Effects of Terrain Attribute Weights on Fuzzy *k*-means Landform Classifications

Yongxin Deng

Department of Geography
University of Southern California
yongxind@usc.edu

Jingfen Sheng

Department of Geography
University of Southern California
sheng@usc.edu

1. Introduction

DEM-based fuzzy *k*-means landform classification seeks to define similarities between continuous spatial patterns of fuzzy landform class memberships and other biophysical properties so as to construct a direct linkage between landform and other components of the biophysical environment (Burrough et al. 2001). This classification method assumes the long-recognized existence of cause-effect interactions between landform and the environment (Moore et al. 1991, Wilson and Gallant 2000a) although these interactions are difficult to define (Florinsky 1998). Fuzzy *k*-means landform classification is thereby useful because the identified landform-environment linkages may represent effects of these landform-environment interactions and can be extrapolated to other places for spatial predictions of the biophysical environment.

It is reasonable to presume that multiple landform properties embodied by corresponding terrain attributes may be involved in the formation of the same biophysical property. However, the involved landform properties may carry different degrees of significances. For example, slope gradient and aspect both influence runoff generation process and solar radiation fluxes, but they display varying extents of effects on these two processes. This logic, nonetheless, is not followed by the common practice of assigning a uniform weight – implying an equal significance – to all the terrain attributes incorporated in landform classification or other terrain-based environmental analysis. Embedded in the above context, this paper attempts to answer the following two questions:

1) How sensitive is the fuzzy *k*-means landform classification to the adjustment of terrain attribute weights?
2) How does the change of attribute weights influence the level of correspondence between the classification results and a soil map?

The answers to these questions are expected to arouse the attention of the related GIS community to variant significances of landform properties in determining a particular biophysical property – an issue that has been largely ignored. We also hope the experiments reported in this paper may lead to more exhaustive tests of the influence of terrain attribute weights. The following sections of the paper will first describe the background of the experiments reported in this paper, then explain the experimental design as well as the data and methods adopted for the experiments, and lastly report the results and conclusions obtained from this research.

2. Background

The crisp options of either including or excluding a terrain attribute are prevalent in most current biophysical applications that incorporate terrain attributes. This is equivalent to assigning weights of either 1 (included) or 0 (excluded) to all attribute. The practical difficulty underlying this situation is that we cannot confidently define the significance of one-to-one correlations between terrain attributes and specific biophysical properties. While this difficulty persists, fuzzy *k*-means landform classification serves as a means to test the effects of varying attribute significances (weights). It also represents an opportunity to circumvent the difficulty because analyzing multiple classification results may potentially identify the "best" combination of attributes and their weights for a given biophysical property.

The above-described crispness in assigning attribute weights is similar to the reasons that raised the need to introduce the fuzzy theory into soil and landform classifications: it was realized that the crisp, human-defined soil and landform boundaries were problematic and spatial continuity of biophysical patterns needed more attention (e.g. Burrough et al. 2000, McBratney and de Gruijter 1992, McBratney et al. 1992). The geographic continuity of biophysical patterns, or gradual change from one geographic location to another, genetically depends on the continuity in the attribute space, shown by gradual change of attribute values from one data point to another. Hence various fuzzy classification algorithms (Burrough and McDonnell 1998) can describe spatial continuity between geographic places (cells, points, etc.) only after defining the continuous attribute transitions between data points in the attribute space.

Furthermore, the continuity in the attribute space depends on how the attribute space is defined. If the attribute selections and weight assignments can be continuously tuned, the concept of fuzziness or continuity can be introduced to define the attribute space itself. In this way, the dimensions of the attribute space (expressed as a group of attributes and their weights), or the methods used to define a data point in attribute space, become continuously adjustable. This allows a conceptual representation when several terrain attributes impose unequal impacts on a specific biophysical process and/or pattern, but we are unable to define these impacts based on a cause-effect interaction mechanism.
The continuous attribute weight adjustments in fuzzy *k*-means landform classification thereby provide an empirical approach to identify the optimal group of attributes/weights that can be used to define the best representation for a particular biophysical pattern.

In this context, the human representation of the biophysical continuity should incorporate three factors: gradual transitions of attribute values between data points, continuous spatial relationships (such as similarities) of each data point with its geographic neighbors, and continuously modifiable attributes (and their weights) used to describe the data points. The biophysical meaning of the weight assignment is that it describes how significant each terrain attribute (or landform property) is to the targeted biophysical property. To test the effects of weight assignments using fuzzy *k*-means landform classification has a practical value because the results can be used to define how to group terrain attributes sensibly to accurately represent a correlated biophysical pattern (Burrough et al. 2001). And the further investigation of this idea may allow us to produce landform classifications that are directly applicable to particular purposes.

The rapid development of computer technology and spatial data coverage offers tremendous opportunities to tackle the problem of varying attribute weights. Hence it is possible to generate numerous classifications on a personal computer rapidly given a relatively complete list of terrain attributes and a small yet computable weight increment. When all these patterns are compared to a known biophysical pattern in a GIS environment, it is possible to identify the best-matching pair of patterns. Extrapolation of this result to other places is expected to predict unknown spatial patterns. This idea is currently under investigation and a second paper examining the sensitivity of the fuzzy classification to the number of used terrain attributes by Deng and Wilson (2004) is currently under review.

