Numerical Modeling of soil-landscape relationships using diversity indices and conditional probability: A case study from an Iranian arid region

Mohsen Bagheri-Bodaghabadi (m.baghery@areeo.ac.ir)
Agricultural Research, Education and Extension Organization (AREEO)
Azam Jafari
Shahid Bahonar University of Kerman
Mojtaba Zeraatpisheh
Henan University
Hamidreza Owliaie
Yasouj University
Peter Finke
Ghent University
Ming Xu
Henan University

Research Article

Keywords: Soil-Landscape relationship, Geomorphology, Pedodiversity, Numerical analysis

Posted Date: September 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2053011/v1

Numerical Modeling of soil-landscape relationships using diversity indices and conditional probability: A case study from an Iranian arid region

Mohsen Bagheri-Bodaghabadi[1*1], Azam Jafari[2], Mojtaba Zeraatpisheh[3], Hamidreza Owliaie4[**],
Peter Finke5, Ming Xu6

[1] Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, m.baghery@areeo.ac.ir

[2] Department of Soil Science, College of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran, a.jafari@uk.ac.ir

[3] Henan Key Laboratory of Earth System Observation and Modeling, Henan University, Kaifeng 475004, China, zeraatpishem@yahoo.com

[4] Department of Soil Science, College of Agriculture, Yasouj University, Yasouj, Iran, owliaie@gmail.com

5 Department of Environment, Research group of soilscape genesis, Ghent University, Coupure links 653, B-9000, Ghent, Belgium, peter.finke@ugent.be

[6] College of Geography and Environmental Science, Henan University, Kaifeng 475004, China mingxu@henu.edu.cn

Abstract

We quantified some mental and qualitative concepts about the soil-landscape relationships by numerical analysis of landforms in soil identification using diversity indices and conditional probability with a given sample size in Darab and Khosuyeh plains (a rural district) in the south of Iran in Fars province. The geomorphology map was prepared based on the Zinck method and used as a basic design for soil sampling. Finally, 200 soil profiles (0-150 cm) were excavated and described. Diversity indices and conditional probability were calculated based on soil taxonomic

---

[*] Corresponding author. Azam Jafari and Mohsen Bagheri Bodaghabadi contributed equally to this work as first authors

[**] Corresponding author.

E-mail addresses: m.baghery@areeo.ac.ir (Mohsen Bagheri Bodaghabadi), owliaie@gmail.com (Hamidreza Owliaie).

and geomorphological hierarchies. The results showed that diversity indices increase from landscape to landform level. The lowest and highest diversity indices were obtained at each geomorphic level for the soil order and soil family. The geomorphic diversity based on the soil taxonomy hierarchy showed that soil orders, including Entisols and Inceptisols, are observed in various landscapes and landforms. In contrast, some soil classes, such as Mollisols and its lower levels (suborder, great group, etc.), did not have geomorphic diversity. The conditional probability based on the geomorphological hierarchy indicated that the presence possibility of specific soil at the higher level (landscape) is less than, the lower level (landform), which indicates the more homogeneity of soils at lower geomorphic levels. However, the probability of observing a certain geoform increased according to the soil classification hierarchy, consistent with the results of diversity indices. The efficiency of diversity indices and conditional probability in showing the distribution and possibility of soil separation depends on the alignment of soil and geomorphological processes and the diagnosis of these processes.

Keywords: Soil-Landscape relationship; Geomorphology; Pedodiversity; Numerical analysis.

Introduction

In arid regions, knowledge of the soil spatial distribution is crucial for land management and sustainable crop production. In soil survey plans, the spatial distribution pattern of soils is determined based on field observations and the usage of landscape features [1]. In these projects, possible groupings of soils are tried out in order to achieve the greatest similarity within the group and to maximize differences between groups. One method of drawing boundaries and separating soils is the use of relationships of soils with landforms, especially in arid regions [2]. That is why a thorough understanding of soils is rooted in a deep understanding of landscape evolution [3]. Soil-landscapes are complex and diverse due to pedological, geomorphological, topographical, and hydrological processes acting over hundreds and thousands of years [4]. Soils are an essential part of the landscape and are basically controlled by topographical, hydrological, and geomorphological processes [5]. These processes operate in soils and are responsible for the differentiation of soils in landscapes. Therefore, soil distribution can be most efficiently identified by separating units with similar topographical, hydrological, and geomorphological processes.

Soils and landforms develop together, and this development has a two-way impact. Soils are affected by landforms, and through their developmental accessions and features, they in turn influence geomorphic evolution [2]. The intrinsic and extrinsic processes operate simultaneously in soils, and the resulting profile as a recorder reflects the time-integrated balance of these processes [4]. In this way, the relationships of factors, processes, and soil properties enable us to recognize and delineate unique landforms [6].

The geopedologic approach assumes that the same landforms with the same evolutional history have more or less the same soil patterns and distribution within their boundaries [7]. In fact, the landform is the most uniform level in the hierarchical system of geopedology [8], which presumes that the hydrological, geomorphological, and pedological processes are constant at this level. On a local scale, most of the variation in soil characteristics is a function of the landform's nature because most factors are relatively constant (climate, parent material, biologic processes, and time) [9]; therefore, the occurrence of similar soils at the landform level is expected.

Given the propensity of environmental studies to develop from a qualitative and mental state to a quantitative state, the mathematical sciences and statistics can greatly help to quantify the two-way relationship of soil-landscape [10]. In this regard, it can be noted that diversity indices have been used to quantify soil variability and analyze spatial soil patterns [11-13]. In an early study, Ibáñez et al. [14 , 15] used the ecological diversity indices as criteria to measure variation in soil types. These measurements show quantitatively how diversity changes within a hierarchical system of soil taxonomy or geomorphology [16]. However, the obtained numbers have meaning only when compared to each other, and alone do not express specifically what is the soil-geomorphological relationship (i.e., a diversity index of 0.7 does not evoke a specific meaning). On the other hand, these indices are primarily used to detect the evolution pathway of soils based on the hierarchical structure of geomorphology, which suggests that only one relationship exists, namely soil detection based on landform detection, while there is a two-way relationship. Toomanian et al. [17] have used this method to show how the soil evolution occurs in Zayandeh-rud valley, central Iran. Therefore, the identification of landforms based on recognized soils has not been studied. Additionally, a concept for and quantitative interpretation of the two-way soil-geomorphology relation have not been studied (unless in the context of integrated soil-landscape modelling, e.g., van der Meij et al. [18]). Here, a question raised is how to quantify the soil-landscape relationship so that it can be numerically shown how much each landform can represent a soil, or conversely, to what extent each soil can be a representative or indicator of a landform.

From this perspective, there is a need for an approach that can not only determine the role of soils in the detection of landforms but also to quantitatively measure the soil-geomorphology relationship and its interpretation. In order to quantify and model soil variability, the use of statistical and mathematical tools are growing [19, 21]. In this regard, the conditional probability approach can be helpful, because the conditional probability is calculated for a set of values that can cover a wide range of soil variables (quantitative and qualitative). The main advantage of this approach is its ability to take into account qualitative data, such as soil classes. The research of Bagheri-Bodaghabadi and Toomanian [2] showed that conditional probability could be a good tool to investigate the soil-landscape relationship. Considering that at the same geomorphic surfaces, soil-forming factors, as well as pedogenic processes, are expected to act in the same way, and the most homogeneity is observed in these units, it can be stated using the theory of statistics and probabilities that the conditional probability of soil can represent a certain geomorphic unit.

Based on the abovementioned statements, it can be expected that the diversity indices with the conditional probability approach can be an effective method to a better understanding and numerically assessment of the soil-landscape relationships. Thus, the present study aimed to use the diversity indices, statistics, and conditional probability (i) to quantify the soil-landscape relationship, (ii) to provide numerical analysis of the importance of soil-forming factors on soil formation, and (iii) to numerically show the role of soils in the identification of landforms numerically and vice versa, that is, landforms in soils identification.

Materials and Methods

Study area

The study area is located in Fars province, south Iran, with an area of 26,500 ha in the Darab and Khosouye plains. Darab and Khosouye plains have an area of 15340 and 11160 ha, respectively. These areas are located between 28° 27' 45.3" to 28° 46' 7.2" N latitudes and 54° 30' 8.4" to 54° 31' 40.4" E longitudes (Fig. 1). This area is surrounded by mountains, that mainly consist of limestone, shale and sandstone rock. Erosion and deposition processes have been the main geo-formation processes in the area. The study area has a hot and dry climate, and hot winds occur in the northwest to southeast direction. The rainfall occurs mainly from December to March, and the average monthly rainfall is 250 mm. The average annual temperature in the area is 23 °C. According to Soil Survey Staff [22], the soil moisture and temperature regimes are Ustic and Hyperthermic, respectively. Agriculture is the most important landuse in the study area, and is found mostly in the alluvial plains.

