Application Of fuzzy membership functions and ISOCluster analysis method for soli-landscape map (Case Study: Damavand region, Iran)

Amir Karam¹, Nazila Yaghoob Nejad Asl²*, Ebrahim Beheshti javid³

¹Associate Professor of Geomorphology Department, Kharazmi University, Iran, Tehran
²PhD student of Geomorphologic Hazards, University of Mohaghegh Ardabili, Ardabil
³PhD student of Geomorphologic Hazards, University of Mohaghegh Ardabili, Ardabil

Article info
Article history:
Received 09.29.2015; Accepted 12.21.2015

Keywords:
Fuzzy membership functions, ISOCluster analysis method, soil-landscape mapDamavand,

Abstract
Soil and land information is vital for land use planning and infrastructure development. Maps of soils landscapes are powerful tools to guide sustainable planning and natural resource management. Fuzzy set theory has been widely used in soil science for soil classification and mapping, land evaluation, fuzzy soil Geostatistics, soil quality indices. In fuzzy logic-based approaches, soil spatial variation is expressed as spatial variation of membership in soil classes, which is then used to produce conventional soil-landscape map and to predict spatial variation of specific soil properties. The aim of this study is using fuzzy membership functions and ISOCluster analysis method for soli landscape maps with high accuracy. The following seven topographic variables (elevation, slope gradient, curvature, profile curvature, Planimeter curvature, flow accumulation and topographic wetness index) were used in this study to characterize the environment conditions. The proposed methods were tested in Damavand region in the Northeastern of Tehran. A set of membership functions were constructed to represent the descriptive knowledge on soil-landscape relationships, then for soil-Landscape map ISOCLUSTER analysis method was used. Using this operator 5 soil-landscape units were determined as follows: 1-Valley bottoms, 2-low land plain, 3-moderate hill slope, 4-steep hill slope, 5-mountain. The soil-landscape map achieved 74% of accuracy. 74% Accuracy is acceptable for an initial soil mapping. This suggests that the membership functions constructed in this study do capture the major pattern of soil-landscape relationships over the region.
INTRODUCTION

Soil and land information is vital for land use planning and infrastructure development. Maps of soils landscapes are powerful tools to guide sustainable planning and natural resource management (http://www.environment.nsw.gov.au). Stable, healthy and productive landscapes and soils are essential for producing most of our food, and for maintaining environmental function, managing water quality, sustaining our primary industries, and supporting rural and urban communities. Our rural landscape is under continuing pressure. Acidification, salinity, erosion, invasive native scrub and loss of native vegetation are seriously affecting the condition of our soil and challenging the communities and industries that depend on the land. Soil-landscape map is a survey of land resources which delineates repeating patterns of landscapes and associated soils. A soil-landscape map unit reflects soil and landscape processes. In addition to the key parameters of soil and landscape, geology plays a part at broad levels through the influence of tectonics on landform, and at more detailed levels through the influence of lithology on soil parent material. Other environmental factors such as climate and native vegetation also contribute to distribution of soil and landscapes and are incorporated into the mapping units (Noel et al. 2004).

