SEMI-AUTOMATED CLASSIFICATION OF LANDFORM ELEMENTS IN ARMENIA BASED ON SRTM DEM USING K-MEANS UNSUPERVISED CLASSIFICATION

ARTAK PILOYAN¹, MILAN KONEČNÝ²

¹Faculty of Geography and Geology, Yerevan State University, Armenia
²Department of Geography, Faculty of Science, Masaryk University, Brno

Manuscript received: January 12, 2017
Revised version: February 3, 2017


ABSTRACT: Land elements have been used as basic landform descriptors in many science disciplines, including soil mapping, vegetation mapping, and landscape ecology. This paper presents a semi-automatic method based on k-means unsupervised classification to analyze geomorphometric features as landform elements in Armenia. First, several data layers were derived from DEM: elevation, slope, profile curvature, plan curvature and flow path length. Then, k-means algorithm has been used for classifying landform elements based on these morphometric parameters. The classification has seven landform classes. Overall, landform classification is performed in the form of a three-level hierarchical scheme. The resulting map reflects the general topography and landform character of Armenia.

Corresponding author: artakpiloyan@ysu.am

Introduction

Information about landforms is crucial, for instance, for studies of landscape evaluation and, assessment of potential erosion, natural hazards and risk prediction, as well as for regional planning of landscapes, cadastral assessment of land use etc. The classical ways to implement (combine) relief units on landscape assessment is to delineate them during ground truthing and/or by using aerial photographs. Such an approach is relatively time consuming and the results are largely dependent on the researcher’s subjective assessment and, therefore, in many cases they may not have sufficient transparency and reproducibility (Dragut, Blaschke 2006). Terrain analysis in environmental assessments of landscape was quite rare until the last decade, while terrain topology is a key indicator in a wide range of environmental processes (Bates et al. 1998, Butler 2001). Nevertheless, such a landscape approach is an important step in a variety of regional studies (Dragut, Blaschke 2006). Particularly, landscape pattern could have a crucial influence on the spread of disturbance (Turner 1990, Forman 1995, Butler 2001), water quality (Hunsaker et al. 1992, Arakelyan, Piloyan 2011), environmental quality and biodiversity (Gordon et al. 1994), as well as substantiate the mechanisms of material flow (Dalrymple et al. 1968).

Terrain analysis is the set of activities which leads to the compilation of terrain characteristics or

The aim of this paper is to perform landform classification via k-means unsupervised classification of geomorphometric parameters computed from the SRTM DEM of the territory of Armenia.

Material and Methods

Geomorphometric parameters

Elevation, slope, plan and profile curvatures according to Evans (1972) and Shary (1995, Shary et al. 2002) schemas and the flow path length according to Horton method (Horton 1945, Mitasova, Hofierka 1993, Zavoianu et al. 2013) were used as a DEM-based geomorphometric layers for semi-automated landform classification of the territory of Armenia. Evans (1972), Lastochkin (1991), and Simonov (1999) point out that elevation, slope and aspect are compulsory for every kind of geomorphometric analysis. But for solving a number of geomorphological and geo-ecological issues plan and profile curvatures are also necessary. Plan curvature is a measure of flow convergence (kh < 0) and divergence (kh > 0) and determines soil water or the deposition of particles. Profile curvature controls water flow acceleration (kh > 0) and deceleration (kh < 0) and therefore the erosion potential of an area (Shary 1995, Shary et al. 2002, Schillaci et al. 2015). For both parameters units are expressed in 1/m. The effectiveness of the flow path length maps is also undisputed, because this parameter is a component for many empirical equations for calculation of surface erosion (Zavoianu et al. 2013).

Digital Elevation Model

To perform terrain analysis a DEM containing information about relief morphology in sufficient accuracy is very useful. During the last two decades the availability of DEM data has been continuously growing. The most common freely available global DEMs are SRTM (Shuttle Radar Topographic Mission) with 90 m and 30 m spatial resolution (Rabus et al. 2003, Berry et al. 2007, Farr et al. 2007) and ASTER GDEM (Advanced Spaceborne Thermal Emission and Reflection Radiometer/Global Digital Elevation Model) with 30 m spatial resolution (Hayakawa et al. 2008). SRTM DEM provides an essentially constant quality, while the quality of ASTER DEM is uneven (Crippen et al. 2007, Huggel et al. 2008). This difference is partly due to the wavelengths used in these models. The ASTER sensor works in the near infrared bandwidth which is disturbed by cloud coverage, while the SRTM works with microwave frequencies which are not affected by clouds (Stevens et al. 2004, Hubbard et al. 2007).