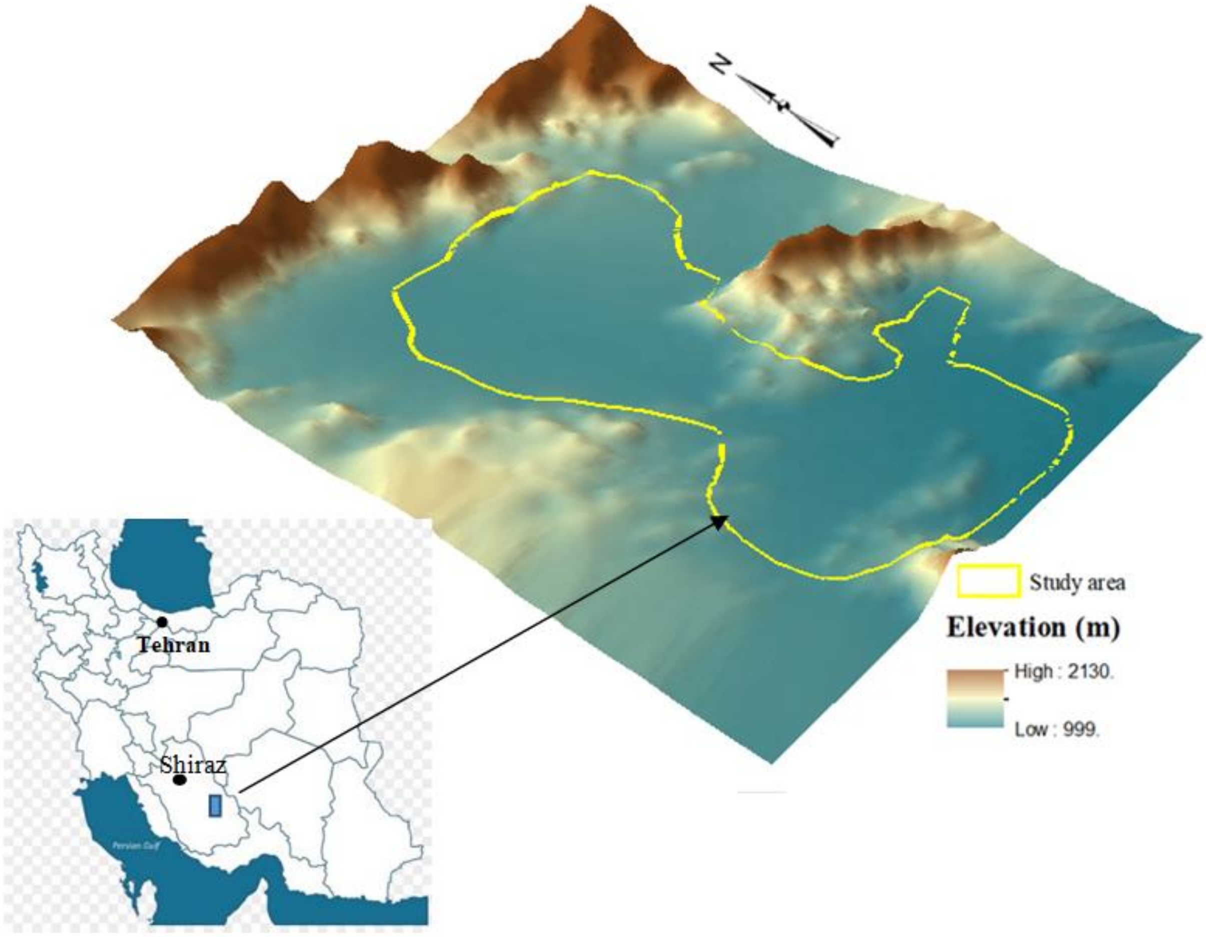

Fig. 1- Study area location in Iran, Fars province

Geomorphology map

The approach of Zinck [8], known as the geopedologic approach, uses geomorphological principles, leading to so-called geoforms, to distinguish landforms having the same history of formation and evolution. In this procedure, six different geoform levels are usually determined, as described in Table 1 [8]. Since applying the sixth (morphogenetic environment) and fifth (geostructure) levels of geopedology structure (Table 1) comprises vast areas, in a standard soil survey by this approach, only the first to fourth levels of this structure are considered.

Table 1- Geomorphology levels based on Zinck [8] method

Level	Category	Generic concept	Short definition
6	Order	Geostructure	Large continental portion characterized by a given type of geologic macro-structure (e.g. cordillera, geosyncline, shield)
5	Suborder	Morphogenic environment	Broad type of biophysical environment originated and controlled by a style of internal and/or external geodynamics (e.g., structural, depositional, erosional, etc.)
4	Group	Geomorphic landscape	Large portion of land/terrain characterized by given physiographic features: it corresponds to a repetition of similar relief/molding types or an association of dissimilar relief/molding types (e.g., valley, plateau, mountain, etc.)
3	Subgroup	Relief/molding	Relief type originated by a given combination of topography and geology structure (e.g., cuesta, horst, etc.)
			Molding type determined by specific morphoclimatic conditions and/or morphogenic processes (e.g., glacis, terrace, delta, etc.)
2	Family	Lithology/facies	Petrographic nature of bedrocks (e.g., gneiss, limestone, etc.) or origin/nature of unconsolidated cover formations (e.g., periglacial, lacustrine, alluvial, etc.)
1	Subfamily	Landform/terrain form	Basic geoform type characterized by a unique combination of geometry, dynamics, and history

The geomorphological hierarchy approach was performed on aerial photographs (1:55,000) [23]. We employed existing knowledge on soil-landscape relationships with geology (associated with variations in parent material), topography (associated with erosional and depositional processes), and geomorphology (indicating effects of erosional and depositional processes) in the delineation of homogeneous units. Stereoscopically interpreted air photos of the study area were imported into a GIS environment and then were geo-referenced using the reference points and google earth image. Finally, boundaries between landforms were checked in the field, and landforms were mapped. A four-level geomorphological hierarchy was prepared, including landscape, relief, lithology, and landform (Table 2). Afterward, a geomorphology map was prepared in the GIS environment (Arcmap 10.2) (Fig. 2).

Table 2- Hierarchic description of geomorphic units in the study area

Landscape	Relief molding	Lithology	Landform	Code	#profiles	Soil classes*
Piedmont (Pi)	Dissected Bajadas	Silty, sandy sediments contain lime, dolomite, volcanic and metamorphic rocks	Lower section of bajada with fin sediments	1111	6	B
	Fan Delta	Silty, sandy, gravelly sediments	Delta with diverging drainage network	1211	12	E, H
	Fans		Lower section of fan	1311	12	H, E
			Cultivated fan	1312	16	D, H, I, M
	Pediment	Dolomite and sandstone remnants	Complex slope	1411	16	C, D, I
Alluvial plain (Ap)	Alluvial plain with constant slope	Silty, sandy alluviums contain lime and dolomite	Complex slope plain	2111	8	F
			cultivated plain	2112	28	D,F,G,H
			Marginal sediments with poor drainage	2113	18	J, K
			Cultivated marginal sediments with poor drainage	2114	18	J, L
		Salty fine alluvium	Complex slope plain	2121	8	H, I
		Silty, clay lime alluviums	Cultivated plain with low slope	2131	32	B,F,G,H,I
	Alluvial plain with complex slope	silty clay loam, high saline alluviums	Low-slope plain with fine sediments	2211	6	N
Flood plain (Fp)	Seasonal flood plain	Silty, sandy, gravelly sediments contain lime, gypsum, and dolomite	River marginal alluviums	3111	20	A, G, I

