Clustering of landforms using self-organizing maps (SOM) in the west of Fars province

Marzieh Mokarram¹ and Dinesh Sathyamoorthy²

¹ Department of Range and Watershed Management, College of Agriculture and Natural Resources of Darab, Shiraz University, Iran.
² Science & Technology Research Institute for Defence (STRIDE), Ministry of Defence, Malaysia

Email: m.mokarram@shirazu.ac.ir

Abstract. The aim of this study is to cluster landforms in the west of the Fars province, Iran using self-organizing maps (SOM). In SOM, according to qualitative data, the clustering tendencies of landforms were investigated using six morphometric parameters, which were slope, profile, plan, elevation, curvature and aspect. First, topographic position index (TPI) was used to prepare the landform classification map. The results of SOM showed that there were five classes for landform classification in the study area. Cluster 5 corresponds to high slope, high elevation but with different of concavity and convexity that consist of ridge landforms. Cluster 3 corresponds to flat areas, possibly plantation areas, in medium elevation and almost flat terrain. Clusters 1, 2 and 4 correspond to channels with different slope conditions.

1. Introduction

Landforms are the result of geomorphologic processes that occur on the earth’s surface. The term landform is used for morphometrically homogeneous land-surface regions due to the effects of common geological and geomorphological processes [1]. As this concept of landforms is an idealized one, it then follows that the closer the studied landform conforms to its definition, the greater the accuracy of the obtained model. A subdiscipline of geomorphology is geomorphometry, which provides quantitative and qualitative description and measurement of landforms [1, 2, 3, 4], and is based on the analysis of variations in elevation as a function of distance. A basic underlying principle of geomorphometry is digital elevation models (DEMs). A number of studies have employed DEMs for the automated mapping of landforms [5, 6, 7, 8, 9]. Derivation of landforms can be carried out using various approaches, including classification of terrain parameters [10, 11], filter techniques [12], cluster analysis [13] and multivariate statistics [14]. A common focus of the study of landforms is to consider them as formed by small and simple elements which are topologically and structurally related. A more complete description of landform may be achieved by using complex terrain parameters calculated from simple ones, e.g., the topographic wetness index [15], stream power index [16], aggradation and degradation indices [17], thresholds [10], automated classification using object-based image analysis multivariate descriptive statistics [10, 18], double ternary diagram classification [1], discriminant analysis [19], fuzzy logic and unsupervised classification [6, 14, 20, 21, 22] and neural networks [23].
Self-organizing map (SOM) is an unsupervised and nonparametric artificial neural network (ANN) algorithm that clusters high dimensional input vectors into low dimensional (usually two-dimensional) output map that preserves the topology of the input data [23]. Preserving topology means that the SOM map preserves the relations between the input neighboring points in the output space [24]. Hosokawat and Hoshit (2001) used SOM to generate a damage distribution map in Kobe city in Japan that corresponds with the actual damage recorded following the 1995 earthquake [25]. Ehsani and Quiel (2008) employed SOM and Shuttle Radar Topography Mission (SRTM) data to characterize yardangs in the Lut desert, Iran [23]. The results demonstrate that SOM is a very efficient tool for analyzing aeolian landforms in hyper-arid environments that provides very useful information for terrain feature analysis in remote regions. Ferentinou and Sakellariou (2010) applied SOM in order to rate slope stability controlling variables in natural slopes, while Ferentinou et al. (2010) used SOM to classify marine sediments [26]. Mokarram et al. (2014) used SOM to study the relationships between geomorphological features of alluvial fans and their drainage basins. The results of the analysis showed that several morphologically different fan types were recognized based on their geomorphological characteristics in the study area [22].

The west of Fars, Iran has the highest points in the Fasa city, with a rich wildlife preserve, known as Kharmankoooh, which is a hunting prohibited region. The area has a variety of landforms that it is suitable for the wildlife. As different animals live in different landforms, clustering of landforms can be used to predict the type of wildlife in the study area. To this end, the aim of this study is to cluster the landforms using SOM based on morphometric characteristics.

2. Material and methods
2.1. Study area
The study area is located in west of Fars province, southwest Iran, which is shown in Figure 1. It located between 29° 00′ to 29° 25′ northern latitude and 53° 19′ to 53° 54′ eastern longitude (Figure 1). Six morphometric parameters were analyzed; slope, profile, plan, elevation, curvature and aspect (Figure 2 and Table 1).

