and Sinai, 1981), slope and aspect (Moore et al., 1991), as well as

slope and plan curvature (Burt and Butcher, 1986). Infiltration is affected by profile and plan convexity. Profile convexity affects the acceleration of surface flow, and plan convergence affects the convergence, divergence, and depth of flow, leading to considerations of slope position and area drained (Schmidt, 1998). Kirkby and Chorley (1967) noted that areas which are most frequently saturated are those adjacent to flowing streams, lines of greatest slope convergence, and local concavities. Slope and curvature have been related to slope instability, fault recognition (Florinsky and Kuryakova, 1996), landslide susceptibility (Pike, 1988; Gorsevski et al., 2003), A-horizon depths, calcium carbonate concentrations, and gleyed horizons (Pennock et al., 1987). Heerdegen and Beran (1982) used terrain statistics that describe overland flow patterns such as gradient and curvature of flow path to quantify source areas of overland flow, such as hollows with strong concavities and convergent flow.

2.2.3 Geomorphometric Objects

Many have advocated a parametric (attribute-based) approach to description of land systems with the objective of facilitating the comparison of land systems based on elevation derivatives Speight, 1974; Bull, 1975). The problem of integrating scale, spatial extent, and resolution makes single object classifications of landscapes unfeasible (Wood, 1996). Hence, a system of landforms tends to be classified either according to homogeneous regions (Speight, 1968; Dikau, 1989), or specific features such as streamhead locations (Tribe, 1991).

Geomorphometric components are grouped into two parts, the ‘geomorphometric point’ and the ‘geomorphometric object’. As the land surface is represented by a continuous and infinite number of geomorphometric points (elevation), the three dimensional space can be characterized by geomorphometric parameters, or terrain attributes, most simply represented by the derivatives

of elevation (relief, aspect, curvature, etc.) (Schmidt and Dikau, 1998). Geomorphometric objects are areal or linear elements of the land surface consisting of an array or matrix of geomorphometric points and given a semantic definition, such as a hillslope, ridge, channel, or crest. Within each object, there may exist indices to translate semantic knowledge to a set of rules to quantify the point-object relationships, such as the internal frequency of point attributes (height, slope, curvature), and their spatial arrangement (geomorphometric structure) (Schmidt and Dikau, 1989). Within geomorphometric objects, Dikau (1989) proposed form facets (units with homogeneous gradient, aspect, and curvature) and form elements (units with homogeneous plan and profile curvature) as basic components of landforms (Figure 4).

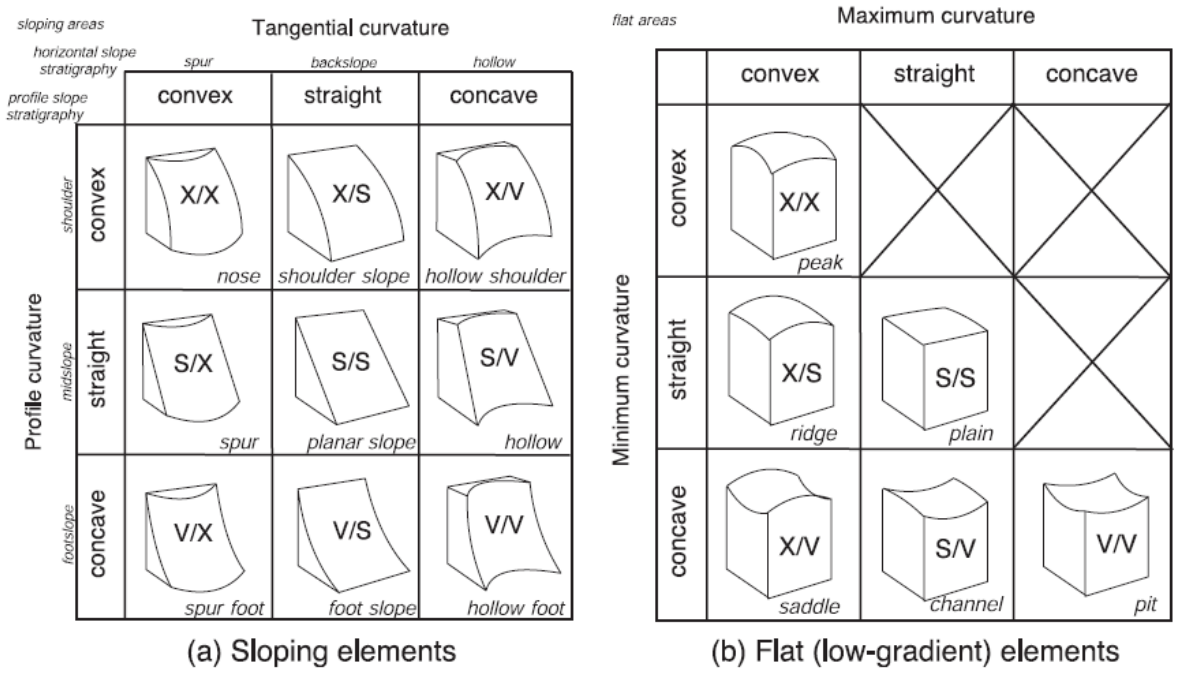


Figure 4. Dikau's 15 form-facets; sloped elements distinguished by combinations of signs of profile and tangential curvature, and flat elements distinguished through combination of signs of maximum and minimum curvature (Dikau 1989).

Primary geomorphometric parameters, or terrain attributes, are defined as indices of the geomorphometric point and can be split into three groups: simple, complex, and compound

parameters (Schmidt and Dikau, 1998). Simple geomorphometric parameters can be generated from a gridded DEM through moving-window operations and include the first five derivatives of elevation, and primary flow direction (Schmidt and Dikau, 1998; Wilson and Gallant, 2000; Jordan, 2003). Complex geomorphometric parameters are derived through the analysis of the whole matrix of the DEM and contain structural information about the surrounding neighborhood, such as drainage area, slope position, and flow length (Schmidt and Dikau, 1998). Compound or combined geomorphometric parameters are calculated from simple or complex analytical functions, and express the effect of landforms in modeling different processes, such as soil-water distribution, surface water saturation, incoming solar radiation, and potential for sheet erosion (Moore et al., 1991; 1993). Prime examples of compound geomorphometric parameters include the Compound Topographic Index or Wetness Index (Beven and Kirkby, 1979), Sediment Transport Capacity (Zhang and Montgomery, 1994), and modifications of the Length-Slope Steepness Factors from the original and Revised Universal Soil Equations (Moore et al., 1991). In areas with small amplitudes of relief, non-local and compound topographic variables are the most significant correlators of phytocoenosis properties (Florinksy, 1996). Table 1 (pg. 26) distinguishes commonly used geomorphometric parameters, and their significance in hydrology, terrain analysis, and soil-landscape modeling.

A geomorphometric object is a clearly defined object or unit, such as a flow path, hillslope, or catchment. The objects are subdivided into two main subclasses, linear objects and areal objects (Schmidt and Dikau, 1998). Analysis of primary geomorphometric parameters is used for the identification and extraction of geomorphometric objects. The geomorphometric object can be viewed as a hierarchical, multi-scaled systems, including slopes, valleys, or drainage basins (Dikau, 1989). Such a hierarchical system allows finer scale geomorphometric objects to

be used for deriving coarser scale objects using their geomorphometric attributes (Chorly, 1984). In addition, processes controlled by attributes with meanings at scales of macrorelief (10^6 , magnitude of meters) such as drainage density, attribute dispersion or variability, may not hold meaning at the scale of microrelief or mesorelief where aspect, curvature, and connectiveness have a higher degree of control over the dominant geophysical processes (Dikau, 1989).

Table 1. Geomorphometric Parameters. (Wilson and Gallant, 2000).

Attribute	Definition	Significance
Slope	Gradient	Climate, vegetation, energy
Catchment Area	Area draining to catchment outlet	Time of concentration and flow prediction
Specific Catchment Area	Upslope area per unit width of contour	Runoff volume, soil moisture
Flowpath Length	Maximum distance of water flow to point in catchment	Sediment yield, time of concentration
Elevation Percentile	Proportion of cells in a user-defined neighborhood lower than the focal (center) cell	Landscape position, flora and fauna distribution, viewshed analysis
Roughness	Magnitude of elevation variability in user-defined neighborhood	Orographic effects, trafficability, water storage, flora and fauna movement
Profile Curvature	Slope profile curvature	Flow acceleration, erosive energy
Tangential Curvature	Contour curvature normalized by the slope	Flow convergence and divergence, soil properties
Wetness Index	Specific catchment area divided by slope	Zones of saturation, runoff generation
Stream Power Index	Measure of erosive power based on unit stream power theory	Erosive power of runoff, locations of detachment and deposition, channel initiation
Radiation Indices	Estimations of incoming or ongoing short- and longwave radiation	Energy balance, snowmelt, flora and fauna distribution

Advances in data resolution such as LIDAR, improved modeling capabilities such as object-based segmentation (Blaschke et al., 2006), and established representations of hierarchical relationships through methods such as artificial neural networks (Moller et al., 2008), create opportunities to carry out research on higher efficiency in algorithms in order to integrate the data wealth generated for GIS, and to incorporate complex objects into operable scenarios.

2.2.4 Uncertainty in Geomorphometric Objects

Traditional landform classification based in geomorphometry has utilized a system of units with maximum internal homogeneity and maximum external heterogeneity (Dikau, 1989). This paradigm is referred to as the 'double-crisp' model (Burrough et al., 1997), explained in more detail later. This model is incorporated into the polygons of the Soil Mapping Units (SMU) of the NRCS Soil Surveys (Figure 5), traditionally the basis for land resource planning in the United States.

Within a GIS environment, elevation is primarily represented through a gridded digital elevation model (DEM) or its derived surfaces such as a Triangulated Irregular Network (TIN). DEMs exist theoretically as a continuous field of elevation represented by a series of discrete objects (pixels or cells), each continuing a unique elevation value, and exhibiting internal homogeneity within the pixel area. Many existing parameters used for surface characterization are dependent upon both the resolution of the DEM and the analysis window used to compute them (Schmidt and Andrew, 2005). Landforms have been described as a mixture of different processes at different scales (Chorley et al., 1984), and likewise surface parameters should be expressed at a variety of scales (Wood, 1996). Despite this, many of the approaches still rely on existing techniques and ignore the incorporation of scale effects (Schmidt and Andrew, 2005).

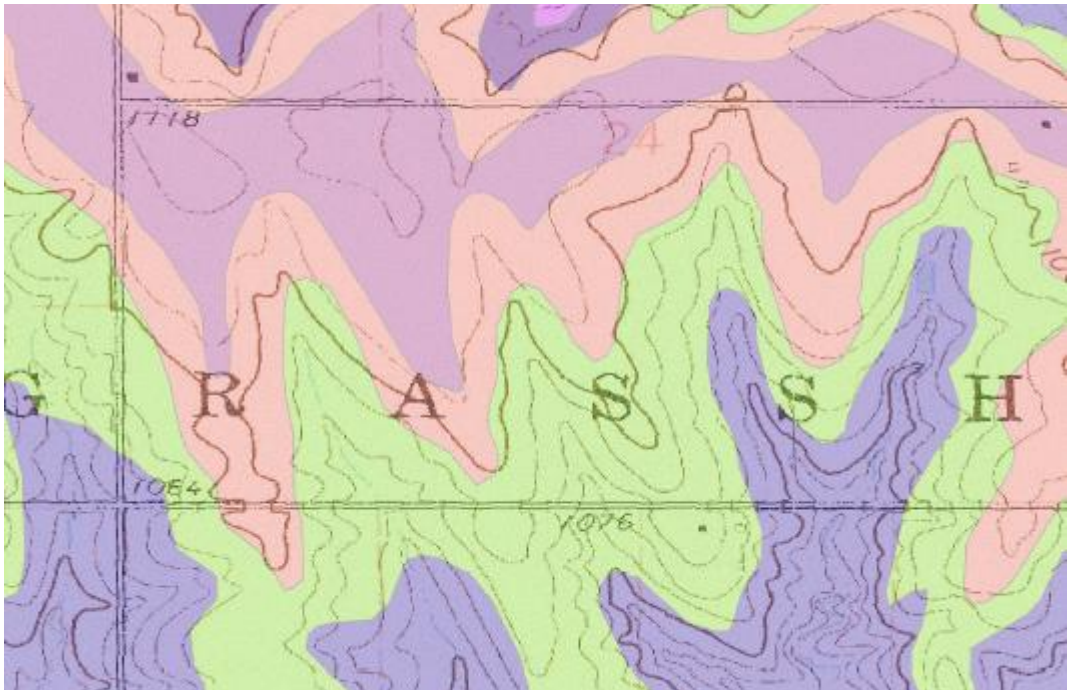


Figure 5. Soil Mapping Units from SSURGO within the study area, each color-coded soil mapping unit (polygon) is composed of several soil types bounded by definite boundaries and conceptualized to have maximum external heterogeneity and internal homogeneity, but fails to take into account spatial variation of biophysical processes, the gradation of terrain units.

With this theoretical context, representation of an elevation surface should take into account (Burrough et al., 1997; Schmidt and Andrews, 2005):

- Vagueness, uncertainties, and transitional states present within a continuous dataset,
- That properties of a single location may relate to or determine a number of classes,
- Precise boundaries cannot be defined,
- Properties of a single location relevant at one scale may change with increasing or decreasing scale.

2.3 Fuzzy Set theory

2.3.1 The use of Fuzzy Set theory in Geomorphometry

Traditionally, the soil-landscape model is characterized by what is referred to as a 'Double-Crisp Model' (Burrough et al., 1997), composed of entities having fixed and uniform conditions within each entity and abrupt boundaries between entities (MacMillan et al., 2000). In early Geomorphometric work, landform classes were defined through the central concepts of each class and the salients of class boundaries. However, this conceptual model fails to take into account the strength of membership of a location to one class, and relative confusion between classes (McBratney and Odeh, 1997). The use of crisp classes is sometimes inappropriate for terrain and soil modeling because of the continuous nature of elevation and its derivatives, soil properties, and the processes that influence them (Burrough, 1989). Problems encountered with previous efforts include fragmented or chaotic spatial patterns, and the inflexibility of Boolean classifications that require adjustment for different types and scale of landforms, and lack of generality across different landscapes (MacMillan et al., 2000).

Due to dependence on the user's perception and semantics of their specific paradigm (Moller et al., 2008), landscape and semantic landforms lack a clear definition have been increasingly conceptualized as vague phenomenon (Burrough et al., 1992; Brandli., 1996; Fisher et al., 2004). With the continuous nature of elevation surfaces, spatial representation of terrain features should also emphasize that the transition between adjacent landforms is gradual, not abrupt (Qin et al., 2009). The use of Fuzzy Set theory allows soft boundaries and uncertainties to be represented as a part of the landscape structure, and specific parts of the landscape to be

partial members to one or more landform classes on a scale of 0 to 1, with 1 indicating prototypical or full membership. Because of the multi-scale variation in feature identification, the use of fuzzy set theory would allow a user represent multi-scale uncertainty.

2.3.2 Fuzzy Set theory

Zadeh (1965) introduced the idea of working with fuzzy sets as a method of dealing with inexact concepts and non-ideal classes. Fuzziness is a type of imprecision characterizing classes that cannot have, or do not have sharply defined boundaries (Burrough, 1989). Crisp sets based on Boolean theory allow only binary membership (true or false), and allow for no overlap of membership (i.e. there is no union between sets A and A') (Robinson, 2003) (Figure 6).

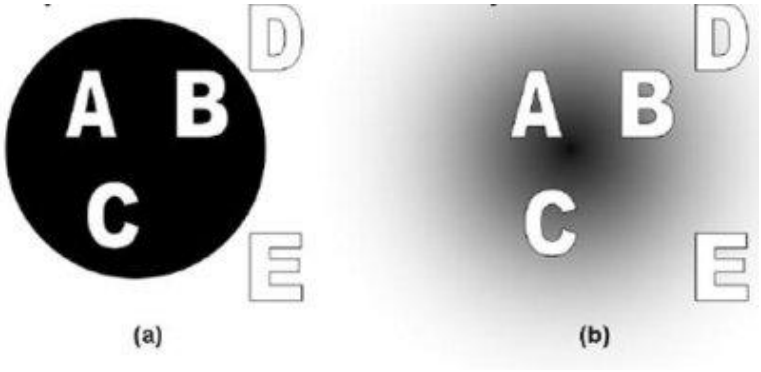


Figure 6. An alphabetical sequence (x:A:E) represented within the attribute space of Classic Boolean theory (a) and Fuzzy Set theory (b). In Classic Boolean theory, membership to the set is restricted to A,B, and C while D and E are non-members. In Fuzzy Set Theory all members of the set have some degree of membership to the set noted by the grayscale (Klir and Yuan, 1995).

This allows for no expression of imprecision, partial membership, or errors of misclassification or measurement. These rules are termed as the Law of the Excluded Middle and the Law of Contradiction. The rules assume that all important changes occur at the defined class boundary

(Burrough et al., 1992), and cannot deal with the inexactness qualitative ambiguous terminology such as 'poorly drained', 'slightly susceptible', or 'very shallow' (Burrough, 1989).

The idea of Fuzzy Set Theory and Fuzzy Logic is to allow entities to belong to more than one class to varying degrees, violating both the Law of the Excluded Middle and Law of Contradiction. A fuzzy set is expressed as:

$$A = \{X, U_A(x)\} \qquad x \varepsilon X \qquad (2.1)$$

where $x \varepsilon X$ denotes a field of objects, and $U_A(x)$ is the grade of membership of x in A for all x in X. The grade of membership is a range from 0 to 1, 1 representing prototypical, or full membership to A, and 0 representing non-membership (Burrough, 1989). It is important to note that the grades of membership of x in A reflect the possibility of membership, not the probability (Klir and Yuan, 1995).

2.3.3 Fuzzy K-Means Algorithm

Fuzzy Classification methods can be divided into two groups; Similarity Relation Models (SR) and the Semantic Import Model (SI) (Burrough et al., 1992). SR models utilize the Fuzzy k-means or c-means algorithm based on the methods of Bezdek et al. (1984). The SR model is based on numerical taxonomy and discriminatory analysis (Burrough, 1989), and is an exploratory approach that utilizes clustering algorithms to find a natural partition of objects in multivariate space into a user-defined number of classes. Within GIS this has been utilized for climatic classification (McBratney and Moore, 1985), soil science (McBratney and de Gruijter, 1992; McBratney and Odeh, 1997), topoclimatic landscape classifications (Burrough et al., 2001), and landslide susceptibility (Gorsevski et al., 2003). The main advantage of this approach

is that it helps the user define natural classes based on an infinite number of datasets. The main disadvantage is that the natural classes may be vague or arbitrarily based on cluster differences (Burrough, 1989). The resulting classes and class definitions are optimized for a particular site, and will never be exactly the same for any two sites (MacMillan et al., 2000).

To construct a Fuzzy K-means clustering classification, the user selects a number of parameters to base the classification on. In geomorphometry and soil-landscape modeling, these are usually multiple geomorphometric parameters (Burrough et al., 2001). Based on each spatial location $X_{ij,}$ and its parameter values, the algorithm iterates through the parameters and partitions the data clusters based on a user-defined number of classes and a fuzzy exponent to determine the amount of fuzziness or overlap between the clusters/classes. Thus, each spatial location, or pixel within a DEM, is given a fuzzy membership to each cluster/class, which sum to 1 for each location. The shorter the attribute distance to each cluster center, the greater similarity of the location X_{ij} to that class. A number of additional algorithms, such as the partition coefficient and classification entropy can be used to determine the optimal number of classes based on variation in the clusters. Figure 7 shows a four-class solution to a Fuzzy K-means classification utilizing slope, plan and profile curvature, compound topographic index, and stream power index for a portion of the study area. The resultant classes can be considered 'natural classes' in that they are partitioned almost exclusively by their attribute distance, and may not correlate with real-world classes (such as soil-mapping units defined previously) defined linguistically or through artificial boundaries.

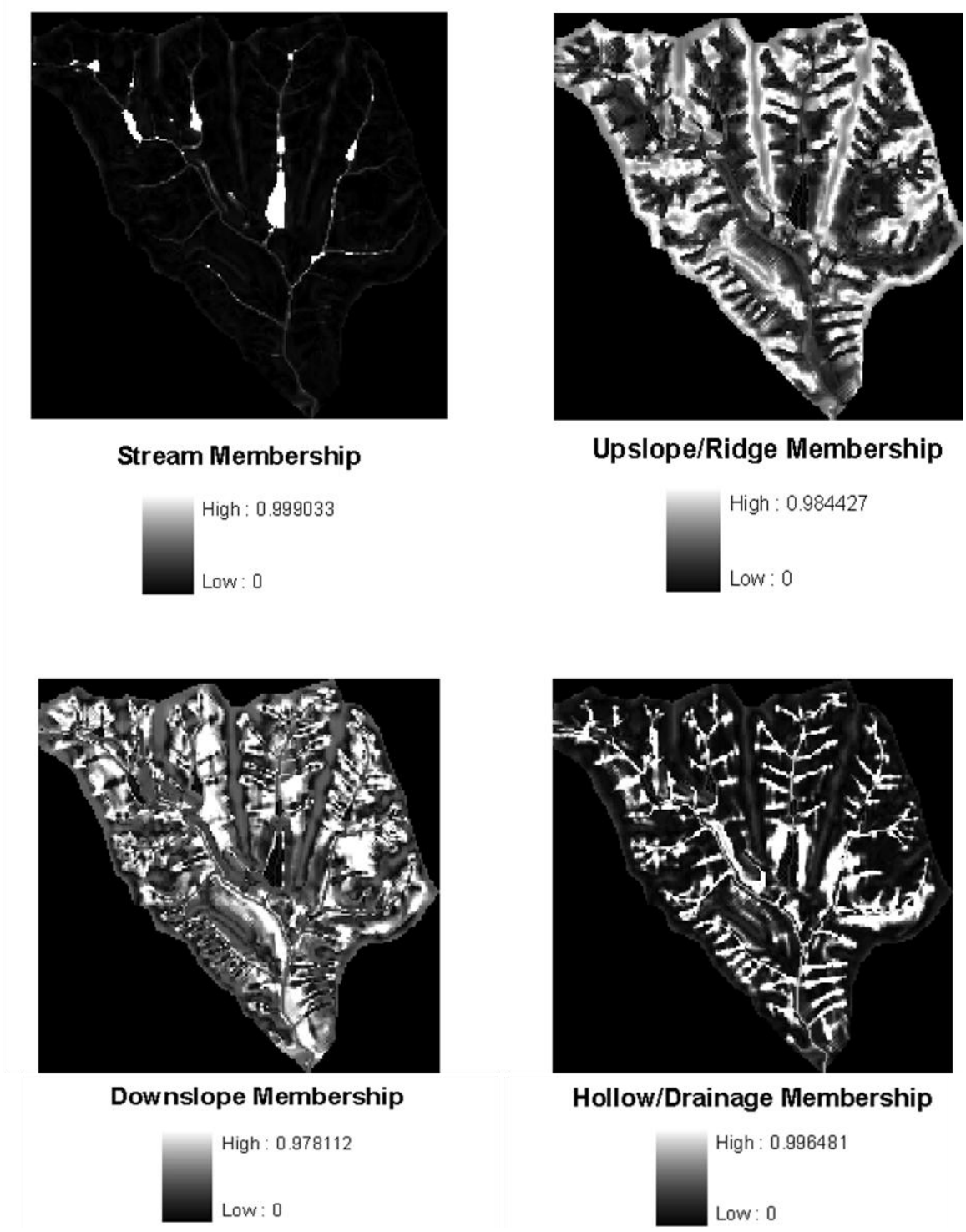


Figure 7. Four-class Fuzzy K-means solution for portion of study area. Lighter areas indicate locations with membership degrees higher to the cluster center, while darker areas represent low degrees of membership to corresponding cluster center. Top of the page corresponds to North, which is repeated for the remaining map figures in the thesis.

2.3.4 Fuzzy Semantic Import Model

The second approach to building a fuzzy set is to specify class-limits on the basis of expert knowledge and construct mathematical functions around these definitions (Burrough, 1989). MacMillan and Shary (2009) characterized three types of knowledge-based models used to infer environmental conditions based on knowledge of environmental processes:

- Very-limited knowledge, where expertise is confined to an expert's ability to correctly identify specific instances or cases of a desired class or outcome,
- Partial knowledge, where an expert has a general idea of what the objects to be predicted look like, where they typically occur in space, and the main conditions, processes, or controls under which they typically develop,
- Exact knowledge, forms of expert knowledge considered to be complete and perfect.

Within the context of partial knowledge, information and data can be input into fuzzy semantic import (SI) models, based on heuristic rules. The fuzzy semantic import model permits formal recognition of imprecise and overlapping semantics used to classify data. This method is useful for scenarios where qualitative definitions of classes exist, but the definitions are inherently uncertain (Burrough et al,. 1992), while allowing incorporation of expert knowledge and judgment (MacMillan et al., 2000). Thus, a membership function can be constructed to represent a vague linguistic concept, or semantic construct, based on underlying quantitative information. For example, the semantic construct 'steepness' may hold a different meaning in the Central Plains than it does in the Rocky Mountains due to the difference in magnitude for slope steepness, but the steepest areas of either landscape would be represented by similar high membership values.

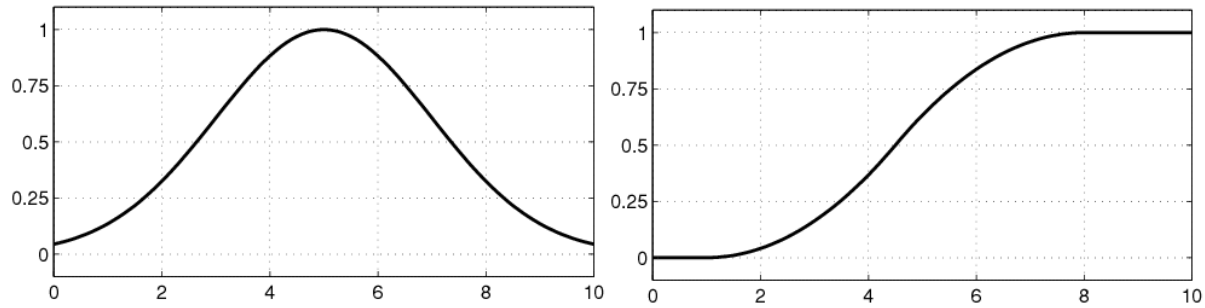
Based on the data distribution, salient criteria, or heuristic rules for each terrain attribute, a fuzzy SI membership function can be fit to the data to produce a fuzzy set, or semantic construct (Robinson, 2003). The SI model utilizes a fuzzy membership function to convert an attribute value from X_A to a fuzzy membership value U_A, and thus defines the degree of membership of a location to a central concept, prototype, or a semantic construct such as 'steepness' or 'highness' (Burrough et al., 1992). For more detailed overviews of the types of functions used in the SI Model, refer to Burrough (1989), Burrough et al. (1992), and Robinson (2003).

General examples of membership functions are the first order polynomials:

$$u_{(a)}x = \frac{1}{1+((x-c)/d)^2} \qquad (2.2)$$

$$u_{(a)}x = \frac{1}{1+((x/c)^d)} \qquad (2.3)$$

where $u_{(a)}x$ is the membership value or similarity of attribute (x) to the central concept, or standard index (b) and d is the dispersion index controlling the shape of the curve (Burrough, 1989; MacMillan and Shary, 2009). The dispersion index determines that the value at the crossover point is the point at which the grade of membership = .5, which is generally considered the lower-bound for membership to a fuzzy set. The central concept is the attribute value at which membership to a class is 100% (x=c), and receives a full membership value (1.0), and decreases from 1.0 on both sides of the central concept. Variables can be modeled as closed or open functions, with closed functions allow non-zero values only in a bounded interval, whereas open functions allow non-zero values along the entire 0-1 range (Robinson, 2003). Equations 2.2 is referred to as a Gaussian Model (Figure 8a), and Equation 2.3 is generally referred to as the S-curve (right, Figure 8b) (MATLAB, 2010).



Figures 8a and 8b. The Gaussian (Figure 8a, left, Equation 2.2) and S-Curve (Figure 8b, right, Equation 2.3) fuzzy membership functions (MATLAB, 2010) where the y-axis is the fuzzy membership value, and the x-axis represents the value of the attribute being fuzzified.