*: Soil classes were defined in Table 3

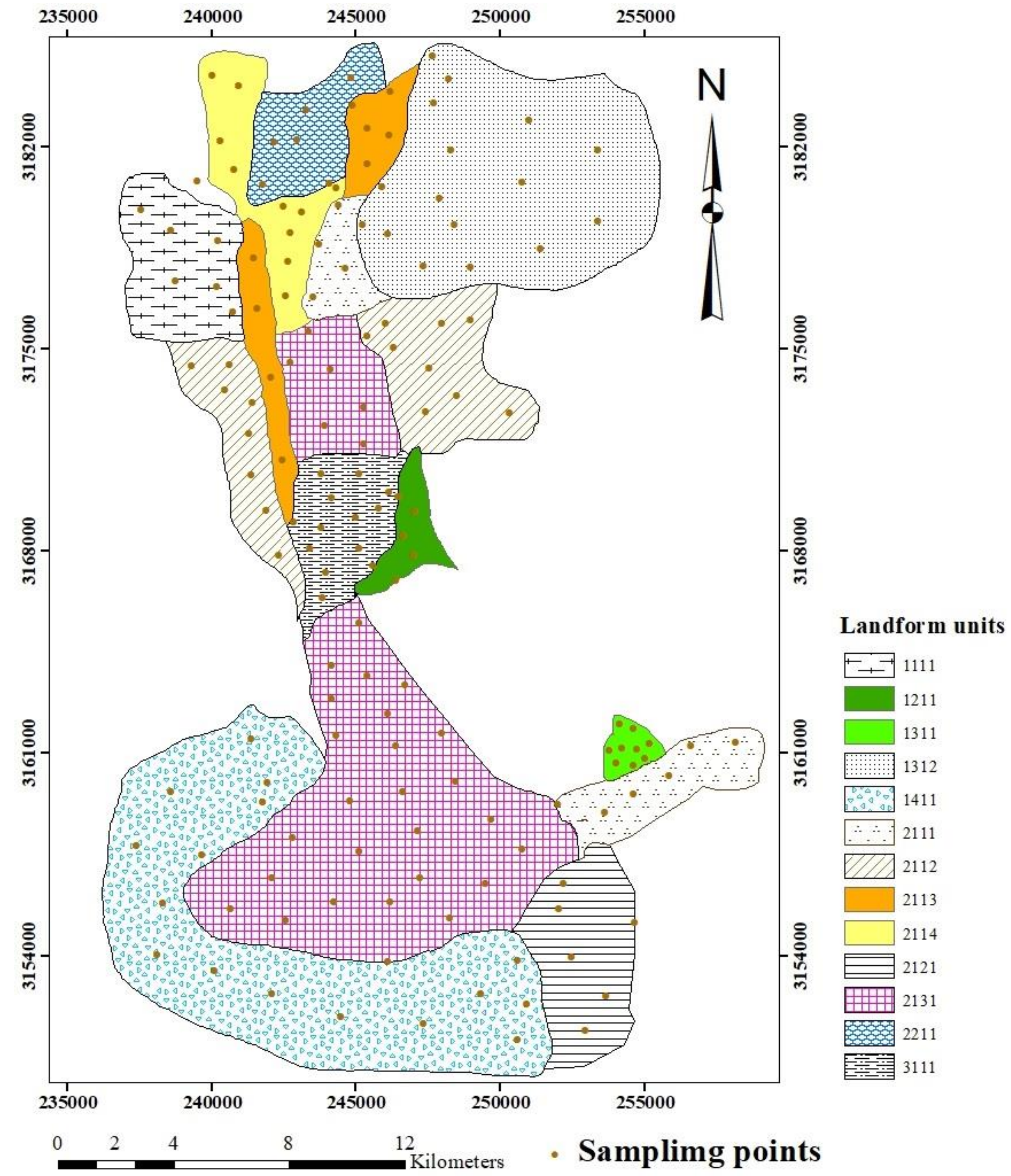

Fig. 2 - Geomorphology map at the landform level (description of units presented in Table 2) and sampling locations

Sampling and laboratory analysis

The geomorphology map was used to select the proper location of sample areas for description and sampling (Fig. 2). The sampling design was linked with landform units and comprised stratified random sampling that the area was divided into sub-areas based on landforms, and within each landform, sampling locations were randomly selected as such that sample size was proportional to the landform area (Fig. 2). This resulted in 200 profiles (0-150 cm), which were then described, sampled, analyzed, and classified up to the soil family level according to the USDA soil classification system [22]. Table 3 presents soil classes and the number of sampling points for each soil. The geographical location of soil profiles was determined using the Etrex Vista Garmin Global Positioning System (GPS) with +/-3 meter accuracy. All excavated soil profiles were described according to the soil description and sampling guidelines in the field [24], and sampling was conducted from all genetic horizons. This resulted in 526 soil samples, which were air-dried, grounded, and passed through a 2-mm sieve. Then, the particle size distribution was determined by the hydrometer method [25] and the coarse fragment (CF) percentage by sieving. Soil pH was determined in saturation paste [26] and electrical conductivity (EC) in saturation paste extract [27]. The content of organic carbon (OC) was determined by means of the Walkley-Black method [28], and calcium carbonate equivalent (CCE) was calculated after treating the samples with HCl and titration by NaOH [29].

Table 3- Soil classes in the study area with the number of observations for each soil.

Code	Family	No.	Subgroup	No.	Great group	No.	Suborder	No.	Order	No.
A	Fine, carbonatic, hyperthermic, Ustic Torriorthents	12	Ustic Torriorthents	22	Torriorthents	22	Orthents	54	Entisols	72
B	Fine loamy, gypsic, hyperthermic, Ustic Torriorthents	10								
C	Coarse loamy, carbonatic, hyperthermic, Typic Ustorthents	8	Typic Ustorthents	32	Ustorthents	32				
D	Fine loamy, carbonatic, hyperthermic, Typic Ustorthents	12								
E	Loamy skeletal, carbonatic, hyperthermic, Typic Ustorthents	12								
F	Fine, carbonatic, hyperthermic, Typic Ustifluvents	18	Typic Ustifluvents	18	Ustifluvents	18	Fluvents	18		
G	Fine, carbonatic, hyperthermic, Typic Calciustepts	16	Typic Calciustepts	56	Calciustepts	56	Ustepts	82	Inceptisols	100
H	Fine loamy, carbonatic, hyperthermic, Typic Calciustepts	40								
I	Fine, carbonatic, hyperthermic, Typic Haplustepts	26	Typic Haplustepts	26	Haplustepts	26				
J	Fine, carbonatic, hyperthermic, Typic Haplaquepts	18	Typic Haplaquepts	18	Haplaquepts	18	Aquepts	18		
K	Fine, carbonatic, hyperthermic, Typic Calciaquolls	6	Typic Calciaquolls	6	Calciaquolls	6	Aquolls	18	Mollisols	22
L	Fine, carbonatic, hyperthermic, Typic Haplaquolls	12	Typic Haplaquolls	12	Haplaquolls	12				
M	Fine loamy, carbonatic, hyperthermic, Typic Haplustolls	4	Typic Haplustolls	4	Haplustolls	4	Ustolls	4		
N	Fine, mixed, hyperthermic, Typic Haplosalids	6	Typic Haplosalids	6	Haplosalids	6	Salids	6	Aridisols	6

Soil diversity indices

In this paper, the taxonomic diversity within the framework of the USDA Soil Taxonomy [22], with five taxonomic categories: order, suborder, great group, subgroup, family are discussed. The soil types identified in each of five soil taxonomic categories are considered as soil individual entities in diversity analysis. Pedodiversity indices including Shannon K-entropy, richness and evenness for each geomorphic category were calculated by summation of indices of all patterns incorporated in each category. To calculate the diversity index, the number of individuals of the objects belonging to i_{th} unit, '*ni*', and the total number of individuals collected, *N*, are considered (for example, the number of profiles observed at a particular geomorphic level is '*ni*' and the total number of soil profiles observed in the region is '*N*'). Diversity indices are calculated based on the relative frequency of soil class towards total observed profiles in the area [30]. The most common relative frequency index is the Shannon k-entropy index that shows soil complexity [30] and is calculated according to the following equation:

$$H' = -\sum_{i=1}^{n} p_i \times ln\, p_i \qquad (1)$$

where 'H'' is the entropy or diversity of society, and 'p_i' is expressed by *ni/N*. The richness index (S), which is the number of different objects or entities such as soil classes in a given ecosystem or predefined territory (e.g., geomorphic category), is used to calculate the evenness index (E) when all the components in the unit have equal probability as if:

$$\mathrm{H}' = \mathrm{Hmax} = \mathrm{lnS} \qquad (2)$$

The relation between entropy (H') and maximum entropy (Hmax) may be used, therefore, as a measure of evenness E and be mathematically expressed by the equation:

$$\mathrm{E} = {}^{\mathrm{H}'}/_{\mathrm{lnS}} \qquad (3)$$

The E index can take any value between 0 and 1, where 1 represents the situation in which all species or objects are equiprobable, and it tends to 0 where there is a highly non-uniform distribution of relative abundance.

The diversity indices were calculated based on soil taxonomic levels (e.g., order, suborder, great group, subgroup, and family) and geomorphic levels (e.g., landscape, relief, lithology, and landform).

Conditional probability

Conditional probability could be applied as an excellent tool to investigate the soil-landscape relationship [2, 19 , 31]. Presume that two events *A* and *B* are given in the similar sample space, and the event probability of *B* is greater than zero, i.e., P(B)>0, then the conditional probability of *A* when *B* has occurred (i.e., P (A|B)) is equal to the ratio of the simultaneous probability distribution of A and B (i.e., P (A∩B)), to the unconditional probability B (i.e., P(B)). In mathematical terms:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{4}$$

Now, if it is supposed that each soil (*Si*) and landform unit (*Gj*) are two events in the same landscape, based on conditional probability, the probability of the presence of soil (*Si*) in the case of landform unit (*Gj*) is observed, can be shown as follows:

$$P(Si|Gj) = \frac{P(Si \cap Gj)}{P(Gj)} \tag{5}$$

It is also possible to show the probability of the presence of landform unit (*Gj*) if soil (*Si*) is observed as follows:

$$P(Gj|Si) = \frac{P(Gj \cap Si)}{P(Si)} \tag{6}$$

When the probability of the existence of soil (*Si*) is 100%, provided that a specific landform unit (*Gj*) is observed, this means that the soil (*Si*) can only be found in that landform (Gj). Therefore, the closer the P(Si|Gj) average of all map units is to 1, the better the soil-landscape model was able to delineate units or soil types. On the other hand, the closer the value of P(Gj|Si) is to 1, means that soil (*Si*) is an indicator soil for landform (*Gj*). Obviously, this does not mean that there is only soil (*S)* in the landform (G), and other soils may also exist in that landform. However, soil (*Si*) can be considered as one of the indicator soils in that landform. As a result, the closer the average value of P(Gj|Si) for all soils in the study area is to 1, means that there is a very strong relationship between soil evolution and related landforms.