Figure 1. Location of the study area.
Figure 2. Morphometric parameters of the study area.
Table 1. Morphometric parameters measured for the determination of landform classification.

| Parameter        | Min  | Max  | Average | STDVE |
|------------------|------|------|---------|-------|
| Slope (°)        | 0.38 | 42.12| 9.94    | 12.11 |
| Profile (1/m)    | -0.40| 0.44 | 0.01    | 0.14  |
| Plan (1/m)       | -0.48| 0.66 | 0.01    | 0.20  |
| Elevation (m)    | 1484.00| 2805.00| 1898.58| 345.93 |
| Curvature (1/m)  | -0.48| 0.98 | 0.01    | 0.26  |
| Aspect (°)       | 7.70 | 358.15| 169.42  | 111.06 |

2.2 Landform classification using topographic position index (TPI)

TPI [27] compares the elevation of each cell in a DEM to the mean elevation of a specified neighborhood around that cell. Positive TPI values represent locations that are higher than the average of their surroundings, as defined by the neighborhood (ridges). On the other hand, negative TPI values represent locations that are lower than their surroundings (valleys). TPI values near zero are either flat areas (where the slope is near zero) or areas of constant slope (where the slope of the point is significantly greater than zero) [27].

TPI (Eq. (1)) compares the elevation of each cell in a DEM to the mean elevation of a specified neighborhood around that cell. Mean elevation is subtracted from the elevation value at the center [27]:

\[ TPI = Z_0 - \frac{1}{n} \sum_{n=1} Z_n \]  

where;

- \( Z_0 \) = elevation of the model point under evaluation
- \( Z_n \) = elevation of grid
- \( n \) = the total number of surrounding points employed in the evaluation.

Combining TPI at small and large scales allows a variety of nested landforms to be distinguished Table 2 and Figure 3. The exact breakpoints among classes can be manually chosen to optimize the classification for a particular landscape. As in slope position classifications, additional topographic metrics, such as variances of elevation, slope, or aspect within the neighborhoods, may help delineate landforms more accurately [27].

Table 2. Landform classification based on TPI [27].

| Classes                                | Description                                      |
|----------------------------------------|--------------------------------------------------|
| Canyons, deeply incised streams        | Small Neighborhood: \( z \leq -1 \)              |
|                                        | Large Neighborhood: \( z \leq -1 \)              |
| Midslope drainages, shallow valleys    | Small Neighborhood: \( z \leq -1 \)              |
|                                        | Large Neighborhood: \(-1 < z < 1\)              |
| Upland drainages, headwaters           | Small Neighborhood: \( z \leq -1 \)              |
|                                        | Large Neighborhood: \( z \geq 1 \)              |
| U-shaped valleys                       | Small Neighborhood: \(-1 < z < 1\)              |
|                                        | Large Neighborhood: \( z \leq -1 \)              |
| Plains small                           | Neighborhood: \(-1 < z < 1\)                     |
|                                        | Large Neighborhood: \(-1 < z < 1\)              |
|                                        | Slope \( \leq 5^\circ \)                        |
| Open slopes                            | Small Neighborhood: \(-1 < z < 1\)              |
|                                        | Large Neighborhood: \(-1 < z < 1\)              |
Slope > 5°

Upper slopes, mesas
Small Neighborhood: \(-1 < z < 1\)
Large Neighborhood: \(z \geq 1\)

Local ridges/hills in valleys
Small Neighborhood: \(z \geq 1\)
Large Neighborhood: \(-1 < z < 1\)

Midslope ridges, small hills in plains
Small Neighborhood: \(z \geq 1\)
Large Neighborhood: \(-1 < z < 1\)

Mountain tops, high ridges
Small Neighborhood: \(z \geq 1\)
Large Neighborhood: \(z \geq 1\)

The classes of canyons, deeply incised streams, midslope and upland drainages, and shallow valleys tend to have strongly negative plane form curvature values. On the other hand, local ridges / hills in valleys, midslope ridges, small hills in plains and mountain tops, and high ridges have strongly positive plane form curvature values.

![Topographic lines diagram](image)

**Figure 3.** Position of each class on topographic lines [28].

### 2.3. Self-organizing map (SOM)

SOMs are unsupervised ANNs formed from neurons located on a regular, two-dimensional regular planar array grid (Figure 4). In fact SOM is based on unsupervised learning, which means that no human intervention is needed during the learning and little needs to be known about the characteristics of the input data. SOM offers a solution to apply a number of visualizations linked together [29]. When several visualizations are linked together, scanning through them is very efficient because they are interpreted in a similar way [30]. The U-matrix produced from SOM visualizes distances between neighboring map units and thus shows the cluster structure of the map. Samples within the same cluster will be the most similar according to the variables considered [30].