Equation 2.2 can be modified by additional conditions to convert it to an S-curve or a Z-curve (the left-shoulder asymmetric version of the S-curve), with Equation 2.4a representing the asymmetric right-shoulder function (S-curve), and Equation 2.4b representing the asymmetric left-shoulder function (Z-curve):

$$u_{(a)}x = \frac{1}{1+((x-c)/d)^2} \text{ for } 0 \le x \le c$$

$$u_{(a)}x = 1 \text{ for } x \ge c \qquad (2.4a)$$

$$u_{(a)}x = 1 \text{for } x \le c \qquad (2.4b)$$

2.3.5 Fuzz Overlay

Within multi-dimensional attribute space, it may be desirable to combine information from several fuzzy sets to produce a final fuzzy surface which represents the outcome of the interaction of multiple fuzzy membership values. Within geomorphometric landform classification, this involves an overlay of surfaces such as elevation, slope, curvature, or other surfaces to create a fuzzy layer representing a single semantic landform. These operations require the use of fuzzy operators, connectives, or aggregations. The main operators in Fuzzy

Set Theory are similar to those in Boolean logic as initially proposed by Zadeh (1965) and later expanded by Kaufman (1975) and Kandel (1986) comprising the intersect, union, and complement (Robinson, 2003).

The initial operations defined for fuzzy sets are union, intersection, and complement. For two fuzzy sets, A and B, the union of two fuzzy sets A and B is denoted by $A \bigcup B$:

$$A \bigcup B = \int_X U_A(x) \vee U_B(x)/x \quad (2.5)$$

where v is the symbol for maximum, and union corresponds to the logical connective OR. The symbol '/' should be interpreted as meaning "with respect to", and the integral sign symbolizes 'for the universe of objects x' (Kandel 1986). The intersection of fuzzy sets A and B is represented through the notation:

$$A \bigcap B = \int_X U_A(x) \wedge U_{(B)}(x)/x \quad (2.6)$$

where $\wedge$ denotes intersection, and the intersection command corresponds to the logical connective AND (minimum operator). Applied to ecology and pedology, this has its theoretical basis in the Limiting Factor Principle of Ecology (Qin et al., 2009). The complement of a fuzzy set A is defined by:

$$A = \int_X (1 - U_A(x))/x \quad (2.7)$$

which corresponds to the logical connective NOT. Because of the Boolean nature of the intersection and union operators in taking only the best/worst case characterizations (Robinson, 2003), the above operators cannot compensate for the degree of similarity or dissimilarity between the two sets being intersected or unioned. However, aggregation operators allow for more complex operations with more expert-driven input (Robinson, 2003).

The algebraic product, of two fuzzy sets, A and B, can be defined as

$$AB = \int_X (U_A(x) * U_B(x))/x \quad (2.8)$$

with the product of two fuzzy sets being interpreted as a soft AND. The algebraic mean of two fuzzy sets is referred to as the cartesian product if the fuzzy sets are unweighted, or the convex combination model if different weights (that sum to 1) are applied to each fuzzy set used in the algebraic mean (Kandel, 1986). Convex-combination is often used because of its abilities to integrate preferences for multiple criteria from multiple layers into the final fuzzy membership. The convex-combination Model is expressed as:

$$F = \sum_{i=1}^{m} W_i * U_i(x_{ij}) \quad (2.11)$$

where $Ui(xij)$ is the membership value of class i to the j^{th} location, and W_i is the weight assigned to layer i, where the sum of weights equals one, and $W_i > 0$ (Robinson, 2003). Thus, relative weights can be assigned based on the relative importance of each layer (MacMillan et al., 2000; Robinson, 2003). The use of convex combinations is especially helpful when linguistic modifiers are used. Oberth et al. (2000) used convex combination to combine sets of soil properties into soil quality indices. Burrough et al. (1992) utilized convex combinations applied to weights of six variables (water supply, nutrients, oxygen, erodibility, soil depth, and slope) to

create a land suitability map. The extensive study of McMillan et al. (2000) calculated fuzzy membership values for landform classification based on a weighted linear average of the individual membership functions. In this case, attributes such as slope and curvature were given more weight because of their use in the Pennock's hillslope model (1987), which the study incorporated.

2.3.6 Measures of Membership and Confusion

Before creating defuzzified or crisp classes, several additional measures of classification variance and confusion are useful to determine areas of ambiguity between the classes. The most commonly used measures are the confusion index and normalized class entropy, which measure class overlap in attribute space and variability of classes. With multiple surfaces representing membership values to different classes, there will be locations where the fuzzy membership value of one class will dominate the other and vice-versa, and others where fuzzy membership values of two classes are very similar. Measurements of class overlap may be used to adjust parameters in the original membership function or logical operators used in fuzzy overlay. Class-specific relationships may be established based on these measures as well.

The Confusion Index (CI) is a measure of class similarity and how well each location has been classified. The CI is calculated as:

$$CI_j = u_{j(max-1)} / u_{j(max)} \quad (2.12)$$

where $u_{j(max)}$ is the fuzzy membership value of the class with maximum membership value at location j, and $u_{j(max-a)}$ is the fuzzy membership value of the class with the second-largest membership at location j (Burrough et al., 1997). If CI is close to zero, one class dominates the

location. As CI edges closer to one, there is a higher degree of similarity between two classes, and hence more confusion. Thus, areas with low confusion indices represent more crisp zones with low uncertainty, and areas with high confusion indices represent more fuzzy zones with higher uncertainty. A Confusion Index created from the previous 4-class Fuzzy K-means landform classification is shown in Figure 9.

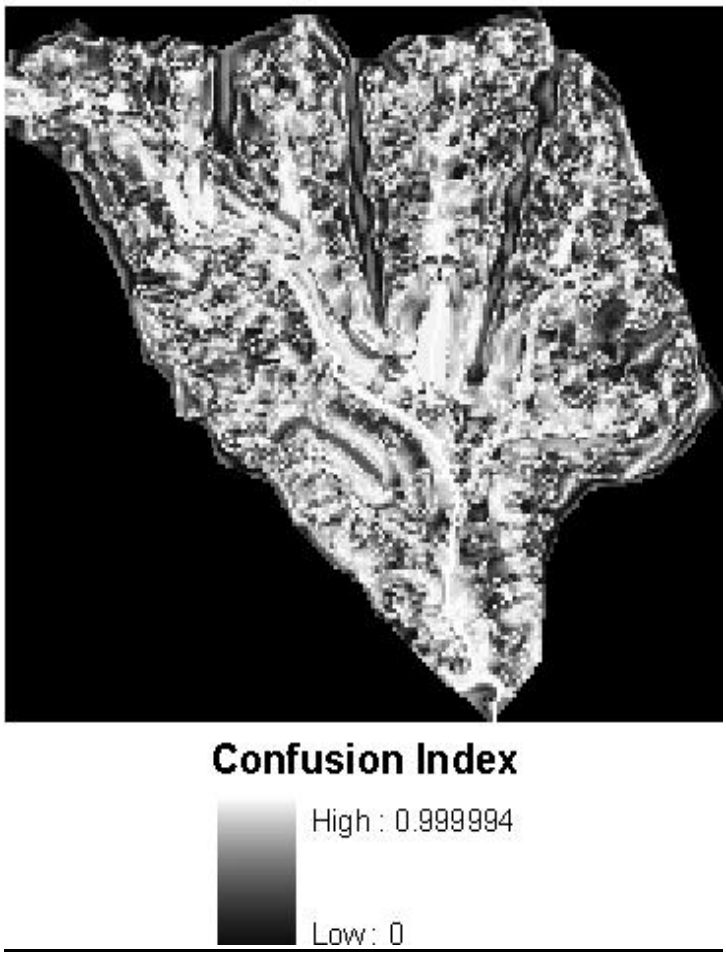


Figure 9. Confusion Index of four-class Landform system derived from Fuzzy K-means clustering. Lighter locations (lowlands and drainages) represent areas of higher confusion, while darker areas (ridges) represent areas of lower confusion.

When considering the variance of the classification or the ambiguity of a spatial location with regards to the diversity of classes and their membership values, classification entropy is a key measurement, and is derived from the Shannon-Weiner Diversity Index. Like fuzzy

membership values or the confusion index, entropy can be standardized between 0 (fuzzy membership value for one class dominates) and 1 (fuzzy membership values fairly equal between all classes) (Wood, 1996). Classification Entropy for a single scale is measured as:

$$H_j = -\sum_{i=1}^{n} u_{ij} * \log[u_{ij}] \qquad (2.13)$$

$$H_{max} = -\log\left(\frac{1}{n}\right) \qquad (2.14)$$

Where H_j is the entropy value at location *j*, u_{Aj} is the standardized fuzzy membership value for class type *i* at location *j*, and *n* is the number of classes (Fisher et al., 2004). Entropy can thus be standardized on a 0 to 1 scale by dividing Classification Entropy by Standardized Entropy (H_j/H_{max}) (Wood, 1996). This is also referred to as the Classification Entropy (H) (Burrough et al., 2001).

The defuzzification of a single fuzzy membership value to a crisp sets is referred to as an alpha-cut (α-cut). The alpha cut is a fuzzy membership value that is used as a threshold to delineate contours of equal value, where locations that are equal to or greater than the alpha-cut value as members of the set (Boolean '1'), while those outside the threshold are coded as non-members (Boolean '0') (Fisher et al., 2004). Alpha-cuts are often used with Fuzzy k-means clustering when there is no expert-knowledge in determining the distribution of fuzzy membership values.

2.4 Incorporating Scale Effects in Geomorphometry

The most important aspect of terrain analysis based on a finite-resolution DEM is the scale-dependency of the relationship under investigation (Hengl and Evans, 2009). The continuous

nature of the Earth's surface contains patterns may occur across a full range of sizes and scales. The extent of the surface that is sampled is important because the land surface is a non-stationary variable (Evans and Cox 1999). The fidelity with which the DEM models the true surface will depend on the surface roughness and DEM resolution (Wood, 1996). Some of the key problems in modeling environmental systems are the effects of scale and accuracy on topographic attributes and the incompatibilities between scales at which many physically-based models operate. Moore et al. (1993) listed four major scale related issues:

- Object (landform) size in which homogeneity is assumed,
- The method of analysis used to derive the terrain attribute values,
- Aggregation of data derived from different resolution, accuracies, and structures,
- Scale differences between model process representation and data available for model

The first point ties into the vagueness of morphometric features. Within quantitative geomorphometry, semantic vagueness is derived from both a lack of well-defined boundary conditions, and epistemic vagueness in which the class to which an object belongs varies with perception (Figure 10) (Fisher et al., 2004). Issues with boundary conditions are related to the 'double-crisp' soil-landscape model mentioned earlier, and can be accounted for using fuzzy set theory and geostatistics. The second form of uncertainty is due primarily to the scale of geographic measurement under which the object, or landform, is derived. Many applied classifications of landforms are specific to a particular scale, and tend to treat variation in the land surface that occurs below a given scale as random noise (MacMillan and Shary, 2009)

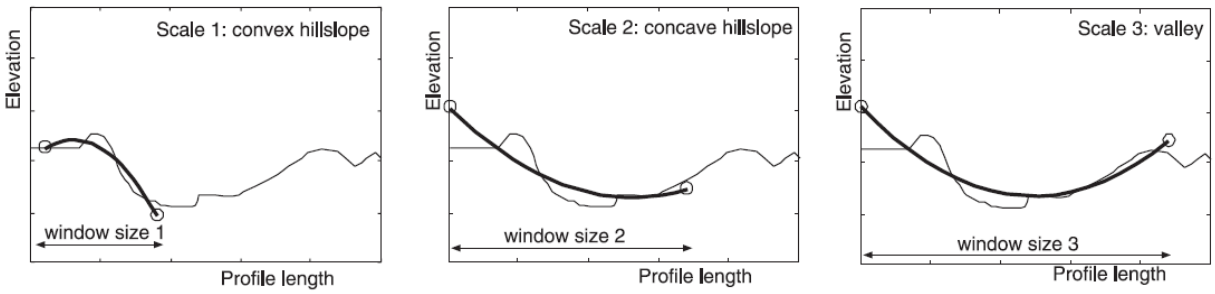


Figure 10. Demonstration of how measurements at various scales can change surface characteristics (Fisher et al. 2004).

The limitations for any grid-based representation of the surface are constrained by the resolution of the DEM and the accuracy of its interpolation (Wood, 1996). Most acknowledged problems in elevation interpolation come from noise at a finer scale than the DEM or variations at a coarser scale than the sampling method. Parameters like gradient and curvature are heavily-dependent on the scale inherent in the attribute calculation technique (Schmidt and Andrews, 2005), and noise amplifies variation and errors in the calculation of surface derivatives. The scale implied by the resolution of the DEM is not always optimal for landscape characterization, and many of the classical approaches to derivative calculation are based on fixed-scale approaches, as represented by the use of Horn's (1981) and Zevenbergen and Thorne's (1987) algorithms to calculate slope, aspect, and surface curvature within moving windows of fixed window size (3 x 3) in ESRI ArcGis toolbox and thus only represent the surface over one window size, and are unable to represent surfaces derived from multi-scale measurements (Figure 11).

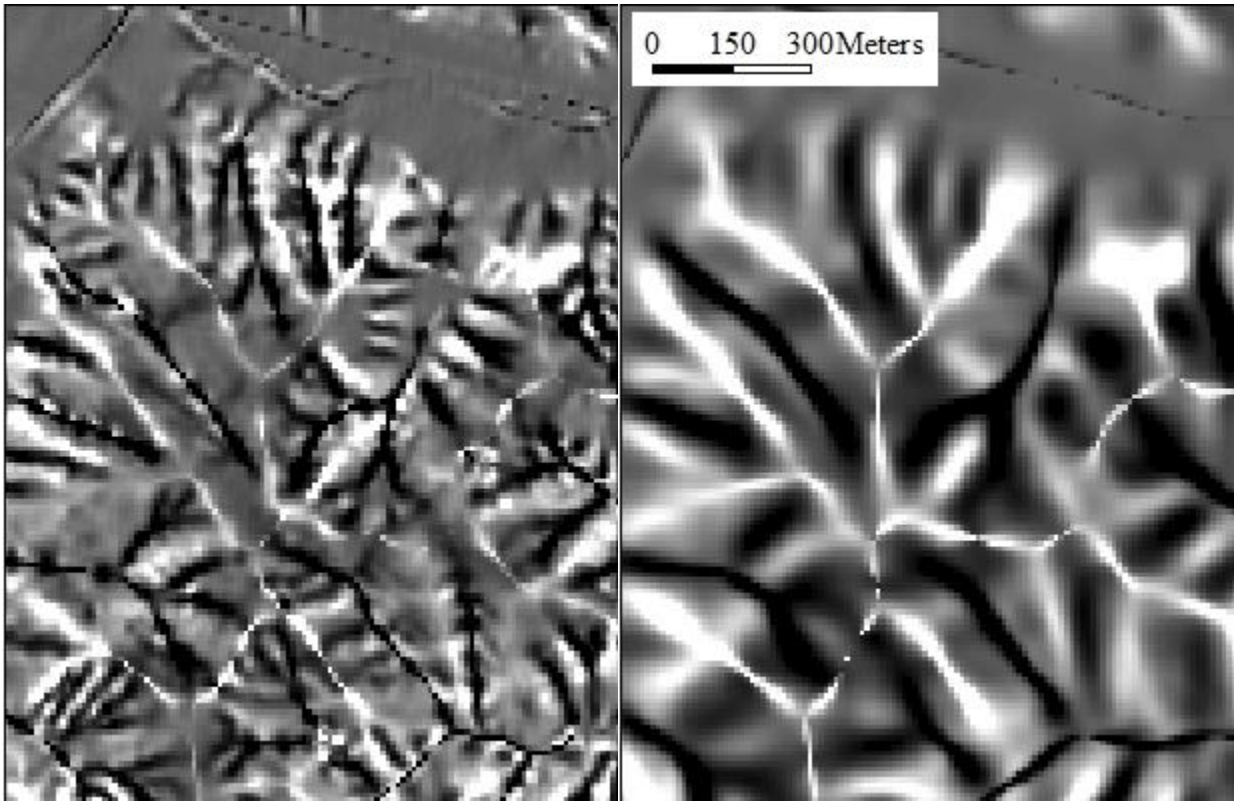


Figure 11. Maps of Tangential Curvature calculated with a 3x3 window (left) and 13x13 window (right); zones of flow convergence are darker in color, and flow divergence lighter

The question then becomes what scale and resolution are appropriate to model the surface? Moore et al. (1993) and Zhang and Montgomery (1994) both argue that the appropriate scale for such modeling processes is somewhat finer than the hillslope scale identifiable in the field. For modeling the surface with direct field applications, there is a consensus that horizontal resolution of 5-10 meters and a vertical resolution of .5 meters is the necessary minimal resolution for describing the surface form for individual hillslopes.

Since landforms are a multi-scale feature, and hillslope position a multi-scale phenomenon (Weiss, 2001), the designation of a point to a single morphometric class can change over different scales, as well as carry more than one piece of information at one scale about biophysical processes and landscape position. Thus there is a need to incorporate techniques that allows computed terrain attributes to show scale-specific variations, and enable a point to be

classified over multiple scales to allow consideration of both its spatial context and uncertainty (Moller et al., 2008).

CHAPTER III

LITERATURE REVIEW

3.1 Introduction

This section covers previous work relevant to the incorporation of scale and fuzzy set theory into Geomorphometry, specifically those studies dealing with landform classification. The problems of scale, spatial extent, and resolution make single objective classifications of landscape unfeasible (Wood, 1996). Landforms tend to be classified either according to homogonous regions (Speight, 1968; Dikau, 1989), or specific features (Werner, 1988; Tribe, 1990; Dietrich and Montgomery, 1989). No single object classification of the land surface exists, and local form, size, neighborhood calculations, surface roughness, and geostatistics have all been used to characterize landforms (Schmidt and Hewitt, 2004). Minar and Evans (2008) recognized three axioms pertinent to land surface characterization:

- Land surface form can be analyzed as a continuum,
- At a given scale, the land surface may exhibit discontinuities,
- These discontinuities result from morphological processes, the majority of which are gravitation work,

In addition to vagueness by definition, a vital problem in defining a semantic term such as landscape or hillslope position is how to spatially represent both local uncertainty and uncertainty across scales. Representing these uncertainties in a quantitative way plays a central role in the understanding of topographic phenomena. Traditional approaches to landform

taxonomy lack both the integration of landform representation over multiple scales and have a tendency to treat semantic classes as double-crisp Boolean boundaries entities.

With each location in a continuous field, a degree of association can be assigned to each class in a set of rules-based classes based on terrain attributes. Since the fuzzy character of real terrain is a result of past cross-scale geomorphic processes, incorporation of multiple membership values may reveal a greater scope of knowledge into potential landscape evolution and development than a traditional crisp system.

3.2 Landform Taxonomy

The modern study of landforms can be traced back to surface and hillslope models presented by Wood (1942), Troeh (1964) Dalrymple et al. (1968), and Ruhe (1969). Since the introduction of digital elevation models in the late 1960s, terrain analysis has diverged into several distinct areas of research (MacMillan et al., 2004). The first area studies the general form of the surface through signs of curvature, drawing from Ruhe's hillslope model (1969) and the six-unit classification of Peucker and Douglas (1974), an early automated landform classification model that divided a DEM into pits, peaks, passes, ridges, flats, and channels based on neighborhood relationships This method was utilized among others by Dikau (1989), Wood (1996), and Schmidt and Hewitt (2004), and the derived classification systems are generally referred to as geometric or morphometric units.

A second area of research is the utilization of surface shape and slope gradient in attempts to quantify relative landform position and interrelate landform position and geomorphic process. These systems attempt to describe taxonomic units, or topographic position, along the continuum

of a single, scale-less hillslope. Pennock et al. (1987), Speight (1990), Zhu (1997), MacMillan et al. (2000), Qin et al. (2009) all utilize such a system that draws from theoretical models such as Wood (1942) Dalympre et al. (1968), and Ruhe (1969). These systems are referred to as geomorphometric units.

The third area of research utilizes explicit calculations such as flow accumulation and flow direction to extract hydrologic objects from a DEM such as stream segments, catchments, and links between them, including work from Mark (1984), Band (1986), Jensen and Dominguez (1988), Dietrich and Montgomery (1989), Quinn (1991), and Tarboton (1997). Many of the ideas are converging towards Sombroek's terrain elements (FAO, 1995), which contended that hydrologic and land use models should be integrated. This review of previous work will focus on the first two groups of research as they are the most commonly used and the most applicable to this study.

3.2.1 Morphometric Units

Taxomony of morphometric units relies on combinations of surface curvature and slope and ignores larger landscape context, landform position, or geostatistical relationships. Peucker and Douglas (1974) introduced one of the first automated methods to detect spurious landforms (peak, pass, pit, plane, channel, and ridge) on the basis of a rules-based detection utilizing positive and negative arrays of elevation differences between a location and its neighbors (Figure 12). Toriwaki and Fukumura (1978) used a similar theory to divide a grayscale image on the basis of connectivity relations between points and structurally thinned connected components into regions. As previously mentioned in Section 2.2.3, Dikau (1989) created a 15-class

landform system to identify form elements of convex, planar, and concave curvatures for profile and tangential curvature for sloping elements, and six combinations of maximum and minimum curvature for non-sloping elements, since curvature is undefined for areas with no slope.

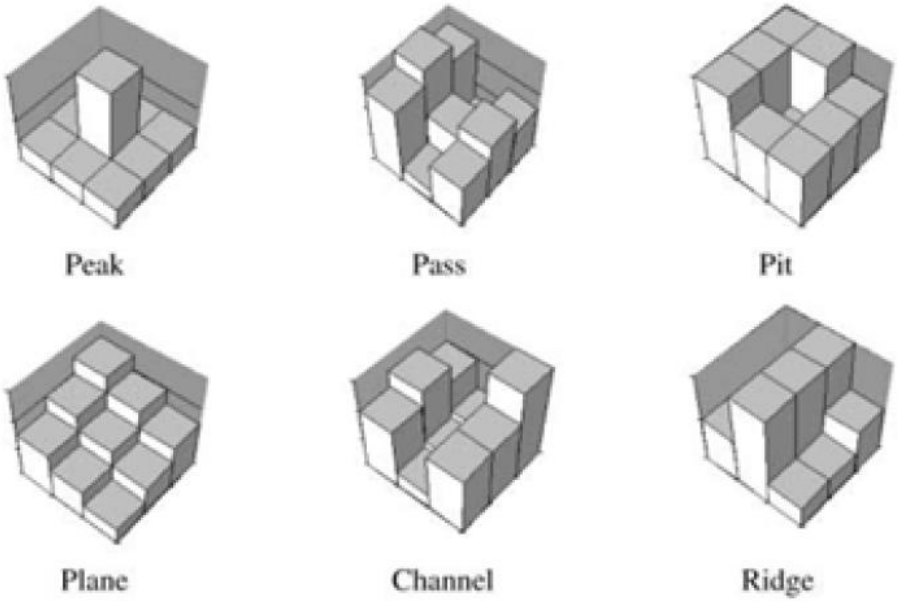


Figure 12. Landform classes developed by Peucker and Douglas (1974).

Using Peucker and Douglas's classes, Wood (1996) derived a multi-scale landform characterization utilizing mean, minimum cross-sectional, longitudinal curvature, and a slope threshold to differentiate the six landforms. More recent work from Shary (2002) is a 12-unit slope model utilizing a combination of Gaussian, tangential, differential, mean, and profile curvature. Several authors have taken the topological relationship between the line features from the Douglas-Peucker model (channel, ridge) in attempts to formulate consistent spatial relationships between them (Werner, 1988; Brandi, 1996). This approach has mainly been a theoretical one, as the lack of relative landscape position or spatial relationships between the units makes them somewhat ambiguous

3.2.2 Geomorphometric Units

Besides identifying landform elements on the basis of surface form, there have been efforts to describe the relative topographic position (upland, midland, lowland), a loosely defined variable that is inherently a multi-scale phenomenon (Weiss 2001) and is often used without any quantitative criteria to discriminate between classes (Coops et al., 1998). Field observations describing topographic position may only capture information relevant at a single scale and ignore finer or coarser scale observations that could have significant control on local process and form (Wood, 1996). The advantages of geomorphometric systems is that they can be combined with relevant physiographic, and geologic models to relate the synergistic relationships between topographic position, abundance and type of flora and fauna, pedological characteristics, and landscape evolution (Figure 13).

Primarily these efforts describe the relative position (upland, midland, lowland) by using additional measures and attributes in addition to slope and curvature to describe surface form. Often single measures such as the Relative Position Index (Skidmore, 1990), Topographic Position Index (Weiss, 2001) and Maximal Dispersion Area (Shary, 2005) may be too dependent on user-defined parameters, meaningful to a specific area or dataset, and too abstract to be used as a generic measure of landform position.

Many existing Geomorphometric taxonomies incorporate theoretical hillslope models (Wood, 1942; Dalympre et al., 1968) into their descriptions to supplement their description of surface form (Figures 14a-b). Examples of these taxonomies are well represented in Pennock et al. (1987), Speight (1990), and Zhu (1997). Pennock et al. (1987) represented different slopes (shoulder slopes, back slopes, footslopes) through hard ranges of plan and profile curvatures and

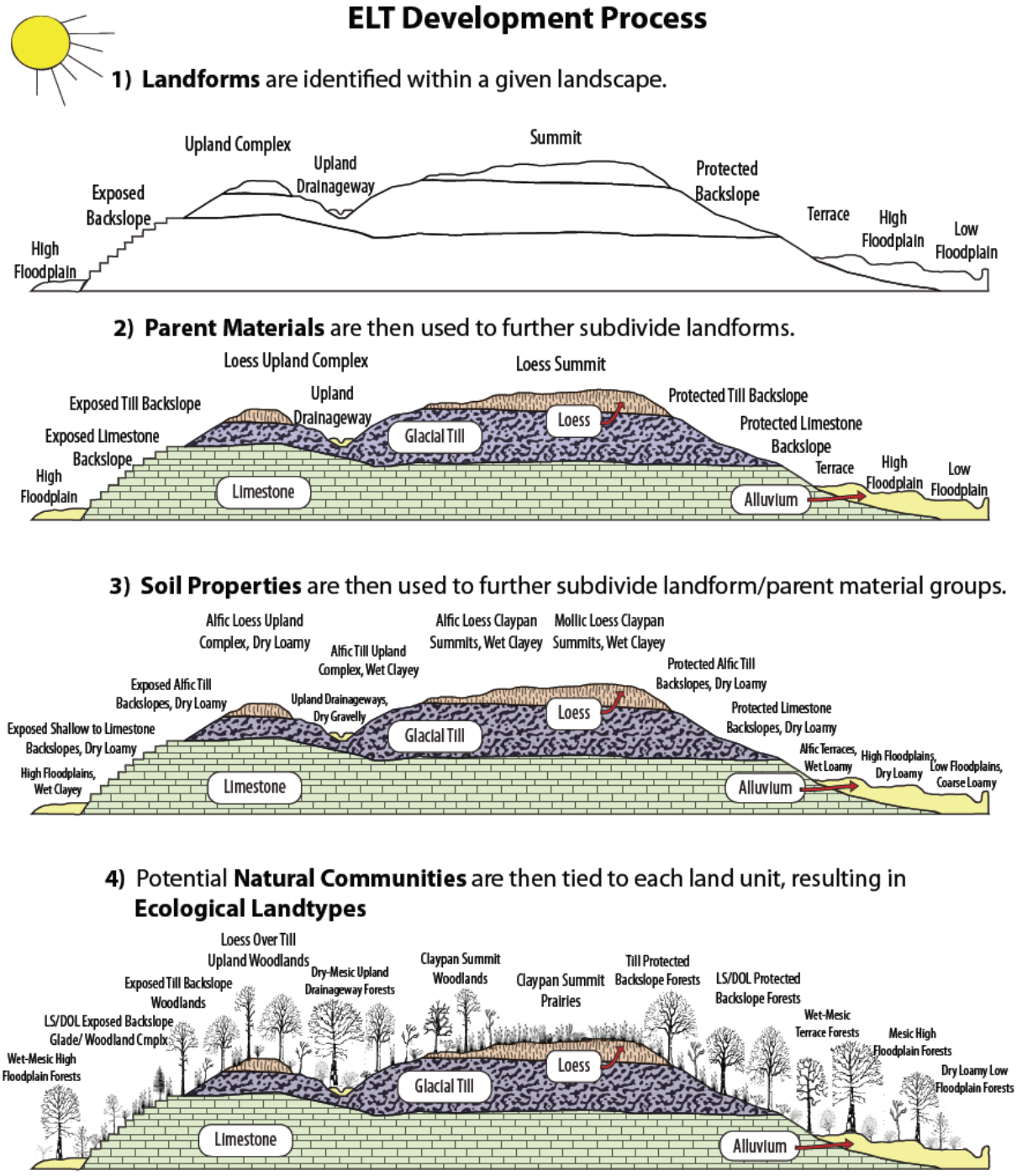


Figure 13. An example of how topographic position (geomorphometric units) provides the framework for detailing more complex ecological relationships from the Missouri Ecological Landtypes (ELT) model (Nigh et al., unpublished)

slope gradient in hummocky terrain in Sasketchawen (Figure 15). One of the more formal systems was developed by Speight (1990), based on previous experience with multi-scale landform mapping and aerial photograph interpretation (Speight 1968; 1974; 1976).