3. Experimental design

Eight terrain attributes – elevation, slope, plan curvature, profile curvature, proximity to the ridgeline, incoming solar radiation, topographic wetness index, and sediment transport capacity index (Wilson and Gallant 2000b, Gallant and Wilson 2000, Burrough et al. 2001) –

calculated from a 10-m DEM were used for two groups of experiments. The first group focused on the relative differences of landform classifications caused by the weight adjustments; the second group explored how the level of correspondence between defuzzified landform class maps (Burrough et al. 2000) and a soil classification map changed due to the adjustment of attribute weights.

3.1 Relative sensitivity test

While all the other attributes were given the same weight (defined as 1) throughout the experiments, the weights of slope and topographic wetness index were adjusted according to the following scheme of weight assignments:

- The weight of topographic wetness index was set to 1 – the same as the other attribute weights – and the weight of slope was varied by assigning slope each of the following weights: 0.25, 0.5, 1, 1.5, 2, 3, 4, 5.
- The weight of slope was set to 1 and the weight of topographic wetness index was varied from 0.25 to 5 (using the same increments as above).
- Topographic wetness index was assigned a weight of 5 and the weight of slope was varied from 0.25 to 5 (using the same increments as above).
- Slope was assigned a weight of 5 and the weight of topographic wetness index was varied from 0.25 to 5 (using the same increments as above).
- The weights of slope and topographic wetness index were simultaneously varied from 0.25 to 5 (using the same increments as above).

A total of 102 fuzzy *k*-means landform classifications were performed at a spatial resolution of 10 m using the abovementioned weight assignments and 3-class, 4-class, or 5-class classification schemes. Classifications with various weight combinations were compared to the classifications with the original weight assignments, where all attributes were assigned the same weights.

3.2 Soil map correspondence test

This group of experiments compared a 5-class soil map, derived from the digital Soil Survey Geographic Database (SSURGO), with hard-class maps for all of the 5-class landform classifications that incorporated the various weight assignments described previously. This approach was adopted to investigate whether there was a trend in terms of the level of correspondence between landform class maps and the soil map when the weights of slope and topographic wetness index were varied over different landform classifications. These two terrain attributes are often cited as key factors for the explanation of soil variability (e.g. Moore et al. 1993).

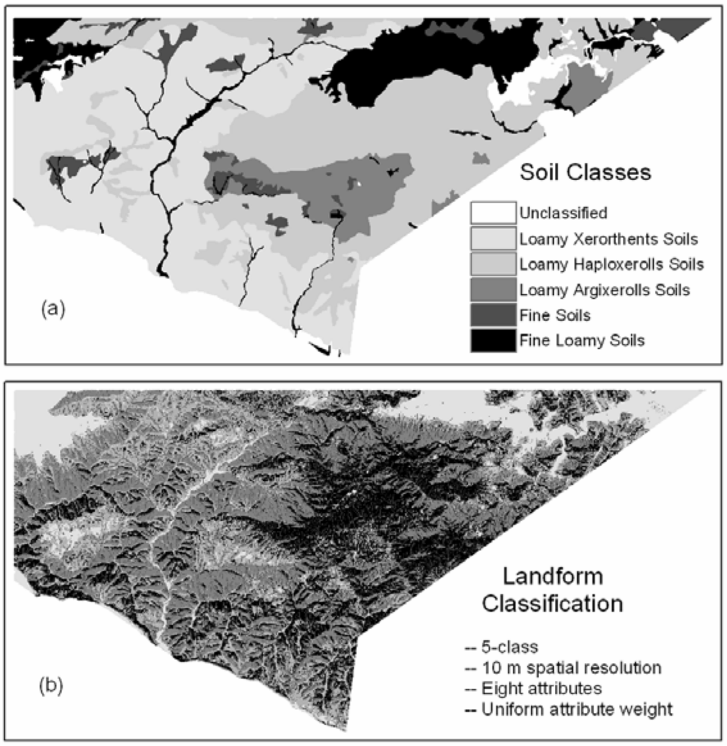

Figure 1 Soil map (a) and one of several possible landform classification maps (b) for study area

A visual comparison of the soil map and a 10-m landform classification map in Figure 1 indicates that a relatively low level of correspondence should be expected. This result may be caused by several factors, involving (1) the general character of the tacit expert knowledge adopted during the soil-mapping process, and (2) the differences between the spatial resolutions of the soil maps (incorporated implicitly during soil survey) and landform classification maps (where spatial resolution was explicitly defined by the cell size or the neighborhood window size used for attribute calculation). However, the heavy reliance of the traditional soil-landscape model (Soil Survey Staff 1951, 1999; Hudson 1990, 1992) on landform interpretations implies that the overlap between a soil class map and a landform class map should be more significant than random overlaps between classification maps, and more significant overlaps should be explained by higher, overall correlations between selected terrain attributes (weights) for landform classifications and factors considered when drawing the soil map.

4. Data Sources and methods

4.1 Study area

A 157-km^2 area located at the southeast corner of Ventura County, California, was chosen as the study area for the reported experiments. This area covers the western end, as well as the highest section, of the Santa Monica Mountains, which extend east-west along the Pacific Ocean coast. The landforms of the study area are characterized (Figure 2) by short, steep and rugged slopes, several high mountain peaks and highland areas, and series of incised stream systems. Table 1 summarizes the statistics of the eight terrain attributes used in this research.