The SOM algorithm consists of two individual stages: the competitive and cooperative stages. In the competitive stage, the best matching neuron is selected, while in the cooperative stage, the weights of the winner are adapted as well as those of its immediate lattice neighbors [31]. Further explanation for each of the steps is as follows:
1. Competitive stage:
Let $A$ be a lattice of $N$ neurons with weight vectors $w_i = [w_{ij}] \in \mathbb{R}^d$. All the neurons receive the same input vector $v = [v_1 ... v_d] \in \mathbb{R}^d$. For each input $v$, we select the neuron with the smallest Euclidean distance ("winner-takes-all", WTA) [32]:

$$i^* = \arg \min_i \| w_i - v \|$$

(2)

where $w_i$ is neuron weights and $v$ is input vector.

2. Cooperative stage
The weight update rule in incremental mode is as follows [32]:

$$\Delta w_i = \eta \Lambda(I, i^*, \sigma \Lambda(r))(v - w_i), \forall i \in A$$

(3)

where $\Lambda$ is the neighborhood function, which is a scalar-valued function of the lattice coordinates of neurons $i$ and $i^*$, $r_i$ and $r_i^*$, mostly a Gaussian:

$$\Lambda(i, i^*) = \exp(-\| r_i - r_{i^*} \|^2 / 2\sigma^2)$$

(4)

with range $\Lambda$ (i.e., the standard deviation). The positions $r_i$ are usually taken to be the nodes of a discrete lattice with a regular topology [32].
Figure 4. The structure of a SOM network: (a) Selection of a node and adaptation of neighboring nodes to the input data. The SOM grid can be (b) hexagonal or (c) rectangular. The black object indicates that the node that was selected as the best match for the input pattern [33].

3. Results and discussion

3.1. Landform classification

The TPI maps generated using small and large neighborhoods are shown in Figures 5. TPI is between -94 to 77 and -203 to 260 for 5 and 15 cells respectively (Figure 5). The landform maps generated based on the TPI values are shown in Figure 6.

Figure 5. TPI maps generated using (a) small (5 cell) and (b) large (15 cell) neighborhood.
Figure 6. Landform classification using the TPI method.
3.2. SOM for landform classification
SOM was applied for the study area to describe the landform classification. The visualizations in Figures 7 and 8 (for landform classification) consist of 16 hexagonal grids, with the U-matrix in the upper left, along with the six component layers (one layer for each morphometric parameter examined in this study). As previously mentioned, the clustering of landform classification used the morphometric parameters of slope, profile, plan, elevation, curvature, aspect (Figure 2).

![SOM visualization through U-matrix (top left) and the six component layers for landform classification.](image)

According to Figure 7, the six figures are linked by position: in each figure, the hexagon in a certain position corresponds to the same map unit. The legend for each of the hexagons shows the degree of color compared to each other. In the SOM method, similar colors show the direct relationship between the parameters. It can be seen that the elevation and slope are closely related to each other. In addition, profile and curvature are closely related to each other. The other parameters are not related together.

One the other hand, in the U-matrix, it is easy to see that the bottom rows of the SOM form an almost clear cluster (class A), as shown in the labels in Figure 8(c). The other classes that consist of C, D, E and H form the other clusters (Figure 8c).

As was shown in Figure 8, the numbers written in the hexagons are data that are absorbed by each of the nodes in the neural network [34]. According to Figure 8, the maximum number of hexagons was 5, indicating that the maximum data in these places is 5. In addition, the minimum number of hexagons is 0, indicating that in these places, there is no data. According to Figure 8, the principal component (PC) projection showed that the study data had high density with good distribution. Finally, using the label map (Figure 8c), the study data was classified into five classes for landforms.
Figure 8. Different visualizations of the clusters obtained from the classification of the morphological variation through SOM: (a) Color code. (b) Principal component projection. (c) Label map with the names of the landform classification.

The characteristics of each group determined by the label map are provided in Table 3. It seems that the five clusters correspond to different terrain forms. From the component planes, the features differentiating the clusters can be seen. In this table, the categorized map units and the corresponding morphometric features are summarized. Cluster 5 corresponds to high slope and elevation, but with different of concavity and convexity that consist of ridge landforms. Cluster 3 corresponds to flat areas, possibly plantation areas, with medium elevation and almost flat terrain. Clusters 1, 2 and 4 correspond to channels with different slope conditions.