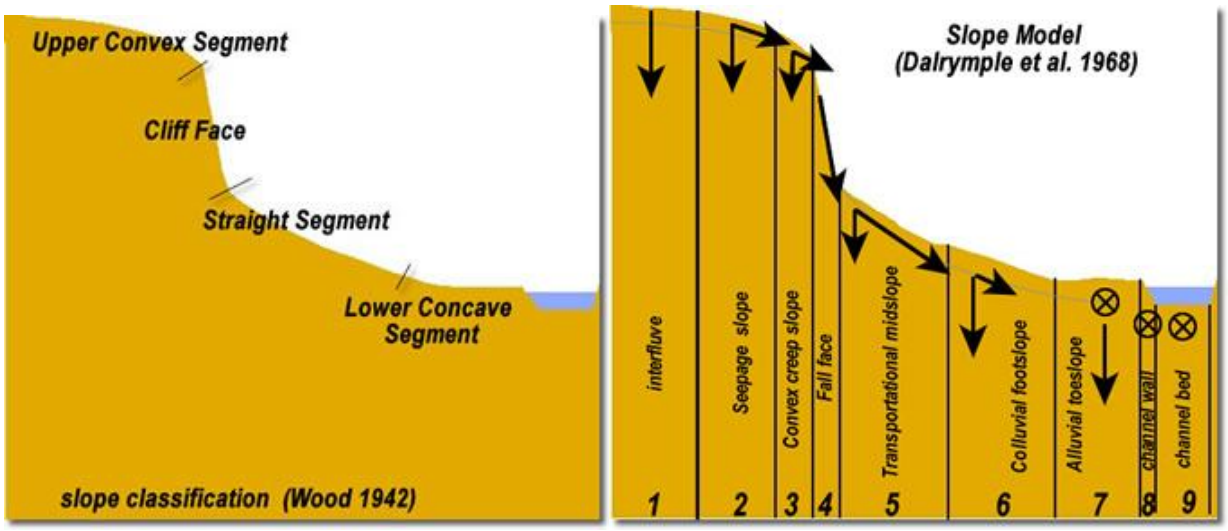


Figure 14a-b. Wood's (left) and Dalrymple et al.'s hillslope models (right)

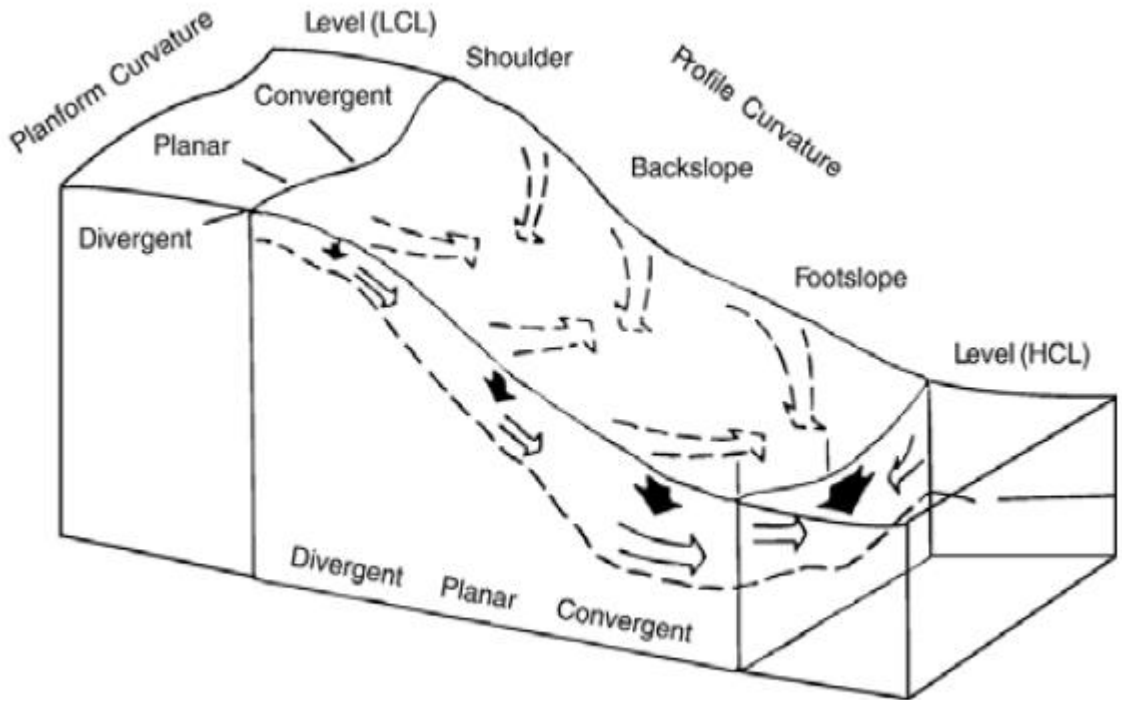


Figure 15. Pennock et al.'s Hillslope Model (1987)

Speight's system has been incorporated into several automated GIS classifications (Coops et al., 1998; klingseisen et al., 2008), and relies on measures of elevation percentile, object adjacency, size thresholds in addition to slope and curvature to characterize landforms and landform position. Speight also incorporated an organizational hierarchy to his system, discussed later.

An area of continued developments is the delineation of slope units and slope break lines, as break lines represent topological links between adjacent landforms. Slope units are defined as "a section of two-dimensional downslope profile having relatively homogeneous form, process, and lithology with upper and lower boundaries located at breaks of slope" (Giles and Franklin, 1998). Klingseisen et al. (2008) use a slope break generation algorithm derived from Giles and Franklin (1998) for locating and segmenting breakpoints within a DEM. More recently, Jordan (2003) suggested the use of hypsometry by identifying breakpoints as class breaks in the area/slope curve to avoid errors associated with the calculations second-order derivatives in identifying slope breaks in the DEM. The use of surface curvatures remains the most popular method of delineating slope units within a gridded DEM (MacMillan et al., 2000; Schmidt and Hewitt, 2004; Qin et al., 2009).

Qin et al. (2009) incorporated ideas from the SOILIM System introduced by Zhu (1997) for representing spatial uncertainty in slope breaks. The system is characterized by five slope units, ridge or summit, slope shoulder, backslope, footslope, and valley. Qin et al. (2009) took the procedure to a finer scale and decomposed the midland section of shoulder slope, backslope, and footslope into 12 finer classes based on Pennock et al. (1987). Central concepts of each landform class are defined for each terrain attribute: elevation, slope, profile curvature, and Relative Position Index (RPI) (Skidmore, 1990). A fuzzy semantic import model for each terrain attribute is then used to convert terrain attributes to fuzzy membership values through a semantic

import model. Thus each terrain attribute, profile curvature for example, is composed of three semantic constructs, profile convexity, planarity, and concavity. The differentiation between slope positions is accomplished through distinguishing between convex, planar, and concave signatures of profile curvature, and selecting heuristic thresholds for both RPI and slope, each using a Fuzzy Membership Function similar to those in Burrough (1989) and Burrough et al. (1992). The minimum operator was applied to designate final membership to a slope position.

3.3 Incorporation of Scale into Landform Taxonomy

Within landform detection in a GIS-environment, the classifications are commonly specific to a particular scale or a narrow range of scales. Scale is accounted for in two ways in landform taxonomy; hierarchical theory integrated into previous morphometric classification to suggest appropriate scales for classifying objects such as nested classification systems (Dikau, 1989; 1990), and explicit methods to calculate surface parameters over a range of scales for a single landform taxonomy (Wood, 1996; Schmidt and Hewitt, 2004).

Dikau (1989; 1990) represented hierarchical units within landform taxonomy, differentiating between hierarchies of time, size, and form, where form facets represent a unit of relatively homogeneous shape nested within a more complicated form element. While slope and curvature may be used to differentiate form at the hillslope scale, attributes such as grain, topographic roughness and variability, ridge density, or bifurcation ratio may be needed to differentiate landform pattern at coarser scales. Different proposed hierarchies are shown in Table 2.

Table 2. Hierarchy in Landform Taxonomy, suggested data sources, and scales of analysis (Dikau, 1989; Dehn et al., 2001; MacMillan et al., 2004)

Scale of Analysis	DEM Resolution	Proposed Unit Name
1:2,000 to 1:5,000	1 x 1 m	Undefined
1:5,000 to 1:10,000	5 x 5 m	Landform Element
1:10,000 to 1:50,000	10 x 10 m	Landform Type
1:50,000 to 1:125,000	25 x 25 m	Undefined
1:125,000 to 1:250,000	100 x 100 m	Physiographic System
1:250,000 to 1:1,000,000	500 x 500 m	Physiographic District
1:1,000,000 to 1:5,000,000	1 x 1 km	Physiographic Region
1:5,000,000 to 1:10,000,000	9 x 9 km	Physiographic Province

A recent approach composes Dikau's original 15 form facets (Section 2.2.3) recognizing eight distinct terrain units (ridge, shoulder, backslope, hollow, spur, terrace, footslope, and valley bottom), then uses a TOPHAT function (Rodriguez, 2002), similar to Weiss's Topographic Position Index (2001) to classify those eight elements into uplands (hill), midland (hillslope), and lowland (valley) elements based on local elevation roughness.

Hierarchy in Landform Taxonomy is also represented in Speight's system, composed of two-level system of landform patterns and landform elements representing the idea of singular landforms. Landforms are a mosaic of tiles, the larger tiles being landform patterns in the order of 300 meters in radius. Landform patterns can be described by topographic roughness, relief, stream occurrence, and drainage pattern. The landform patterns are represented by a series of smaller landform elements approximately 20 meters in radius and are described by slope,

topographic position, morphological type, and geomorphological agent. There are nine elements in Speight's system, although several are rarely used. Ridge, crests, and hillocks comprise the upland elements, and simple, upper, mid, and lower slopes comprising the sloped elements. Two additional elements, flats and depressions may occur at any part of the landscape. Flats occurring on ridge-tops and valleys may be separated by a slope or elevation percentile threshold. Methods based on Speight's system initially identified crests, flats, and depressions before slopes and created a sequence of precedence rules where:

Crests > depressions > flats > slopes

A separate slope classifier command built on pre-existing algorithms was used for breaking the slopes into upper, middle, and lower slopes based on their profile curvature. Versions of Speight's models have incorporated categorical similarity (Haagen, 2004) as a way of representing the uncertainty and fuzziness of landform boundaries (Klingseisen et al., 2008), but don't account for specific attribute similarity.

The second method to incorporate scale is to calculate surface parameters across a range of different scales This method was originated by Wood (1996), who used the multi-scale measurements to extract Peucker and Douglas's six landform classes from a DEM over multiple scales by signs of secondary curvatures (cross-sectional, maximum, minimum curvature, and a slope threshold) through a set of rule-based procedures. This was later expanded to include fuzzy set theory, with fuzzy membership values for each class at a location equaling the percentage that a location was classified as a class type divided by the total number of multi-scale interpolations. Final membership was determined by the modal membership value of all observations. Many studies have shown that predictive values of profile curvature for estimating

soil properties such as moisture or solum depth in a regression model vary with scale (Schmidt and Hewitt, 2004; Smith et al., 2006).

In landform taxonomies based on both morphometric and geomorphometric systems, it is inherent that some of the classes are either more crisp, or more fuzzy and ambiguous than others depending on variables being modeled, or fuzzy conditions being set. If a single semantic landform class displays high ambiguity in measures of classification stability (confusion index or classification entropy) across different scales and different fuzzy operators, it can be considered an unstable class. Thus with an unstable class, there should be a larger amount of errors of commission and a higher potential to misclassify several different classes to this single, unstable, ambiguous class. Such errors have critical implications to the accuracy of soil and vegetation mapping, and operational assumptions about a location's role in an ecosystem, and geomorphological and biological agents at operation, especially at scales related to field or parcel-specific agronomic and natural resource management decisions.

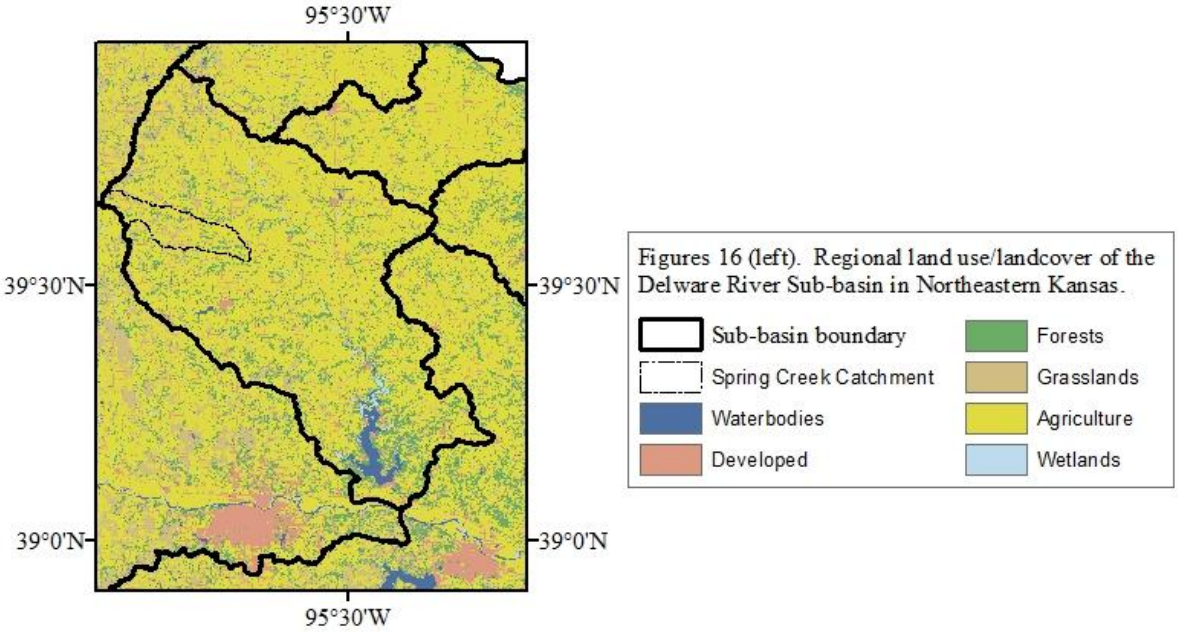
CHAPTER IV

METHODS

4.1 Description of Study Area

The dataset of CRP enrollments used for the study spans the upper portion of the Delaware River Basin (Hydrologic Unit Code 10270103) in Northeast Kansas (Figures 16). For computing and time constraints Spring Creek, a smaller tributary within Jackson and Nemaha counties, was selected as the study area due to a fair amount of CRP enrollments within the watershed, and a prevalence of steeper slopes (Figure 17). The area is part of the Glacial Drift and Loess Hills physiographic region, a dissected glacial drift plain covered by a thick layer of loess, with soils that are high in clay content. Local relief can reach up to 200 feet, with stream valleys that are narrow with steep banks in some places, and the ridgetops can vary from level to strongly sloping over a short distance (USDA, 2006).

The Catchment covers approximately 153 km^2 in area. The area has a mean elevation of 361 meters A.S.L., ranging from 298 to 428 meters. Soils are deep with high clay content and moderately to highly erodible (USDA, 2006). 77% of the area are soil capability classes 1-3 (slight to severe limitations for agriculture) and the remainder soil capability classes 4-8 (not suitable for agriculture) (NRCS, 2008). Mean annual precipitation for the area ranges from 73-99 cm (USGS, 1995).



Figures 16 (left). Regional land use/landcover of the Delware River Sub-basin in Northeastern Kansas.

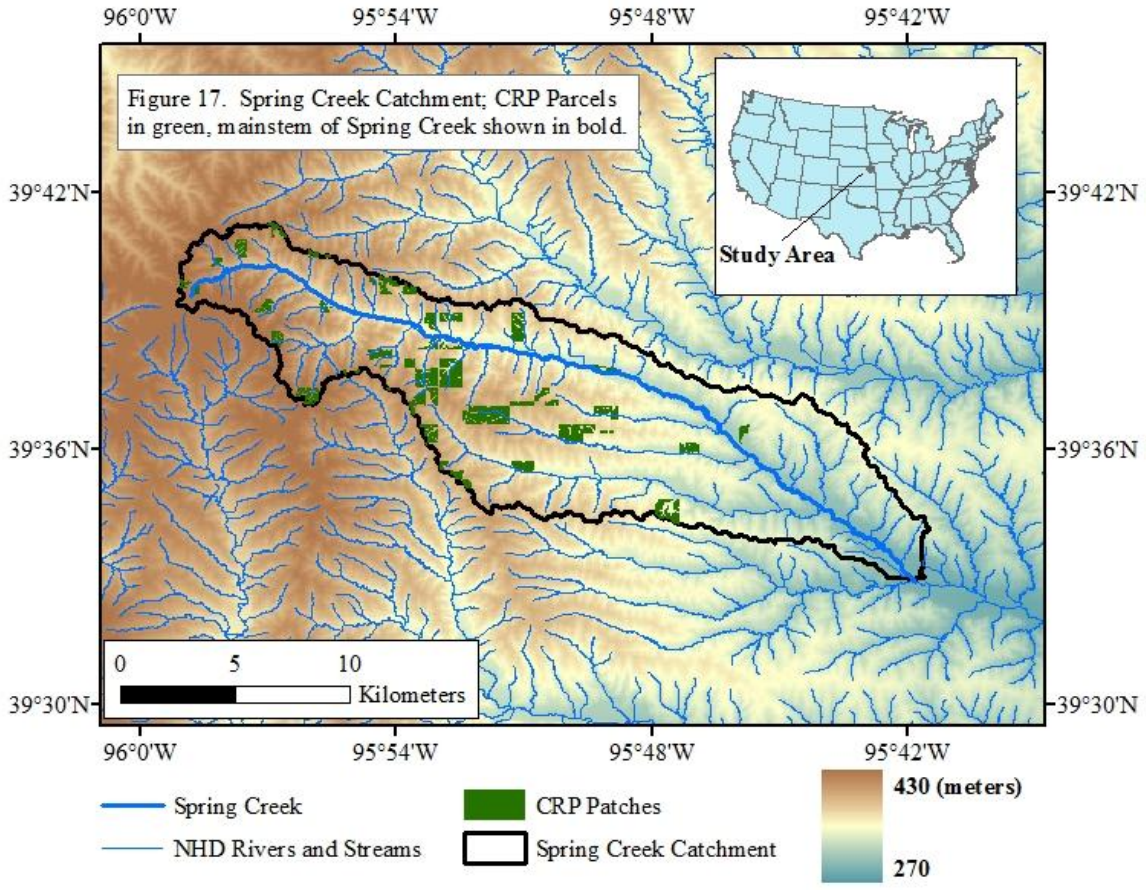


Figure 17. Spring Creek Catchment; CRP Parcels in green, mainstem of Spring Creek shown in bold.

Pre-settlement vegetation for the area was tallgrass prairie, Oak-Hickory Savannahs in the uplands, and Cottonwood, Maple, and other Riparian species along the riverways and valley bottoms (USDA., 2006). Agriculture is common on the flatter loess hills and bottoms, while rangeland is supported on many of the deep clay loams. The U.S.D.A. (2006) lists the specific resource concerns for the Delaware River Basin pertaining to soil erodibility:

- Classic gully and streambank erosion in pasture and hayland are concerns due to overutilization of the plant resource and unabated livestock access to stream corridors and banks,
- Residue, nutrient, and pest management; vegetative practices, and structural practices are necessary to control erosion, protect water quality, and improve soil conditions,
- For cropland, sheet and rill erosion is greater on steeper slopes,
- Over application of nutrients and organics has created surface water quality concerns.

Many of the surface water bodies in the area have fair to poor surface water quality conditions, with 52% of the waterbodies within the Delaware River sub-basin not meeting their designated uses under the Clean Water Act (USDA, 2006). The primary pollutants for streams within the watershed is Fecal Coliform Bacteria and pollutants of concern in the lakes and ponds include ammonia (NH_3) that can cause eutrophication, and atrazine, a broad applicant chemical herbicide used in agriculture. Atrazine's moderate solubility and slow rate of degradation make it highly prone to transport from surface runoff (Stamer, 1995). This is notable as the Delaware River Basin was one of the first areas in the nation where Atrazine was noticed as a major pollutant of surface waters. As a result, the basin became the first pesticide management area in 1994. The primary source of surface water pollutants within the Delaware Basin is row crop agriculture.

4.2 Data Preparation

A polygon feature class of conservation reserve program tracts was attained from the USGS Fort Collins Data Center (http://www.fort.usgs.gov/about/). The layer has parcels from the first twelve enrollment periods of the program spanning eight years (1986-1993). The data also included ancillary data for enrollment year, cover type, previous crop plantings, farm information, and common land unit information. The data was hand-digitized from NRCS County Parcel/Common Land Unit Maps at 1:24,000. Spring Creek Catchment, which spans parts of Jackson and Nemaha Counties, contains the entirety or portions of 139 parcels, encompassing a total area of 10.12 km^2, an average area of .072 km^2 per enrollment parcel, approximately 7% of the total area.

10-meter digital elevation models (DEMs) were downloaded from the USGS Seamless data server (http://seamless.usgs.gov/). The DEM covered the 7.5-minute quadrangles Goff, Soldier, Wetmore, Circleville, Horton Northwest, and Netawaka, Kansas. The DEMs are Level II, created through complex linear interpolation and have been processed or smoothed for consistency and edited to remove identifiable systematic errors that may have been derived from hypsographic and hydrographic data digitizing, such as contour line bias or striping. This procedure produces a maximum root mean square error (RMSE) of ½ a contour, or three meters for height errors in the vertical plane (McMasters, 2002).

After the DEMs were tiled, the ArcGIS Hydrology Toolset (ESRI, 2008) was used to extract the basin boundary. Sinks within the DEM were filled to create continuous flow accumulation and flow direction grids based on the D8 flow direction algorithm and steepest slope descent through FLOW DIRECTION and FLOW ACCUMULATION commands. With

the exception of recently glaciated or karst landscapes, pits in an elevation surface may be considered data errors or local noise (Band, 1986). The filled DEM was only used to extract the basin boundary and not on surface characterization. A point feature class was created, representing the basin pour point or seed point, and located to extract the watershed boundary using the WATERSHED tool.

Many algorithms employed in terrain analysis operate on the basis of the focal neighborhood, an odd-numbered N x N dimension of grid cells whose expression is then transferred to the focal cell at the center. Cells near the edge may have null values ('NaN' or '-9999'), which in some algorithms may be inadvertently incorporated as part of their focal neighborhood due to the size of the moving window, known as edge contamination (Olaya 2009). To account for this, the watershed boundary grid was converted to a polygon without the simplification of the polygon boundary preserving the pixilated edge. The polygon was buffered by 200 meters to account for the edge effects beyond the largest anticipated window size (130 meters by 130 meters). This allows the source cells that define the headwaters or 'walls' of the watershed to be properly represented within their landscape context. The buffered polygon was then used as a mask to extract the basin elevation grid.

A large source of error when calculating higher order elevation derivatives is the sensitivity to DEM noise. Sensitivity is greatest in functions of second derivatives (curvature), and least sensitive in first derivatives (slope, aspect). DEM pre-processing is a recommended procedure to remove gross errors and artifacts and make a better approximation of the land surface (Reuter et. al., 2009). Shary et al. (2002) suggested a parametric isotropic smoothing operation before calculating derivatives and terrain attributes off a DEM surface to reduce the noise and non-systematic errors within the DEM. Statistical properties of point attributes are more stable if the

DEM is smoothed, while reducing noise and discrepancies between various gradient calculation algorithms (Evans, 1972). The initial isotropic smoothing algorithm for a 3 x 3 neighborhood is:

z1	z2	z3
z4	z5	z6
z7	z8	z9

$$z5^* = h * \frac{z2 + z4 + z6 + z8}{9} + \left[1 - h * \frac{4}{9}\right] * z5 \qquad (3.1)$$

Where z1-z9 represent the 9 elevation points of a local 3 x 3 neighborhood, z5* is the filtered elevation value at the focal cell of the neighborhood, z2,z4, z5,z6,z8 represent local elevation values from the local neighborhood, and h is the smoothing factor, which used the value of .5 as recommended by Shary et al. (2002). After simplification, the algorithm breaks down to a simpler form:

$$z5^* = \frac{(41 * z5) + z2 + z4 + z6 + z8}{45} \qquad (3.2)$$

As recommended for the appropriate resolution, the isotropic smoothing algorithm was iterated twice over the DEM before surface characterization was performed as suggested for interpolation type and resolution (Wilson and Gallant, 2000).

The simplified workflow for the methodology is shown in Figure 18. Conceptual definitions of the landform classes, semantic constructs (*highness, steepness,* etc.), and appropriate terrain attributes to describe the constructs were defined for each class were created based on relevant geomorphometric models (Figure 18, Step I). Terrain attributes *(TA(x)))* relevant to the classes and semantic constructs were calculated from the DEM (Step II). Fuzzy membership functions were used to convert terrain attributes *(TA(x)))* to fuzzy membership values ($u_{(a)}x$) on a 0 (non-membership) to 1 (full membership) scale (Step III) to express the semantic constructs. Fuzzy overlay was performed on the semantic constructs to derive fuzzy landform layers for each semantic landform class (Step IV).

For each set of fuzzy landform classes derived, the maximum values were taken to create a crisp layer of final membership (Step V). The entire operation was repeated for all six window sizes (90 m^2 through 1690 m^2) over which the terrain attributes are calculated.

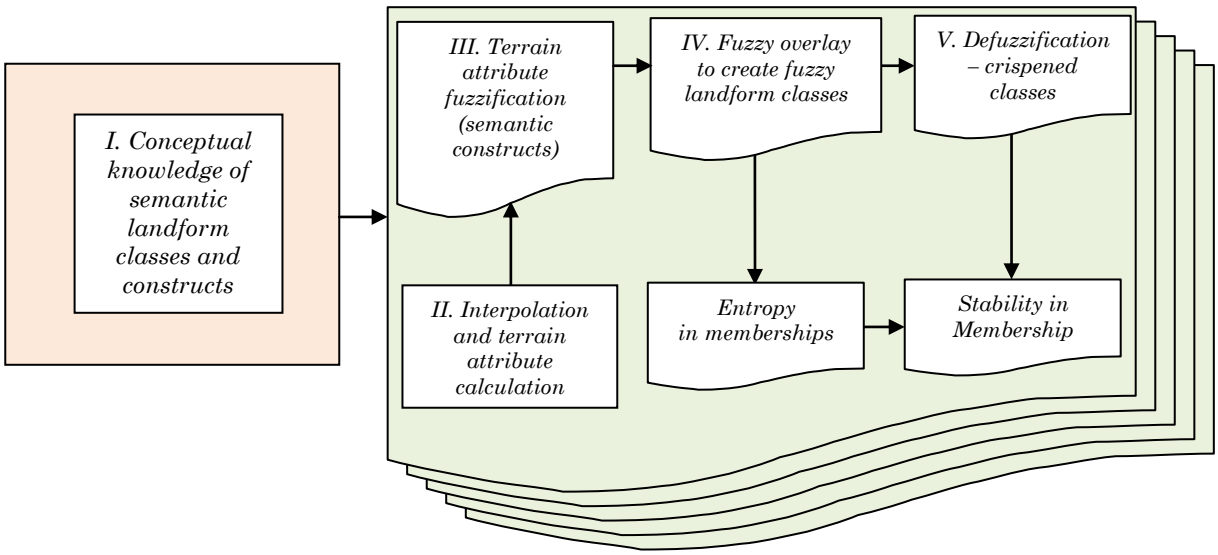


Figure 18. Basic flowchart of methodological steps

4.3 Building Semantic Classes (Step I)

For landform classification procedures directly relevant to soil-landscape modeling and applicable to agriculture, McMillan et al. (2000) list a number of criteria to be met:

- The defined units should exhibit meaningful differences in soil properties, moisture regimes, and crop yield,
- The procedures should apply to a wide range of types and scales of agricultural landscapes,
- The procedures should be based on a single model or protocol and should produce simple output consisting of a limited number of landform classes with defined characteristics,

- The landform classes should support a limited number of protocols for assigning standard management prescriptions for each defined landform unit.