4.2 Data and tools

Several USGS 10 m DEMs were downloaded from the GIS Data Depot

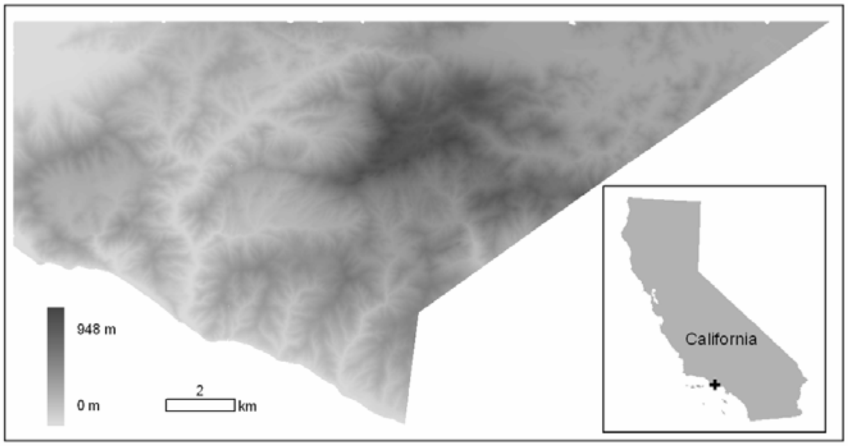

Figure 2 Study area DEM

Table 1 Terrain attribute statistics

	Mean	Maximum	Minimum	Standard deviation
Elevation (m)	317.5	948.2	0.3	185.6
Slope (%)	39.9	309.0	0.0	22.8
Plan curvature (×100 m/m^2)	-0.05	26.03	-50.35	1.48
Profile curvature (×100 m/m^2)	-0.04	51.91	-41.25	1.60
Proximity to ridgeline (m)	16.4	386.3	0	17.2
Incoming solar radiation (×10^3 MJ/m^{2}year)	9.59	13.03	0.48	2.21
Topographic wetness index	7.64	26.22	3.50	2.29
Sediment transport capacity index	3.25	-7.21	10.63	1.85

(http://www.geocomm.com) and used as the original data source for terrain analysis and landform classifications. The digital copy of the Ventura County Soil Survey Geographic Database (SSURGO) developed by the Natural Resources Conservation Service (NRCS) of the U.S. Department of Agriculture was obtained from the California Spatial Information Library (http://gis.ca.gov).

The ArcGIS software was used for data preprocessing, map calculations, and spatial comparisons. The PCRaster software was used to calculate the terrain attributes (Karssenberg et al. 2001; see PCRaster Version 2 Manual at http://pcraster.geog.uu.nl for additional details about attribute algorithms). Specific catchment area and proximity to ridgeline are calculated in PCRaster based on the D8 flow-routing algorithm (O'Callaghan and Mark 1984). The topographic wetness index and sediment transport capacity index are calculated from slope and specific catchment area using simplified methods described in Wilson and Gallant (2000b). Incoming solar radiation is calculated in PCRaster with the assistance of a group of programs that first calculate relief and terrain shading effects from elevation, slope gradient, aspect, and sun height (adjusted for latitude), and then use this information to calculate and accumulate daily incoming solar radiation for each grid point. The FNX730 fuzzy *k*-means classification program based on the fuzzy *k*-means algorithms described by Bezdek et al. (1984), McBratney and de Gruijter (1992), Burrough and McDonnell (1998) was used to produce the fuzzy landform classes. Microsoft Excel was used to perform statistical analysis and to produce output tables and figures.

4.3 Implementation of attribute weights

The following procedure was adopted to implement adjustments of attribute weights in the fuzzy *k*-means landform classification. First, the attribute values in the sample attribute table, which was used as input to the fuzzy classifier, were normalized to have the same sample variance. Second, attribute weights were adjusted by multiplying attribute values of the sample attribute table with new weights. Third, the Euclidian distance function and the new, weighted sample attribute table were used for the classification procedure (Irvin et al. 1997, Burrough et al. 2000). Fourth, the diagonal norm distance function (similar to Equation (1) in Section 4.5.1) and an attribute variance table were used in PCRaster to assign memberships to unsampled cells according to class centers calculated with the fuzzy classifier, but the variance table was modified by multiplying the variances with the new weights. The effect of this procedure is that attribute weights were incorporated into the computation of both the class centers and the membership assignments of cells (based on similarities or attribute distances between cells and class centers).

4.4 Soil map preparation

The Santa Monica Mountains part of the Ventura County SSURGO soil map – 210 polygons belonging to 61 soil map units – was selected for this research. A total of 171 polygons (46 soil map units) were chosen, aggregated, and analyzed further. The remainder of the polygons and soil map units were excluded because they constituted water bodies, bare rocks, urban areas, or landscaped areas (10 soil map units; 33 polygons), or it was difficult to aggregate them into other soil classes (five soil map units; six polygons). The selected soil map units were first grouped into nine categories according to their great group names and soil texture classes (Soil Survey Staff 1999) and later the nine categories were grouped into five classes based on soil texture classes and/or soil profile development (Figure 1a). The final five classes were:

- Loamy Haploxerrolls characterized by loamy soil textures and poor soil horizon development;
- Loamy Argixerolls characterized by loamy soil texture and high sub-surface clay contents;
- Loamy Xerothents characterized by loamy soil textures and soil profiles with relatively low soil moisture content;
- Fine loamy soils (including fine loamy Haploxerolls, Argixerolls, Xerorthents, and Xerofluvents) characterized by fine loamy soil textures; and
- Fine soils (including fine Haploxererts and Argixerolls) characterized by fine soil texture classes.