Results and Discussion

Soil diversity

In the study area, three landscapes, including piedmonts, alluvial plain, and flood plain, were identified (Table 2). Geomorphic hierarchical diversity indices for the study area are presented in Table 4. For calculation of these indices, '*ni*' is the number of profiles observed at each geomorphic level (for example, at the landscape level, '*ni*' is 62, 118, and 20 for landscapes Pi, Ap, and Fp, respectively) and the total number of soil profiles observed in the region is '*N*'=200. The '*ni*' for each geomorphic level can be found in Table 2.

As the level of classification changes from the landscape to the landform, the diversity indices increase. Saldana and Ibáñez [32] noted that an increase of soil heterogeneity through this hierarchical method confirms the divergent evolution of the studied soils.

Landscape and landform have the lowest and highest diversity indices among the geomorphic levels, respectively (Table 4). The presence of more different soils at the landform level make more diversity while reducing diversity indices at the landscape level are due to more uniformity on a small scale.

Table 4- Diversity indices based on geomorphological hierarchy.

Geomorphology level	N	S	H'	Hmax	E
Landscape	200	3	0.90	1.09	0.82
Relief molding	200	7	1.61	1.95	0.73
Lithology	200	9	1.88	2.20	0.86
Landform	200	13	2.43	2.56	0.95

The diversity indices, including richness (S), evenness (E), and Shannon (H) according to soil taxonomic hierarchy, were calculated for each geomorphic level and are shown in Tables 5 to 8. Table 5 shows soil diversity at the landscape level. To determine the soil order diversity at the landscape level, we should find the different types of soil orders at each landscape. For example, three different soil orders, including Entisols, Inceptisols and Mollisols, are observed at landscape 1. In this way, four and two different soil orders are found at landscapes 2 and 3, respectively. Therefore, the richness index (S) for soil order at the landscapes 1, 2, and 3 are 3, 4, and 2, respectively (Table 5). Therefore, it seems that the formation conditions are provided for more different soils in landscape Ap.

For calculation of other diversity indices (H', Hmax and E), the parameters “*ni*” and “N” should be determined, that at the landscape level, “N” is the total number of soil profiles observed at each landscape (62, 118, and 20 profiles for landscapes 1, 2, and 3, respectively, Table 5). The parameter “*ni*” is different depending on soil category (from soil order to soil family) at each geomorphic level, e.g., for three soil orders at landscape Pi, “*ni*” is 34, 24, and 4 for Entisols, Inceptisols and Mollisols, respectively. The ‘*ni*’ for each soil class in certain geomorphic level can be found from combining Tables 2 and 3.

The indices can be calculated by specifying the parameters and using formulas 1, 2, and 3. The calculation of the indices in Tables 6, 7 and 8 is similar to that described.

Table 5- Soil diversity in landscape level based on soil taxonomic hierarchy.

Landscape*	profiles	Order				Suborder				Great group				Subgroup				Family			
		S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E
1	62	3	0.87	1.09	0.79	3	0.87	1.10	0.79	5	1.37	1.61	0.85	5	1.37	1.61	0.85	7	1.86	1.95	0.96
2	118	4	1.05	1.39	0.76	6	1.56	1.79	0.87	9	1.95	2.20	0.89	9	1.95	2.20	0.89	10	2.15	2.30	0.93
3	20	2	0.67	0.69	0.97	2	0.67	0.69	0.97	3	0.95	1.10	0.86	3	0.95	1.10	0.86	3	0.95	1.10	0.86

*: Lnadscapes 1, 2, and 3 are Pi, Ap, and Fp, respectively, that were defined in Table 2.

Table 6- Soil diversity in relief level based on soil taxonomic hierarchy.

Relief molding*	profiles	Order				Suborder				Great group				Subgroup				Family			
		S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E
11	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
12	12	2	0.69	0.69	1.00	2	0.69	0.69	1.00	2	0.69	0.69	1.00	2	0.69	0.69	1.00	2	0.69	0.69	1.00
13	28	3	0.99	1.10	0.90	3	0.99	1.10	0.90	4	1.29	1.39	0.93	4	1.29	1.39	0.93	5	1.53	1.61	0.95
14	16	2	0.56	0.69	0.81	2	0.56	0.69	0.81	2	0.56	0.69	0.81	2	0.56	0.69	0.81	3	1.04	1.10	0.95
21	112	3	0.94	1.10	0.85	5	1.43	1.61	0.89	8	1.85	1.95	0.95	8	1.85	1.95	0.95	9	2.05	2.20	0.93
22	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
31	20	2	0.67	0.69	0.97	2	0.67	0.69	0.97	3	0.95	1.10	0.86	3	0.95	1.10	0.86	3	0.95	1.10	0.86

*: Relief levels were defined in Table 2.

Table 7- Soil diversity in lithology level based on soil taxonomic hierarchy.

Lithological*	profiles	Order				Suborder				Great group				Subgroup				Family			
		S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E
111	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
121	12	2	0.69	0.69	1.00	2	0.69	0.69	1.00	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1.00
131	28	3	0.99	1.10	0.90	3	0.99	1.10	0.90	4	1.29	1.39	0.93	4	1.29	1.39	0.93	5	1.53	1.61	0.95
141	16	2	0.56	0.69	0.81	2	0.56	0.69	0.56	2	0.41	0.69	0.59	2	0.41	0.69	0.59	3	1.04	1.10	0.95
211	72	3	1.02	1.10	0.93	5	1.51	1.61	0.94	6	1.67	1.79	0.93	6	1.67	1.79	0.93	7	1.85	1.95	0.95
212	8	1	0.00	0.00	0.00	1	0.00	0.00	0.00	2	0.69	0.69	1.00	2	0.69	0.69	1.00	2	0.69	0.69	1.00
213	32	2	0.62	0.69	0.89	3	0.83	1.10	0.76	4	1.31	1.39	0.94	4	1.31	1.39	0.94	5	1.54	1.61	0.96
221	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
311	20	2	0.67	0.69	0.97	2	0.67	0.69	0.97	3	0.95	1.10	0.86	3	0.95	1.10	0.86	3	0.95	1.10	0.86

*: Lithological levels were defined in Table 2.

Table 8- Soil diversity in landform level based on soil taxonomic hierarchy.

Landform*	profiles	Order				Suborder				Great group				Subgroup				Family			
		S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E
1111	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
1211	12	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1.00
1311	12	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1.00
1312	16	3	1.04	1.10	0.95	3	1.04	1.10	0.95	4	1.39	1.39	1	4	1.39	1.39	1	4	1.39	1.39	1
1411	16	2	0.56	0.69	0.81	2	0.56	0.69	0.81	2	0.56	0.69	0.81	2	0.56	0.69	0.81	3	1.04	1.10	0.95
2111	8	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
2112	28	2	0.59	0.69	0.86	3	0.79	1.10	0.72	3	0.79	1.10	0.72	3	0.79	1.10	0.72	4	1.28	1.39	0.92
2113	18	2	0.64	0.69	0.92	2	0.64	0.69	0.92	2	0.64	0.69	0.92	2	0.64	0.69	0.92	2	0.64	0.69	0.92
2114	18	2	0.64	0.69	0.92	2	0.64	0.69	0.92	2	0.64	0.69	0.92	2	0.64	0.69	0.92	2	0.64	0.69	0.92
2121	8	1	0	0	0	1	0	0	0	2	0.69	0.69	1.00	2	0.69	0.69	1.00	2	0.69	0.69	1.00
2131	32	2	0.62	0.69	0.89	3	0.83	1.10	0.76	4	1.31	1.39	0.94	4	1.31	1.39	0.94	5	1.54	1.61	0.96
2211	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
3111	20	2	0.67	0.69	0.97	2	0.67	0.69	0.97	3	0.95	1.10	0.86	3	0.95	1.10	0.86	3	0.95	1.10	0.86

*: Landform levels were defined in Table 2.

In each category of geomorphological hierarchy, the diversity indices from the great group to the subgroup level are constant due to the lack of soils in the new classes and the lack of increase of different soils from the great group to the subgroup level (Table 3).

The increasing trend of diversity indices based on the taxonomic hierarchy is quite clear in landscape 2, relief 21, lithology 211, and landform 2131. These units cover a significant area of the study area, and the number of their soil profiles is more than that of other units (Tables 5 to 8). The relationship of entropy -area or the number of sampling locations is a good indicator of area complexity [33, 34]. A positive linear relationship between entropy and the sampling intensity (R^2=0.60; Fig. 3) obtained in the study area agrees with the results of Toomanian et al. [17] and Gue et al. [33]. Richness- sampling density analysis (at the family level) showed richness increases with increasing sample numbers in the study area (Fig.3). A positive relationship (R^2=0.62) was observed between richness and the number of samples (Fig. 3). This implies that different soil types can appear in a larger area.