Field survey with sites being selected using the free survey technique (Noel et al. 2004). The preliminary mapping and ease of access influenced site selection. Fieldwork included the description of sites and soil profiles (mainly from hand auger borings) using the terminology of McDonald et al. (1990). Site locations were either marked on aerial photographs or recorded using a global positioning system (GPS). Site data recorded manually on site cards and later entered onto the Soil Profile Database (Purdie 1993). Soil profiles classified using the Australian Soil Classification (Isbell 1996) and/or the Soil Groups of Western Australia (Schoknecht 2002). Descriptions of map units and main soil types written up in a standard format and correlated with units and soils identified in other surveys. The description includes the proportion of Soil Groups of Western Australia (Schoknecht 2002) within each map unit. This information was then entered into the Map Unit Database. The development of fuzzy logic-based digital soil mapping techniques has attracted much attention in the digital soil mapping community due to its ability to capture and represent the continuous nature of soil spatial variation (Zhu and Band 1994; Burrough 1996; Dobermann and Oberthur 1997; McBratney and Odeh 1997; Zhu 1997; Zhu et al. 2001; Yang et al. 2007). In fuzzy logic-based approaches, soil spatial variation is expressed as spatial variation of membership in soil classes (Zhu 1997; McBratney et al. 2000; Qi et al. 2006), which is then used to produce conventional soil class maps and to predict spatial variation of specific soil properties (Zhu et al. 1996). Qi et al. (2006) developed a prototype-based fuzzy soil mapping approach to represent soil-environment knowledge as fuzzy membership functions, which were also constructed based on the knowledge obtained from soil experts. Liu and Zhu (2009) developed a mapping with words approach based on computational theory of perceptions to define membership functions. There have been many attempts to correlate soil properties with various factors, such as parent material and topography (Wilding et al. 1994; Cook et al. 1996; McBratney et al. 2000). This approach, frequently cited as soil-landscape analysis, was initiated for more accurate and easily obtainable information on the spatial distribution of soils for detailed environmental modeling and site specific land management (McBratney et al. 2000). This emphasis on using geomorphological variables to predict spatial variations in soil properties can be linked to both theoretical and practical considerations. Theoretically, landforms may be the best indicators of soil attributes in places where the impact of other environmental factors is relatively small (Moore et al. 1993). In terms of practical considerations, a topographic map is still the most easily available source of information in many parts of the world, particularly in developing countries where relatively expensive soil surveys have not yet been carried out. And such soil-landscape analysis is considered as a technique in natural resource surveys (Gessler et al. 1995). Despite the recent developments in analytical methodologies, some theoretical questions of systematic correlation between soil properties and landform geometry have not been fully investigated in the study area. The aim of this study is using fuzzy membership functions and ISOCluster analysis...
method in GIS environment for soil-landscape maps with high accuracy in Damavand region.

**Research questions**

1. Can digital terrain parameters be of help in soil survey?
2. What is the relationship between terrain parameters and landforms?

**Hypothesis**

1. The use of digital terrain parameters and geography information system in classifying and mapping soil in sloping areas can increase the accuracy and speed up soil survey.
2. There are strongly relationships between geopedology and terrain parameters.

The study area is located in Northeastern of Tehran in Damavand township of Iran (Fig. 1). Its area is 485 km$^2$ with elevation ranging from 1536 m to 3833 m and slope gradient mostly under $5^\circ$, which is indicative of a generally gentle environmental gradient (Fig. 2). The original vegetation is hedysarum, succor, mallow, and acanthus. Crop in the region at present is generally limited to apple, potato, gladiolus, wheat, grain and soya. The soils in the region are formed on two Colluvial fans and Gravelly Alluvio and Platuux physiographic units. In this region the aquatic and thermal regimes are Xeric and Mesic. (Source: Pedology Report of Damavand Region, Soil& water Research Institute, 1991, 3).

**MATERIALS AND METHODS**

The following seven topographic variables (elevation, slope gradient, curvature, profile curvature, Plan curvature, flow accumulation and topographic wetness index) were used in this study to characterize the environment conditions. Information of (elevation, slope gradient, curvature, profile curvature, Plan curvature, flow accumulation and topographic wetness index) were derived from a 25 M resolution DEM which was created from the Aster satellite radar images.

![Fig. 1. Location and DEM of the study area](image-url)
The selection of seven topographic variables is based on the fact that the area is large (485 km²). We are proposing new methods to automate this process and extend the utility of the classification for scientists and land managers. Continuous classification (fuzzy logic) methods and unsupervised (isodata) classification techniques were used to classify each pixel of a 25 meter resolution digital elevation model (DEM) according to its membership in a landform class. These classes were determined by the natural clustering of the data in attribute space. Attributes used for the classification were elevation, slope gradient, curvature, profile curvature, Planimetric curvature, flow accumulation and topographic wetness index. Because these factors are also important to soil forming processes, soil classes should nest within landforms.