4.5 Comparison methods

4.5.1 Comparison for relative sensitivity test

Three sequential methods were used to compare the spatial patterns of results for the relative sensitivity test described in Section 2.2. The first method identified similar class-pairs by comparing attribute distances between class centers of two classifications *i* and *j* as follows:

$$d_{ij} = \sum_k [(a_{ik} - a_{jk})^2 \times \frac{1}{\sigma_k^2}] \qquad (1)$$

where d_{ij} is the attribute distance between two respective class centers in two classifications (*i* and *j*), a_k is the *k*th common attribute used by the two classifications, and σ_k^2 is the sample variance of a_k. The variance σ_k^2 was used to convert various attribute values to the same magnitude so that all attributes have the same weight in the calculation of attribute distances. Thus each class in one classification was matched to a certain class in another classification to form a pair of most similar classes that have the minimum d_{ij} or the minimum accumulated attribute distance between the two classes in the attribute space.

The second method included four steps that: (1) calculated hard classes from the fuzzy *k*-means landform classifications using the “defuzzification” procedure described by Burrough et al. (2000); (2) calculated overlapped areas between the hard classes derived from the similar class-pairs that were identified with the previous method; (3) defined the sum of these overlapped areas as the overall overlap of hard classes between the two classifications; and (4) calculated the ratio of the overall overlap of hard classes in the entire classification area. By use of hard classes as employed in this method, the similarity between fuzzy classes can be linked to the similarity between conventional hard classes, and it is assumed that more similar classifications tend to have larger areas of hard class overlap.

The third and final method incorporated three steps as follows. It first generated a relative membership residual surface for similar class-pairs in two classifications (*i* and *j*) by calculating:

$$\mu_{ij} = \frac{|\mu_i - \mu_j|}{\mu_i + \mu_j} \times 100\ \% \qquad (2)$$

where μ_{ij} is the relative residual between two memberships μ_i and μ_j. As a result, the value of μ_{ij} ranges from 0 to 100%, indicating the percentage difference between these two memberships. In the second step, the mean and standard deviation of this residual surface were calculated over the entire study area to represent how statistically different the two membership surfaces were. Lastly, the means of the membership residual surfaces for all classes in the two classifications were averaged to represent the general difference between the two classifications. This method explicitly identified the most similar classes and the most similar classifications.

4.5.2 Comparison for soil map correspondence

The method used to compare the landform classifications and the soil map incorporated several steps as well. First, a table was calculated to list how each of the five classes on the soil map overlapped with each of the classes in a 5-class landform classification. Second, the areas of the maximum and second-largest overlaps for each soil class were identified. Third, the percentages of the maximum overlapped area and the sum of the maximum and second-largest overlapped areas to the total area of each soil class were calculated. Fourth, the two sets of percentages were averaged over the five soil classes using the area of each soil class as the weight, and two averaged percentages were obtained to indicate how much of each soil class area was covered by the single- and two-most representative landform classes.

5. Results

5.1 Change of hard class overlaps

Figure 3 shows change of hard class overlaps between the classification with uniform weight assignments and the classifications with various weight adjustments. The following