SRTM DEM with 90 m spatial resolution reduces the possibilities for study of small elements of relief, large elements are represented properly (Wright et al. 2006, Dragut, Eisank 2012), which in our opinion is sufficient for a terrain analysis covering the whole territory of Armenia. For this reason our analysis is based on the SRTM elevation dataset.

SRTM consisted of a specially modified radar system that flew on board the Space Shuttle Endeavour during 11 days in February 2000. SRTM utilized the dual Space-borne Imaging Radar (SIR-C) and dual X-band Synthetic Aperture Radar (X-SAR) configured as a baseline interferometer, acquiring two images at the same time, which could produce a DEM (Rodriguez et al. 2005). Many researchers used SRTM data for applications such as analysis of terrain characteristics (Rabus et al. 2003, Falorni et al. 2005) including volcanic morphology and large aeolian bedforms (Wright et al. 2006, Blumberg 2006), vegetation studies (Kellndorfer et al. 2004),
hydrologic modelling (Ludwig, Schneider 2006), morphotectonic analysis (Grohmann et al. 2006) and identification of circular structures of terrain (Piloyan, Avagyan 2016).

Software

Extraction of geomorphometric parameters and all the analyses in this article were performed in the open source SAGA GIS, version 2.12 a software package licensed under the General Public License (GNU) and software called Whitebox GAT (Geospatial Analysis Tools) version 2.2 Iguazu, developed by Lindsay (2014). ESRI ArcGIS, version 10.2 was used for map design and presentation.

**K-means unsupervised classification algorithm**

Historically it so happened that the k-means algorithm was discovered by several researchers: first of all by Steinhaus (1956) and the same time by Lloyd (1957, 1982), but it has received wide recognition especially after McQueen’s (1967) work. The initial aim of the k-means approach was the reduction of data value to support the transfer of information in the area of pattern recognition (Bezdek 1981).

K-means algorithm is one of the simplest and well known clustering algorithms. This algorithm determines the cluster centers and the elements belonging to them by minimizing the squared error based objective function. The aim of the algorithm is to locate the cluster centers as much as possible far away from each other and to associate each data point to the nearest cluster center. Euclidean distance is usually used as the dissimilarity measure in k-means algorithm. The objective function $J$ is described as follows:

$$ J = \sum_{j=1}^{k} \sum_{i=1}^{N} \left\| x_{ij} - c_j \right\|^2 $n$

where $\left\| x_{ij} - c_j \right\|^2$ is a chosen distance measure between a data point $x_{ij}$ and the cluster centre $c_j$, is an indicator of the distance of the $N$ data points from their respective cluster centres (Orhan et al. 2008).

It should be emphasized that k-means algorithm historically is one of the most important data mining algorithms. We are not the first who use k-means or its soft version fuzzy k-means algorithms to classify geospatial phenomenon. It has been successfully used in geohydrology, soil science and vegetation mapping (Vriend et al. 1988, de Bruin, Stein 1998, Burrough, McDonnell 1998, Burrough et al. 2000, 2001, Schmidt, Hewitt 2004).

Selected terrain parameters (elevation, slope, plan and profile curvature and flow path length) were treated as a single band images and using a k-mean unsupervised algorithm in the Whitebox GAT 2.2 software, the landforms were classified. K-means uses an iterative procedure that finds statistically similar groups in multi spectral space during its analysis. The algorithm starts by randomly locating $k$ clusters in spectral space (Burrough et al. 2000, Moravej et al. 2012). Each pixel in the input image group is then assigned to the nearest cluster centre and the cluster centre locations are moved to the average of their class values. This classification is then repeated until a stopping condition is reached. It is supposed that no two clusters have the same cluster representative. In the case that two cluster centers coincide, a cluster center should be perturbed to avoid coincidence in the iterative process (Chang et al. 2011). The stopping condition may either be a maximum number of iterations or a tolerance threshold which designates the smallest possible distance to move cluster centers before stopping the iterative process (Moravej et al. 2012). In our case, the determinative parameters were 7, 25 and 2% respectively, for the number of predictive classes, maximum iteration and pixel class change tolerance.