Topographic wetness index was calculated according to the following equation (Beven and Kirkby 1979):

\[ w = \ln(a / \tan \beta) \]  

Where \( w \) is the topographic wetness index and \( a \) is the cumulative upslope area draining through a point (per unit contour length), \( \beta \) is the slope gradient at the point. Max down slope was used as \( \beta \) because it is considered better to express the effect of the relief on surface water distribution than the average slope (Hjerdt et al. 2004). This formula is used to calculate \( \ln \):

\[ \ln = \left( \frac{A_s}{\text{slope}} \right) \]  

Where \( A_s \) is the area. This formula is used to calculate \( A_s \):

\[ A_s = \left( a \times c^2 \right) / c \]  

Where \( a \) is the pixel area and \( C \) is the pixel size.

U.S.D.A soil taxonomy is chosen as the soil taxonomy system (Table1). U.S.D.A Soil Taxonomy developed by United States Department of Agriculture and the National Cooperative Soil Survey provides an elaborate classification of soil types according to several parameters (most commonly their properties) and in several levels: Order, Subgroup, Family and series. Soil series is currently used as the basic soil unit for soil mapping in this study. The subgroup is the auxiliary unit of soil Group and is defined according to whether the soils deviate from the central concept of a Group, or if they have some characteristics resulting from additional processes, or have remnant features inherited from the parent materials.
Table 1. Soil Classification and correlation table of Damavand region (source: Pedology Report of Damavand Region, Soil& water Research Institute, 1991, 12)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Calcaric fluvisols</td>
<td>Entisols</td>
<td>Typic Xerofluvents</td>
<td>Loamy skeletal, mixed, mesic</td>
<td>Damavand</td>
<td>1.1</td>
</tr>
<tr>
<td>2</td>
<td>Calcaric fluvisols</td>
<td>Entisols</td>
<td>Typic Xerofluvents</td>
<td>Loamy skeletal, mixed, mesic</td>
<td>Damavand</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>Calcaric fluvisols</td>
<td>Entisols</td>
<td>Typic Xerofluvents</td>
<td>Loamy skeletal, mixed, mesic</td>
<td>Damavand</td>
<td>1.3</td>
</tr>
<tr>
<td>4</td>
<td>Calcaric fluvisols</td>
<td>Entisols</td>
<td>Typic Xerofluvents</td>
<td>Fine loamy, mixed, mesic</td>
<td>Jaban</td>
<td>2.1</td>
</tr>
<tr>
<td>5</td>
<td>Calcic cambisol</td>
<td>Inceptisols</td>
<td>Calcixerollic Xerochrepts</td>
<td>Clayey skeletal, carbonatic, mesic</td>
<td>Sarbandan</td>
<td>3.1</td>
</tr>
<tr>
<td>6</td>
<td>Calcic cambisol</td>
<td>Inceptisols</td>
<td>Calcixerollic Xerochrepts</td>
<td>Fine, carbonatic, mesic</td>
<td>Rostam abad</td>
<td>4.1</td>
</tr>
<tr>
<td>7</td>
<td>Calcic cambisol</td>
<td>Inceptisols</td>
<td>Calcixerollic Xerochrepts</td>
<td>Fine, mixed, mesic</td>
<td>Sorkheheh</td>
<td>5.1</td>
</tr>
<tr>
<td>8</td>
<td>Calcic cambisol</td>
<td>Inceptisols</td>
<td>Calcixerollic Xerochrepts</td>
<td>Fine loamy, carbonatic, mesic</td>
<td>Absard</td>
<td>6.3</td>
</tr>
<tr>
<td>9</td>
<td>Calcic cambisol</td>
<td>Inceptisols</td>
<td>Calcixerollic Xerochrepts</td>
<td>Fine, mixed, mesic</td>
<td>Istgah</td>
<td>7.1</td>
</tr>
<tr>
<td>10</td>
<td>Calcic cambisol</td>
<td>Inceptisols</td>
<td>Fluventic Xerochrepts</td>
<td>Fine, mixed, mesic</td>
<td>Hesar</td>
<td>8.1</td>
</tr>
</tbody>
</table>