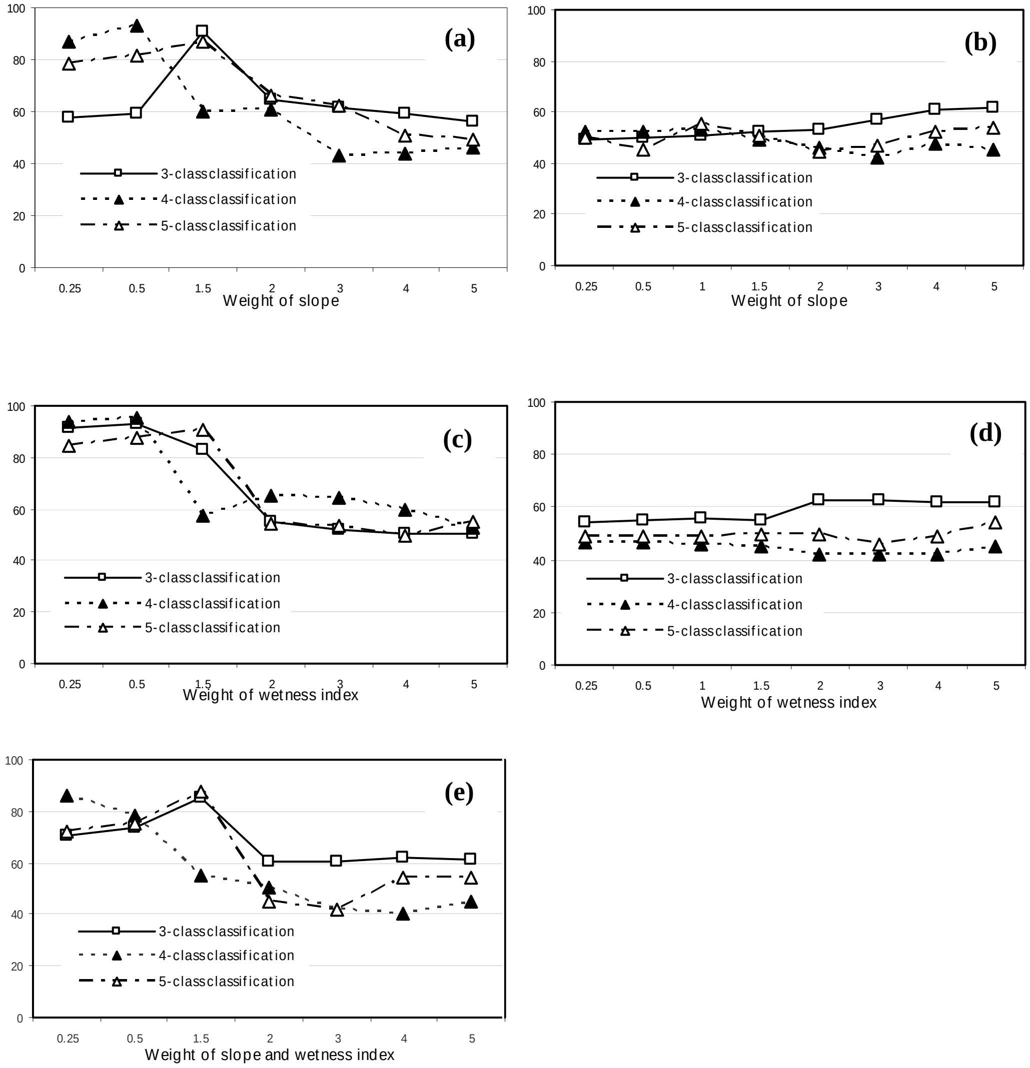

Figure 3 Results of relative sensitivity test (see Section 2.1) – hard-class overlap (in %, see Section 4.5.1) between classifications with original weight assignments (where all attributes have the same weight) and classifications where: (a) topographic wetness index has the same weight (equal to 1) as other attributes but the weights of slope are adjusted; (b) topographic wetness index has a high weight (equal to 5) and the weights of slope are adjusted; (c) slope has the same weight (equal to 1) as other attributes but the weights of topographic wetness index are adjusted; (d) slope has a high weight (equal to 5) and the weights of topographic wetness index are adjusted; (e) weights of both slope and topographic wetness index are adjusted

experimental results can be summarized from these diagrams:

- Overall, the hard class overlaps changed from 40 to 93% when various weights were used for slope and/or topographic wetness index during the fuzzy *k*-means landform classification. A similar range of variation was observed when the weights were varied in all three classification schemes (i.e. the 3-, 4-, and 5-class classifications). This indicates that the results of fuzzy *k*-means landform classifications are sensitive to the change of attribute weights.
- The fastest rate of change in hard class overlaps occurred when the weights of slope and/or topographic wetness index were first increased; hence, the 4-class classifications showed the fastest rate of change when the weights of the two attributes were adjusted from 1 to 1.5 and the 3- and 5-class classifications showed the fastest rate of change of hard class overlaps when the weights of the two attributes were adjusted from 1.5 to 2 (Figures 3a, c, and e). These weight ranges constituted the most sensitive ones for the fuzzy *k*-means landform classifications that were created for this study.
- The hard class overlaps changed very little in the 40-60% range when the weights of the two attributes were increased beyond 2 (or occasionally 1.5 or 3, see Figures 3a, c, and e). In other words, there seems to be a threshold of 2 (or occasionally 1.5 or 3) for slope and topographic wetness index weights beyond which the rate of change is minor.
- There was a much smaller variation when the weights of the two attributes were decreased from 1 to 0.5 and then to 0.25. Most of the hard class overlaps were >70% in these instances and the variation of the hard class overlaps was less than when the weights were increased. The above observations indicate that, at least for

the weight range considered, the fuzzy *k*-means landform classifications are less sensitive to reduction in slope and topographic wetness index weights than they are to weight increases.

- When the weights of one of the two attributes were adjusted while constantly assigning the other attribute a high weight (Figures 3b and d), hard class overlaps fell mostly within a range of 40-60%. There was no obvious trend and the variation within each classification scheme (i.e. the 3-, 4-, or 5-class classifications) was often no more than 10%. This result confirms the third point summarized above and indicates that there might be a range of weights where weight variations cause little or no change in hard class overlaps.
- More or less, variations of hard class overlaps caused by weight adjustments of slope, topographic wetness index, or both produced similar patterns. This may be partly explained by the correlation between slope and topographic wetness index (Moore and Wilson 1992). Alternatively, these similar patterns may indicate that similar weight adjustments for different terrain attributes have similar impacts on fuzzy *k*-means landform classifications. Additional research is required to verify this result.