**Case Study**

A typical characteristic of the hypsometry of the territory of Armenia is a sequence of four major geomorphological provinces (from north to south) which are different in origins and altitude statistics: Those provinces are 1) The Province of Northern Fold-block Ranges and Intermountain Valleys, 2) The Volcanic Highland, 3) The Province of Southern Fold-block Ranges and Intermountain Valleys and 4) Middle Araks Depression (Fig. 1). Based on the hypsometric curve, we can say that the territory of Armenia is a medium-altitude
mountainous country dominated by absolute altitudes of 1300–2500 m (Fig. 2). Visual analysis of the hypsometric curve shows that 10.49% of the territory of Armenia is situated at altitudes up to 1000 m, and almost half (47.84%) – from 1000 to 2000 m. Only 14.41% of the total territory is above 2500 m (Avagyan et al. 2010). According to orographic and morphographic zones (nomenclature according to Zohrabyan 1979) the territory of the Republic is divided as follows: 68.23% of the area is occupied by high and medium mountains (1500–2500 m) and more than a quarter (28.56%) by low mountains (500–1500 m). Extreme altitudes constitute only small areas: 3.21% are the highest peaks and summits of highest ranges and only 0.12% of Armenia is below 500 m. The highest point in the territory of Armenia is the northern peak of the mount Aragats – 4090 m, and no point is below 390 m: the lowest points are in the valley of the Araks River at the state border with Iran. The average elevation of the territory is 1853 m (Avagyan et al. 2010).

It is worth mentioning that hypsometry of fold-block mountains usually develops by neotectonic (Late Neogene – Quaternary) movements which are differentiated in Armenia. Within the individual blocks of the Earth’s crust, they are expressed by different degrees of intensity and different

Fig. 1. Map of the geomorphological provinces of Armenia.
amplitudes. This phenomenon creates various elevation levels of young mountain relief. During the formation of this neovolcanic (Quaternary) relief, along with the neotectonic movements lavas and tuff were deposited. However, in the territory of Armenia, their thickness is no more than a few hundred meters. For this reason, the main hypsometric levels in volcanic relief basically repeat the hypsometry of the relief below the lava cover (Zohrabyan 1979).

The relief of Armenia has repeatedly served as an object for morphological delineation, both alone and as part of larger spatial units. This is evidenced by the relief zoning schemes by Balyan (1969), Gevorgyan, Poghosyan (1970), Zohrabyan (1979), Boyagryan (2007). Most authors recognize that interzonal landscape differentiation is significant. Often the decisive role belongs to the geological and geomorphological factors, which are in fact non-zonal factors. This phenomenon is particularly noticeable at the lower levels of landscape classification, where the location factor is a decisive criterion for the identification of such taxonomic units as landforms (Isachenko 1991).

As landforms forma conjugate system of facies that is combined by the general orientation of physico-geographical processes, they are most clearly expressed in the conditions of classified relief with alternating convex and concave shapes of mesorelief, such as summits and hollows, ridges and depressions, shoulder slopes and footslopes, gullies and ravines, etc.

The relief of most of the territory of Armenia is a succession of such meso forms. The provinces of Northern and Southern Fold-block Ranges and Intermountain Valleys and most part of The Volcanic Highland are alternations of valleys, ridges, crests, summits, passes, channels, pits, hollows and other geomorphological middle level forms.

**Landform Classification**

There are many approaches and classification schemes for the landform classification in digital geomorphology (Pennock et al. 1987, Dikau 1989, Wood 1996, Irvin et al. 1997, Ventura, Irvin 2000, Macmillan et al. 2000, Burrough et al. 2001, Shary et al. 2002, Schmidt, Hewitt 2004, Dragut, Blaschke 2006, Ehsani, Quiel 2008, MacMillan, Shary 2009). Here a combination of the classification criteria of Dikau (1989), and Dragut, Blaschke (2006) with small modifications is used. And to determine a landform classes we were mainly focused on the landform classification scheme of Ventura, Irvin (2000). As noted, the following spatial layers of preliminary information were used for landform classification: altitude, plan curvature, profile curvature, slope and an additional layer flow path length. Table 1 gives the basic statistical information for these parameters and Table 2 their correlation matrix.