Table 3. Characteristics of the clusters from the SOM for the landform classification.

| Group   | Parameters | Slope (°) | Profile (1/m) | Plan(1/m) | Elevation (m) | Curvature(1/m) | Aspect (°) |
|---------|------------|-----------|---------------|-----------|---------------|----------------|------------|
| Cluster 1(A) | Min | Low (<5) | - | - | Medium (1500-2500) | - | North |
|          | Max | High (>23) | + | + | Medium (1500-2500) | + | North |
| Cluster 4 (C) | Min | Medium (5-16) | - | - | Medium (1500-2500) | - | North |
|          | Max | Medium to high (16-23) | + | + | High (>2500) | + | North |
| Cluster 2 (D) | Min | Low (<5) | - | - | Medium (1500-2500) | - | East |
|          | Max | Medium (5-16) | + | + | Medium (1500-2500) | + | North |
| Cluster 3 (E) | Min | Low (<5) | - | - | Low (<1500) | - | East |
|          | Max | Low (<5) | + | + | Medium (1500-2500) | + | Northwest |
| Cluster 5 (H) | Min | Low (<5) | - | - | Medium (1500-2500) | - | Northeast |
|          | Max | High (>23) | + | + | High (>2500) | + | Northwest |

4. Conclusion
The purpose of the study was to determine the effectiveness of SOM as a clustering tool for landform classification. In SOM, according to qualitative data, the clustering tendencies of the landforms were investigated using six morphometric parameters (slope, curvature, aspect, elevation, plan and profile). The U- matrix showed that some of the data are closely related to each other, such as elevation and slope. In addition, considering that PC projection represents the amount of data relationship with each other, PC projection was used to determine the study’s data had high density. The results showed that the data had high density and had correlation with each other. Finally, using the labels in the SOM method, five classes for the landforms were detected. Cluster 5 corresponds to high slope and elevation but with different of concavity and convexity that consist of ridge landforms. Cluster 3 corresponds to flat areas, possibly planation areas, with medium elevation and almost flat terrain. Clusters 1, 2 and 4 correspond to channels with different slope conditions. Based on the clusters, we can predict that there are five types of main wildlife in the study area.