5.2 Change of membership differences

Figure 4 shows change of membership differences (defined in Section 4.5.1) between the classification with uniform weight assignments and the classifications with varying weights. Although the method of hard class overlaps (see Section 4.5.1) incorporate a closer linkage with conventional crisp classification systems, membership differences represent a more accurate way of evaluating differences between fuzzy *k*-means landform classifications because no aggregation of data points in attribute and geographic space is needed for cell-by-cell calculations of membership differences. Therefore, the membership differences can be

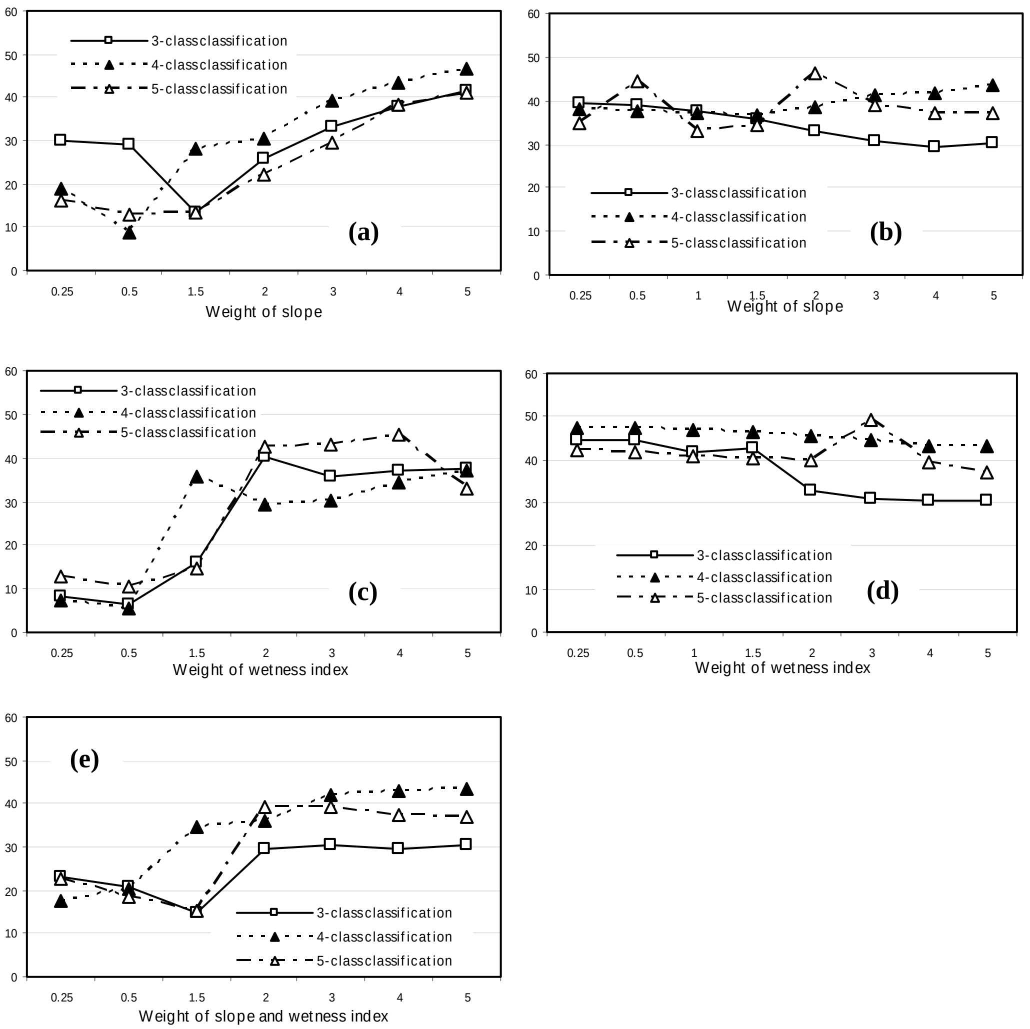

Figure 4 Results of relative sensitivity test (see Section 3.1) – membership differences (in %, see Section 4.5.1) between classifications with original weight assignments (where all attributes have the same weight) and classifications where: (a) topographic wetness index has the same weight (equal to 1) as other attributes but the weights of slope are adjusted; (b) topographic wetness index has a high weight (equal to 5) and the weights of slope are adjusted; (c) slope has the same weight (equal to 1) as other attributes but the weights of topographic wetness index are adjusted; (d) slope has a high weight (equal to 5) and the weights of topographic wetness index are adjusted; (e) weights of both slope and topographic wetness index are adjusted.

used to verify the results obtained with the hard class overlap method and to assess whether hard classes can be used to evaluate and compare fuzzy classifications. The conclusion so obtained is also an important reference for the validity of the research results obtained by comparing the hard-class landform maps with a soil map.

In general, Figure 4 shows that there was a good correspondence between the results obtained by calculating membership differences and the results obtained from the hard-class overlap method. The specific results can be summarized as follows:

- An overall 5-50% membership difference was consistently observed when the attribute weights were varied upwards or downwards from the original uniform weight assignments. This result verifies that fuzzy *k*-means landform classifications are sensitive to the weights used for the terrain attributes.
- The fastest rate of change of membership differences occurred when the weights were changed from 1 to 1.5 for the 3- and 5-class classifications, and from 1.5 to 2 for the 4-class classifications in most instances.
- Similar to the patterns observed with the hard class overlap method, 2 (or occasionally 1.5 and 3) seems to be a threshold for weight adjustments in most cases. However, Figure 4a showed a more or less consistent rate of change of membership differences even when the slope weights had been adjusted beyond 1.5, 2, or 3 times. Such an obvious or consistent pattern was not obtained with the hard class overlap method (cf. Figures 3a and 4a), meaning that the hard class overlap method may not always be reliable in detecting differences between fuzzy classifications.
- Figures 4a, c, and e also show that in most cases (the most obvious exception is the 3-class classification in Figure 4a), the fuzzy *k*-means landform classifications were not sensitive to the weight decrease of slope and topographic wetness index. In almost all cases, membership differences between classifications with the original weight assignments and classifications with the decreased weights are below 30%, and mostly below 20%.