Because of their basis in agricultural landscapes and incorporation of geomorphometric context, a combination of elements in selected previous work was utilized to establish classes and semantic constructs (Speight, 1990; Coops et al., 1998; Qin et al. 2009). These allow for the differentiation of upland, midland, and lowland sections, and provide further criteria for the delineation of the midlands to shoulder slope, back slope, and foot slopes. Building upon these systems allows one to distinguish shedding and receiving areas of the landscape on the basis of surface shape and landscape position. More complicated systems with a higher number of classes (MacMillan et al., 2000; Schmidt and Hewitt, 2004) would impede an analysis of cross-scale behavior of landforms without more specific knowledge of the study area.

Semantic meanings of the landforms based on specific terrain attributes were explored before quantitatively defining the suite of characteristics that compose each landform class. Definitions of the landform classes at their simplest should reference both relative elevation and shape of the local land surface, and reflect surface shape to infer the dominant geomorphological process, and landscape position to distinguish the semantic membership of each location (Coops et al., 1998; Klingseisen et al. 2008). Although the thresholds of terrain attributes or distinguishing characteristics will differ in applications to different regions, a general concept based on generic semantics should be able to be applied repeatedly with consistent results across other landscapes. Semantic geomorphometric classes are represented in the schematic diagram in Figure 19.

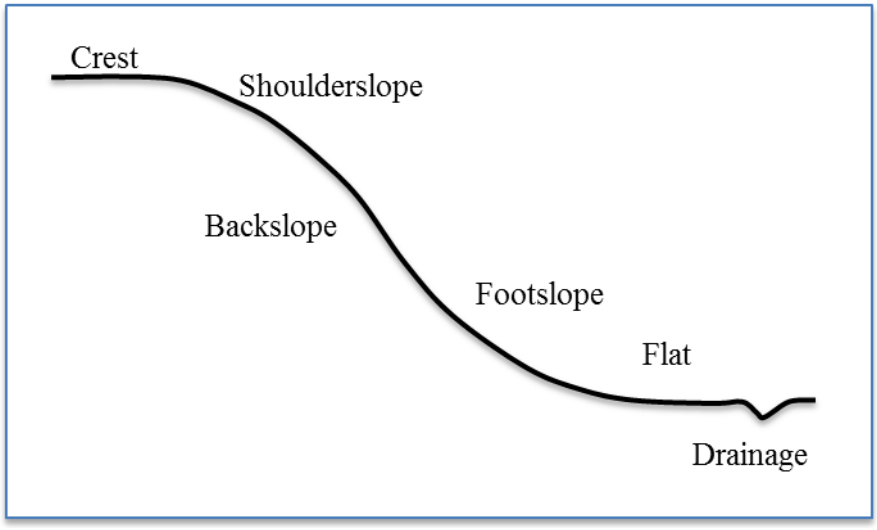


Figure 19. Schematic diagram of six semantic landform classes

- *Crests* – crests are the highest areas in the landscape; generally flat with little slope, convex in tangential curvature (shedding area) with planarity in profile curvature.
- *Slopes* – Slopes are the transitional areas between uplands and lowland, and are broken into three general classes, although further discriminations exist; *shoulderslopes*, *backslopes*, and *footslopes*. The three classes are generally separated on thresholds of curvature and landscape position. *Shoulderslopes* represent areas higher in the landscape adjacent to crests that are areas of greatest erosion and detachment potential from overland flow. Profile curvature is positive, denoting an increasing rate of slope in the downhill direction (higher erosion potential). *Backslopes* can be described as mid-slopes where the profile curvature is relatively planar and detachment potential is relatively equal to deposition potential. *Footslopes* are the lowest slope class and are characterized by having concave profile curvature where sediment deposition is greater than erosion potential.
- *Flats* – Flats are depositional areas of flatness, low-gradient, and planar or undefined curvature, although flats may not be necessarily associated with lowland parts of a landscape.

Expansive flats within valleys are represented by terraces formed by floods and other alluvial activity, and have a positive correlation between their size and stream order of adjacent streams and rivers.

- *Drainages* – In other taxonomies this may be represented by hollows or open depressions, where overland flow converges during precipitation events. Drainages are the antithesis of ridges, with concave tangential curvature and profile curvature, and are the lowest locations in the landscape. These features can be associated with obligatory vegetation species, or hydric soils.

4.4 Calculating Terrain Attributes (Step II)

Terrain attributes to define the semantic constructs based on relevant models (Speight, 1990; Coops et al., 1998; Qin et al., 2009) were calculated from the DEM. Slope gradient and curvature were calculated using the second-order polynomial finite-difference method first introduced by Evans (1972). Least-squared fit using second-order polynomials are regarded as the most robust technique for approximating surfaces and hillslope elements (Schmidt et. al., 2003; Shary, 2002). The bivariate form was later modified by Young (1978), Wood (1996), and most recently Shary (1995; 2002). The method used here is the Wood procedure, a modification of Evans-Young. This method has been shown to be less sensitive to local variations of data than other methods, most notably the one by Zevenbergen and Thorne (Florinsky,1998; Schmidt et al., 2003; Schmidt and Andrews, 2005).

Through the use of normal equations and matrix algebra, Wood (1996) developed a method to approximate surfaces using different window sizes in regularly-gridded DEMs. The

procedure fits a local coordinate system of an odd-numbered N x N window into a series of normal equations. The unknown coefficients from the polynomial expression are simplified to just four terms needed to solve for the six coefficients through the use of matrix algebra. This allows slope and curvature to be calculated over different window sizes and from the same set of coefficients, a significant advantage over the 3x3 neighborhood methods of Zevenbergen and Thorne (1987) and Horn (1981) present in many geospatial analysis programs.

In Wood's modification of the Evans-Young algorithm, a local neighborhood of N x N dimensions is fitted with the bivariate quadratic function:

$$z = ax^2 + by^2 + cxy + dx + ey + f \qquad (4.1)$$

where :

$$a = \frac{\partial^2 z}{\partial x^2} \qquad b = \frac{\partial^2 z}{\partial y^2} \qquad c = \frac{\partial^2 z}{\partial x \partial y} \qquad d = \frac{\partial z}{\partial x} \qquad e = \frac{\partial z}{\partial y} \qquad f = z \quad (4.2)$$

The six coefficients are then represented within the matrix expression as:

$$\begin{pmatrix} \sum x_i^4 & \sum x_i^2 y_i^2 & 0 & 0 & 0 & \sum x_i^2 \\ \sum x_i^2 y_i^2 & \sum x_i^4 & 0 & 0 & 0 & \sum x_i^2 \\ 0 & 0 & \sum x_i^2 y_i^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sum x_i^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sum x_i^2 & \\ \sum x_i^3 & 0 & 0 & 0 & 0 & N \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} \sum z_i x_i^2 \\ \sum z_i y_i^2 \\ \sum z_i x_i y_i \\ \sum z_i x_i \\ \sum z_i y_i \\ \sum z \end{pmatrix} \qquad (4.3)$$

The four constants of the matrix ($x_i^4, x_i^2 y_i^2, x^2, N$) are solved once for a given window size (N) (Wood, 1996). The coefficient matrix (a through f) is then solved using matrix algebra.

Once the six coefficients (a-f) are solved for each window size, a number of surface parameters can be easily calculated. The rate of elevation change, or slope magnitude in both the x and y direction is used to identify the direction and magnitude of the steepest gradient (Wood, 1996). The slope in degrees is represented as:

$$\text{Slope}(^\circ) = \arctan\sqrt{d^2 + e^2} \qquad (4.4)$$

The two secondary derivatives Profile and Plan Curvature separate curvature into two orthogonal components where the effects of gravitational process are either maximized down a hillslope profile (profile curvature), or minimized across a hillslope profile (contour or plan curvature). Profile curvature is the curvature of the normal plan section in the gradient vector direction, and plan curvature is the curvature of the horizontal plan section in the direction perpendicular to the gradient vector. Profile Curvature (4.5) and Plan Curvature (4.6) are represented as (Wood, 1996; Jordan, 2003):

$$\text{ProfCurv} = \frac{ad^2 + be^2 + 2cde}{\left(d^2 + e^2\right)\sqrt{1 + d^2 + e^2}} \qquad (4.5)$$

$$\text{PlanCurv} = \frac{bd^2 + ae^2 - 2cde}{\sqrt{d^2 + e^2}} \qquad (4.6)$$

Tangential curvature was first proposed as an alternative to plan curvature by Shary and Stepanov (1991) and later by Mitasova (1993). Tangential and Profile curvature are both curvatures of normal sections and exhibit similar statistical distributions, whereas plan curvature has markedly different statistical distributions, with a standard deviation up to ten times greater than that of profile curvature (MacMillan and Shary, 2009). Tangential Curvature (4.7) is the curvature of the line formed by intersection of surface with plane normal to the flow line (Jordan, 2003):

$$\text{TanCurv} = \frac{ad^2 + bd^2 - 2cde}{\left(d^2 + e^2\right)\sqrt{1 + d^2 + e^2}} \qquad (4.7)$$

Units of curvature from equations 4.5-4.7 are initially dimensionless (Z units/Z units). To present more meaningful values and differentiate values of convex curvature (positive) from

values of concave curvature (negative), curvature values are multiplied by -100 to represent Z units/100 Z units:

$$\text{FinalCurv} = \text{Curv} * (-100) \quad (4.8)$$

Besides derivatives of elevation, several other terrain attributes were calculated. Elevation Percentile (EP) is the ratio of elevation values less than the elevation of a reference point. Within a gridded DEM of constant resolution, elevation percentile can be computed for the focal center of local neighborhoods of dimension n x n. Elevation Percentile is used as a distinguishing concept for crests and valleys, and is a key component of landform classification in Speight's system (1990) as well as ridgetop and valley bottom flatness indices (Gallant and Dowling, 2003). Elevation Percentile is expressed as (Klingseisen et al., 2008):

$$\text{Elevation Percentile} = \underset{i \in N}{\text{count}} \left(\frac{z_i < z_0}{n_c} \right) \quad (4.9)$$

where z_i is the elevation at location i within, z_0 is the focal center of, and n_c the of cells in a neighborhood sized N. number

Because secondary derivatives are sensitive to small errors in the DEM quality and may contain noise, several other measures of relative curvature are useful in characterizing the mass balance or erosion potential of a location. One alternative attribute is Relative Profile Curvature (RPC) (Behrens et. al 2006). Relative Profile Curvature is a measure of local mass balance, and is a ratio of the average slope of all cells within a neighborhood with a higher elevation of the focal cell to the average slope of cells within the neighborhood with a lower elevation of the focal cell. With the magnitudes of slopes and curvature differing greatly, relative profile curvature allows a standardized measure of balance to be created and compared across scales. Relative Curvature Profile is calculated as:

$$\text{Relative Profile Curvature} = \sum_{i=1}^{N} \frac{u_1^o \in z_i < z_0}{u_2^o \in z_i > z_0} \qquad (4.10)$$

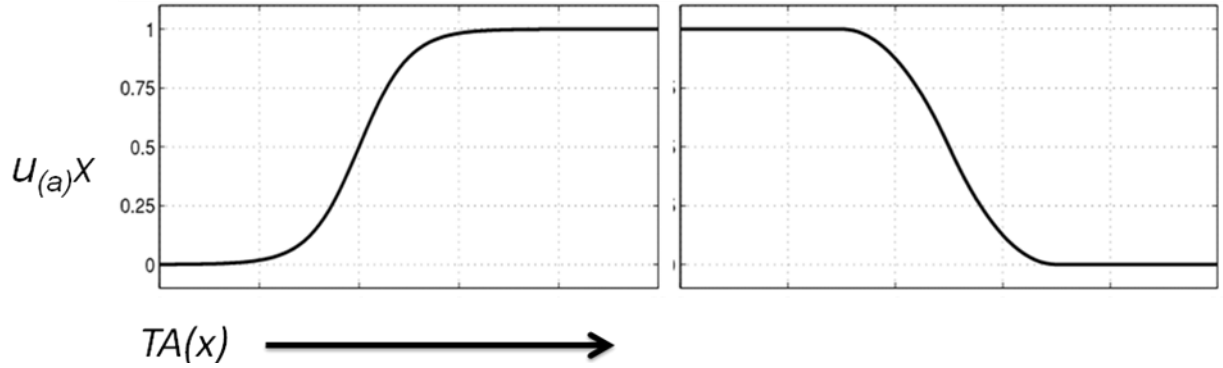
where u_1^o is the mean slope of cells z_i in focal neighborhood of sized N with a greater elevation than cell z_0 at the focal center, and u_2^o is the mean slope of cells z_i in focal neighborhood of sized N with an greater than cell z_0 at the focal center.

Several post-processing operations needed to be completed to correct minor errors and anomalies in the Relative Profile Curvature Layer. Cells with an empty numerator (pits) or an empty denominator (peaks) were assigned a '-1' or '1.#INF' value. To replace the '1.#INF' values, a neighborhood focal mean was calculated for each neighborhood size while ignoring '-1' or 'NaN' values. This new value replaced the original '1.#INF' within each layer.

After deriving appropriate algorithms, terrain attributes were computed for each spatial scale within Python and MATLAB computing environments (Van Rossum, 2005; MATLAB, 2010). Spatial scales over which to derive the landform classes were partly based on literature (Zhang and Montgomery, 1994; McMilllan and Shary, 2009), and feasibility of incorporating relevant information within the units of study, CRP easements, which occur over a range of sizes and shapes. Predictive capabilities of DEMs decrease rapidly at details coarser than the hillslope scale (Pike, 2000). If average hillslope lengths are in the order of 100-200 meters, a scale greater than 200 meters may not be able to capture within-hillslope variability of hydrologic processes (Moore et al., 1993). The spatial scales were thus composed of dimensions of length of 3, 5, 7, 9, 11, and 13 cells representing areas of 90, 250, 490, 810, 1210, and 1690 square meters. Within each spatial scale, terrain attributes were calculated for the focal cell at the center of the neighborhood, before moving to the next cell.

4.5 Converting terrain attributes to fuzzy semantic constructs (Step III)

Semantic constructs to define the landform classes were derived with the use of fuzzy semantic import models (SI). The S-curve (Figure 20a, below-left) and the Z-curve (Figure 20b, below right), were used to convert terrain attributes (TA(x)) to fuzzy membership values ($u_{(a)}x$) on a 0 (non-membership) to 1 (full membership) continuous scale.



Figures 20a-b. Fuzzy membership functions used to convert terrain attributes (TA(x)) to fuzzy membership values ($u_{(a)}x$) (MATLAB, 2010)

The primary forms of SI models used to convert terrain attributes to fuzzy membership values for each semantic construct were first-order polynomials:

$$u_{(a)}x = \frac{1}{1+((x-c)/d)^2} \text{ for } 0 \le x \le c$$

$$u_{(a)}x = 1 \text{ for } x \ge c \qquad (4.11a)$$

$$u_{(a)}x = 1 \text{for } x \le c \qquad (4.11b)$$

$$u_{(a)}x = \frac{1}{1+((x/c)^d)} \text{ for } 0 \le x \le c \qquad (4.12)$$

where $u_{(a)}x$ is the membership value of terrain attribute *x* to semantic construct *a*, with *c* as the central concept and *d* as the dispersion index, and equations 4.11a and 4.11b act as modifiers to

equation 4.11 in order to create the S-curve (4.11a) and Z-curve functions (4.11b). Multiple semantic constructs (Table 3, column 2) were derived to represent linguistic properties of landforms from terrain attributes (Table 3, column 1) through semantic import models (Table 3, column 3) and the use of fuzzy operators. Central concepts, dispersion indices, and bounding criteria (Table 3, columns 4-5) were based on parameters within the literature, and refined based on visual results and statistical distributions of terrain parameters within the entire study area. Thus, for each location (pixel), a membership grade to each semantic construct is calculated. These semantic constructs are then overlaid to compute the degree of membership of a considered location to a semantic landform class. This is known as the heuristic rules approach (Schmidt and Hewitt, 2004). A schematic diagram of landform classes and associated semantic constructs (number from Table 3, column 2) is shown in Figure 21.

Terrain Attribute	Semantic Constructs	Fuzzy Model	Central Concept	Dispersion Index
Elevation Percentile	(1) Highness	4.11a	.65	.2
	(2) Lowness	$u_{(Lowness)}x = 1 - u_{(Highness)}x$		
Slope	(3) Steepness	4.12	6	2
	(4) Flatness	$u_{(Steepness)}x = 1 - u_{(Flatness)}x$		
Profile Curvature	(5) Profile Convexity	4.11a	.10	.05
	(6) Profile Planarity	4.11b	-.10	.05
	(7) Profile Concavity	$u_{(Planarity)}x = 1 - \lvert u_{(convexity)}x - u_{(concavity)}x \rvert$		
Tangential Curvature	(8) Tan. Convexity	4.11a	.10	.05
	(9) Tan. Concavity	4.11b	-.10	.05
Relative Profile Curvature	(10) – Balanced	4.11a	1.45	.3
	(11) Balanced	$u_{(Balanced)}x = 1 - \lvert u_{(-Balanced)}x - u_{(+Balanced)}x \rvert$		
	(12) + Balanced	$u_{(+Balanced)}x = 1 - u_{(-Balanced)}$		

Table 3. Summay of terrain attributes, semantic constructs, and fuzzy membership function parameters.

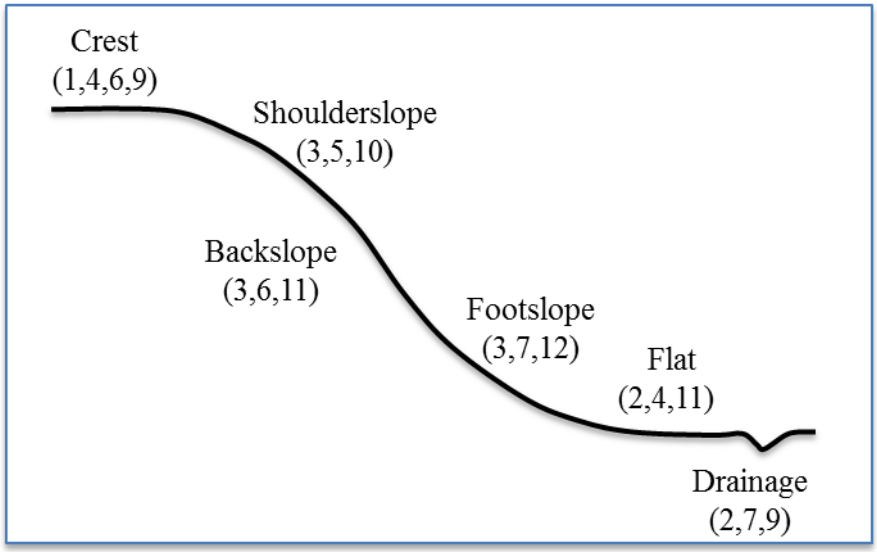


Figure 21. Semantic landforms and the constructs used to define them (in parentheses)

Additional information on the distinguishment of the parameters for the Fuzzy Membership functions area listed below:

- Flatness and steepness - Slope is utilized for the Semantic Constructs flatness and steepness. A Slope threshold is needed to separate flat areas (crests and flats) from steep areas (slopes). Because the distribution of slopes approaches the Gaussian ideal with minor positive skew, a cumulative mean of the slope distribution across all scales of analysis was used to distinguish flat and steep areas and serve as the crossover point, which has equal membership to both flatness and steepness. Means of slope distribution ranged from 6.68^0 at the 9-cell neighborhood to 5.21^o at the 169-cell neighborhood, with a mean of 5.94^o, so 6^o was utilized as the central concept to separate the two constructs of flatness and steepness. Equation 4.12 was used to derive steepness, with flatness derived using the complementary operator. The

dispersion index was visually tested, and found in agreement with existing sources (Gallant and Wilson, 2003).

- Curvature constructs - Profile Curvature is used to construct three semantic constructs, profile convexity, profile planarity, and profile concavity, with tangential convexity and concavity the two constructs for tangential curvature. Two functions are used to define the constructs profile convexity and profile planarity with values from either side of zero, negative values representing concavity, and positive values representing convexity. Equation 4.11 was used to model profile convexity with the modification of equation 4.11a. Central concepts and diversion indices were tested, beginning with the use of statistical distributions of the terrain attribute over scales of analysis (c = $\bar{x}$ (profile curvature) = .085, d = ½σ (profile curvature) = .042). A central concept of .10 was determined as optimal, with the dispersion index of .05 based upon this threshold identifying 1/3 of the landscape as convex, and being in agreement with central concepts utilized in previous studies in similar terrain (Pennock et al., 1987; MacMillan et al., 2000; Qin et al., 2009). The same form was used for deriving semantic construct tangential convexity from terrain attribute tangential curvature. Because of the dichotomous relationship of convexity and concavity in their identical distribution on both sides of zero, equation 4.11b with a sign reversal in the denominator was utilized to derive the constructs for profile and tangential concavity. To derive profile planarity, a simple equation was derived to produce full membership of the construct where curvature is close to 0. The difference between the constructs profile convexity and profile concavity was subtracted from 1 (Construct 7 in Table 3). Thus the crossover point for both profile convexity and profile concavity is the location of full membership to profile planarity.

- Highness and lowness - Elevation Percentile is used to define two semantic constructs, highness and lowness. Elevation Percentile has been utilized in several studies (Coops et al., 1998; Gallant and Dowling, 2003; Klingseisen et al., 2008) to define areas high and low in the landscape. Existing thresholds for defining a crest are an elevation percentile of at least .65. Equations 4.11a was used to derive the constructs, where .65 is the central concept, c, for highness, and .2 is the dispersion index. For the semantic construct lowness, the complement operator was used (Construct 2, Table 3)
- Balance constructs - Relative profile curvature is used to produce three constructs; negatively balanced for areas like crests and shoulderlsopes (implying sediment detachment), balanced for areas where the net sediment loss is relatively equal to deposition, and positively balanced for areas of net deposition. The distribution of Relative Profile Curvature is relatively stable across the scales of analysis ($\bar{x}$ =1.14 to 1.17 between 9-cell neighborhood and 169-cell neighborhood, σ =.68 to .66 in same range). The same membership function model was used for balance as the profile curvature constructs with Equation 2.4 and modification of Equation 2.5. Negatively balanced was initially modeled with the central concept (1.45) derived from the mean relative profile curvature across scales (1.15) + ½ the approximate standard deviation (rounded down to .3). ½ the standard deviation was also used for the dispersion index, giving a fuzzy membership value of the semantic construct a crossover point at approximately the mean of the dataset across all scales (1.15). For the additional semantic constructs, positively balanced was created through the complement of Negatively balanced, and balanced was created the same way as profile planarity, by subtracting difference of fuzzy membership values between negatively balanced and positively balanced from one.

4.6 Fuzzy Overlay and Defuzzification (Step IV and V)

Fuzzy overlay was performed to derive surfaces of all semantic landform classes (*crest, shoulderslope, backslope*, *footslope*, *flats*, and *drainage*) using the corresponding set of semantic constructs (Table 3, Figure 21) and applying three different operators. The products of this step are six fuzzy landform layers, with cell values indicating the degree of membership to a particular class. Previous discussion of overlay and operators can be found in section 2. The six fuzzy landform layers generated by one of the operators (Fuzzy Gamma) at the 49-cell window size can be seen in Figure 20.

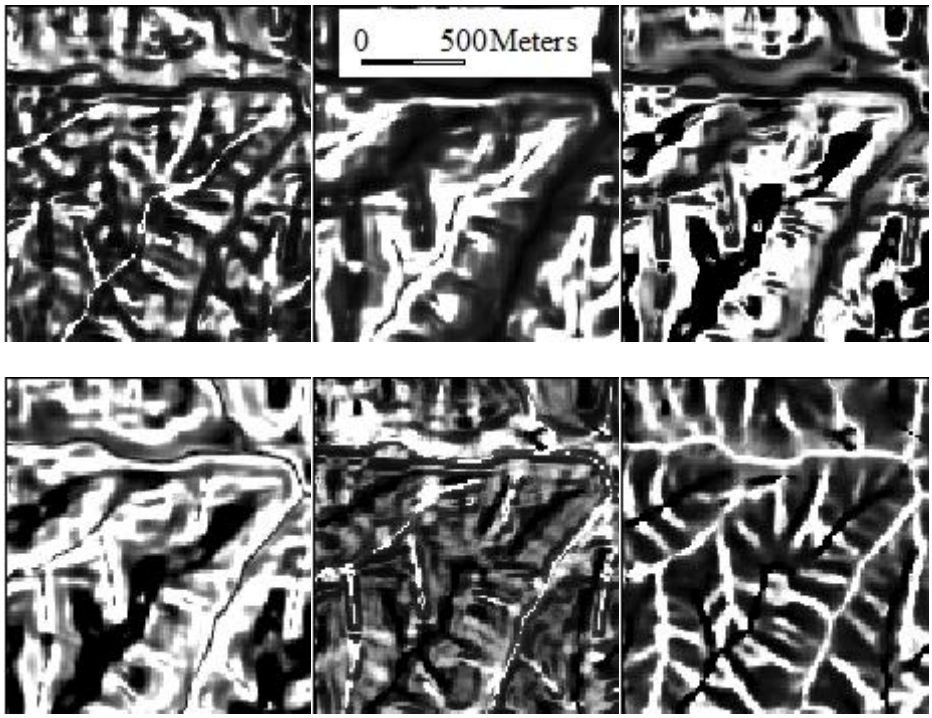


Figure 22. Fuzzy landforms generated by the fuzzy gamma operator at the 49-cell window size. Lighter areas represent higher membership values to that class (upper left-crests, upper center-shoulderslopes, upper right-backslopes, lower left-footslopes, lower center-flats, lower right-drainages).

The three operators used for fuzzy overlay, intersect, algebraic mean, and fuzzy gamma, have been used extensively in soil-landscape modeling(MacMillan et al., 2000; Qin et al., 2009), risk analysis related to soil erosion (Metternicht and Gonzalez, 2005), ordination of vegetation (Roberts et al., 1986), and other geomorphometrically-related subjects (landslide susceptibility, subsurface mineral deposits) (Tangestani et al., 2004). These operators vary in levels of constraint, liberation, and mathematical complexity:

- Intersect – The intersect operator assigns a location the minimal value of multiple fuzzy sets (semantic constructs) and responds to the logical connective 'AND' between semantic constructs A,B,C, and D:

$$U_{A\cap D} = \min(.76,.94,.36,.45) = .36 \qquad (4.13)$$

  The intersect operator is the least liberal and most constraining of the three operators tested, with no interacting or compensating effects between the fuzzy sets that are overlaid.

- Algebraic Mean – The algebraic mean, is calculated as the mean of multiple fuzzy sets, and responds to the relational algebra operator 'Cartesian Product':

$$U_{(mean)} = \left(\prod_{i=1}^{n} u_i\right) = \frac{1}{4}\sum_{i=1}^{4}(.76,.94,.36,.45) = .63 \qquad (4.14)$$

  Algebraic mean is the middle ground between the constraining effects of the intersect operator and the compensatory effects of the fuzzy gamma operator (see below).

- Fuzzy Gamma – The fuzzy gamma operator is a combination of the algebraic product and the algebraic sum (An et al., 1991). Within two fuzzy sets, A and B, the Algebraic Sum is written as the complement of the Algebraic Product:

$$u_C x = \int u_A x \wedge u_B x / 1 - (u_A x * u_B x) \qquad (4.15)$$

where membership to a new set u_c, is defined via intersection $\wedge$ as the difference between sum of two membership values and their product. The fuzzy gamma operator is then stated as:

$$U_{(gamma)} = \left(1 - \prod_{i=1}^{n}(1 - u_i)\right)^{1-\lambda} * \left(\prod_{i=1}^{n} u_i\right)^{\lambda} \qquad (4.16)$$

where λ is the fuzzy gamma coefficient, and a coefficient value of 0 equals the fuzzy algebraic product and a value of 1 equals the fuzzy algebraic sum. Because the outputs of a fuzzy algebraic sum are increasive, and outputs of the fuzzy algebraic product are decreasive, the fuzzy gamma operator is seen as a flexible compromise between the two where there is no indication which of the semantic constructs may be the limiting factor of. Among the fuzzy operators, the fuzzy gamma operator can be viewed as having the highest amount of compensatory effects between the set of semantic constructs. Fuzzy gamma coefficient values (λ) closer to 0 give heavier weight to the fuzzy algebraic sum, while values closer to 1 give heavier weight to the fuzzy algebraic product. Because of a lack of background research on expert-driven effects of changing gamma operator threshold, arbitrarily chosen fuzzy gamma coefficients of .25, .5, and .75 were tested. The range of the derived membership values varied widely, although less than .05% of cells changed crisp classes between the three coefficients within the same scale when the six fuzzy layers were defuzzified. It was decided to use the fuzzy landform layers derived from the fuzzy gamma coefficient at the .5 level, and use it as a basis to compare against landforms derived from algebraic mean and intersect operator.

The most liberal and least constraining operator and direct opposite of the intersect operator, union, was not used because it produced the least realistic visual results. Figure 23 demonstrates results of defuzzified landforms derived from the fuzzy gamma and union operators at the 49-cell window size.