Continuous classification is based on fuzzy set concepts (Zadeh 1965), which provide a mathematical method of dealing with continuous data. These classification procedures have been used to classify natural resource phenomenon such as climate data (McBratney and Moore 1985), geologic data (Bezdek et al. 1984), and soils data (McBratney and DeGruijter 1992; Odeh et al. 1992; Slater 1994). Continuous classification is accomplished through an iterative process in which individual data points are allocated class memberships using simple statistical measures of distance in attribute space (Ward et al. 1992). The membership value indicates how closely related an individual is to the centroid (attribute mean) of that class. We used MacFuzzy (Ward et al. 1992), which performs continuous classification using fuzzy k-means. It provides the option of allowing extragrades, which are data points poorly represented by any class; therefore, they have a low membership in most or all of the classes. Their characteristics are different enough from other points to warrant not being classified and are placed into an outlier class (http://proceedings.esri.com). Fuzzy membership function describes how similarity between a local soil and the typical case of the given soil type will change as environmental conditions change. The similarity value varies from 0 (which means that local soil is very different from the given soil type) to 1 (which means that local soil is exactly the same with the given soil type). Relationships between soil and its environment can be captured using some combination of three basic forms: bell-shaped, s-shaped and z-shaped (Zhu 1999) (Figure 3). We used a Gaussian-like function as the basic form of fuzzy membership (Eq. (4)) (Zhu 1999; Shi et al. 2004; Qi et al. 2006):

\[
S_{ij,v} = e^{-((z_{ij,v} - v)^2 / \sigma^2)} D^v_{ij} (4)
\]
where $S_{k,v}^{ij}$ is the similarity of the local soil at point $(i,j)$ to soil type $k$ based on environmental variable (factor) $V$; $Z_{0,v}^{ij}$ is the value of environmental variable $v$ at the point; $Z_{0,v}$ is the typical value of environmental variable $v$ when the similarity of local soil to soil type $k$ is 1.0; and $D_{v}^k$ is the difference between $Z_{0,v}$ and the value of environmental variable.

**UNSUPERVISED CLASSIFICATION**

The unsupervised classification of the data was accomplished using ISOCLUSTER and MLCLASSIFY in ArcInfo GRID. The ISOCLUSTER function uses an isodata clustering algorithm for defining natural groupings of data points in attribute space. It is a commonly used algorithm in satellite image classification in which spectral signatures from multiple wavebands (equivalent to attributes) are used to determine classes. In our case, we used the five defined attributes for each pixel to determine class compositions. The ISOCLUSTER function returns a signature file which can be used as input to a maximum likelihood classifier (MLCLASSIFY); (http://proceedings.esri.com).

**RESULTS AND DISCUSSION**

Seventy fuzzy membership maps, one for each soil series, were generated for the study area. Fig. 4 shows three of them as example.

---

Fig3. Fuzzy membership functions: a. Gaussian, b. Fuzzy Linear membership function
Fig. 4. Fuzzy membership maps for each soil series: a. Topographic wetness index (Absard), b. Slope gradient (Absard), c. Planimetric curvature (Absard)