- Similar to the observations in Figure 3, membership differences changed very little and did not show obvious trends when one of the two attributes retained a high weight and the weights of the other attribute were adjusted. Almost all the membership differences fall in a range of 30-50% and indicate the possibility that a threshold of weight change exists.
- Figures 4a and c show that change of membership differences caused by weight adjustments of slope and topographic wetness index produced very different results. Figure 4e displays an intermediate pattern between those evident in Figures 4a and c. Because slope and topographic wetness index are genetically correlated (Moore and Wilson 1992), this difference should be sufficient to reject the conclusion that similar patterns of change in fuzzy *k*-means landform classifications can be produced by modifying the weights of different terrain attributes in similar ways.

5.3 Results of soil map correspondence test

Figure 5 shows the results of comparing a soil map (Figure 1a) with hard class maps of the fuzzy *k*-means landform classifications produced using various weight assignments (see Section 4.5.2). The following results can be summarized from these diagrams:

- In almost all of the instances summarized in Figure 5, 30-40% of the area of each of the five soil classes was covered by the landform class that overlapped most with the soil class, and 50-70% of the area was covered by the two landform classes that showed the most overlap with each of the soil classes.
- The variation of the percentage of each soil class captured by one or two landform classes was < 15% for landform classifications produced with different weight assignments. This result indicates that adjustments of terrain attribute weights did

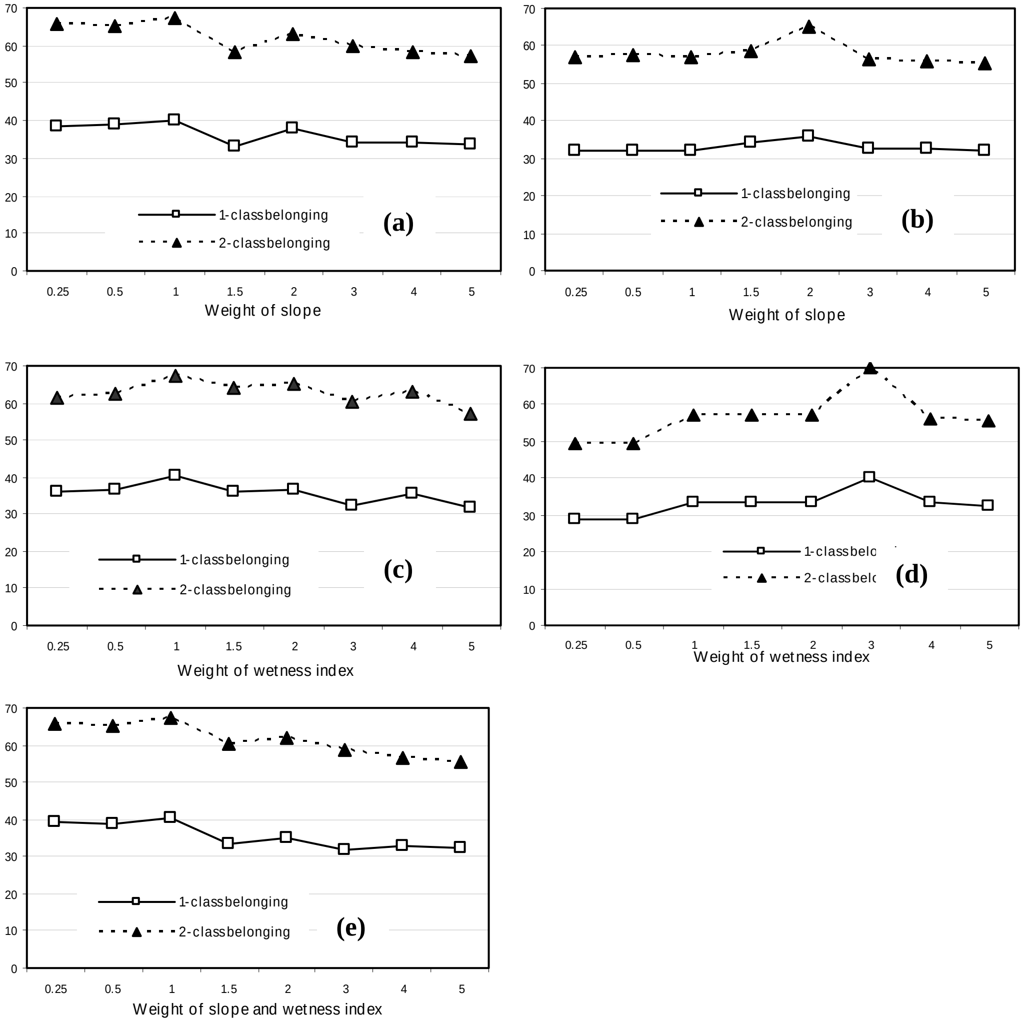

Figure 5 Results of soil map matching test (see Section 3.2) – area of each soil class falling into one or two landform classes (in %, see Section 4.5.2) when landform classes are derived with the following weight adjustments: (a) topographic wetness index has the same weight (equal to 1) as other attributes but the weights of slope are adjusted; (b) topographic wetness index has a high weight (equal to 5) and the weights of slope are adjusted; (c) slope has the same weight (equal to 1) as other attributes but the weights of topographic wetness index are adjusted; (d) slope has a high weight (equal to 5) and the weights of topographic wetness index are adjusted; (e) weights of both slope and topographic wetness index are adjusted

not produce dramatic variations in the degrees of correspondence between the soil and landform classifications maps.