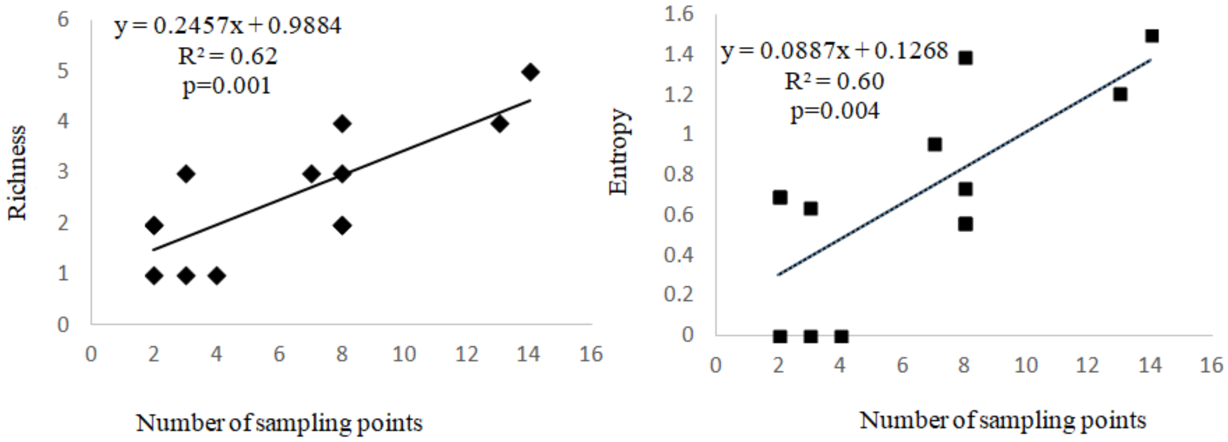

Fig. 3- Relationships between richness (left figure) and entropy (right figure) and number of sampling points at the soil family and landform levels

In landforms 1111, 1211, 1311, 2111, 2113, 2114, 2211, the general trend of increasing the diversity indices from order to the family is not the same as the others (Table 8). The salinization process is the main soil-forming process leading to local convergence soil development in landform 2211, while irregular sedimentation is the dominant factor responsible for diverging soil types on the other landforms. Consequently, the number of different soils and soil diversity in these landforms does not increase from the order to lower categories.

Geomorphic diversity

We also calculated the geomorphic diversity of soils based on the geomorphological hierarchy (Table 9). Geomorphic diversity indicates whether soil is observed on different geomorphic surfaces.

To calculate the diversity indices for each soil category, the number of profiles belonging to a given soil in particular geomorphic level, *ni*, and the total number of profiles belonging to a given soil, N, are considered. For example, the total number of profiles described as Entisols (N) is 72 and or for Inceptisols is 100, and the number of Entisols and Inceptisols observed at landscape 1 is 34 and 24 (*ni*), respectively. The richness index for Entisols at the landscape level is 3 because this soil is observed at three different landscapes (landscape 1 with *ni*=34, landscape 2 with *ni*=26, and landscape 3 with *ni*=12) in the study area. In this way, the richness index for Entisols is 6, 7, and 9 at the relief, lithology and landform levels. According to Table 9, the other soil orders are observed at fewer landscapes and landforms; therefore, Entisols have more geomorphic diversity.

Table 9 – Geomorphic diversity of soils based on the geomorphological hierarchy

		Landscape				**Relief**				**Lithology**				**Landform**			
Soil classes	Pr.	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E	S	H'	Hmax	E
Entisols	72	3	1.02	1.10	0.93	6	1.65	1.79	0.92	7	1.89	1.95	0.97	9	2.14	2.20	0.97
Inceptisols	100	3	0.81	1.10	0.73	5	1.04	1.61	0.64	7	1.67	1.95	0.86	10	2.15	2.30	0.93
Mollisols	22	2	0.47	0.69	0.68	2	0.47	0.69	0.68	2	0.50	0.69	0.68	3	0.99	1.10	0.91
Aridisols	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Orthents	54	3	0.91	1.10	0.83	6	1.75	1.79	0.98	7	1.85	1.95	0.95	8	1.98	2.08	0.95
Fluvents	18	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92	3	1.06	1.10	0.97
Ustepts	82	3	0.89	1.10	0.81	5	1.13	1.61	0.70	7	1.79	1.95	0.92	8	1.91	2.08	0.92
Aquepts	18	1	0	0	0	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92
Ustolls	4	1	0	0	0	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92
Aquolls	18	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Salids	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Torriorthents	22	3	0.99	1.10	0.91	3	0.99	1.10	0.91	3	0.99	1.10	0.91	3	0.99	1.10	0.91
Ustiorthents	32	2	0.38	0.69	0.54	4	1.31	1.39	0.94	4	1.31	1.39	0.94	5	1.52	1.61	0.94
Ustifluvents	18	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92	3	1.06	1.10	0.96
Calciustepts	56	3	0.83	1.10	0.76	4	1.02	1.39	0.74	4	1.62	1.39	1.17	7	1.74	1.95	0.89
Haplustepts	26	3	0.98	1.10	0.89	4	1.20	1.39	0.86	5	1.52	1.61	0.94	5	1.52	1.61	0.94
Haplaquepts	18	1	0	0	0	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92
Calciaquolls	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Haplaquolls	12	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Haplustolls	4	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Haplosalids	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
Us. Torriorthents	22	3	0.99	1.10	0.91	3	0.99	1.10	0.91	3	0.99	1.10	0.91	3	0.99	1.10	0.91
T. Ustiorthents	32	2	0.38	0.69	0.54	4	1.31	1.39	0.94	4	1.31	1.39	0.94	5	1.52	1.61	0.94
T. Ustifluvents	18	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92	3	1.06	1.10	0.96
T. Calciustepts	56	3	0.83	1.10	0.76	4	1.02	1.39	0.74	4	1.62	1.39	1.17	7	1.74	1.95	0.89
T. Haplustepts	26	3	0.98	1.10	0.89	4	1.20	1.39	0.86	5	1.52	1.61	0.94	5	1.52	1.61	0.94
T. Haplaquepts	18	1	0	0	0	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92
T. Calciaquolls	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
T. Haplaquolls	12	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
T. Haplustolls	4	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
T. Haplosalids	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
A[1]	12	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
B	10	2	0.67	0.69	0.97	2	0.67	0.69	0.97	2	0.67	0.69	0.97	2	0.67	0.69	0.97
C	8	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
D	12	2	0.64	0.69	0.92	3	1.10	1.10	1	3	1.10	1.10	1.00	3	1.10	1.10	1
E	12	1	0	0	0	2	0.69	0.69	1	2	0.69	0.69	1	2	0.69	0.69	1
F	18	1	0	0	0	1	0	0	0	2	0.67	0.69	0.99	3	0.96	1.10	0.87
G	16	2	0.56	0.69	0.81	2	0.56	0.69	0.81	3	1.04	1.10	0.95	3	1.04	1.10	0.95
H	40	2	0.67	0.69	0.97	3	0.94	1.10	0.85	5	1.54	1.61	0.96	6	1.72	1.79	0.96
I	26	3	0.98	1.10	0.89	4	1.20	1.39	0.86	5	1.52	1.61	0.94	5	1.52	1.61	0.94
J	18	1	0	0	0	2	0.69	0.69	1	2	0.69	0.69	1.00	2	0.69	0.69	1
K	6	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92	3	1.06	1.10	0.96
L	12	1	0	0	0	1	0	0	0	1	0	0	0	2	0.64	0.69	0.92
M	4	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
N	6	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0

Generally, as the category changes from landscape to landform, the geomorphic diversity increases. Descending the taxonomic system toward lower categories, a landscape is split into different landforms according to the geomorphological hierarchy. The higher categories like landscape are delineated in the small scale associated with high heterogeneity and break down to more homogeneous units at lower categories (landforms) related to local soil formation conditions.

The highest geomorphic diversity in the study area was observed for Entisols and Inceptisols at the soil order level, Orthents and Ustepts at the suborder level, Ustiorthents, Calciustepts, and Haplustepts at the great group level; and Typic Ustiorthents, Typic Calciustepts, and Typic Haplustepts at the subgroup level (Table 9). In other words, these soils are observed in various landscapes and landforms, while some soil classes, including Mollisols and their lower categories, did not show geomorphic diversity. In fact, Mollisols and their lower categories are observed in certain conditions that occur in specific landscapes and landforms and or certain geomorphic units.

It seems that the geomorphic diversity can be an appropriate indicator of the conditions of the formation of different soil types. For example, Entisols can be presented in any conditions in terms of climate, topography, and parent materials, which is shown in Table 9, with high geomorphic diversity for Entisols. This is likely caused by their young age, and is why these soils are often called azonal soils. For Aridisols, the general trend of increase in the diversity indices from the landscape to the landform is not similar to that observed for other soils. This is probably due to a lack of suitable conditions for the formation of Aridisols in different geomorphic units. Entisols may occur at various landscapes because they are too young to be impated by all soil forming factors; Aridisols occur at selected places because they are confined to the driest conditions; Mollisols are at preferent locations because they are confined to the more moist conditions; and finally Inceptisols are characterized by redistribution of matter over the profile and are thus confined to the more leaching environments.

The decreasing trend of geomorphic diversity from order to soil family in the study area indicates that the soil classification criteria for the higher soil categories are easily provided at more geomorphic units than the lower soil categories. It means that soil orders could be found in different landforms and landscapes, while certain processes must have occurred in the landforms and landscapes to contain the lower soil categories, e.g., Inceptisols compared to Calciusteps.