**Table 1. Statistics of geomorphometric parameters.**

<table>
<thead>
<tr>
<th>Geomorphometric parameter</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude [m]</td>
<td>371</td>
<td>1852</td>
<td>4054</td>
<td>604</td>
</tr>
<tr>
<td>Flow path length [m]</td>
<td>0</td>
<td>2534</td>
<td>33411</td>
<td>3061</td>
</tr>
<tr>
<td>Plan curvature [1/m]</td>
<td>-3.07</td>
<td>0.0057</td>
<td>3.43</td>
<td>0.23</td>
</tr>
<tr>
<td>Profile curvature [1/m]</td>
<td>-4.17</td>
<td>0.0062</td>
<td>4.1</td>
<td>0.19</td>
</tr>
<tr>
<td>Slope [º]</td>
<td>0</td>
<td>11.1</td>
<td>65.1</td>
<td>8.6</td>
</tr>
</tbody>
</table>

**Table 2. Correlation matrix of geomorphometric parameters.**

<table>
<thead>
<tr>
<th>Geomorphometric parameter</th>
<th>Altitude</th>
<th>Flow path length</th>
<th>Plan curvature</th>
<th>Profile curvature</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Flow path length</td>
<td>0.022</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Plan curvature</td>
<td>0.033</td>
<td>-0.030</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Profile curvature</td>
<td>-0.079</td>
<td>0.033</td>
<td>-0.449</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Slope</td>
<td>0.235</td>
<td>-0.212</td>
<td>0.051</td>
<td>0.000</td>
<td>1</td>
</tr>
</tbody>
</table>
It is worth mentioning that aspect have not been included in basic information sources in spite of the fact that it is an important derivative of elevation (Evans 1972). The reason for not using is that ...aspect produces additional zonation, for instance when the aspect of hillslopes changes from south to west or north to east. This zonation makes the outputs too confusing. Moreover, north-facing slopes are artificially split due to the great difference between pixel values (e.g. 1 and 360°) (Dragut, Blaschke 2006). But information about the aspect could easily be utilized for specific geoecological and geomorphological studies, in particular, in the evaluation of potential erosion, slope stability assessment, identification of the linear and circular morphostructures.

Dikau (1989) distinguished nine landform classes based on combinations of convex, straight and concave profiles (Fig. 3). In theory these are all possible combinations of landform classes which can be derived from the values of plan and profile curvature. However, three of these classes occur more frequently than the others in real relief. Those three classes are: “shoulder” – convex element between summit and backslope, “backslope” – the typically linear, inclined part of hillslope and “footslope” – the concave part of the hillslope that welds the linear segments to lower terrain and is both an erosional and depositional surface. Taking into account the resolution of the preliminary DEM and the territory of case study the other six classes have been attached with different possible degrees of membership to one or more of the main classes (as shown by arrows in Fig. 3).

However, it is clear that landform classification based only on a curvature values would be incomplete. In the methodology of Dragut,

Table 3. Geomorphometric parameters used in landform classification.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Description</th>
<th>Altitude [m]</th>
<th>Flow path length [1/m]</th>
<th>Plan curvature</th>
<th>Profile curvature</th>
<th>Slope [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summits, Ridges</td>
<td>An upland surface with an inclination which differs distinctly from the hill</td>
<td>Higher than</td>
<td>&lt;500</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>2</td>
<td>Cliffs, Steep</td>
<td>Surface areas with slope gradient higher than 35°</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>&gt;35°</td>
</tr>
<tr>
<td>3</td>
<td>Shoulders</td>
<td>Convex element between summits and backslope</td>
<td>ND</td>
<td>ND</td>
<td>±0</td>
<td>±0</td>
<td>ND</td>
</tr>
<tr>
<td>4</td>
<td>Backslopes</td>
<td>The typically linear, inclined part of hillslope</td>
<td>ND</td>
<td>ND</td>
<td>±0</td>
<td>±0</td>
<td>ND</td>
</tr>
<tr>
<td>5</td>
<td>Footslopes</td>
<td>The concave part of the hillslope that welds the linear segments to lower</td>
<td>ND</td>
<td>ND</td>
<td>-</td>
<td>-</td>
<td>ND</td>
</tr>
<tr>
<td>6</td>
<td>Toeslopes</td>
<td>The region which extends away from the base of the hillslope and composed</td>
<td>Lower than</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>&lt;2°</td>
</tr>
<tr>
<td>7</td>
<td>Flat areas,</td>
<td>Surface areas with slope gradient less than 2°</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>&lt;2°</td>
</tr>
</tbody>
</table>

Fig. 3. Fundamental local landform elements based on plan and profile curvatures.
Blaschke (2006) Dikau-based classes are combined with another four, irrespective of curvature values which were derived from slope gradient and an additional parameter, a local dominance criterion. In our approach, as an additional parameter we used the value of flow path length, which is essentially close to relative altitudes, but easier to derive from elevation. Areas and relief elements for which the value of flow path length is 500 m and less and are situated higher than the neighbours are defined as “summits” and “ridges”. Parts of slopes which are lower than the neighbour slopes are “toeslopes”. The areas with slope gradient less than 2° are defined as “flat areas”, “floodplains”, while higher than 35° as “cliffs” and “steep slopes”. The last two classes are the only which were classified based on constant values. Thus, the result was seven landform classes (Table 3).