References
[1] Crevenna AB, Rodríguez VT, Sorani V, Frame D and Ortiz MA 2005 Geomorphometric analysis for characterizing landforms in Morelos State, Mexico. Geomorphology, 67, 407–422.
[2] Dikau R, Brabb EE, Mark RK and Pike RJ 1995 Morphometric landform analysis of New Mexico. Zeitschrift für Geomorphologie, N.F., Suppl.-Bd., 101, 109-126.
[3] Pike RJ 2002 A Bibliography of Terrain Modeling (Geomorphometry), the Quantitative Representation of Topography-Supplement 4.0. Open-File Report 02-465, U.S. Geological Survey.
[4] Dehn M, Gritter H and Dikau R 2001 Principles of semantic modeling of landform structures. Computers & Geosciences, 27 (8), 1005–1010.
[5] Macmillan R, Pettapiece W, Nolan S and Goddard T 2000. A generic procedure automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. Fuzzy sets and Systems, 113, 81–109.
[6] Burrough PA, van Gaans PFM and MacMillan RA 2000 High resolution landform classification using fuzzy-k means. Fuzzy Sets and Systems, 113, 37-52.
[7] Meybeck M, Green P and Vorosmarty CJ 2001. A New Typology for Mountains and Other Relief Classes: An Application to Global Continental Water Resources and Population Distribution, Mount. Research & Development, 21, 34 - 45.
[8] Schmidt J and Hewitt A 2004 Fuzzy land element classification from DTMs based on geometry and terrain position. Geoderma, 121, 243–256.
[9] Saadat H, Bonnell R, Sharifi F, Mehuys G, Namdar M and Ale-Ebrahim S 2008 Landform classification from a digital elevation model and satellite imagery. Geomorphology, 100, 453–464.
[10] 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.
[11] Dikau R, Brabb EE, Mark RK and Pike RJ 1995 Morphometric landform analysis of New Mexico. Zeitschrift für Geomorphologie, N.F., Suppl.-Bd., 101, 109-126.
[12] Sulebak JR, Etzelmüller B and Sollid JL 1997 Landscape regionalization by automatic classification of landform elements. Norsk Geografisk Tidsskrift, 51 (1), 35–45.
[13] Etzelmüller B and Sulebak JR 2000 Developments in the use of digital elevation models in periglacial geomorphology and glaciology. Physica Geographie, 41.
[14] Adediran AO, Parcharidis I, Poscolieri M and Pavlopoulos K 2004 Computer-assisted discrimination of morphological units on north-central Crete (Greece) by applying multivariate statistics to local relief gradients. Geomorphology, 58, 357–370.
[15] Moore ID and Nieber JL 1989 Landscape assessment of soil erosion and non-point source pollution. Journal Minnesota Academy of Science, 55, 18–24.
[16] Moore ID, Gessler PE, Nielsen GA and Peterson GA 1993a Soil attribute prediction using terrain analysis. Soil Science Society of America, 57, 443–452.
[17] Moore ID, Turner AK, Wilson JP, Jenson SK and Band LE 1993b GIS and Land–Surface–Subsurface Process Modeling. In: Goodchild, M.F., Parks, B.O., Steyaert, L.T. (Eds.), Environmental Modeling with GIS. Oxford University Press, New York, pp. 196–230.
[18] Evans IS 1979 An integrated system of terrain analysis and slope mapping. Final report on grant DA-ERO-591-73-60040, University of Durham, England, 192pp.
[19] Giles PT 1998 Geomorphological signatures: classification of aggregated slope unit objects from digital elevation and remote sensing data. Earth Surface Processes and Landforms, 23, 581-594.
[20] Mokarram M, Roshan G and Negahban S. 2015. Landform classification using topography position index (case study: salt dome of Korsia-Darab plain, Iran). Model. Earth Syst. Environ. (2015) 1:40
[21] Mokarram M and Sathyamoorthy D. 2015. Modeling the relationship between elevation, aspect and spatial distribution of vegetation in the Darab Mountain, Iran using remote sensing data. Model. Earth Syst. Environ. (2015) 1:30
[22] Mokarram M, Seif A and Sathyamoorthy D. 2014. Use of morphometric analysis and self-organizing maps for alluvial fan classification: case study on Ostorankooh altitudes, Iran. IOP Conference Series: Earth and Environmental Science.
[23] Ehsani AH and Quiel F 2008 Application of Self Organizing Map and SRTM data to characterize yardangs in the Lut desert, Iran. Remote Sensing of Environment 112 (2008) 3284–3294.
[24] Kohonen, T 2001 Self-organizing maps, 3rd ed. Springer-Verlag.
[25] Hosokawat M and Hoshit T 2001 Landform Classification Method Using Self-Organizing Map and its application to Earthquake Damage Evaluation. 0-7803-7031-7/01/$10.00 (C) 2001 IEEE
[26] Ferentinou M, Hasiotis T and Sakellariou M 2010 Clustering of geotechnical properties of marine sediments through self – organizing maps: An example from the ZakynthosCanyon–Valley system, Greece. In: Submarine Mass Movements and their consequencesIV , Mosher, D., Shipp, C., Moscardelli, Chaytor J, Baxter C, Lee H and Urgeles R, (Ed.), Advances in Natural and Technological Hazards Research, v. 28, 43-54, Springer, The Netherlands.
[27] Weiss A 2001 Topographic Positions and Landforms Analysis (Conference Poster). ESRI International User Conference. San Diego, CA, pp. 9-13.
[28] Jenness J. 2006 Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x, v. 1.3a. Jenness Enterprises, http://www.jennessent.com/arcview/tpi.htm, (last accessed on February 2010).
[29] Buza A, Mcdonald JA, Michalak J and Stuetzle W 1991 Interactive data visualization using focusing and linking. Proceedings of IEEE conference of visualization, 156-63.
[30] Dhubkarya DC, Nagariya D and Kapoor R 2010 Implementation Of A Radial Basis Function Using VHDL. Global Journal of Computer Science and Technology, 16-19.
[31] Kohonen T (1995). Self - Organizing Maps. Springer, Berlin.Lambeck, K., 1996. Sea-level change and shore-line evolution in Aegean Greece since UpperPalaeolithic Time. Antiquity 70, 588-611.
[32] Hulle MMV 2012 Self-organizing Maps. G. Rozenberg et al. (eds.), Handbook of Natural Computing, DOI 10.1007/978-3-540-92910-9_19, Springer-Verlag Berlin Heidelberg 2012.
[33] Dykes J 2005 Exploring Geovisualization. Elsevier.
[34] Venna J and Kaski S 2001. Neighborhood preservation in nonlinear projection methods: An experimental study. Lecture Notes in Computer Science 2130: 485-491.