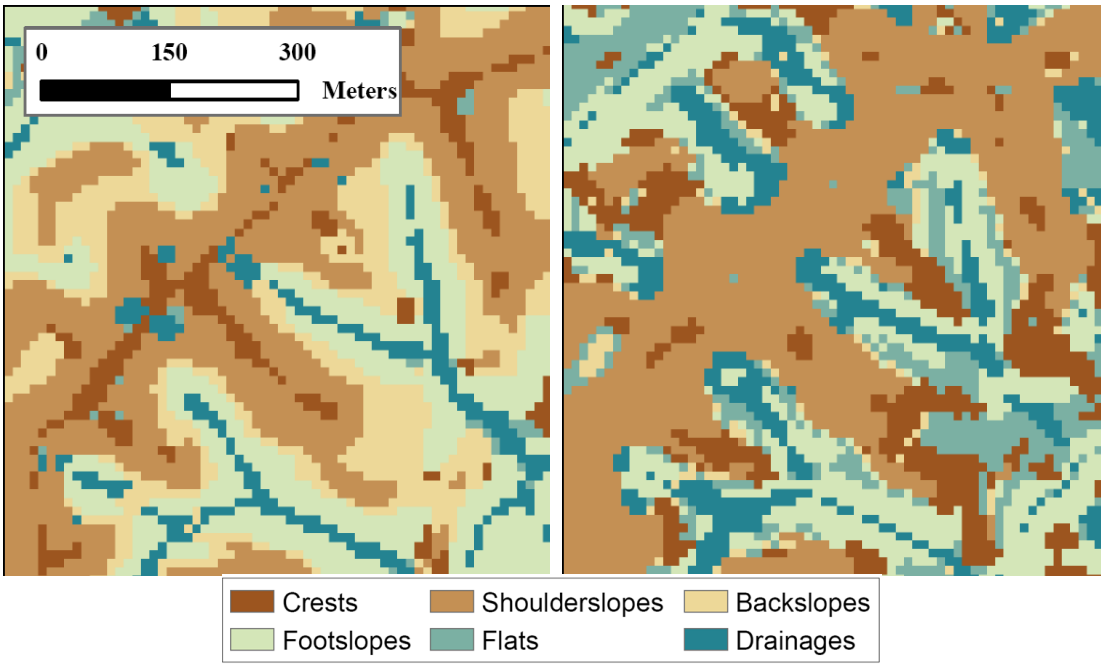


Figure 23. Differences in crisp landforms derived from the fuzzy gamma operator (left) and the union operator (right) at the 49-cell window size. The high incoherency of the union landforms precluded it from further analysis.

To derive a final, crisp layer of six landform classes, a final overlay operation was applied to the six fuzzy landform layers utilizing the maximum of membership (MOM) method (Zadeh, 1965). This is referred to as defuzzification, where a final, quantifiable result is mapped to a crisp set from given multiple fuzzy sets and membership (Klir and Yuan, 1995). Many other defuzzifcation methods are available for inference or rule-based system, which do not apply here. For each set of classes derived from the three operators, the semantic landform class with the highest membership value was taken to create a boolean, crisp layer of final membership at each scale. Crisp layers derived from fuzzy gamma (left), algebraic mean (center), and intersect (right) operators (49-cell window size) are shown in Figure 24 underlying a single CRP tract.

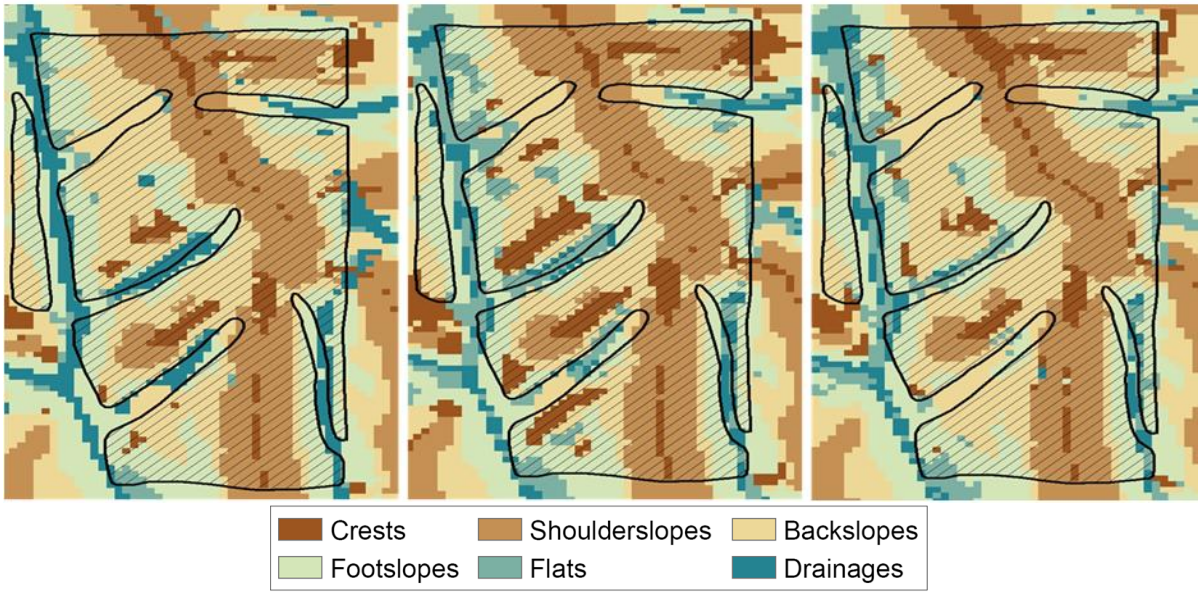


Figure 24. Defuzzified landform layers derived from the Fuzzy Gamma (left), Algebraic Mean (center), and Intersect (right) operators underlying a single CRP Tract.

4.7 Analysis of Stability, Fuzzy Membership, and Entropy

To understand the distribution of crisp landforms, the effect of overlay operators on the fuzzy membership values, and the resultant classification stability of the crisp landforms, the analysis is split into three sections. The first section covers the relationships between the crisp class areas across scales, utilizing gross cell counts and the cross-tabulation matrix to measure specific categorical trajectories between different scale-steps. The second section analyzes the distribution of fuzzy membership values for each of the three operators using a Wilcoxon test for non-parametric data. The final part combines the crisp and fuzzy portions of the analysis by examining the underlying classification stability of each semantic landform class using the measure of classification entropy. This framework allows for separate analysis of crisp landforms, fuzzy membership values, and a combination of crisp and fuzzy.

Cross-tabulation matrices were constructed to compare changes in semantic landform distributions between spatial scales and between operators. These tables are commonly-used tools for multi-dimensional change analysis or accuracy assessment in land cover/land use with multiple time steps (Pontius et al., 2004). An example of a cross-tabulation matrix is shown below in Table 4.

	b=1	**b=2**	**….**	**b=j**	**Total (b)**	**Losses**
a=1	P_{11}	P_{12}	…	P_{1j}	$P_{1}+$	$P_{1+} - P_{11}$
a=2	P_{21}	P_{22}	…	P_{2j}	$P_{2}+$	$P_{2+} - P_{22}$
…	…	…	…	…	…	…
a=j	P_{j1}	P_{j2}	…	P_{jj}	$P_{j}+$	$P_{j+} - P_{jj}$
Total (a)	P_{+1}	P_{+2}	…	P_{+j}	1	
Gains	$P_{+1} - P_{11}$	$P_{+2} - P_{22}$	…	$P_{+j} – P_{jj}$		

Table 4. Template cross-tabulation matrix (based on Pontius et al., 2004), where *a* and *b* represent two datasets with *j* classes, *a* representing the reference map, and *b* representing the comparison map. The proportional values of agreements between two sets for a given class (*a* =1 and *b* = 1 at a given location) are located on the diagonal. Comparison map total for class *a* = 1, etc. are located in the column Total (b) and reference map total in row Total (a). Totals for both column and row equal 1 (100% of the landscape). P_{ab} denotes portion of the landscape experiencing a change of category between the reference map and the comparison map.

The resulting matrix can be used to summarize total agreement between two datasets, and the categorical trajectory from one class to another (the direction of change). Within land cover change analysis, the terms gain and loss are used to describe additions and subtractions from or to class j, while swap refers the location of category changing over time while the overall class quantity remains the same (Pontius et al 2004). In this case for geomorphometry, gains, losses, and swaps use coarsening scales rather than change across time as frame of reference in determining directions of change between crisp classes.

Analysis of cross-tabulation matrices allows only insight into the Boolean memberships of the landform classes. To test the similarity or dissimilarity of the fuzzy membership values for each fuzzy landform class are between values derived from the intersect, algebraic mean, and fuzzy gamma overlay operators, the Wilcoxon ranked-sum test for paired samples was performed. The Wilcoxon ranked-sum test is a non-parametric method for determining if two independent data vectors come from the same distribution, even if one vector is shifted to the left or right of the other. Hypotheses for Wilcoxon ranked-sum test for two populations, x_1 and x_2, are as follows:

$$H_{o:}\, u_{1/2}(x_1) = u_{1/2}(x_2) \qquad (4.17)$$

$$H_a: u_{1/2}(x_1) \neq u_{1/2}(x_2) \qquad (4.18)$$

where $u_{1/2}(x)$ is the median of population x (MATLAB, 2010). Unlike parametric methods like Student's t or paired t-tests, there are no general assumptions for the test; though not required, it is often assumed that the populations have similar shapes and dispersions (Wuensch, 2010). In addition, the test has shown to perform with only slight losses with vastly different sample sizes (Sawilosky, 2005).

In the Wilcoxon Ranked-sum test, both populations are ranked numerically from greatest to smallest, and the ranks are summed. Under the null hypothesis, the sampling distribution or the sum of ranks, T, has the following statistical distribution:

$$u_t = \frac{n_1(n_1 + n_2 + 1)}{2} \qquad (4.20)$$

$$\sigma_t^2 = \frac{n_1 n_2}{12}(n_1 + n_2 + 1) \qquad (4.21)$$

where u_t is the mean and σ_t^2 the standard deviation of the two vectors being compared, n_1 and n_2. If the rank of sums, T, is higher or lower than u_t beyond the upper or lower bounds of the

rejection regions of the critical value of level of significance tested, the null hypothesis is false and the two distributions are not equal. If random locations share similar distributions of membership values for each class derived from the three overlay operators, it could be said there is no significant difference in how the overlay operators assign fuzzy memberships.

To quantitatively define locations of high classification ambiguity or uncertainty and determine if specific classes displayed high ambiguity, classification entropy (*H,* equations 2.13-2.14) was utilized. Classification entropy is a form of Shannon-Weiner Diversity Index which measures the variance in fuzzy membership in a dataset with a finite number of classes (Foody, 1997). Theoretrical application of the classification entropy was previously discussed in section 2.3.6.

Previous examples of the use of entropy in measuring in categorical classification stability include Brown (1998) to measure boundary vagueness in presettlement forest types. It has been used extensively in Soil-landscape modeling to assess the results of Fuzzy K-means classification of landform classes. Wood (2001) used entropy measures in combination with analysis of fuzzy membership values to determine pits at valley confluences are an artifact of a DEM.

Similar to the Wilcox Ranked-sum tests, the fuzzy membership values of the six fuzzy landforms derived from each overlay operator at each scale were normalized on a 0-1 scale (equation 4.19). Classification entropy was then calculated at each pairing of scales and operator using MATLAB (MATLAB, 2010). To compare the classification entropy of different landforms, the proportional relationships of the six Boolean classes at different levels of entropy were examined, an idea similar to exploring the effect of different $\alpha-$cut levels (section 2.3.6) on the areas of a defuzzified fuzzy set.

Proportional percentages of the range of classification entropy for each operator were calculated. Entropy levels to examine class proportionality were chosen at 0.99, 0.95, 0.90, 0.85, and 0.75 percentile to give more weight to identifying areas with highest classification entropy (high uncertainty and ambiguity). The specific values of entropy to examine in each pairing of scales/overlay operators at which to examine class proportions were determined by the range of each entropy surface:

$$E_P = E_{x(max)} - P(E_{x(max)} - E_{x(min)}) \qquad (4.22)$$

where E_p is the entropy value at each proportional percentage, P (0.99, 0.95, 0.90, 0.85, 0.75), $E_{x(max)}$ is the maximum classification entropy of the entropy surface, and $E_{x(min)}$ is the minimum classification entropy. The proportion of semantic classes at each entropy level was examined for each pairing of scales/overlay operators. Because of the variety of operators and scales being tested, a semantic landform class could be considered 'ambiguous' if it is a high proportion of the noted locations of high entropy between operators and across scales. Many vegetation and soil-mapping applications that rely on relative landscape position to predict community composition based on moisture availability, dominant processes, such as surface runoff versus groundwater recharge regimes.

CHAPTER V

RESULTS

5.1 Variance of Terrain Attributes and Landform area over Scales

Some of the basic distributions of terrain attributes have been discussed earlier in building Fuzzy Membership Functions. Each terrain attribute showed a markedly different behavior across the different scales of measurement. Bar plots of the mean and standard deviation of slope are shown in Figure 25. Means of terrain attributes across all scales of measurement are shown below in Table 4. With the exception of elevation percentile, the standard deviations of all terrain attributes decreased with decreasing scale, which is what would be expected as local noise is smoothed out as the range of measurement increased. Based on simple linear trends of the terrain attributes across scales, slope was the most volatile attribute with the steepest slope (m = .295).

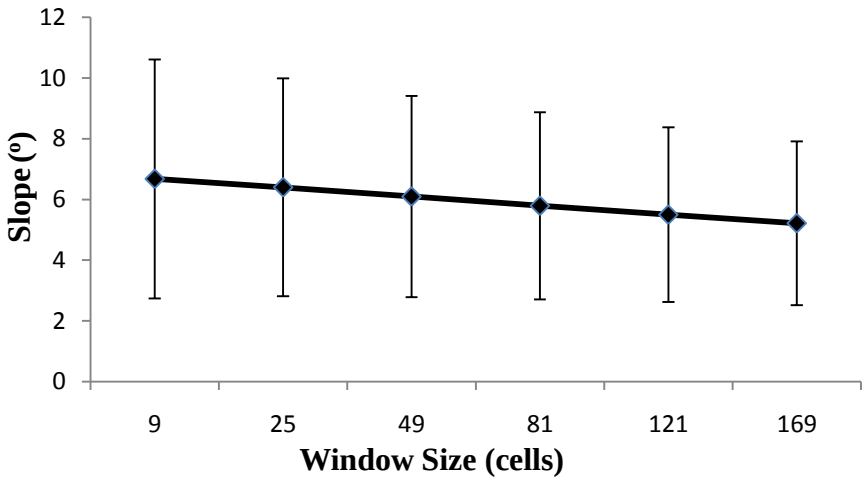
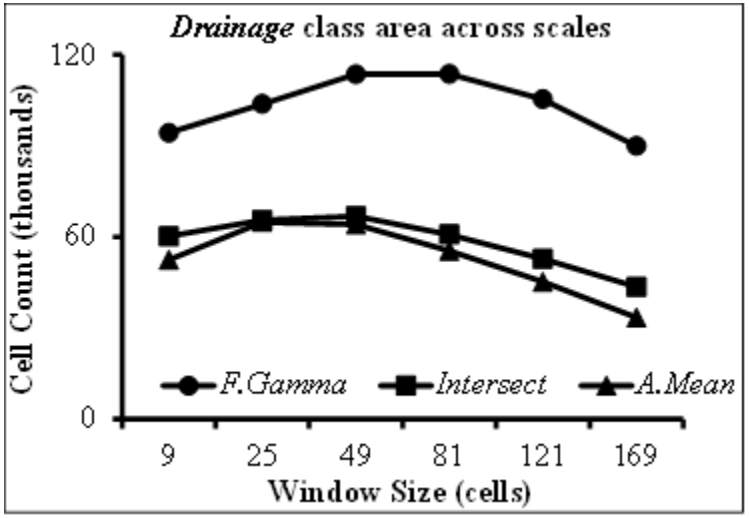


Figure 25. Change in mean (points) and standard deviation (bars) of slope measurements across window sizes.

	Window Size					
Attribute	**9-cell**	**25-cell**	**49-cell**	**81-cell**	**121-cell**	**169-cell**
Slope	6.6794	6.4021	6.0969	5.792	5.4975	5.2178
Prof. Curv.	-0.0015	-0.0011	-0.0008	-0.0008	-0.0007	-0.0007
Tan. Curv.	0.0026	0.0021	0.0018	0.0016	0.0015	0.0015
RPC	1.1127	1.1197	1.1138	1.1029	1.0893	1.0755
EP	0.4404	0.4777	0.4887	0.4937	0.4965	0.4984

Table 5. Mean of terrain attributes calculated over window sizes (Prof. Curv. = profile curvature, Tan. Curv. = tangential curvature, RPC = relative profile curvature, EP = elevation percentile). Units of slope are degrees and profile and tangential curvature Z units/100 Z units. Relative profile curvature is unit-less and elevation percentile is a proportion.

Cell counts for the six classes were plotted to compare the non-spatial distributions and behavior of landform areas across scales between the three overlay operators (intersect, algebraic mean, and fuzzy gamma). Percentages of each landform class across scales for each operator are shown in Table 6 through 8. Three general trends could be observed for the three operators. First, as scale decreased, the area (pixel counts) of *slopes* decreased while the area of *flats* and *crests* increased for each fuzzy operator. Second, *slopes* showed similar areas for all three overlay operators; *backslopes* showed larger area than either *backslopes* or *shoulderslopes*. Third, the only landform class which displayed a clear optimal scale for area or feature detection was *drainages*, which exhibited a parabolic-like behavior for each overlay operator, with the vertex of the parabola occurring at either the 49 or 81-cell neighborhoods and decreasing towards smaller and larger scales (Figure 26). *Drainages* over three scales are shown in Figures 27a-c.



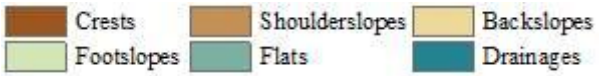


Figure 26. Area of landform class *drainages* over window sizes

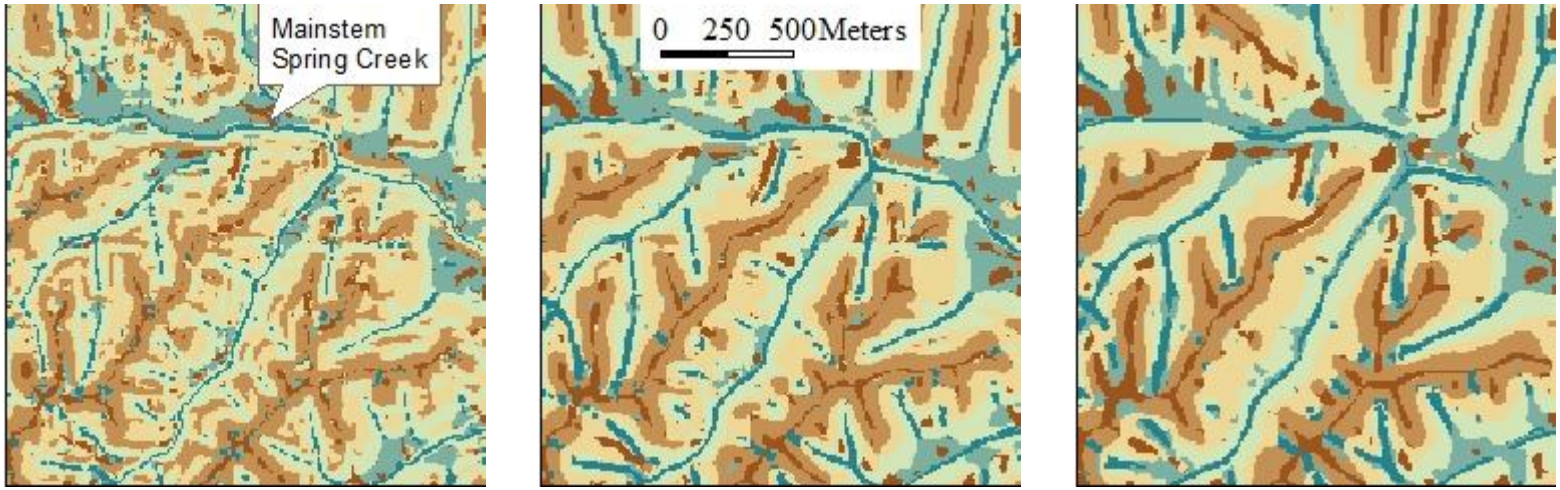


Figure 27a-c. Landfrom classes derived from the Fuzzy Gamma operators represented at the 25-cell (Figure 27a, left); 81-cell (Figure 27b, center) and 169-cell (Figure 27c, right) window sizes.

Table 6-8. Proportions of crisp classes derived over window sizes from intersect (Table 6), algebriac mean (Table 7), and fuzzy Gamma (Table 8) overlay operators (Cr. – *crests*, SS – *shouldeslopes*, BS – *backslopes*, FS – *footslopes*, Fl. – *flats*, Dr. – *drainages*)

Table 6. Proportions of crisp classes generated from intersect operator

Window Size (cells)	Cr.	SS	BS	FS	Fl.	Dr.
9	13.09	22.60	27.65	16.54	16.21	3.92
25	12.81	21.98	26.42	16.25	18.27	4.26
49	15.03	20.28	26.49	14.83	19.02	4.35
81	16.92	19.22	25.02	13.55	21.33	3.95
121	18.66	18.17	23.38	12.25	24.11	3.43
149	20.20	16.99	21.89	10.82	27.26	2.83

Table 7. Proportions of crisp classes generated from algebraic mean operator

Window Size (cells)	Cr.	SS	BS	FS	Fl.	Dr.
9	11.56	23.16	26.07	17.62	18.17	3.41
25	14.88	21.65	24.76	16.62	17.87	4.22
49	18.48	19.52	23.41	14.92	19.50	4.17
81	21.33	17.85	21.41	13.36	22.45	3.61
121	23.87	16.11	19.46	11.78	25.84	2.94
149	26.21	14.28	17.76	10.06	29.51	2.18

Table 8. Proportions of crisp classes generated from fuzzy gamma operator

Window Size (cells)	Cr.	SS	BS	FS	Fl.	Dr.
9	5.41	21.73	25.19	19.38	22.17	6.13
25	8.89	20.50	26.86	18.16	18.83	6.75
49	11.52	18.81	26.37	16.74	19.17	7.39
81	13.77	17.20	24.74	15.53	21.37	7.39
121	15.81	15.44	22.95	14.28	24.66	6.85
149	17.67	13.6	21.34	12.83	28.71	5.85

The main difference in distribution of landforms between operators occurred when considering the classes in two groups. Upland classes (*crests*, *shoulderslopes,* and *backslopes)* were more frequently represented in crisp landform layers derived from the intersect and algebraic mean overlay operators as compared to the fuzzy gamma operator, for which a higher proportion of lowland classes occurred (*footslopes*, *flats*, and *drainages).*

Within CRP parcels, there was a disproportionate representation of classes semantically representing both upland and lowland areas. A greater proportion of upland classes (*crests*, *shoulderslopes,* and *backslopes*) and a lesser proportion of lowland classes (*footslopes, flats,* and *drainages)* were represented in CRP parcels as compared to the rest of the study area across all three fuzzy overlay operators (Figure 28). This discrepancy in upland and lowland proportions was greatest at the intermediate scales and lowest towards both coarser and finer spatial scales.

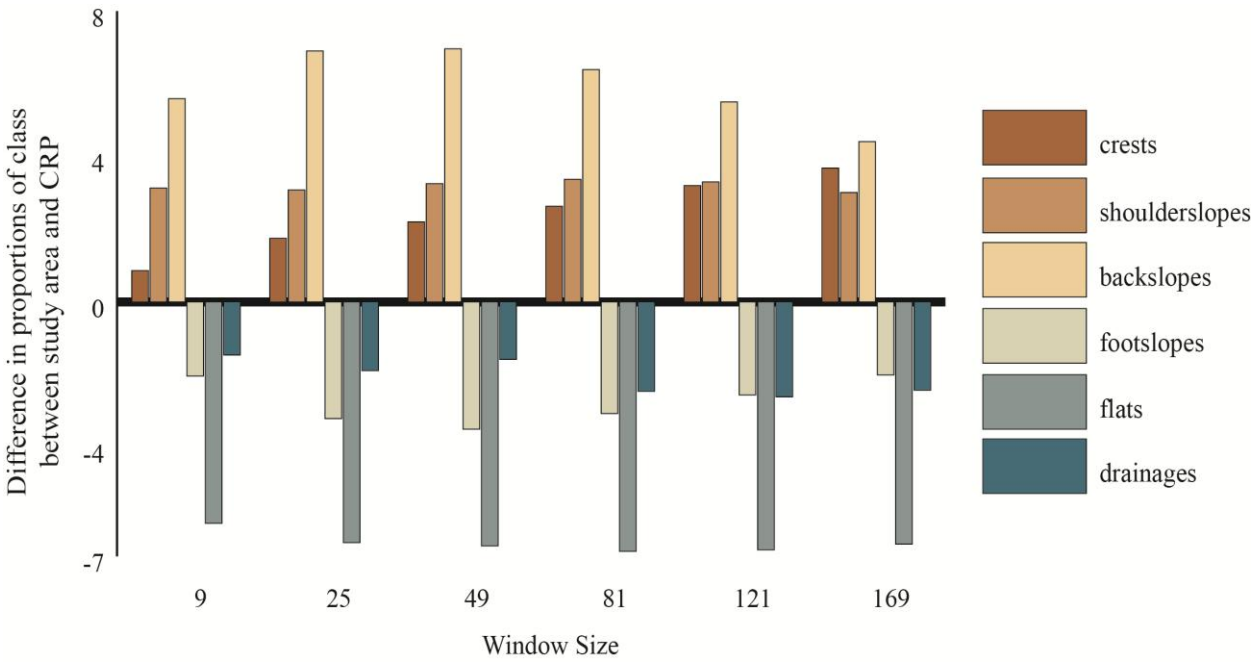


Figure 28. Proportional differences of landform classes generated by fuzzy gamma operator between CRP parcels and study area

Cross-tabulation matrices were calculated between pairs of scales for each of the three operators resulting in a total of 63 tables. Tables 9 and 10 (below) represent cross-tabulation matrices between the 25-cell and 81-cell window sizes (Table 9) and 81-cell and 169-cell window sizes (Table 10) for fuzzy landforms generated by the fuzzy gamma operator. Only two examples are given, as the majority of change directions between crisp landform classes were similar in all three overlay operators. Gains of classes between the 25-cell and 81-cell window sizes are noted by the red italicized cells, losses by the green-colored italicized cells, and agreement between scales is noted by the bolded diagonal.