When we compare the soil diversity (Tables 5, 6, 7, 8) and the geomorphic diversity (Table 9), it can be observed that the soil diversity mostly does not change from the relief level to the lithology level. In contrast, the geomorphic diversity increases from the relief level to the lithology level, particularly for higher soil categories (order and suborder levels). It means that a given soil might be observed at different lithologic units (geomorphic diversity) but remains the same soil at all of them; in other words, lithogenic processes were not aligned with pedologic processes.

Increasing or decreasing diversity indices make sense when presented according to a specific classification system; otherwise, each of the calculated indices alone has no specific meaning. For example, the richness and Shannon indices of the soil great group at the 1312 landform are 4 and 1.39, respectively (Table 8). To interpret these numbers, it is necessary to calculate the values of these indices at the levels of lithology, relief, and landscape. In other words, the interpretation of these indices is based on the trend of changes in the specific classification system, and even a number in a specific class cannot indicate high diversity or low diversity in that class. This is also observed in geomorphic diversity based on the geomorphological hierarchy (Table 9). For instance, the richness and Shannon indices of the landform level for Entisols were calculated as 9 and 2.14, respectively. It does not specify whether these numbers represent various geomorphic conditions for Entisols order in the study area, and an explanation can be only provided compared to other levels of classification and determining the trend of changes. In fact, diversity indices are a tool for quantitative expression of soil-landscape relationships, but this quantitative expression is just a number and has no meaning without a specific classification structure.

Conditional probability

The conditional probability was used to clarify the soil-landscape relationship independent of the classification system. Based on the conditional probability, it can be determined to what extent each geoform can represent a soil or, conversely, to what extent each soil can be a representative or indicator for a geoform. Conditional probability calculations were performed for some soil classes from the order to the soil family and are presented in Table 10. Soils with less than eight observations at the soil family level were not included in the calculations (Table 3). For this reason, eleven soil families were used, which accounted for 80% of the observations. To calculate the conditional probabilities of each soil, the relative frequency of the effective area of each soil pedon belonging to a given landform was considered. The effective area of each soil pedon was calculated based on the method employed by Bagheri Bodaghabadi et al. [35]. Following this method, a primary map was obtained using the Thiessen method based on the information of soil types and their locations in the study area. It is almost equal to the relative frequency of the soil profiles, i.e., the number of observations of each soil divided by the total number of soil profiles.

Table 10 - Conditional probability values P (x) for some soil classes based on geomorphological hierarchy (P(Si | Gj)).

Soil Class	Landform	P(x)	Lithology	P(x)	Relief molding	P(x)	Landscape	P(x)
Fine loamy, carbonatic, hyperthermic, Typic Calciustepts (Family H)	1211	0.50	121	0.50	12	0.50	1	0.26
	1311	0.50	131	0.35	13	0.35		
	1312	0.25						
	2112	**0.43**	211	0.17	21	0.23	2	0.19
	2121	0.50	212	0.50				
	2131	0.25	213	0.25				
Fine, carbonatic, hyperthermic, Typic Haplustepts (Family I)	1312	0.25	131	0.14	13	0.14	1	0.12
	1411	0.25	141	0.25	14	0.25		
	2121	0.5	212	0.50	21	0.13	2	0.11
	2131	0.31	213	0.31				
	3111	**0.20**	311	0.20	31	0.20	3	0.20
Typic Calciustepts	1211	0.50	121	0.50	12	0.50	1	0.26
	1311	0.50	131	0.35	13	0.35		
	1312	0.25						
	2112	**0.71**	211	0.28	21	0.30	2	0.26
	2121	0.50	212	0.50				
	2131	0.25	213	0.13				
Typic Haplustepts	1312	0.25	131	0.14	13	0.14	1	0.12
	1411	0.25	141	0.25	14	0.25		
	2121	0.50	212	0.50	21	0.13	2	0.11
	2131	0.31	213	0.31				
	3111	**0.20**	311	0.20	31	0.20	3	0.20
Calciustepts	1211	0.50	121	0.50	12	0.50	1	0.26
	1311	0.50	131	0.35	13	0.35		
	1312	0.25						
	2112	**0.71**	211	0.28	21	0.30	2	0.26
	2121	0.50	212	0.50				
	2131	0.25	213	0.25				
Haplustepts	1312	0.25	131	0.14	13	0.14	1	0.12
	1411	0.25	141	0.25	14	0.25		
	2121	0.50	212	0.50	21	0.13	2	0.11
	2131	0.31	213	0.31				
	3111	**0.20**	311	0.20	31	0.20	3	0.20
Ustepts (Inceptisols)	1211	0.50	121	0.50	12	0.50	1	0.39
	1311	0.50	131	0.50	13	0.50		
	1312	0.50						
	1411	0.25	141	0.25	14	0.25		
	2112	**0.71**	211	0.28	21	0.45	2	0.43
	2121	1.00	212	1.00				
	2131	0.69	213	0.69				
	3111	**0.40**	311	0.40	31	0.40	3	0.40

Table 10 shows the probability of soil presence (Si) provided the probability of the existence of geoform (Gj) ($\mathrm{P}(\mathrm{Si} \mid \mathrm{Gj})$) for the suborder to soil family categories based on geomorphological hierarchy. In general, at each geomorphic level, the conditional probability of the existence of soils increases at higher categories of soil taxonomy. For example, the probability of the existence of the soil family of *H* in the landform of 2112 is 0.43, or as follow:

$$\mathrm{P}(\text{Family-H} \mid \text{Geo2112}) = 0.43$$

This number is calculated as follows (similar calculations are applicable for other levels):

The probability of the existence of the landform 2112 (Geo2112) in the whole region is equal to the number of observations of this landform (28) compared to the total observations (200), so we have:

$$\mathrm{P}(\text{Geo2112}) = 28 \div 200 = 0.14$$

And the simultaneous probability of the soil family *H* (*Family-H*) in landform 2112 is equal to the number of observations of this soil in landform 2112 (12 profiles) compared to the total observations (200), so we have:

$$\mathrm{P}(\text{Family-H}) = 12 \div 200 = 0.06$$

Thus, the probability of the presence of soil family *H* in landform 2112 is equal to the quotient of the unconditional probability division:

$$\mathrm{P}(\text{Family-H} \mid \text{Geo2112}) = 0.06 \div 0.14 = 0.43$$

For the subgroup level (SubG) related to soil *H* or *SubG-H* (i.e., Typic Calciustepts subgroup), the probability of the existence of soil *SubG-H* in Geo2112 is 0.71:

$$\mathrm{P}(\text{SubG-H} \mid \text{Geo2112}) = 0.71$$

For the great group category (*GG*) related to soil *H* or *GG-H* (i.e., Calciustepts great group), *P*(GG-H | Geo2112) is expected to increase, but its value remains constant at 0.71 because, for landform 2112, there is no distinction between subgroup, great group, suborder (SubO) and order (Ord) categories in the study area. As another example, we can express *P* (Family-I | Geo3111), which for landform 3111 is not a distinction between soil family, subgroup, and great group categories and value is equal, i.e.:

P (Family-I | Geo3111) = P (SubG-I | Geo3111) = P (GG-I | Geo3111) = 0.20

However, for the suborder and order, this value has increased to 0.40, which means:

P (SubO-I | Geo3111) = P (Ord-I | Geo3111) = 0.40

Such results, which are observed for almost all soil classes and geomorphic levels in the study area, implicitly show that different soils have different evolutionary pathways in different conditions and, it should not necessarily be expected to occur soil diversity according to the soil classification hierarchy. For example, a soil great group is divided into several subgroups, and each subgroup is divided into several soil families. As the results showed, it is quite possible that in a region or even a geoform unit, a suborder has a great group, a subgroup, and a family, while at the same time through the classification hierarchy, no distinction is made at different levels of the taxonomy. However, some soil diversity indices did not change based on the geomorphological hierarchy and even remained constant up to the landform level. For example, in landform 2111, there is no distinction or change for soil *F* from the soil order to lower categories. In other words, the probability of the presence of soil *F* from the soil order to soil family remained constant for landform 2111:

P (Family-F | Geo2111) = P (SubG-F | Geo2111) = P (GG-F | Geo2111) = P (SubO-F | Geo2111) = P (Ord-F | Geo2111) = 1.00

Based on the probability value equal to one, it can be shown that the landform 2111 is an indicator geomorphic unit for soil *F*, and this geomorphic unit has 100% purity for soil *F*. Therefore, if we are looking for soil *F*, this soil can be observed in landform 2111 with 100% probability. Therefore, it can be reported that soil *F* is typical soil in the geomorphic unit 2111. Of course, this does not mean that soil *F* does not exist in other places or geomorphic units because the geomorphic diversity of soil *F* based on geomorphological hierarchy (Table 9) indicates that soil F is also observed in other landforms (2112 and 2131). On the other hand, it can be stated that the geomorphological processes in landform 2111 were exactly in line with the soil-forming processes and acted the same in the whole landform 2111; this is leading to the formation and evolution of specific soil (soil *F*). Bagheri Bodaghabadi and Toomanian (2019) achieved similar results in studying soil diversity in Isfahan, Iran.