According to (Fig. 5.), 5 soil-landscape units have been determined as follows: 1-Valley bottoms, 2-low land plain, 3-moderate hill slope, 4-steep hill slope, and 5-mountain.Fig. 5. Soil-landscape map using ISOCLUSTER analysis method. A catenary sequence together with the environment conditions was calculated. This catenary sequence together with the environment conditions of the study area for each unit constituted the descriptive knowledge about the soil-landscape relationships over the region and formed the basis for constructing membership functions (Table 2). According to Table (2), soil-landscape characteristics of the study area have been determined as follows: 1.Valley bottoms soils: area: 86.87 Klm², these soils are distributed in the western and southeastern parts of the study area, topographic characteristics: elevation: 1817 meter, slope: 28%, curvature: -0.07, profile curvature: 0.99, planimetric curvature: 0.99, flow accumulation: 0.99, topographic wetness index: 6.53, pedologic characteristics: U.S.D.A soil taxonomy: mesic Typic Xerofluvents, Loamy Skeletal, mixed, order: Inceptisols, FAO classification: Calcic cambisol. 2. Low land plain soils: area: 202.91 Klm², topographic characteristics: elevation: 2000 meter, slope: 17%, curvature: 0.008, profile curvature: 0.99, planimetric curvature: 0.99, flow accumulation: 1, topographic wetness index: 7.12, pedologic characteristics: U.S.D.A soil taxonomy: mesic, Calcixerollic Xerochrepts fine, order: Inceptisols and Entisols, FAO classification: Calcic cambisol. 3. Moderate hill slope soils: area: 84.13 Klm², topographic characteristics: elevation: 2226 meter, slope: 34%, curvature: -0.002, profile curvature: 0.99, planimetric curvature: 0.99, flow accumulation: 0.98, topographic wetness index: 3.6, pedologic characteristics: U.S.D.A soil taxonomy:
mesic, mixed, Loamy Skeletal, Typic Xerofluvents, order: Entisols, FAO classification: Calcaric fluvisols. 4. Steep hill slope soils: area: 62.05Km², topographic characteristics: elevation: 2523 meter, slope: 42%, curvature: 0.05, profile curvature: 0.99, plani metric curvature: 0.99, flow accumulation: 0.99, topographic wetness index: 6, undefined soil classification. 5. Mountainous soils: area: 48.13Km², topographic characteristics: elevation: 2961 meter, slope: 61%, curvature: 0.02, profile curvature: 0.99, plani metric curvature: 0.99, flow accumulation: 0.97, topographic wetness index: 5.6, undefined soil classification.

Unsupervised and fuzzy classification methods both produce results that can aid soil scientists and others interested in soil-landscape processes. A few caveats apply. 1. Some knowledge of the landforms in question is necessary to evaluate the efficacy of the results of either method. Also, care should be taken in the choice of attributes; they should reflect the nature of the landscape and the phenomenon being studied. 2. The quality of the descriptive knowledge is dependent on the implementation of the purposive sampling approach such as the selection of the environmental variables, number of environmental clusters, the selection of classification algorithm and parameters which are discussed in another paper (Zhu et al. 2008).

The continuous (fuzzy) classification provides much additional information on each point, which may be useful for detailed understanding of landscape processes. The unsupervised classification provides summary information about landforms in a relatively quick manner; results are easy to view and transform into polygonal form recognized by soil scientists (http://proceedings.esri.com).

Table 2. Soil-landscape relationship

<table>
<thead>
<tr>
<th>Soil-Landscape unit</th>
<th>Area (Km²)</th>
<th>Elevation (meter)</th>
<th>Slope (%)</th>
<th>Curvature</th>
<th>Profile curvature</th>
<th>Plan curvature</th>
<th>Flow accumulation</th>
<th>Topographic wetness index</th>
</tr>
</thead>
<tbody>
<tr>
<td>valley</td>
<td>86.87</td>
<td>1817</td>
<td>28</td>
<td>-0.07</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>6.53</td>
</tr>
<tr>
<td>low land</td>
<td>202.91</td>
<td>2000</td>
<td>17</td>
<td>0.008</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
<td>7.12</td>
</tr>
<tr>
<td>plain</td>
<td>84.13</td>
<td>2226</td>
<td>34</td>
<td>-0.002</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>3.6</td>
</tr>
<tr>
<td>moderate hill slope</td>
<td>62.05</td>
<td>2523</td>
<td>42</td>
<td>0.05</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>6</td>
</tr>
<tr>
<td>steep hill slope</td>
<td>48.13</td>
<td>2961</td>
<td>61</td>
<td>0.02</td>
<td>0.99</td>
<td>0.99</td>
<td>0.97</td>
<td>5.6</td>
</tr>
</tbody>
</table>