Table 9. Cross-tabulation matrix between 25-cell and 81-cell window size derived from the fuzzy gamma operators (Cr. = crests, BS – backslopes, SS – shoulderslopes, FS – footslopes, Fl. – flats, Dr. – drainages)

81-cell window size	25-cell window size: Cr.	SS	BS	FS	Fl.	Dr.	Total	Losses
Cr.	**7.47**	0.42	0.45	0.23	0.30	0.04	*8.89*	***1.43***
SS	2.10	**12.89**	3.92	0.57	0.84	0.18	*20.50*	***7.61***
BS	1.43	3.08	**16.65**	2.68	2.68	0.35	*26.86*	***10.21***
FS	0.35	0.11	2.60	**11.15**	1.66	2.29	*18.16*	***7.01***
Fl.	2.34	0.53	0.84	0.58	**14.23**	0.30	*18.83*	***4.59***
Dr.	0.09	0.18	0.27	0.32	1.66	**4.23**	*6.75*	***2.52***
Total	13.77	17.20	24.74	15.53	21.37	7.39	100.00*	**66.97**
Gains	***6.31***	***4.31***	***8.08***	***4.38***	***7.14***	***3.16***		

*Cross-scale agreement: 66.62%

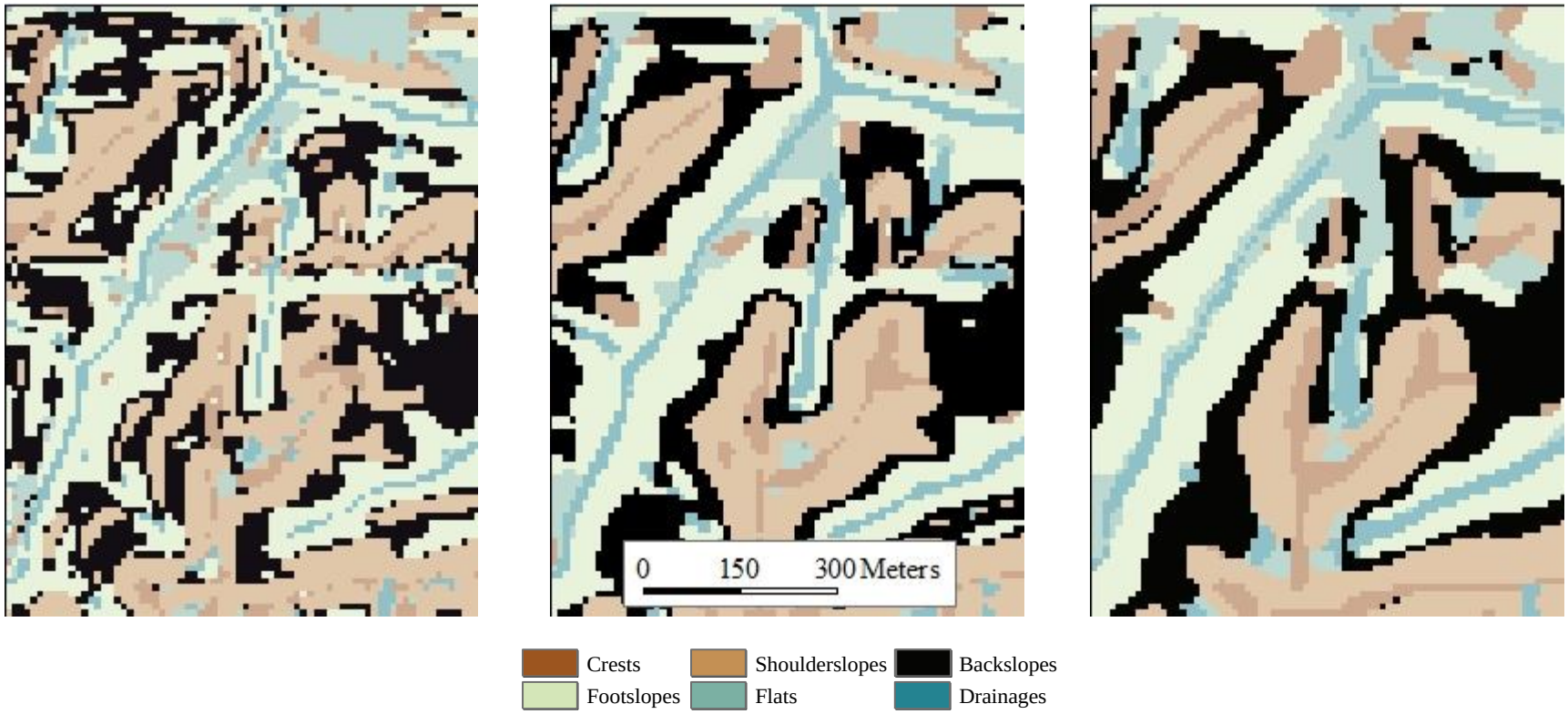
Table 10. Cross-tabulation matrix between 81-cell and 169-cell window size derived from the fuzzy gamma operator.

169-cell window size	81-cell window size: Cr.	SS	BS	FS	Fl.	Dr.	Total	Losses
Cr.	**12.13**	0.25	0.67	0.29	0.42	0.01	*13.77*	***1.64***
SS	2.33	**11.64**	2.25	0.06	0.84	0.07	*17.20*	***5.56***
BS	1.48	1.31	**16.59**	1.81	3.42	0.12	*24.74*	***8.15***
FS	0.22	0.01	1.14	**10.19**	2.71	1.26	*15.53*	***5.34***
Fl.	1.46	0.33	0.46	0.34	**18.61**	0.17	*21.37*	***2.76***
Dr.	0.05	0.06	0.23	0.14	2.70	**4.21**	*7.39*	***3.18***
Total	*17.67*	*13.60*	*21.34*	*12.83*	*28.71*	*5.85*	100.00*	
Gains	***5.54***	***1.36***	***4.75***	***2.64***	***10.10***	***1.64***		

*Percent Agreement: 73.37%

The matrix gives more information on specific trajectories across scales. Areas of both *crests* and *flats* continually increase in area with coarsening scale, though *flats* is the only class where one can see that both gains increase and losses decrease as scale becomes coarser (losses for flats decrease from 4.59 to 2.76 between the two sets of scale comparisons, while gains increase from 7.14 to 10.10 (Table 9 and Table 10). Correspondingly, all three *slope* classes decrease in area as scale becomes coarser. Figures 29a-c exhibit how a single class (*backslopes)* changes between sets of scales.

Specific change directions between classes are also observable. There is an increasing association in gains and losses between *drainages* and *flats* at coarsening scales. Losses for the other four semantic classes (*crests, shoulderslopes, backslopes, footslopes)* are split more evenly among the other classes. For example, in Table 9, a cumulative total of 10.21% of losses that occurred in the *backslopes* class between semantic landforms generated from the 25-cell window size and those generated in the 81-cell window size. Of the losses that occurred in the *backslopes* class, 1.43% were classified as *crests*, 3.08% as *shoulderslopes*, 2.68% as *footslopes*, and 2.68% as *flats* at the 81-cell window size. For losses in the *drainages* class at the same scale comparison, over 60% of the losses for *drainages* were locations reclassified as *flats* at the 81-cell window size. This proportion increased to over 80% as window size increased from 81 cells to 169 cells in Table 10.



Figures 29a-c. Change of class distribution over window sizes. Backslopes are symbolized with black and other classes dropped back in visual hierarchy to help highlight the behavior of a single semantic landform class. Window sizes are 25-cell (Figure 29a, left), 81-cell (Figure 29b, middle), and 169-cell (Figure 29c, right) and crisp landforms are derived from the Fuzzy Gamma operator. Based on visual results, one would expect the shape and perimeter indices for each class to reflect a greater magnitude of change at the finer scales; characteristics that may not be visible through analysis of the gross, crisp cell counts or the cross-tabulation matrix.

To visually compare the differences of cross-scale agreement between the three operators, the percent total agreement for each cross tabulation matrix within each operator was plotted against the ratio of the comparison map (larger window size) to the reference map (smaller window size), as shown in Table 4. This ratio decreased as the scales being compared coarsened. For example, percent agreement between the 9-cell (reference map) and 25-cell window sizes (comparison map) would result in a scale ratio of 2.78 (25/9), and a comparison of landforms generated at the 121-cell (reference map) and 169-cell (comparison map) window sizes would give a ratio of 1.4 (169/121). This ratio thus takes into account the size of the window sizes at the two scales being compared as well as the relative sizes of the scales. Figure 30 shows the relationship between percent agreements for each pair of scales plotted against a ratio of window size of the two scales for comparisons within each operator.

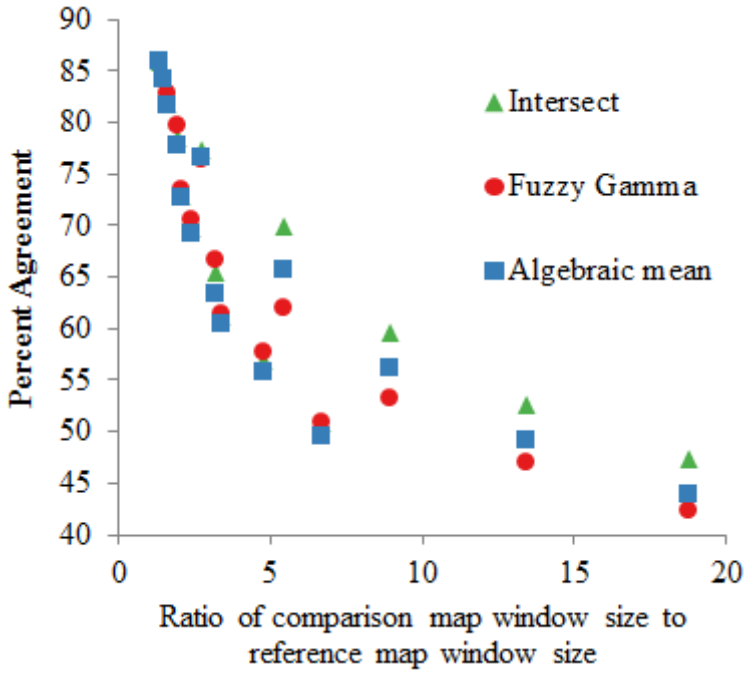


Figure 30. Percent agreement between the ratio of comparison map window size to reference map window size across all scales.

Percent agreement was related to both magnitudes of window sizes for scales being compared and the magnitude of the difference between them. For example, there was 77.27% agreement comparing landforms derived from the 9-cell and 25-cell window sizes (ratio of window sizes ((x) = 2.87), 78.39% between 25-cell and 49-cell window sizes (x=1.96), up to 85.87% agreement between landforms derived at the 121-cell and 169-cell window sizes (x=1.40). The percent agreement between the three operators was similar at small ratio sizes, (e.g. 86.35% agreement for the Gamma Operator, 85.91% for the Algebraic Mean, and 85.87% for the Intersect Operator between 121-cell and 169-cell window sizes, x=1.40). There is less variation between the levels of agreement of the three operators as both the scales coarsen and the size difference between the scales being compared decreases. Agreement between different overlay operators at the same scale was also examined (Figure 31). In this instance, the two datasets (reference maps and comparison maps) being analyzed represent crisp layers generated from overlay operators at the same scale. Highest agreement between the three operators at the same scale could be observed at the 25-cell window size.

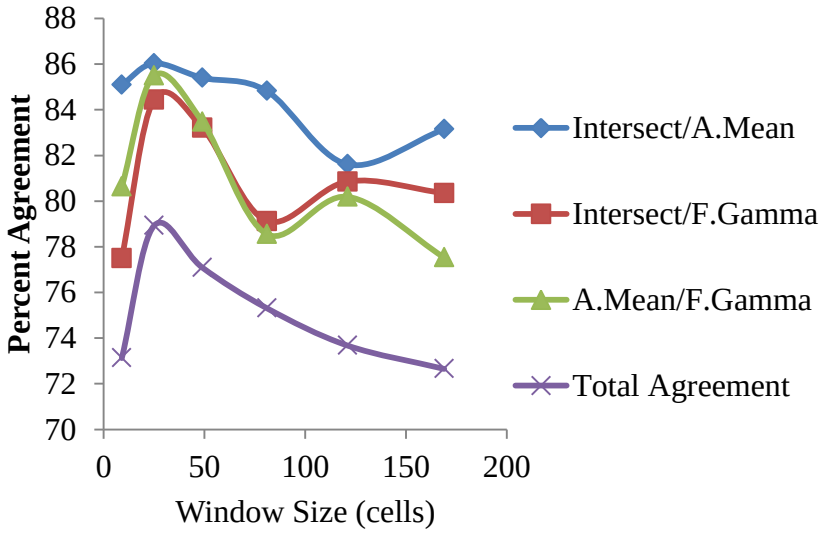


Figure 31. Agreement between the three operators over window sizes.

5.2 Similarity in the distribution of fuzzy membership values

To determine whether the distributions of fuzzy membership values come from similar distributions between the three different overlay operators, the distribution of fuzzy membership values of each semantic class were compared with the distributions of the same class derived from the other overlay operators within each scale. A 5% level of significance was used for the test. 1000 random points were generated across the study area and fuzzy membership values were extracted for each of the fuzzy landform classes derived from each overlay operator at every scale. Fuzzy membership values generated by the three overlay operators at the 49-cell window size for *crests* and *flats* are shown in Figure 32a-c and 33a-c. To attempt to scale the membership values for equal dispersion (range), the membership values were normalized to a 0-1 scale so the fuzzy membership values of the six fuzzy landforms derived for the same overlay operator summed to unity:

$$u_{i(r)} = \sum_{i=1}^{6} \frac{u_i}{\sum u_n} \qquad (4.19)$$

where $u_{i(r)}$ is the normalized fuzzy membership value for class *i*, u_i is the unscaled fuzzy membership value, and $\sum u_n$ is the sum of unscaled fuzzy membership values for the six fuzzy landform layers derived by an overlay operator for a single location. Distributions of normalized fuzzy membership values for the semantic landform class *flats* derived from the fuzzy gamma operator at the 49-cell window size is shown in Figure 34.

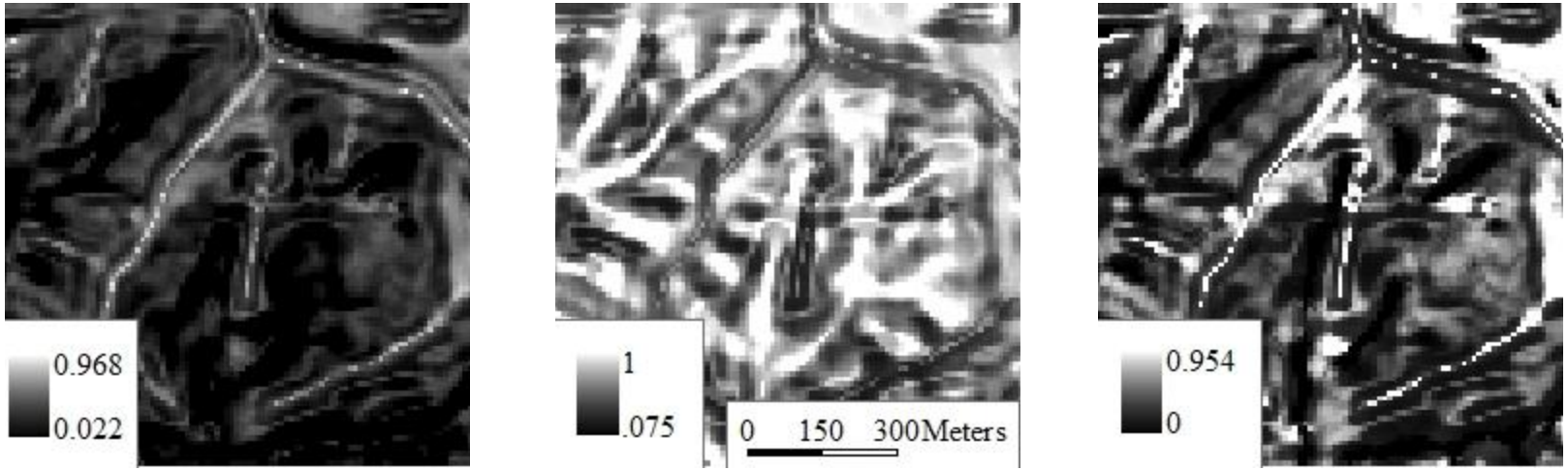


Figure 32a-c. Fuzzy layers for the semantic class *crests* from the intersect (32a, left), algebraic mean (32b, center), and fuzzy gamma (32c, right) overlay operators at the 49-cell window size. Lighter locations indicate a higher degree of membership to *crests*.

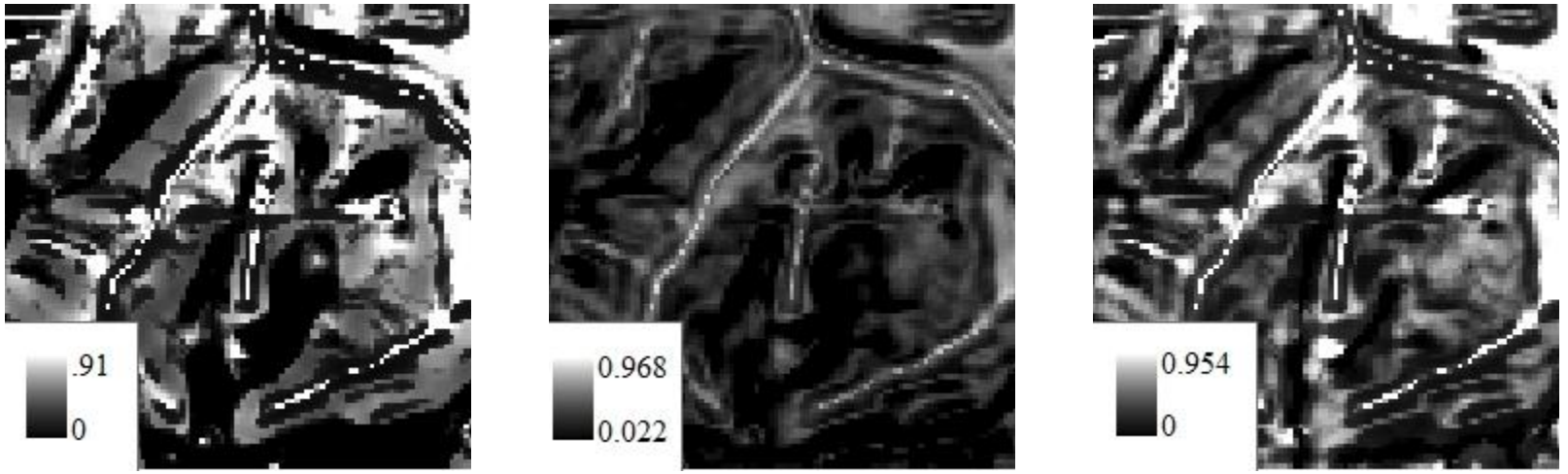


Figure 33a-c. Fuzzy layers for the semantic class *flats* generated by the intersect (33a, left), algebraic mean (33b, center), and fuzzy gamma (33c, right) overlay operators. Lighter locations indicate a higher degree of membership to *flats*.

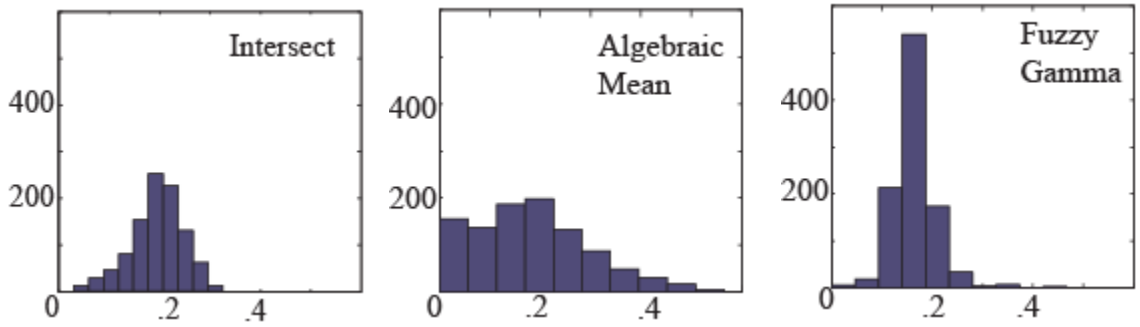


Figure 34. Normalized distributions of fuzzy membership values of the semantic landform *flats*. *Flats* was the only class to show close to normal statistical distributions.

Once the fuzzy membership values for the six fuzzy landform layers were normalized, MATLAB was used to run the Wilcoxon ranked-sum test. At each scale, fuzzy landform layers for each class were compared between operators, resulting in a total of 108 paired tests. Results of the tests are shown in Table 11, and scales were the distributions were significantly similar noted below.

Table 11. Times the null hypothesis for Wilcoxon Ranked-sum test is rejected (p =5%)

	Crests	**Shoulderslopes**	**Backslopes**	**Footslopes**	**Flats**	**Drainages**
Intersect-Gamma	0	0	0	0	0	0
A.mean-Gamma	0	0	1[a]	0	3[b]	1[c]
A.mean-Minimum	0	0	0	0	0	0

[a]= 25-cell window size

[b]=9-cell, 81-cell, 121-cell window sizes

[c]=25-cell window size

While there is similarity in the distribution of membership values for *flats* between the algebraic mean and fuzzy gamma operators, it is clear that the three operators distribute fuzzy membership values in significantly different ways independent on the underlying scale. However, this analysis has two significant shortcomings. First, it is independent of spatial arrangement due to the assumptions that were made in choosing the test on the underlying datasets. Secondly, there is no interaction or comparison with the other fuzzy membership values at each location. While we can assume that each operator distributes fuzzy membership values for each class in distinctly different fashions, the results of agreement between overlay operators at a each scale (72-76%, Figure 32) indicates there is a large amount of overlap in the maximum of membership method (MOM) in determining which semantic class has the dominant fuzzy membership value at each location. To compensate for the lack of spatial significance and interaction with other fuzzy membership values of semantic classes, a dual approach is needed to incorporate spatial significance, and crisp and fuzzy membership values.

5.3. Entropy and Classification Stability

To determine the uncertainty and ambiguity of crisp classes and the interaction between crisp classes and underlying fuzzy membership values, a combined crisp-fuzzy approach was used. This approach utilized three components to compensate for weaknesses of the previous analysis:

- The crisp membership value of each location,
- The classification entropy to incorporate the FMVs of every semantic class at a given location,

- Crisp class proportions at different levels of entropy, an extension of the idea of utilizing $\alpha - cuts$ for determining crisp memberships from a single fuzzy set; instead determining class proportions at different levels of uncertainty or vagueness

The approach outlined in section 4.6 was utilized to examine if there were specific semantic crisp classes that exhibited high levels of entropy, how the proportions of classes changed at different entropy levels, and if these relationships between class proportions and entropy levels were consistent across scale and between operators.

Normalized fuzzy membership values derived from the Wilcoxon ranked-sum test were used to calculate classification entropy (H) (equation 2.13-2.14). Entropy grids were calculated for each pairing of scales/operators. Entropy surfaces for the three overlay operators generated at the 49-cell and 169-cell window sizes are shown in Figures 35a-c and 36a-c. The increased smoothness in crisp semantic landforms layers with coarsening scale is reflected in a similar fashion in the spatial patterns of entropy. As scale coarsens, lower entropy values (areas of less ambiguity and uncertainty), were increasingly located in ridgetops (*crests* and *shoulderslopes)* and higher entropy values (areas of high ambiguity and uncertainty) were increasingly located in the areas in between (*slopes* classes*).* However, at the finer scales, these trends are less obvious based on visual inspection of the entropy surfaces.

Even after normalization of the FMV used to calculate entropy, there were distinct differences in the distributions of entropy between the three overlay operators at the same scale. Although the distribution of all three entropy surfaces were left-skewed, entropies generated by the intersect operator had a much smaller range and standard deviation than either distribution of entropy values generated by the algebraic mean or fuzzy gamma operators (Figure 36b and 37b). The frequency distributions of entropy values produced by the algebraic mean and fuzzy gamma

operators were very similar in shape, and had similar peaks and sub-peaks in frequency curves. This made analyzing the proportion of classes at standard entropy values (0.99, 0.95, 0.90, 0.85, 0.75) unfeasible. To attempt to analyze the three operators at the noted entropy levels while taking into account the different distributions of entropy, equation 4.22 was utilized to calculate a range-weighted entropy value specific to each pairing of scales and operators. The locations (grid cells) within each entropy level were identified, and proportional relationships of the six semantic classes were determined from at each entropy level. From these simple extractions, the frequency at which specific classes occurred at specific entropy levels was gauged for pairing of operators and scales. Classes with a natural degree of high or low uncertainty should exhibit similar proportional rankings compared to the other classes at each pairing of scales and overlay operators.

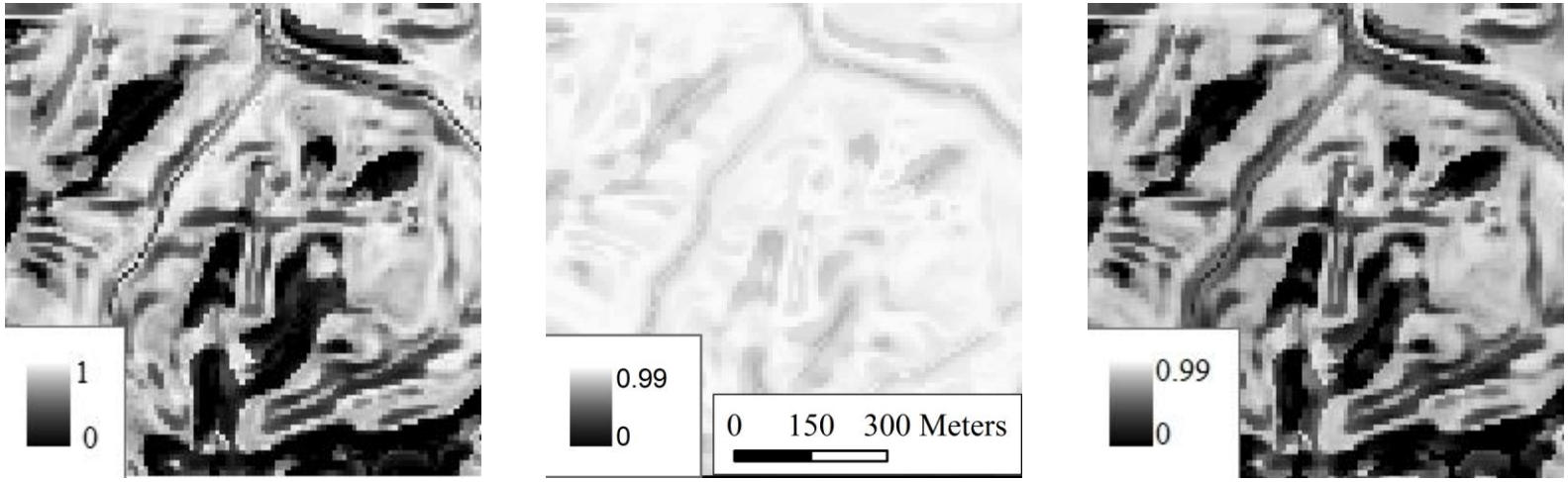


Figure 35a-c. Entropy surfaces for the intersect (35a, left), algebraic mean (36b, center), and fuzzy gamma (37c, right) overlay operators at the 49-cell window size. Higher values (lighter colors) indicate more ambiguous areas where fuzzy memberships of two of more classes are similar, while lower values (darker colors) show crisper areas dominated by few classes. Differences in range of entropy values between the algebraic mean and other two is clearly seen.

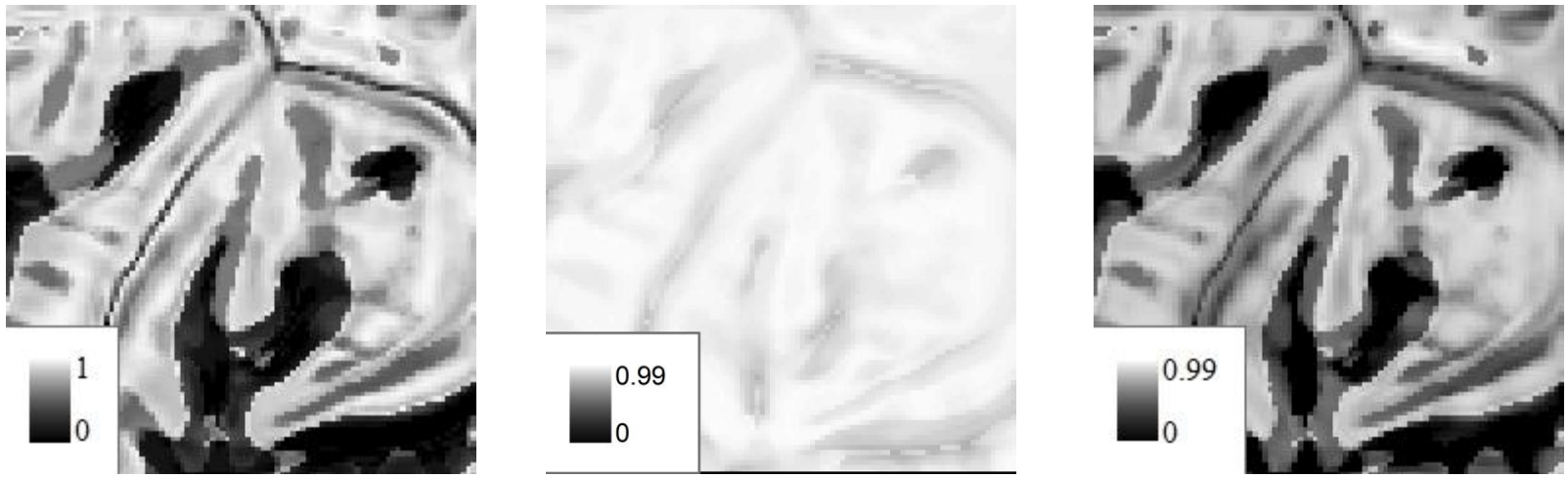


Figure 36a-c. Same entropy surfaces for the three operators from landforms generated at the 169-cell window sizes.

The surfaces representing locations within entropy levels at 0.95, 0.85, and 0.75 are shown for the three operators at the 49-cell window size in Figures 37a-c (intersect operator), 39a-c (algebraic mean operator), and 41a-c (fuzzy gamma operator). Accompanying graphs for each set of entropy figures (Figures 40a-c for the intersect operator, 42a-c for the algebraic mean operators, and 44a-c for the fuzzy gamma operator) show the proportional relationship of semantic classes at the entropy levels for the 9-cell, 49-cell, and 169-cell window size.