The conditional probability for some soils does not change between two or more geomorphic levels, for example, between the lithology level and the relief level. It means that the applied factors and processes for breaking the geomorphic categories are not in line with the occurred processes for soil evolution. In other words, the increased subdivision of the geomorphic levels and the increase of the study scale do not affect soil distinction. For example, for soil *I* and soil *H* at the relief level (13) and the lithology level (131), the probability values remain equal (0.14 and 0.35, respectively, Table 10). Therefore, it can be concluded that soils *I* and *H* are less affected by the soil-forming factor of "parent materials" or maybe that other acting factors were more effective. Esfandiarpoor Borujeni et al. [36] reported the same conclusions for the Borujen region, in the southwest of Iran.

The conditional probability of the presence of some soils remains constant through the geomorphological hierarchy (Table 10). For example, soil family *I* have a constant probability value at all geomorphic levels:

P (Family-I | Geo3111) = P (Family-I | Geo311) = P (Family-I | Geo31) = P (Family-I | Geo3) = 0.20

The dominant environmental conditions defining the soil-forming factors did not lead to more diversity for soil *I* in landform 3111; however, the soil-forming processes in this unit were subject to other features that have not been considered during the geomorphological hierarchy. Therefore soil diversity has developed in the landform 3111, so that three soils, including *A*, *G,* and *I*, were formed in this geomorphic unit (Table 2).

P (Family-G | Geo3111) = P (Family-I | Geo3111) = 0.2

P (Family-A | Geo3111) = 0.6

These results indicate that other features should be included in the hierarchy of this approach to achieve better soil delineation by using the geopedology method. For example, dividing the landform into different phases makes it possible to separate the soils better. Several studies have suggested adding more levels or using the phase level to separate more soils [36, 37].

The average conditional probability values for each geomorphic level based on the soil taxonomic hierarchy was obtained (Table 11). As expected, the probability values increase from the soil family to the order levels at each geomorphic category. For example, the average conditional probability value increases from 0.448 to 0.706 from the soil family to order categories at the landform level. This confirms that heterogeneity increases at the more detailed taxonomic levels (soil family) in the study area. Therefore, what was suggested qualitatively by soil diversity (or derived soil diversity indices, [12, 17, 37]), can also be shown quantitatively by using the conditional probability.

Table 11 - Mean conditional probability values for different geomorphic units based on soil taxonomic hierarchy .

Geomorphic units	Order	SubOrder	Great Group	SubGroup	Family
Landscape	0.154	0.150	0.124	0.124	0.111
Relief molding	0.291	0.287	0.257	0.257	0.228
Lithology	0.415	0.405	0.319	0.319	0.286
Landform	0.706	0.653	0.592	0.592	0.448

It can also be observed (Table 11) that for each soil taxonomic category, the mean conditional probability decreases from the landform to the higher levels. For example, at the family category, the conditional probability value decreases from 0.448 to 0.111 from the landform to the landscape levels. This quantitatively proves it is more likely to find a particular soil at lower geomorphic levels (landform level). Several studies have obtained the same results (e.g., Keshtkar et al. [38]).

The mean conditional probability values (Table 11) can be interpreted as the percentage of the soil taxonomic classes separated by each geomorphic unit or, the percentage of soil classes differentiated and delineated by certain geomorphic unit. For example, at the landform level, only 44.8% of the soil family is separated. Since none of the soil classes reach 100% homogeneity at any geomorphic level, the units or boundaries separated by the geopedological method do not involve the complete spatial distribution of the soils in the study area. The landform level is conceptually and theoretically equal to the subfamily level, or the lithology level is equal to the family level (Table 1). However, in the most homogeneous separated units in this system, the landform level, there is not enough accuracy in separating the soil family, and about 55% of the existing soil families are being neglected. In the best scenario, it can be mentioned that the landform level is suitable for delineating soils at the subgroup category, which has a 60% separation probability. These results are consistent with the findings of Esfandiarpour Borujeni et al. (2010). According to the studies by Nazari et al. [39] and Toomanian and Esfandiarpoor Borujeni [7], none of the geomorphic levels defined by Zinck [8] are compatible with soil classes in the Soil Taxonomy (Table 1), and these geomorphic units are not capable of homogeneous separation of soil classes attributed to those geomorphic units.

The probability of observing geomorphic units (*Gj*) provided the presence of soil (*Si*), i.e., P (Gj | Si) for some geomorphic units was calculated (Table 12). The other geomorphic units, due to the limited size of the article, were not provided. For each geomorphic level, the conditional probability increases as taxonomic hierarchy levels decrease. For example, the probability of observing landform 2112 is 0.50 if the soil *G* (FamilyG) is present, or as follow:

P (2112| FamilyG) = 0.50

For the subgroup category related to soil *G* or *SubG-G* (i.e., Typic Calciustepts subgroup) and the great group category *GG-G* (Calciustepts great group),

P (2112| GG-G) =P (2112| SubG-G) = 0.36, while:

P (2112| SudO-G) = 0.24 and

P (2112| Ord-G) = 0.20

Table 12 Conditional probability values P (x) for some geomorphic units (*Gj*) based on soil taxonomic hierarchy (P (Gj | Si)).

Geomorphic levels	Soil Family	P(family)	P(SubG)	P(GG)	P(SO)	P(O)
2112	D	0.33	0.13	0.13	0.07	0.11
	F	0.22	0.22	0.22	0.22	
	H	0.30	0.36	0.36	0.24	0.20
	G	0.50				
2121	I	0.15	0.15	0.15	0.10	0.08
	H	0.10	0.07	0.07		
2211,221,22	N	1	1	1	1	1
211	D	0.33	0.13	0.13	0.07	0.06
	F	0.67	0.67	0.67	0.67	0.17
	G	0.5	0.36	0.36	0.24	0.38
	H	0.3				
	J	1	1	1	1	
	K	1	1	1	1	0.82
	L	1	1	1	1	
212	I	0.15	0.15	0.15	0.10	0.08
	H	0.10	0.07	0.07		
21	B	0.4	0.18	0.18	0.15	0.36
	D	0.33	0.13	0.13		
	F	1	1	1	1	
	G	0.75	0.64	0.64	0.44	0.68
	H	0.6				
	I	0.54	0.54	0.54	0.17	
	J	1	1	1	1	
	K	1	1	1	1	0.82
	L	1	1	1		
2	All 21+N					

The existence of such a trend for soil taxonomic levels and geomorphic levels well indicates the compatibility of the hierarchical structure of the geopedological approach with the soil classification structure. On the other hand, as expected, these findings show the diversity of the soils over the levels in the geomorphological hierarchy. The findings obtained by Toomanian et al. [17] using the soil diversity indices for the eastern region of Isfahan are in line with our findings. Several studies have been presented similar findings in investigating the diversity of soil map units in other regions [13, 36].

For some soils such as soil *N* (Haplosalids at landform level 2211) or soils *J*, *K,* and *L* (at the lithology level 211), due to high values of conditional probability and a large number of relevant observations, it is possible to introduce these soils as indicator soils for these geomorphic levels.

In other words, the soil-forming processes that led to the formation of these soil families were, to a large extent, closely related to the geomorphological processes of the geomorphic unit. In some cases, this correlation has been so strong that even at the suborder level, it is possible to detect the occurrence of the landform 2211 with a 100% probability. Soils on the landform 2211 in the alluvial plain landscape are classified as Aridisols order (Typic Haplosalids) due to high salinity, sedimentation of fine materials, and flooding. For this reason, a high correlation is established between the soil and the relevant geomorphic unit. Therefore, it can be mentioned that the low geomorphic diversity of some soils (i.e., soils *K*, *L, M,* and *N*) can be related to the specific formation conditions or the specific geomorphic level of these soils (Table 9). This confirms that the lower the geomorphic diversity, the higher the probability of observing geomorphic units (Gj) provided the presence of soil (Si), (P (Gj | Si)). This is based on our results using a given sample size in the study area. The mean conditional probability values based on all observations for soil taxonomic categories based on geomorphological hierarchy were calculated (Table 13).

Table 13. Mean conditional probability values for soil taxonomic categories based on geomorphological hierarchy.

soil taxonomic category	Landscape	Relief Molding	Lithology	Landform
Order	0.508	0.438	0.122	0.106
SubOrder	0.649	0.568	0.393	0.249
Great Group	0.683	0.587	0.372	0.337
SubGroup	0.683	0.587	0.372	0.337
Family	0.746	0.682	0.540	0.495

These results show that what percentage of the geomorphological hierarchy levels assigned to a specific soil is separated at each level of soil taxonomy. For instance, by observing a soil family, it is possible to differentiate the landform assigned to that soil family with an average of 49.5%

probability. This value increases to 74.6% for the landscape of that soil family. These outcomes are consistent with the results of other studies that have examined the relationship between soil and landscape qualitatively or by soil diversity indices [12, 17, 36].