- The best matches between the soil map and landform class maps captured in Figures 5(a), (c), and (e) occurred when the weights of the slope and topographic wetness index were set to 1. This particular result (40.2% for the overlap with one landform class and 67.1% for the overlap with two landform classes) constituted the second best pair of result generated in all 102 tests.

- The single best match between the soil map and landform class maps (70.2% for the overlap with two landform classes) was obtained when weights of 5 and 3 were used for slope and topographic wetness index, respectively (Figure 5d). The match measured with the overlap of each soil class with one landform class in this instance (39.9%) was similar to the best result obtained when weights for all terrain attributes were set to 1.

6. Conclusions

This paper seeks to reject the traditional Boolean logic in selecting terrain attributes for biophysical analysis and to introduce the concept of continuity into the determination of the dimensions (and their relative significance) of the attribute space. The reported experiments indicate that results of fuzzy *k*-means landform classifications are sensitive to the weight adjustments of adopted terrain attributes. When the variations are expressed as hard class overlaps and fuzzy membership differences, the adjustments of terrain attribute weights within a range of 0.25-5 can generate classification differences of up to 50-60%. However, the patterns of variations vary for different ranges of weights. The most rapid rate of change occurred within a weight range of 1.5 to 3. Decreasing the weights of the slope and topographic wetness index terrain attributes produced fewer impacts than increasing their weights. Threshold weights beyond which the classification results change little may exist, and the results from this study suggest that a weight of two for slope and topographic wetness index might serve as such a threshold.

Generally low levels of correspondence were observed between the soil map generated from the SSURGO database and the fuzzy landform classes. This may be attributed mostly to different paradigms, procedures, and tools (including data) used to produce these two kinds of classifications. However, the results show that certain combinations of terrain attributes and weights may produce landform classifications that better fit the soil map.

The weight combinations tested by this research are far fewer than a more complete set of weight combinations. For example, the weight increments for this research – 0.5-1 – is very large but the test terrain attributes are relatively few. However, it took about 200 person hours to complete the experiments reported here and more complete tests would require the automation of the numerous, complicated analysis and data processing steps.

Acknowledgements

This paper is completed under the guidance of Professor John Wilson. Thanks are also owed to Professor Peter Burrough, who provided PCRaster, and his colleague, Dr. Pauline van Gaans, who provided the FNX fuzzy classification program, at the University of Utrecht in the Netherlands.

References

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

Burrough P A and McDonnell R A 1998 *Principles of Geographical Information Systems*. Oxford, Oxford University Press

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

Burrough P A, Wilson J P, van Gaans P F M, and Hensen A J 2001 Fuzzy *k*-means classification of topo-climatic data as an aid to forest mapping in the Greater Yellowstone Area, USA. *Landscape Ecology* 16: 523-46

Deng Y X and Wilson J P 2004 The Effects of Attribute Space Dimensions on Fuzzy *k*-means Landform Classifications. *Annals of Association of American Geographers*: In review

Florinsky I V 1998 Combined analysis of digital terrain models and remotely sensed data in landscape investigations. *Progress in Physical Geography* 22: 33-60

Gallant J C and Wilson J P 2000 Primary topographic attributes. In Wilson J P and Gallant J C (eds) *Terrain Analysis: Principles and Applications*. New York, John Wiley and Sons: 51-86

Hudson B D 1990 Concepts of soil mapping and interpretation. *Soil Survey Horizons* 31 63-73

Hudson B D 1992 Soil genesis, morphology and classification: The soil survey as paradigm-based science. *Soil Science Society of America Journal* 56: 836-41

Irvin B J, Ventura S J and Slater B K 1997 Fuzzy and isodata classification of landform elements from digital elevation data in Pleasant Valley, Wisconsin. *Geoderma* 77: 137-54

Karssenberg D, Burrough P A, Sluiter R, and de Jong K 2001 The PCRaster software and course material for teaching numerical modeling in the environmental sciences. *Transactions in GIS* 5: 99-110

McBratney A B and de Gruijter J J 1992 A continuum approach to soil classification by modified fuzzy *k*-means with extragrades. *Journal of Soil Science* 43: 159-75

McBratney A B, de Gruijter J J, and Brus D J 1992 Spatial prediction and mapping of continuous soil classes. *Geoderma* 54: 39-64

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

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

Moore I D and Wilson J P 1992 Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation. *Journal of Soil and Water Conservation* 49: 174-80

O'Callaghan J F and Mark D M 1984 The extraction of drainage networks from digital elevation data. *Computer Graphics and Image Processing* 28: 323-44

Soil Survey Staff 1951 *Soil Survey Manual.* Washington DC, US Department of Agriculture

Soil Survey Staff 1999 *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys.* Washington DC, United States Department of Agriculture Natural Resources Conservation Service

Wilson J P and Gallant J C 2000a Digital terrain analysis. In Wilson J P and Gallant J C (eds) *Terrain Analysis: Principles and Applications*. New York, John Wiley and Sons: 1-28

Wilson J P and Gallant J C 2000b Secondary terrain attributes. In Wilson J P and Gallant J C (eds) *Terrain Analysis: Principles and Applications*. New York, John Wiley and Sons: 87-132