Before discussing the class proportions at each entropy level, it's worth noting some of the differences and similarities between the three operators in the area identified within each entropy level. The intersect operator generated larger areas (number of pixels) at the three highest entropy levels (0.99, 0.95, 0.90) than either the algebraic mean or fuzzy gamma operators. For example, the area identified at the highest level of entropy (0.99) for fuzzy surfaces generated by the intersect operator (from .2 to .4 % of the total area from the 9-cell to 169-cell window sizes) was three times the order of magnitude larger than the area at the same entropy-level generated by the fuzzy gamma operator (3.2×10^{-3} to 3.9×10^{-4} % of the total area) and two times the order of magnitude larger than generated by the algebraic mean operator (1.3×10^{-2} to 1.2×10^{-3} % of the total area). In contrast to the two other operators, the area (pixel count) of locations within the three highest entropy levels (0.99, 0.95, and 0.85) within the intersect operator increased as scale was coarsened. This effect of scale on this process was generally decreasive in the other two operators.

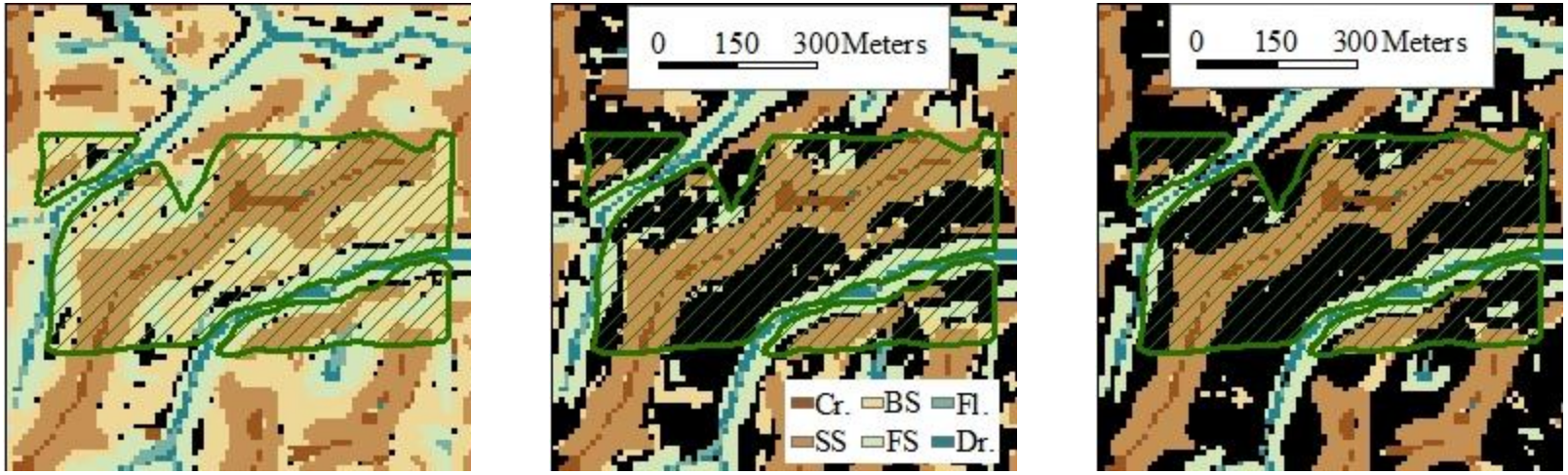
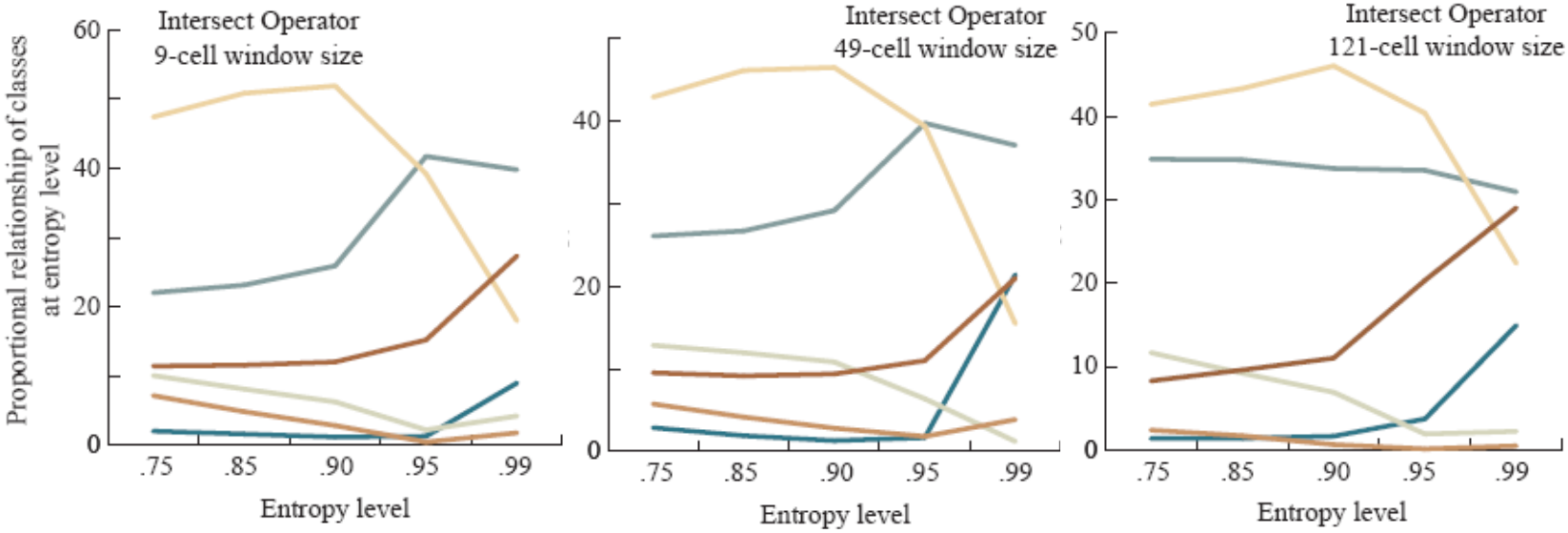


Figure37a-c. Levels of Entropy (0.95 - Figure 37a (left), 0.85- Figure 37b (center), 0.75 - Figure 37c (right)) calculated from fuzzy membership values of semantic landforms derived by the intersect operator at the 49-cell window size. Locations within each entropy level are shown in black. The green polygon indicates a CRP parcel



Figures 38a-c. Proportional relationships of classes at different levels of entropy in landforms generated by the intersect operator.

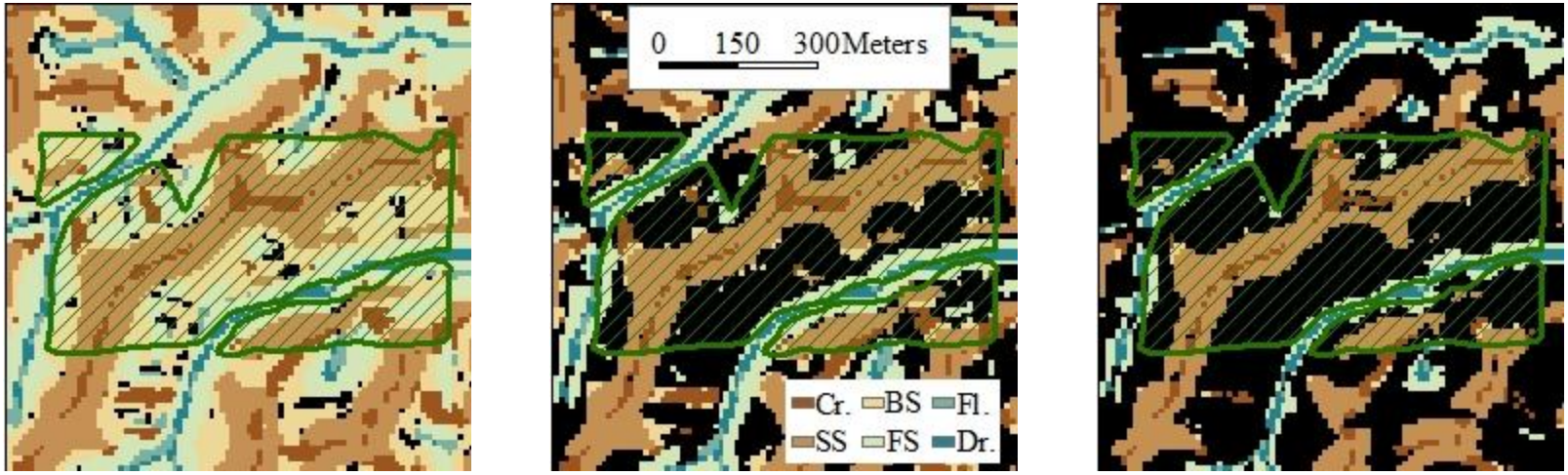
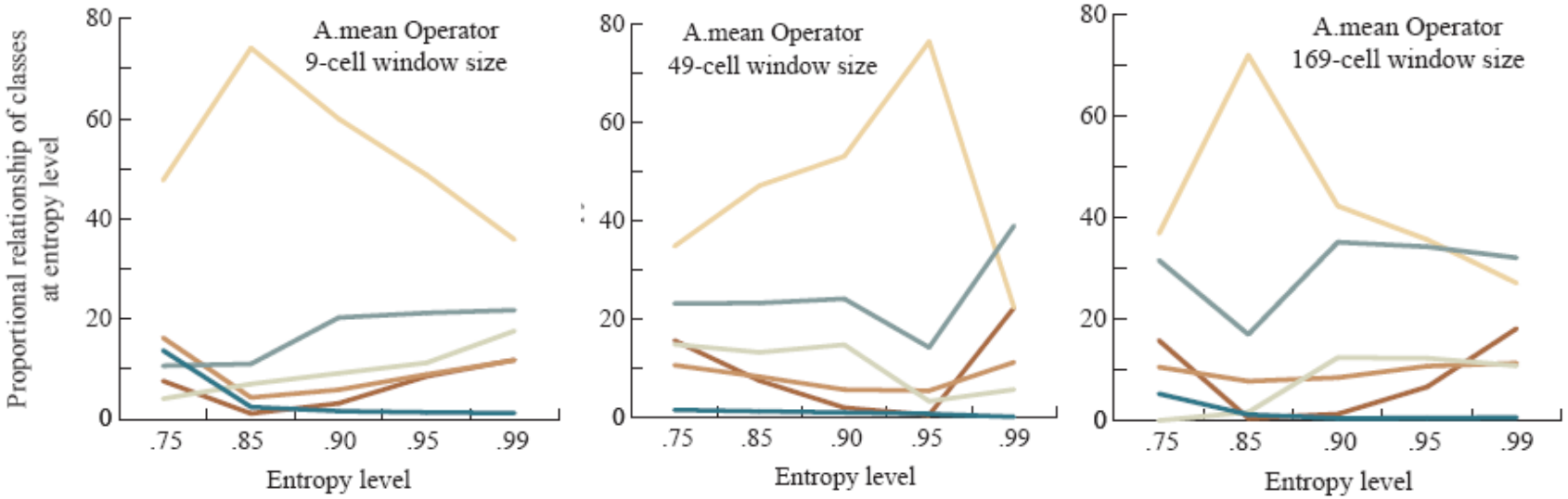


Figure39a-c. Locations of entropy levels 95-75% calculated from the algebraic mean operator at the 49-cell window size.



Figures 40a-c. Proportional relationships of crisp classes derived from the algebraic mean operator over multiple window sizes.

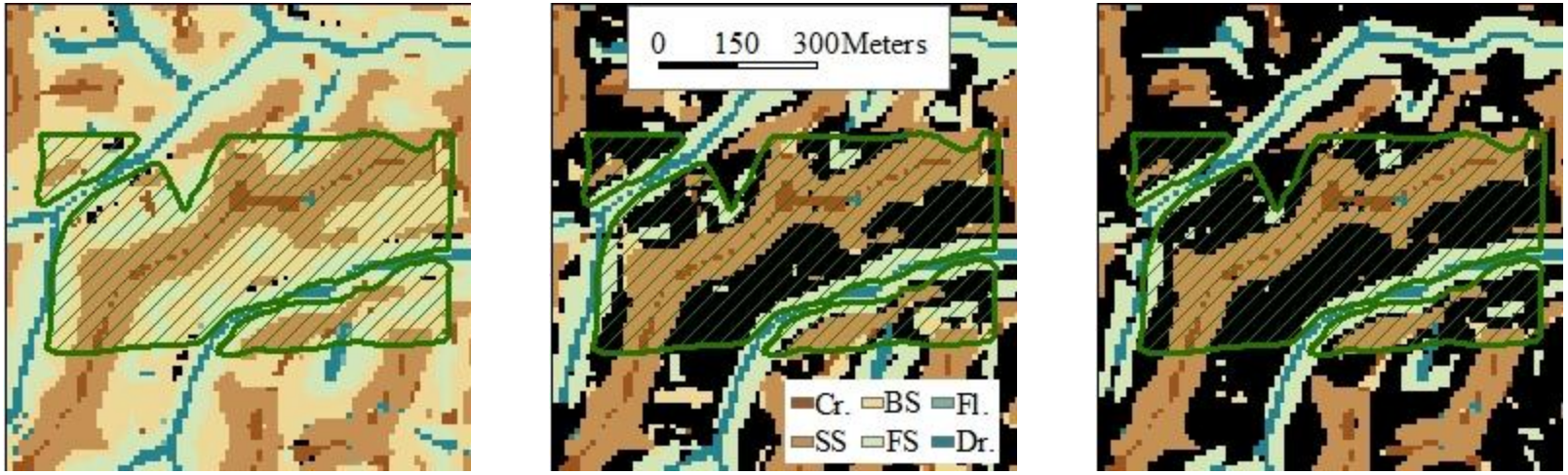
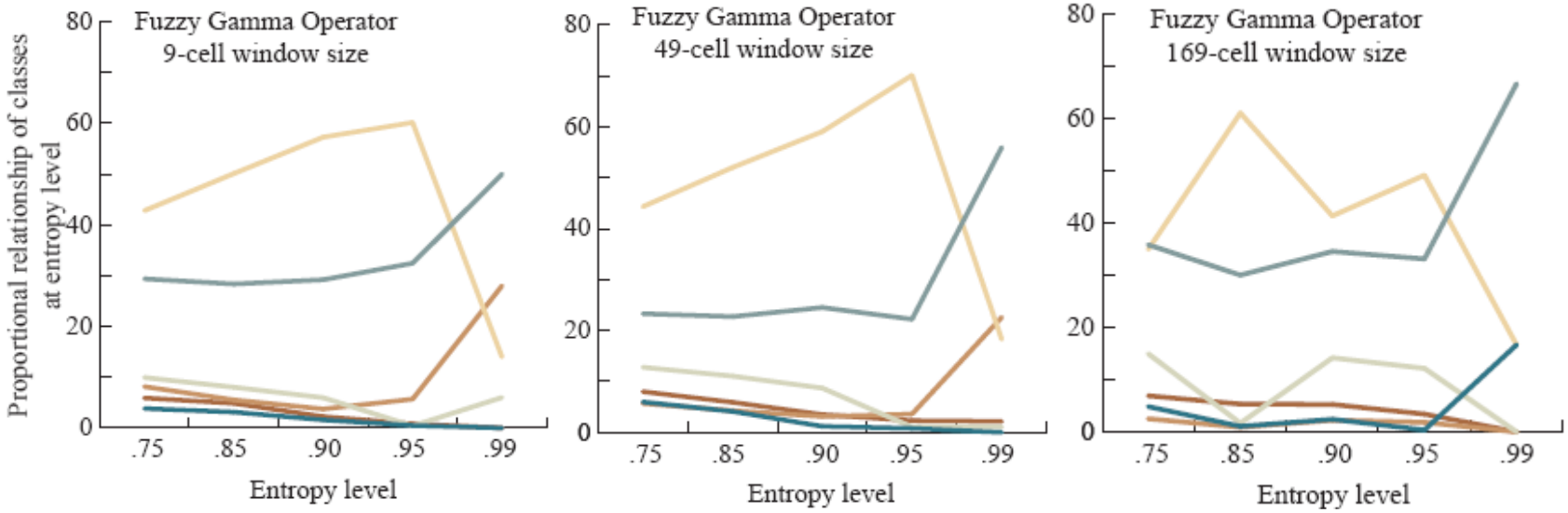


Figure41a-c. Locations of entropy levels 95-75% derived from the algebraic mean operator at the 49-cell window size calculated from fuzzy gamma operator.



Figures 42a-c. Proportional relationships of classes at the five levels of entropy generated from the fuzzy gamma operator.

The spatial patterns of entropy vary between the overlay operators at the 0.95 level, although this may be a function of the different number of locations within that entropy level between the operators (pg. 102). However, at the lower entropy levels tested (0.85 and 0.75), all three overlay operators exhibited similar spatial patterns of entropy locations; these locations were increasingly located within the *slopes* classes, with only scattered locations within the extreme top and bottom of the hillslope profiles (*crests, shoulderslopes, drainages).*

There were three major similarities that were exhibited with this procedure. First, across all operators, scales, and levels of entropy, the semantic classes *backslopes* and *flats* had the highest proportions at each entropy level. This was remarkably consistent across all scales, operators, and level of entropies, with one exception (Figures 39, 41, and 43). This was the proportion of classes located at the highest level of entropy (0.99). At the highest level of entropy, there were a larger variety of classes with large proportions than at any other entropy level. Within the intersect operator, *crests* and *drainages* had much higher proportions at the 0.99 entropy level across different scales (8.96% to 32.62% for *drainages*; 18.10% to 29.17% for *crests*) than at the other entropy levels (<4% for *drainages*; <11% for *crests*). Within entropy levels derived from the algebraic mean operator, it was *crests, footslopes,* and *shoulderslopes* that exhibited this same behavior. In those derived by the fuzzy gamma operator, it was primarily *crests* at the finer scales, and *drainages* at the coarser scales. Finally, *backslopes* exhibited a much lower proportion at the 0.99 entropy level than at the remaining entropy values. As scale was coarsened, the proportion of *backslopes* gradually decreased, *flats* increased, with inconsistent trends in the other four classes.

The final aspect was to determine the locations of high entropy jointly shared between the three operators, and the resulting proportions in different crisp classes. Pixel locations common

to all three operators at two entropy levels (0.90, 0.95) were determined, and used to determine class proportions. This was done across three window sizes (9-cell, 49-cell, and 169-cell). Results are shown in Table 12.

Table 12. Entropy locations (0.95, 0.90) and their crisp proportions that are common to all three operators (Cr. – *crests*, SS –*shoulderslopes*, BS – *backslopes*, FS – *footslopes*, Fl. – *flats*, Dr. – *drainages*).

	Entropy >= 0.95			**Entropy >=.90**		
	9	**49**	**169**	**9**	**49**	**169**
Total Pixels	542	2530	533	290879	324890	173429
Proportions						
Cr.	0	1.98	0.19	0.37	7.53	0.12
SS	0	1.23	5.25	3.26	1.90	1.58
BS	50	69.76	57.79	71.39	60.46	56.72
FS	0	0.20	0.19	6.81	7.78	11.34
Fl.	50	26.84	36.59	17.92	25.78	30.22
Dr.	0	0.08	0.00	0.25	0.06	0.02

The proportion of high entropy locations common to all three operators was greatest in the mid-scale range (2530 cells for locations in top 95% of entropy levels and 324,890 for locations in the 0.90 level). The proportions of *backslopes* were just as in these common entropy locations (50 to 71.39%) as they were within a single operator (39 to 78%, Figures 39 through 43).

PART VI

DISCUSSION AND CONCLUDING REMARKS

The goal of this project was to 1) analyze change trajectories of crisp geomorphometric classes across scales derived from fuzzy set theory, 2) determine how different fuzzy logic operators impact the distribution of crisp and fuzzy membership values, 3) determine if spatial locations of high entropy represent transitional states between semantic geomorphometric classes and if high degrees of entropy are related to specific classes, and 4) speculate implications of 1-3 on applications of multi-sale geomorphometric models to vegetative and pedological mapping, or land use-land cover parcels such as the Conservation Reserve Program. The discussion presented here will follow the train of thought of the research goals.

Each terrain attribute showed a markedly different behavior across the different scales of measurement (Figure 25, Table 5), with slope shown to be the most volatile attribute across scales. Implications for the magnitude of change of terrain attributes are different within each operator. This is an important consideration, as within the intersect overlay, low membership values of a single terrain attribute can act as the constraining factor in assigning the final membership from the overlay operation to a location. In contrast, with the algebraic mean or fuzzy gamma operators, the values of all semantic constructs are weighted or taken into account. Within a physical process-based investigation, the use of a minimum operator might be more relevant than other fuzzy operators, as many physical and ecological processes are steered based on single limiting factors (Qin et al, 2009).

The crisp layers derived from the fuzzy surfaces from the three overlay operators demonstrated the same general change trajectories between the six crisp classes across scales

(Tables 6-8). The semantic classes *crests* and *flats* were increasive in area as scale was coarsened, while the three *slopes* classes were decreasive in area. *Drainages* was the only class that showed a single optimal scale for detection area (Figure 26, 27a-c). This is more evidence that the most accurate representation of specific classes is optimized by a specific combination of scale of measurement and DEM resolution, which are then dependent on the terrain characteristics of the study area, such as terrain roughness. In features such as streams, the size of these adjacent flats can be related to stream order.

This is unique in several ways. Being primarily a linear feature, drainages showed the greatest connectivity at the larger scales (9-cell and 25-cell window sizes). The mainstem of Big Spring Creek is clearly shown delineated in Figure 27. As scale decreased, areas and widths of the drainage segments increased as connectivity between them decreased (Figures 26 & 27). The greatest increase in area for drainage features occurred in areas considered as zero-order basins or first-order streams, while the larger order streams, such as features representing Big Spring Creek, became increasingly dissipated and disconnected as scale was coarsened.

Within higher-order stream features, terrain attributes and semantic constructs take on more of the characteristics of adjacent floodplains (*flats)* than the handful of cells defining the channel features as the window size over which the terrain attributes are measured increases. The extent and degree of flatness of these features are correlated with stream order, so this effect is not seen in small-order drainage features, such as 'zero-order' catchments in headwater and channel initiation areas. Utilizing the theory of stream order, membership to semantic constructs such as steepness, and tangential and profile concavity increase as stream order decreases. This represents a problem in attempting to link feature hierarchy, such as stream ordering, and detection scale. Stream initiation thresholds could have been utilized to create stream networks

based on flow accumulation values, but it was decided not to incorporate global geomorphometric parameters into the landform models based on lack of specific knowledge of the study area, and instead to base the classes strictly on surface-shape parameters to create an approach that was as generic as possible.

Based on the cross-tabulation matrices, the percent agreement of crisp landforms between two scales increased as scale became coarse. When comparing the largest window sizes, the three operators had similar agreement (Figure 30). Change trajectories generally reflected the trends seen in gross cell counts between scales, and specific trajectories between *drainages* and *flats* mentioned was evidenced (Table 9 and 10). One missing aspect in the examination of multi-scale landform change is the absence of any change analysis on the spatial properties of individual landform classes or objects (patches), and the importance of defining adjacency and neighborhood relationships between the objects (Arrell et. al., 2007; Moller et. al 2008). Object-based segmentation has accounted for some of these shortcomings. However, a greater variety of geospatial procedures, such as landscape metrics, or shape and pattern descriptors may be required to completely understand change patterns in landform characteristics across scales above and beyond non-spatial methods that deal only with categorical change such as the cross-tabulation matrix. For example, the percent agreement between the two scale comparisons (66.62% Table 9 and 73.37% for Table 10) are clearly not reflective of the complete picture of semantic landform class changes in the two scale steps when the crisp landforms are examined visually. There is clearly a larger amount of noise, patchiness, and perforation exhibited by the semantic landforms at the 25-cell window size that is absent from landforms at the other two scales.

To analyze distribution of the fuzzy membership values for each class between the three operators, the Wilcoxon ranked-sum test was utilized. The three overlay operators distributed fuzzy memberships in significantly different ways, with only six significant results out of 108 tests (Table 11). When computing fuzzy membership values for a single landform (Shi et. al, 2005; Gallant and Wilson, 2003), or testing for crisp areas generated by an alpha-cut, this test would have more meaning. The lack of consideration of spatial relationships and exclusion of the fuzzy membership values of other classes placed limitations on the spatial interpretations of the Wilcoxon test.

An alternative to the Wilcoxon test would have been a paired mean difference t-test to test if the difference of means between FMV generated from a pair of overlay operators were significantly different from each other. Such a result would have given more weight to the spatial location since it compare a pair of FMVs for a single class/different operators at a single location, rather than the two populations ranked numerically prior to comparison. There were two assumptions made regarding the fuzzy membership values that prevented use of such a test. First, because the overlay operators generated fuzzy membership values in different ways based on the underlying mathematics and logical operators, it was assumed the distributions of fuzzy membership values (FMV) were the most part not normal. Few classes exhibited distributions close to Gaussian (Figure34), and most did not. Secondly, because the overlay operators utilized at least one identical semantic construct in each of their logical operations, it could not be assumed the three different FMVs derived for a class by the three operators came from independent distributions.

Another weakness behind analyzing fuzzy membership values distributions via the Wilcoxon ranked-sum test is not accounting for diversity of all six FMV at a single location.

Many previous approaches to analyzing fuzzy geomorphometric landforms have constituted approaches at examining effects of multiple window sizes or resolutions (Gallant and Dowling, 2003), attribute distance-similarity (Shi et al., 2005) or alpha-cuts (Minglian, 2009) on singular fuzzy morphometric classes, so there such a measure is still useful in understanding how the parameters of the semantic import model (central concept, dispersion index) or different overlay operators impact the range of FMV for a single semantic class.

However, contrarily to the Wilcoxon results, we can assume some similarity in the spatial distribution of maximum membership values between the operators based on the results of the Boolean analysis. The main evidence for this is the percent agreement of classification ranging from 72% to 78% (Figure 31) when classes from all three operators were compared at each scale. Thus, the trajectories of maximum of membership overlay method in assigning final crisp membership were consistent between the three operators.

The final aspect of the results to discuss is the proportional relationships of classes at the different level of entropies (0.99, 0.95, 0.90, 0.85, and 0.75). This portion of the results shows how levels of vagueness and ambiguity can be higher in specific classes independent of scale and overlay operators. As with the crisp landforms and terrain attributes, noise and variance decreased within entropy surfaces calculated at coarser scales, with areas of low entropy increasingly concentrated around *crests* and *shoulderslopes* and *drainages,* and with areas of high entropy in the *slopes* in between (Figures 35-36). This may be construed as an artifact of the semantic constructs and specific terrain attributes used to define each class, but the approach of creating landform classes with different parameters is a well-defined in its theoretical and conceptual foundation (Speight, 1990; Coops et al., 1998; MacMillan et al., 2000; Qin et. al. 2009).

For four out of the five entropy levels within each operator (0.95, 0.9, 0.85, and 0.75), *backslopes* and *flats* consistently had the highest proportion of all classes (Figures 37, 39, and 41). This was remarkably consistent across all scales. The only inconsistency is the proportional classes identified at the highest level of entropy (0.99). The dominance of *backslopes* decreased, and there were greater proportions of other classes. This can be construed in two ways:

1) the semantic classes *backslopes* and *flats* can truly be interpreted as transitional classes, in that *backslopes* act as the transition between uplands and lowlands area (sediment detachment and sediment deposition), and *flats* primarily as the transitional zone between *slopes* and *drainages*, the area of the highest level of hydrologic activity (figure 21).

2) the areas of highest uncertainty (0.99-level entropy locations) are less tied to specific classes and are more tied specific locations of high instability, due to high variation in the land surface. The class proportions identified in these lie outside of patterns shown at other levels of entropy, and are independent of scale or overlay operator.

3) Locations at the highest level entropy may be independent of the overall uncertainty and ambiguity of the classification system and more as a result of artifacts from DEM processing or measurement. In considering previous work (Wood, 2001; Shary et al., 2002), these locations may exist purely as a result of DEM artifacts or random errors, which are not a natural consequence of scale or overlay operation. However the fact

that these conditions exist even at entropy locations at the largest window size lends doubt to this idea, as it was previously shown how visually and quantitatively local nose and variation was removed by increasing the scale of measurement

The evidence for point one is fairly clear. Besides those two classes generally having the highest entropy levels across scales and between operators, *backslopes* as noted on p. 93, had the highest proportion of losses within the cross-tabulation matrices (Tables 9 & 10), which were more evenly proportioned among four other classes (*crests, shoulderslopes, footslopes,* and *flats*) than any other of the semantic classes.