VALIDATION AND EVALUATION OF THE SOIL-LANDSCAPE MAP

To calculate the Accuracy of the Soil-landscape map regression models is used. Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable). The variables we are using to predict the value of the dependent variable are called the independent variables (or sometimes, the predictor, explanatory or regressor variables). Multiple regression also allows you to determine the overall fit (variance explained) of the model and the relative contribution of each of the predictors to the total variance explained (https://statistics.laerd.com). The most appropriate model could be a straight line, a higher degree polynomial, a logarithmic or exponential. The strategies to find an appropriate model include the forward method in which we start by assuming the very simple model i.e. a straight line:

\[ Y = a + bX \text{ or } Y = b_0 + b_1X \]  \hspace{1cm} (5)

Then we find the best estimate of the assumed model. If this model does not fit the data satisfactorily, then we assume a more complicated model e.g. a 2nd degree polynomial \( Y = a + bX + cX^2 \) and so on. In a backward method we assume a complicated model e.g. a high degree polynomial, we fit the model and we try to simplify it. We might also use a model suggested by theory or experience. Often a straight line relationship fits the data satisfactorily and this is the case of simple linear regression. The simplest case of linear regression analysis is that with one
predictor variable (http://www.ncbi.nlm.nih.gov). In this study elevation, slope gradient, curvature, profile curvature, Plani metric curvature, flow accumulation and topographic wetness index variables are independent variables and the dependent variable is soil-landscape map.

Based on the above description, we can say that the accuracy of the Soil-landscape map generated using ISOCLUSTER analysis method was at about 74%. (Table 3). The results indicated that the soil-landscape map can capture the local variation of soil and the overall spatial distribution of soil well. 74% accuracy is acceptable for an initial soil mapping. This suggests that the membership functions constructed in this study do capture the major pattern of soil-landscape relationships over the region.

**Conclusions**

The unsupervised classification may be best suited to identifying landforms, or areas of rapid change, which may require more intensive sampling to obtain a complete picture of the soils in an area. Because the fuzzy classification provides more complete information about the landscape characteristics, it may be better suited for the prediction of soil properties in sparsely sampled areas and to further understanding of underlying soil-landscape processes. From the results of the case study we concluded that:

1. The constructed fuzzy membership functions were able to produce good quality soil spatial information.
2. Accuracy of the Soil-landscape map generated using ISOCLUSTER analysis method was at about 74%. 74% accuracy is acceptable for an initial soil mapping.
3. The proposed methods provide an effective way to quantify knowledge on soil-environment relationships for soil-landscape map, especially for those areas with limited data.
4. The membership functions constructed in this study do capture the major pattern of soil-landscape relationships over the region.

**Table 3. Regression statistics between environment conditions and soil-landscape map**

<table>
<thead>
<tr>
<th>Std. Error of the Estimate</th>
<th>Adjusted R Square</th>
<th>R Square</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.804</td>
<td>0.557</td>
<td>0.557</td>
<td>0.74</td>
</tr>
</tbody>
</table>

**Acknowledgements**

This research is supported by Basir remote sensing institute, Tehran, Iran. Our thanks go to the Basir remote sensing institute.

**References**


Beard, J.S., (2005). Pre-European Vegetation - Western Australia (NVIS Compliant version), Department of Agriculture Western Australia, Metadata Statement, Metadata_ID: ANZWA1050000123


Yang, X., Chapman, GA., Gray, JM., Young, MA., (2007). Delineating soil-landscape facets from digital elevation models using compound


http://en.wikipedia.org/wiki/Regression_analysis
http://www.google.com/earth/
www.basir-rsi.ir : 25 M resolution DEM which was created from the Aster satellite radar images provided via Basir rem