The seven classes represented in Table 3 are structured in a hierarchy and grouped into similar groups. Overall, landform classification is performed in the form of a three-level hierarchical scheme (Fig. 4). On the first level they are distinguished according to altitudinal zonation (Zohrabyan 1979): Lower Mountain Zone (<1500 m), Middle Mountain Zone (1500–2500 m) and Higher Mountain Zone (>2500 m).

On the middle level of the class hierarchy classification was performed according to slope mesorelief elements: These elements are: “summits” and ridges, “flat areas” and the actual slope. It should be noted that on this level, for instance, the “summits” are separated not only in the Higher Mountain Zone, but also in the others, where they can be expressed on the relief as a tops of local small hills. The same logic applies also for the flat areas.

Finally on the last level, only slopes from the second level were classified. The classes are five: “cliffs” and “steep slopes”, “shoulders”, “backslopes”, “foot slopes”, “toeslopes”.

**Results**

Thus, k-means unsupervised classification methodology has been used for classifying landform elements from DEM of Armenia, based on morphometric parameters of relief (Fig. 5). The resulting map reflects the general topography and landform character of the territory. The advantage of the approach is that it not only produces a landform classification, but the final map can also be used to measure the uncertainty in the classification. The method is simple with respect to computational and parameter requirements and is therefore easy to apply. However, it should be mentioned, that the results of DEM processing are difficult to quantitatively verify because of the lack, at this stage of the research, of ground truth data for geomorphologic features beyond altitude.

Visual analysis of the map obtained through classification shows that “peaks”, “steep slopes” and “flats” have been most clearly distinguished, which is logical, given the fact that they have been classified based on constant threshold values (Table 3).

Table 4 presents the percentage frequency of the spatial distribution of landforms in mountain zones for the whole territory of Armenia. The fact that “foot slopes”, “backslopes” and “shoulders” are the dominant forms indicates the erosion-fragmented nature of relief, which is characterized by the alternation of concave and convex forms of relief. Steep slopes are mainly situated.

![Fig. 4. Landform class hierarchy.](image-url)
Fig. 5. Map of landforms in Armenia.
in the Zones of Northern and Southern Fold-block Ranges and are expressed as slopes of large canyons and gorges.

“Flat areas” (except the Middle Araks Depression) are common in The Volcanic Highland, where they are represented as floors of intermountain valleys, volcanic plateaus, river floodplains and large erosional terraces.

Discussions and Conclusions

Combination of the classification criterias of Dikau (1989), and Dragut, Blaschke (2006) was used to create a map of Armenia with seven landform classes using k-means landform classification algorithm in the Whitebox GAT software. However, the number of classes can be modified, based on study area characteristics and research objectives.

The results show that it is possible to derive landforms for mountain regions based on geomorphometric parameters using a relatively simple algorithm. The advantage of the algorithm is that it requires a relatively limited set of parameters, which are directly related to landform features (Table 3). Finally, automated and semi-automated landform classification methodologies are important parts for many ecological applications, soil resource modelling, landslide hazards, sea-floor and coastal geomorphology (Dragut, Blaschke 2006). Taking into account the very complicated and mountainous characteristics of the topography of Armenia we argue that this methodology can be important in applications dedicated to disaster risk assessment and management.

Automated and semi-automated methods of landform classification based on GIS and Remote Sensing technologies using geomorphometric parameters in scientific environments started over 25 years ago, but this is the first attempt for the territory of Armenia or parts. Therefore, it cannot be without the obvious and not obvious drawbacks. First of all it is worth mentioning that results accuracy assessment was not carried out based on ground truthing data and auxiliary aerial images for the results of this study. It is important to note that in the case of classification of k-means no object is left unclassified. This is achieved through overlapping every pixel in case study in such a way, that as a result of this classification, each point belongs to only one k cluster. Still, the accuracy assessment in a classification is crucial, and we believe that it will be performed at further stages of study. At this stage, the main purpose was to test one of the existing methods for landform classification of the territory of Armenia and evaluate its behaviour in a territory with such heterogeneous terrain.

References


Lindsay J. B. 2014. The Whitebox GeoSpatial Analysis Tools project and open-access GIS. In: *Proceedings of the GIS Research UK 22nd Annual Conference*, The University of Glasgow, 16–18 April. DOI: 10.13140/RG.2.1.1010.8962