The most direct application for the knowledge gained in this thesis is its application in accounting for accuracy and uncertainty in landform mapping. Since assumption of topographic position is inherent to vegetative and pedological mapping, a key component in accuracy assessment efforts should be to account for different levels of uncertainty in different classes, and specific classes may display an inherent higher of level of uncertainty, and systems will behave differently depending on what scale landform-terrain attribute relationships are derived, and the overlay or cumulative operators use to map the crisp sets.

For specific application within the CRP, the crisp classes *backslopes* (at finer scales) (Figure 28) and flats (at coarser scales) were shown to be dominant classes, and *backslopes* were shown to be present at proportions higher in CRP than in the entire study area. Because of the high level of uncertainty in crisp representation of these two classes, targeted land management, conservation practices, or ecological modeling incorporating topographic position (such as the ELT (figure 13)) as an assumed variable may be introducing an unintended source of error into their models. It is never encouraged to imply process from surface shape (Gallant and Wilson,

2000), but remote sensing and GIS applications utilizing supervised or semi-automated machine-learning approaches for large-scale ecological mapping and monitoring could benefit from incorporating and weighing uncertainty of different objects and classes in underlying data layers.

Within this project, one source of uncertainty, boundary vagueness (The Law of the Excluded Middle, section 2.3.2), was accounted for through the use of fuzzy set theory that allows defining boundaries in datasets defined as continuous fields. By testing the effects of three different overlay operators on fuzzy membership values, and repeating the analysis over multiple scales we accounted for thematic vagueness and allowed for locations to have multiple fuzzy memberships (The Law of Contradiction, section 2.3.2), and a location to take on different crisp class values under the use of different operators and across scales. By accounting for these measures of uncertainty, it was possible to determine that several classes in a geomorphometric hillslope model are naturally unstable independent of the operator and scale of analysis utilized.

There are some constraints and potential gaps. The first of which is defining the parameters of the semantic import models: the central concept, dispersion index, and crossover point (2.3.4). The approach taken was to begin to model the variable based on their statistical distribution throughout the study area and gradually tweak them until they seemed a good visual fit. With no expert opinion or opportunities to verify the parameters, an approach based on generality appeared to be the best procedure. As previously mentioned for the same reason, it was decided not to use global geomorphmetric parameters, such as the wetness or compound topographic index, stream initiation thresholds, or stream-to-ridge distance indices.

Another issue with the approach used here is the semantic import models do not take into account changes in the distribution of terrain attributes from headwaters to the outlet. For a catchment-based study are, the rigidness of the semantic import parameters (central concept,

dispersion index) do not allow for flexibility when the character of terrain changes. Previous approaches to solve this problem included calculating terrain attributes up to the longest wavelength of terrain features (ridge to ridge distance, so to speak) (Pike, 1988), or use of geostatistical interpolations to model neighborhood relationships and distance-decay of variance of terrain attributes (Wood, 1996). However, with time and computing constraints it was infeasible to perform kriging or calculate semi-variograms at every single cell for every single semantic construct over a large scale, but it is an approach to consider in the future.

The decision on what terrain attributes to use/not to use provided another source of uncertainty. Besides *crests* and *drainages*, none of the other classes made accommodations for tangential curvature for converging or diverging flow patterns, like Dikau or Pennock's hillslope models (Figures 14 and 15). As mentioned in introduction to the methods, it was decided to leave these measures out to simplify the hillslope model and be able to draw some meaning to cross-scale change trajectories.

Finally, there was decision to use the CRP as an exemplary unit of analysis. The results and implications have demonstrated that natural classes of geomorphometric objects exhibit varying levels of uncertainty when captured as objects in a GIS environment, and this would affect the use of the variable topographic position in any kind of land surface or land cover mapping. Because of the large variety of ecological benefits that CRP provides, along with the vast array of management strategies to optimize the benefits of land that is currently enrolled (references), a multi-scale approach to delineating geomorphometric units provides a flexible approach to quantifying landscape position. This is an important point, as many of the current uses for CRP and similar programs deal with environmental issues that occur at a variety of scale from micro-habitat of endangered species, carbon sequestration at the field-scale, preservation of seasonal

habitat for migratory wildfowl along intercontinental flyways, aquifer recharge, and original program goals of reduction of soil erosion and improvement of water quality.

References

Aandahl, A. R., 1948. The characterization of slope positions and their influence on the total nitrogen content of a few virgin soils of western Iowa, Soil Science Society of America Proc. 13:449-454.

Allen, A.W., Vandever, M.W. (eds.), 2005. The Conservation Reserve Program: planting for the future. Proceedings of a National Conference, Fort Collins, Colorado, June 6-9, 2004: U.S. Geological Survey, Fort Collins Science Center Scientific Investigations Report 2005-5145. 248 p.

ArcGIS 9.3:Environmental Systems Research Institute. ArcGIS: Release 9.3 [software]. Redlands, California: Environmental Systems Research Institute, 1999-2008.

Arrell, K., Fisher, P.F., Tate, N.J., Bastin, L., 2007. A fuzzy k-means classification of elevation derivatives to extract the morphometric classification of landforms for Snowdonia, Wales, Computers and Geosciences, 33 (10): 1366-1381.

Babcock, B.A., Lakshminarayan, P.G., Wu, J., 1995. The economic, environmental, and fiscal impacts of a targeted renewal of conservation reserve program contracts. Working paper 95-WP 129. Ames, IA: Center for Agricultural and Rural Development, Iowa State University.

Baker, J.S., Galik, C.S., 2009. Policy options for the conservation reserve program in a low-carbon economy. Working Paper. Dunham, NC: Climate Change Policy Partnership, Duke University. http://www.nicholas.duke.edu/ccpp/ccpp_pdfs/low.carbon.policy.pdf.

Band, L.E., 1986. Topographic partitions of watersheds with digital elevation models. Water Resources Research. 22(1), 15-24.

Batie, S. S., Schulz, M.A., Schweikhardt, D.B., 1997. A continuation of environmental conservation policy: the conservation reserve program. Staff Paper 97-16. Department of Agricultural Economics, Michigan State University, East Lansing.

Behrens, T., 2003. Digitale Reliefanalyse als Basis von Boden-Landschafts-Modellen –

am Beispiel der Modellierung periglaziärer Lagen im Ostharz. Boden und Landschaft 42 42, 189 pp. Giessen.

Beven K.J., Kirby M.J., 1979. A physically based variable contributing area model of basin Hydrology. Hydrologic Sciences Bulletin. 24, 43-69.

Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: the fuzzy c-means clustering algorithm. Computers and Geosciences. 10(2), 191-203.

Brandli, M., 1996. Hierarchical Models for the Definition and Extraction of Terrain Features. in: Burrough, P.A., Frank, A.U. (Eds), Geographic Objects with Indeterminate Boundaries, PA: Taylor & Francis Inc., Bristol, pp. 257-270.

Brown, D.G., 1998. Classification and boundary vagueness in mapping presettlement forest types, International Journal of Geographical Information Science. 12(2):105-129.

Bull, W.B., 1975. Allometric change of landforms. Geological Society of American Bulletin 86(11), 1489-1498.

Burrough, P.A., 1989. Fuzzy mathematical methods for soil survey and land evaluation. Journal of Soil Science. 40, 477-492.

Burrough, P.A., MacMillan, R.A., Van Deursen, W., 1992. Fuzzy classification methods for determining land suitability from soil profile observations and topography. Journal of Soil Science. 43, 193-210.

Burrough, P.A., van Ganns, P.F.M., Hootsmans, R., 1997. Continuous classification in soil survey: spatial correlation, confusion, and boundaries. Geoderman. 77, 115-135.

Burrough, P.A., Wilson, J.P., van Gaans, P.F.M., Hansen, A. 2001. Fuzzy k-means classification of topo-climatic data as aide to forest mapping in the Greater Yellowstone Area. Landscape Ecology. 16, 523-646.

Burt, T.P., Butcher, D.P., 1986. Development of topographic indices for use in semi-distributed hillslope runoff models. Zeitschrit fur Geomorphologie, N.F. Supplementband 58, 1-19.

Chorley, R.J., Malm, D.E.G., Pogorzelski, H.A., 1957. A new standard for estimating drainage basin shape. American Journal of Science. 255, 138-141.

Chorley, R.J., Schumm, S.A., Sugden, D.E., 1984. Geomorphology. Methuen, London, 607 p.

Coops, N.C., Gallant, J.C., Loughhead, A.N., Mackey, B.J., Ryan, P.J., Mullen, I.C., Austin, M.P., 1998. Developing and testing procedures to predict topographic position from digital elevation models (DEMs) for species mapping (Phase 1). Report to Environment Australia, Client Report No. 271, CSIRO FFP, Canberra, 56 pp.

Dalrymple, J.B., Blong, R.J, Conacher, A.J., 1968. A hypothetical nine unit landsurface model. Zeitschrit fur Geomorphologie. 12: 60-76.

Dao, T.H., Stiegler, J.H., Banks, J.C., Bogle-Boerngen, L., Adams, B., 2000. Post-contract grassland management and winter wheat production on former CRP fields in the Southern Great Plains. Agronomy Journal. 92, 1109–1117.

Dehn, M., Gärtner, H., Dikau, R., 2001, Principles of Semantic Modeling of Landform Structures. Computers & Geosciences. 27, 1005-1010.

Desmet, P.J.J. and Govers, G., 1996. A GIS-procedure for automatically calculating the USLE LS-factor on topographically complex landscape units. Journal of Soil and Water Conservation. 51 (5), 427-433.

Dikau, R., 1989. The application of a digital relief model to landform analysis in geomorphology. In: Raper, J. (Ed), Three-dimensional applications in Geographical Information Systems. Taylor & Francis, London, pp. 51-77.

Dikau, R., 1990. Geomorphic Landform Modeling Based on Hierarchy Theory. In 4th International Symposium on Spatial Data Handling, Zurich, Switzerland. pp.230-239.

Dragut, L., Blaschke, T., 2006. Automated classification of landform elements using object-based image analysis. Geomorphology. 81(3-4): 330-344.

Evans, I.S., 1972. General geomorphometry, derivatives of altitude and descriptive statistics. In: Chorley, R.J. (Ed), Spatial Analysis in Geomorphology, Methuen, London, pp. 36-80.

Evans 1980. An integrated system of terrain analysis and slope mapping. Zeitschrift fur Geomorphologie, N.F. Suppletmentband 36, 274-295.
Evans, I.S., 1984. Properties of the Wessex land surface: an investigation of dimensionality and the choice of key variables. Durham Geomorphometry Report 8, Department of Geography, University of Durham, England, 131 pp.

Evans, I.S., Cox, N.J., 1999. Relations between land surface properties: altitude, slope and curvature. In: Hergarten, S., Neugebauer, H.J. (Eds), Process Modelling and Landform Evolution. Springer, Berlin, pp. 13– 45.

Falcidieno, B., Spagnuolo, M., 1991. A new method for the characterization of topographic surfaces, International Journal of Geographic Information Systems. 5, 397-412.

FAO, 1995. Global and national soils and terrain digital databases (SOTER): Procedures manual.van Engelen V.W.P., and Ting-tiang W. (EDS.), World Soil Rsources report 74 Rev 1. Land and Water Division, Food and Agriculture Organization of the United Nations. ISSN0532-0488. ISRIC, Wageningen, 128 p.

Ferris, J., Siikamäki, J., 2009. Conservation Reserve Program and Wetland Reserve Program. RFF Backgrounder, Outdoor Resources Review Group.
http://www.rff.org/News/Features/Pages/OutdoorResourcesReviewGroup-Pubs.aspx

Fisher P., Wood, J., Cheng, J., 2004. Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. Transactions of the Institute of British Geographers. NS 29, 106–128.

Florinsky, I.V., Kuryakova, G.A., 1996. Influence of topography on some vegetation cover properties. Catena. 27, 123-141.

Florinsky, I., 1998. Accuracy of local topographic variables derived from digital elevation models. International Journal of Geographical Information Science. 12 (1), 47– 61.

Foody, G.M., 1997. Fully fuzzy supervised classification of landcover from remotely sensed imagery with an artificial neural network. Neural Comput. Appl. 5, 238–247.

Franklin, J., McCullough, P., Gray, C.. 2000. Terrain variables used for predictive mapping of vegetation communities in southern California. In Terrain Analysis: Principles and applications. Wilson, J.P. and J.C. Gallant (EDS). Wiley and Sons, New York, 331-353.

Franti, T. G., Christiansen, A.P., Holshouser, D.L., Klein, R.N., Roeth, F., Schild, J.A., Zoubek, G.L., 1996. Agricultural Management Practices to Reduce Atrazine in Surface Water. Lincoln, NE: Cooperative Extension, Institute of Agriculture and Natural Resources, University of Nebraska–Lincoln. A-20.

Gallant, J.C., Dowling, T.I., 2003. A multi-resolution index of valley bottom flatness for mapping depositional areas. Water Resources Research. 39 (12), 1347-1360.

Gorsevski P.V., Gessler, P.E., Jankowski, P., 2003. Integrating a fuzzy k-means classification and a Bayesian approach for spatial prediction of landslide hazard. Journal of Geographical Information Systems. 5, 223-251.

Giles, P.T., Franklin, S.E., 1998. An Automated Approach to the Classification of the Slope Units Using Digital Data. Geomorphology. 21, 251-264.

Heerdegen, R. G., Beran, M.A., 1982, Quantifying Source Areas through Land Surface Curvature and Shape. J. Hydrol. 57, 359-373.

Hengl T., Evans, I.S., 2009. Mathematical and digital models of the land surface. In: HenglT and Reuter HI (EDS), Geomorphometry-Concepts, Software, Applications. Developments in Soil Science, vol. 33, Elsevier, Amsterdam, 31-63.

Horn, B.K.P., 1981. Hill shading and the reflectance map, Proceedings of IEEE. 69(1), 14–47.

Jenson S. K., Domingue, J.O., 1988. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis, Photogrammetric Engineering and Remote Sensing. 54(11), 1593-1600.

Jordan, G., Csillag, G., Szucs, A., and Qvarfort, U., 2003. Application of digital terrain modelling and GIS methods for the morphotectonic investigation of the Káli Basin, Hungary. Zeitschrift für Geomorphologie. 47, 145-169.

Kandel, A., 1986. Fuzzy Mathematical Techniques with Applications. Addison Wesley. New York.

Kaufman, A., 1975. Theory of Fuzzy Subsets, vol. I. Academic Press, New York.
Kirkby, M.J., Chorley, R.J., 1967. Throughflow, overland flow and erosion. Bull. Int.Assoc. Sci. Hydrol. 12(3), 5--21.

Klingseisen, B., Metternich, G., Paulus, G., 2008. Geomorphometric landscape analysis using a semi-automated GIS-approach. Environmental Modelling & Software. 23(1), 109-121

Klir GJ, Yuan, B, 1995. Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, New Jersey.

MacMillan, R.A., Shary, P.A., 2009. Chapter 9 Landforms and landform elements in geomorphometry.In: Hengl T and Reuter HI (EDS), Geomorphometry-Concepts, Software, Applications. Developments in Soil Science, vol. 33, Elsevier, Amsterdam, 227-254.

Macmillan, R.A., Goddad, T.W., Nolan, S.C., Pettapiece, W.W., 2000. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. Fuzzy Sets and Systems. 113, 81– 109.

Mark, D.M., 1975. Geomorphometric parameters: a review and evaluation. Geografiska Annaler. 57(3-4), 165-177.

Mark, D.M., 1984. Automatic detection of drainage networks from digital elevation models. Cartographica. 21, 168-178.

MATLAB version 6.5.1. Natick, Massachuesetts: The Mathworks Inc., 2010.

Metternicht, G., Gonzalez, S., 2005. FUERO: Foundations of a fuzzy exploratory model for soil erosion hazard prediction. Environ. Model. Softw. 20 (6), 715-728.

McBratney, A., de Gruijter, J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Sciences. 43, 159-175.
McBratney A.B., Moore, A.W., 1985. Application of fuzzy sets to climatic classification. Agricultural and Forest Meteorology. 35(1-4), 165-185

McBratney, A.B., Odeh, I.O.A., 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma. 77, 85 -114.

McMaster, K.J., 2002. Effects of digital elevation model resolution on derived stream network positions. Water Resources Research. 38(4), 131- 138.

Minar J., Evans, I.S., 2008. Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping. Geomorphology. 95, 236-259.

Mingliang, L., Bin, Z., Guoan, T., Youfu, D., 2009. Fuzziness and multi-neighborhood strategy for DEM-based ridgeline extraction. ESIAT '09 Proceedings of the 2009 International Conference on Environmental Science and Information Application Technology - Volume 03. *Pg.* 679-682. IEEE Computer Society Washington, DC, *USA.*

Mitasova, H., Hofierka, J., 1993. Interpolation by regularized spline with tension: II.Application to terrain modelling and surface geometry analysis. Mathematical Geology. 25, 657-655

Mitasova, H., Hofierka, J., Zlocha, M., Iverson, L.R., 1996, Modeling topographic potential for erosion and deposition using GIS. Int. Journal of Geographical Information Science. 10(5), 629-641.

Möller M., Friedrich, K., Lymburner, L., Volk, M., 2008. Placing soil genesis and transport processes into a landscape context: A multiscale terrain analysis approach. Journal of Plant Nutrition and Soil Science. 171, 419–430.

Montgomery, D., Dietrich, W.E., 1989. Channel initiation, drainage density and slope, Water Resources Research. 25(8), 1907-1918.

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

Moore ID, A.Lewis, J.C. Gallant, 1993. Terrain attributes: estimation methods and scale effects. In: Jakeman AJ, Beek MJ, McAleer MJ (EDS). Modelling change in environmental systems. John Willey and Sons, London

Moore, I., Burch, G., 1986. Physical basis of the length-slope factor in the Universal Soil Loss Equation. Soil Society of America Journal. 50, 1294 – 1298.

Moore, I. D., Nieber, J.L., . 1989. Landscape assessment of soil erosion and nonpoint source pollution. J. Minn. Acad. Sci. 55, 18-24.
Moore, I.D. Wilson, J.P., 1992. Length-slope factors for the revised universal soil loss equation: simplified method of estimation. Journal of Soil and Water Conservation. 47(5), 423-428.

Moore, I.D., Gessler, P.E, Nielsen, G.A., Peterson, G.A., 1993, Soil attributeprediction using digital terrain analysis. Soil Science Society of America Journal. 57, 443-452.

National Research Council, 1986. Soil Conservation: Assessing the National Resource Inventory, Vol. I. National Academy Press, Washington, DC.

Nigh, T., Kabrick, J., Meinert, D., Steele, A., Steele, K., Young, F., 2010. Ecological Landtypes of Missouri: The Springfield Plain. Missouri Ecological Landtype Assocation.

Oberth, T., Aylward, M., Doberman, A., 2000. Using auxiliary information to adjust fuzzy membership functions for improved mapping of soil qualities. INT.J. Geographical Information Science. 14(5), 4-28.

Odeh, I.O.A., 1994. Spatial prediction of soil properties from a digital elevation model. Geoderma. 63, 197–214.

Park, S., Egbert, S.L., 2002. Assessment of Soil Erodibility Indices for Conservation Reserve Program (CRP) Lands in Southwestern Kansas Using Satellite Imagery and GIS Techniques. Environmental Management. 36(6), 886-898.

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

Pennock, D.J., 2003. Terrain attributes, landform segmentation, and soil redistribution. Soil Tillage Res. 69, 15–26.

Peucker, T.K., Douglas, D.H., 1975. Detection of surface-specific points by local parallel processing of discrete terrain elevation data. Computer Graphics and image processing. 4, 375-387

Pike, R.J., 1988. The geometric signature: Quantifying landslide terrain types from digital elevation models, Mathematical Geology. 20 (5), 491-511.
Pike, R. J., Rozema R.J., 1975. Spectral analysis of landforms. Ann. Assn. Amer. Geog. 82, 1079-1084.

Pike, R., 2000. Geomorphometry - diversity in quantitative surface analysis. Progress in Physical Geography. 24, 1-21.

Pontius Jr., R. G., Shusas, E. and McEachern, M., 2004, Detecting important categorical land changes while accounting for persistence. Agriculture, Ecosystems & Environment. 101(2-3), 251-268.

Qin, C., Li, B., Pei, T., Shi., X., Zhu, A., Zhou, C., 2009. Quantification of spatial gradation of slope positions. Geomorphology. 110, 152-161.

Quinn, P.F., Chevallier, P., Planchon, O., 1991. The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. Hydrologic Processes. 5, 59-79.

Rahman, S., Arneson, C., Munn, L.C., Vance, G.F., 1997. Wyoming rocky mountain forest soils mapping using an ARC/INFO Geographical Information System. Soil science society of America journal. 61, 1730-1737.

K.G. Renard, Foster, G.R., Lane, L.J., Laflen, J.M., 1996. USLE and its revisions since 1985. In: Soil Loss Estimation in Soil Erosion, Conservation and Rehabilitation. Agassi, M. (eds). New York: M.Dekker, p.169-202.

Reuter, H.I., Wendroth O., Kersebaum, K.C., 2006. Optimisation of relief classification for different levels of generalisation.Geomorphology. 77, 79-89.

Ribaudo, M.O., Heimlich, R., Hoag, D.L., Smith, M.E., 2001. Environmental Indices and the politics of the Conservation Reserve Program. Ecological indicators. 1, 11-20.

Roberts, D.W., 1986. Ordination on the basis of fuzzy set theory. Vegetation. 66, 123-131.

Robinson V.B., 2003. A perspective on the fundamentals of fuzzy sets and their use in Geographical Information Systems. Transactions in GIS. 7(1), 3-30.

Rodriguez, F., Maire, R., Courjault-Rade, P., Darrozes, J., 2002. The Black Top Hat function applied to a DEM: a tool to estimate recent incision in a mountainous watershed (Estibe`re Watershed, Central Pyrenees). Geophysical Research Letters. 29(6), 91-94.

Ruhe, R.V. 1969. Quaternary landscapes in Iowa. Iowa State University Press, Ames.

Schmidt, J., Merz, B., Dikau, R., 1998. Morphological structure and hydrological process modelling. Zeitschrift fu¨r Geomorphologie, N.F., Supplement Band. 112, 55–66.

Schmidt, J., Dikau, R., 1999. Extracting geomorphometric attributes and objects from digital elevation models—semantics, methods, future needs. In GIS for Earth Surface Systems—Analysis and Modelling of the Natural Environment. Dikau, R. and H. Saurer (EDS), Berlin Stuttgart, pp. 153–173.

Schmidt, J., Evans, I.S., Brinkmann, J., 2003. Comparison of polynomial models for land surface curvature calculation. International Journal of Geographical Information Science. 17 (8), 797–814.

Schmidt, J., Hewitt, A., 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. Geoderma. 21:243–256.

Schmidt, J., Andrew, R., 2005. Multi-scale landform characterization. Area. 37, 341–350.

Secchi, S., Gassman, P.W., Williams, J.R., Babcock, B.A., 2009. Corn-Based Ethanol Production and Environmental Quality: A Case of Iowa and the Conservation Reserve Program. EnvironmentalManagement. 44, 732-744.

Shary, P.A., 1995. Landsurface in gravity points classification by a complete system of curvatures. Mathematical Geology. 27, 373-390.

Shary, P., Sharaya, L.S., Mitusov, A.V., 2002. Fundamental quantitative methods of land surface analysis. Geoderma. 107, 1-32.

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

Shary, P.A., Stepanov, I.N., 1991. On the second derivative method in geology. Doklady AN SSSR, 319: 456-460.

Shi, X., Zhu, A.-X., Wang, R.-X., 2005. Fuzzy representation of special terrain features using a similarity-based approach. In: Cobb, M., Petry, F., Robinson, V. (Eds.), Fuzzy Modeling with Spatial Information for Geographic Problems. Springer-Verlag, New York, pp. 233–251.

Skidmore, A.K., 1990. Terrain position as mapped from a digital elevation model. International Journal of Geographical Information Systems. 4(1), 33-49.

Smith, M.P. Burt, J.E., Stiles, C., Zhu, A. , 2006. The Effects of DEM resolution and neighborhood size on digitial soil survey. Geoderma. 137, 58-69.

Speight, J. G., 1968. Parametric description of land form, in Stewart, G. A., ed., Land evaluation: McMillan, London, p. 239-250.

Speight, J.G., 1974. A parametric approach to landform regions. In Progress in Geomorphology. Institute of British Geographers special publ. No.7, Oxford:Alden & Mowbray Ltd. at the Alden Press, 213–230.

Speight, J.G., 1976. Numerical classification of landform elements from air photo data, Zeitschrift fur Geomorphologie. 154-168.

Speight, J.G., 1990. Australian Soil and Land Survey: Field Handbook, 2nd ed. Inkata Press, Melbourne.

Stamer, J.K., Gunderson, K.D., Ryan, B.J. , 1995. Atrazine Concentrations in the Delaware River, Kansas. USGS Fact Sheet 001-94.

Strahler, A.N., 1952. Hypsometric (area-altitude) analysis of erosional topography. Bulletin of the GeologicalSociety of America. 63, 1117– 1141.

Tangestani M.H., 2004. Landslide susceptibility mapping using the fuzzy gamma approach in a GIS, Kakan catchment area, southwest Iran. *Australian Journal of Earth Sciences*. 51, 439–450.

Tarboton, D.G., 1997. A new method for the determination of flow directions and upslope areas in grid Digital Elevation Models, Water Resources Research. 33, 309–319.

Tobler, W.R., 1969. An analysis of a digitized surface in Davis, C.M. (ed), A Study of the LandType, Report DA-31-124-ARO-D-456, pp.59-76, Department of Geography, University of Michigan.

Toriwaki J., Fukumura, T., 1978. Extraction of structural information from greypictures. Computer Graphics and Image Processing. 7, 30–51.

Tribe, A., 1991. Automated recognition of valley heads from digital elevation models. Earth Surfaces Processes and Landforms. 16, 33-49.

Troeh, F. R., 1964. Landform parameters correlated to soil drainage: Soil Sci. America Proc., 28(6), 808-812.

U.S.D.A. , K.D.H., 2006. Kansas Rapid Watershed Assessment: Delaware River. ftp://ftpfc.sc.egov.usda.gov/KS/Outgoing/Web_Files/Technical_Resources/rwa/Delaware_RWA.pdf.

U.S.G.S., 1996. Fort Collins Science Center GIS Data Site. Accessed June 2009. http://www.fort.usgs.gov/Products/GIS/GIS_abstract.asp?GISID=5

van Rossum, G. et al., 2005. Python Language Website, http://www.python.org/.

Weiss, A., 2001. Topographic position and landforms analysis. Poster Presentation, ESRI User Conference, San Diego, CA.

Werner, C., 1988. Formal analysis of ridge and channel patterns in maturely eroded terrain. Annals of the American Association of Geographers. 78, 253–270

Wilson JP and JC Gallant, 2000. Terrain analysis. Principles and applications. John Wiley & Sons, London.

Wischmeier, W.H., Smith, D.D., 1968. Predicting Rainfall Erosion Losses: A Guide to Conservation Planning. USDA Handbook, No.537. US Department of Agriculture, Washington, DC.

Wood, A., 1942. The development of hillslopes. Geologists Association Proceedings. 533, 128-138.

Wood, J., 1996. The geomorphological characterisation of Digital Elevation Models. PhD Thesis Department of Geography, University of Lancaster, UK.

Wood, J., 2001. Visualising the Structure and Scale Dependency of Landscapes, in: Fisher, P. and Unwin, D. (eds.) Virtual Reality in Geography, London: Taylor and Francis163-174.

Wood, W.F., Snell, J.B., 1960. A Quantitative System for Classifying Landforms, U.S. Army Quartermaster Research & Engineering Command Report EP-124, Natick, Massachusetts. 20 p..

Wu, J., Egbert, S.L., Nellis, M.D., Price, K.P., Ranson, M.D., 1997. Evaluating soil properties of CRP land using remote sensing and GIS in Finney County, Kansas. Journal of Soil and Water Conservation. 52(5), 352–358.

Young, M., 1978. Terrain analysis: program documentation. Report 5 on Grant DA-ERO-591-73-G0040, 'Statistical characterization of altitude matrices by computer'. Department of Geography, University of Durham, England, 27 pp.

Zadeh, L., 1965. Fuzzy sets. Information Control. 8, 338-353.

Zaslavsky, D., Sinai, G., 1981. Surface Hydrology: I – Explanation ofPhenomena. ASCE Journal of the Hydraulics Division. 107, 1–16.

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

Zhang, W., Montgomery, D.R., 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations. Water Resour. Res. 30(4), 1019-1028.

Zhu, A.-X., 1997. A similarity model for representing soil spatial information. Geoderma. 77, 217–242.

Zinn, J., 2000. Conservation Reserve Program: Status and Current Issues. Congressional Research Service Report 97–673. Congressional Research Service, Washington.