Note that the probability values for the great group and subgroup levels are exactly the same (Table 13). In the study area, there was no further soil delineation from the great group level to the subgroup level and no new soil discriminated, while from the subgroup level to the soil family level with the presence of different soils, separation and differentiation occurred between soils; thus the dispersion and diversity have increased. In order to differentiate the soils at the great groups into different subgroups, they must follow certain rules and conventions in the soil classification system. This means that they have to comply with specific conditions; thus, it is possible that none of these exist in the study area, while to be in multiple families, it is enough to make changes in one of the nine characteristics defined for the soil family. As shown in Table 3, the soil families of a subgroup are defined and separated only by differences in soil texture. Therefore, if there is more flexibility in separating the great groups into subgroups in the soil taxonomic system, both more diverse soils will be identified and help separate more uniform units.

The probability values are sometimes equal for different taxonomic levels (i.e., subgroup and great group levels in the previous example). Such cases indicate that the soil-forming processes that lead to the differentiation of classification levels, for example, subgroup level from great group level, were not diverse in the study area, and then, the distinction between classification levels was not obtained. Therefore, the information value of those classification levels in the soil-landscape relationship is the same, and the lower levels do not provide more information about the impact of soil-forming processes. Descending the taxonomic system toward lower category levels does not introduce more properties that may be related to local conditions and natural selection for the recognition of geomorphic units.

No change from the great group to subgroup level was also observed in the values of diversity indices (Table 14), which is consistent with the results of conditional probability. These results indicated a close relationship between conditional probability values and diversity indices.

Table 14. Mean diversity indices values for different geomorphic units based on soil taxonomic hierarchy

Geomorphic units	Diversity indices	Order	SubOrder	Great Group	SubGroup	Family
Landform	S	2.64	3.76	5.68	5.68	6.46
	H’	0.82	1.09	1.40	1.40	1.58
	Hmax	0.94	1.22	1.49	1.49	1.70
	E	0.82	0.85	0.87	0.87	0.88
Lithology	S	2.40	3.28	4.08	4.08	4.82
	H’	0.76	0.97	1.19	1.19	1.37
	Hmax	0.83	1.08	1.30	1.30	1.45
	E	0.83	0.79	0.85	0.85	0.89
Relief molding	S	1.94	2.24	2.62	2.62	3.00
	H’	0.57	0.64	0.80	0.80	0.94
	Hmax	0.63	0.75	0.89	0.89	1.00
	E	0.79	0.75	0.81	0.81	0.85
Landscape	S	3.49	4.67	7.16	7.16	8.37
	H’	0.96	1.26	1.67	1.67	1.94
	Hmax	1.23	1.47	1.91	1.91	2.07
	E	0.79	0.86	0.87	0.87	0.93

For this purpose, the correlation of diversity indices with conditional probability for geomorphic units was obtained (Table 15).

Table 15. Correlation of diversity indices and conditional probability for different geomorphic units

Geomorphic units	Diversity indices	Conditional probability

		Landscape	Relief molding	Lithology	Landform
Landform	S	-.981**	-.946*	-.981**	-.899*
	H	-.970**	-.942*	-.966**	-.915*
	Hmax	-.964**	-.946*	-.958*	-.932*
	E	-.923*	-.882*	-.921*	-0.867
Lithology	S	-.962**	-.951*	-.954*	-.944*
	H	-.969**	-.955*	-.962**	-.941*
	Hmax	-.956*	-.930*	-.951*	-.914*
	E	-0.867	-.898*	-0.861	-0.816
Relief molding	S	-.976**	-.976**	-.965**	-.965**
	H	-.989**	-.991**	-.980**	-.967**
	Hmax	-.974**	-.960**	-.967**	-.942*
	E	-0.867	-.898*	-0.861	-0.816
Landscape	S	-.991**	-.964**	-.989**	-.918*
	H	-.980**	-.960**	-.975**	-.933*
	Hmax	-.981**	-.943*	-.983**	-.892*
	E	-0.862	-.887*	-0.839	-.946*

In many cases, the relationship between diversity indices and conditional probability values at the 99% level has a very significant correlation. Based on these results, it can be expected that the regression relationship between the mean conditional probability values with the diversity indices is significant for each geomorphic unit (7 to 10 equations). Based on the analysis of variance, the regression relationships obtained for each geomorphic unit at the level of 99% are significant and concerning the value of R^2, are very accurate.

$$P(x,\text{landscape}) = 0.070 - 0.027*S(L)+0.092*Hmax(L)+0.083*E(L) \quad (R^2 = 1) \tag{7}$$

$$P(x,\text{Relief}) = 0.557 - 0.085*S(R) + 0.109*Hmax(R) -0.213\ E(R) \quad (R^2=1) \tag{8}$$

$$P(x,\text{Litology}) = 1.329 + 0.128*S(Li) - 0.625\ *Hmax(Li) -0.845*E(Li) \quad (R^2=1) \tag{9}$$

$$P(x,\text{Landform}) = -4.780 + 0.073*S(La) - 1.338\ *Hmax(La) +7.989*E(La) \quad (R^2=1) \tag{10}$$

where L, R, Li, and La are Landscape, Relief, Lithology and Landform, respectively.

Based on these results, some mental and qualitative concepts about soil-landscape relationships with mathematical models became quantitative and numerical, which can be very attractive and practical for researchers who are investigating in this field.

Conclusion

In the present study, we tried to show how diversity indices and conditional probability can be used to express soil diversity and the probability of soil observation and delineation based on a given sample size in an arid region. The results of the study are summarized:

1- If the diversity indices and conditional probability for a certain soil are constant over the geomorphological hierarchy and are obtained as zero and one, respectively, it means that that specific soil is typical of a particular region. In other words, the soil-forming processes that led to soil formation, to a large extent, were closely related to the geomorphology processes.

2- The probability of observing a certain soil (*Si*) in the presence of a certain geomorphic level (*Gj*) (P (Si | Gj)) decreased from soil order to soil family, while the diversity indices increased based on the soil taxonomic hierarchy because the convergence increases from the soil family to the soil order. For each soil taxonomic level, the conditional probability and diversity indices decreased from the landform level to the higher levels (landscape). This proves that there are more homogeneous soils at the lower geomorphic levels.

3- The probability of observing a certain geomorphic surface (*Gj*) in the presence of a specific soil (*Si*) (P (Gj | Si)) and also diversity indices increased based on the soil taxonomic hierarchy (from soil order to soil family), while the conditional probability (P

(Gj | Si)) decreased based on the geomorphological hierarchy (from the landscape to the landform) for each soil taxonomic level.

4- It is possible that the probability value and the values of diversity for different levels of classification (such as subgroup and great group levels) are equal. This indicates a lack of diversity in soil-forming processes at different levels of classification and, therefore, between these levels, soil diversity has not developed. The information value of these classification levels in the relationship between soil-landscapes is constant, and the lower levels will not present more information about soil-forming processes.

5- Based on the high values of the conditional probability for some soils at a certain geomorphic level and their very low geomorphic diversity (such as some Mollisols under poor water drainage conditions, Aquolls), these soils can be introduced as indicator soils for that geomorphic level and occur in specific landscapes, landforms, and geomorphic levels. It seems the geomorphic diversity shows the conditions for the formation of different soil types.

Declarations:

All authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors.

Conflict of Interest:

The authors have no conflicts of interest to declare. All authors have seen and agree with the contents of the manuscript. We certify that the submission is original work and is not under review at any other publication.

Data Availability:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding:

None.

Author Contributions:

Bagheri-Bodaghabadi, Jafari and Zeraatpisheh wrote the main manuscript text, Owliaie prepared the data and Xu and Fikne edited the English language. All authors reviewed the manuscript.

Ethics declarations

Ethics Approval

Not applicable in this MS.

Consent to Participate

Not applicable in this MS.

Consent for Publication

Not applicable in this MS.

Conflict of Interest/Competing Interests

The authors declare no competing interests.

References

1. Grinand, C., Arrouays, D., Laroche, B. & Martin, M.P. (2008). Extrapolating regional soil landscapes from an existing soil map: sampling intensity, validation procedures, and integration of spatial context. *Geoderma*, 143, 180–190.
2. Schaetzl, R. and S. Anderson. (2005). *Soils: genesis and geomorphology*. Cambridge University Press, New York. 817 p.
3. Ma, Y., Minasny, B., Malone, B.P., Mcbratney, A.B. (2019). Pedology and digital soil mapping (DSM*). European Journal of Soil Science*, 70(2): 216-235.
4. Grunwald, S. (2006). What Do We Really Know about the Space–Time Continuum of Soil-Landscapes?. in: Grunwald, S. (Ed.), *Soil-Landscape Modeling, Geographic Information Technologies and Pedometrics*. CRC Press, Boca Ratón, Taylor and Francis Group, pp. 3–37.
5. Schoorl, J.M., Veldkamp, A. (2001). Linking land use and landscape process modelling: a case study for the Alora region (south Spain). *Agriculture Ecosystems and Environment* 85(1–3), 281–292.
6. Arnold, R.W., 2006. Soil Survey and Soil Classification. in: Grunwald, S. (Ed.), Soil-Landscape Modeling, Geographic Information Technologies and Pedometrics. CRC Press, Boca Ratón, Taylor and Francis Group, pp. 38–60.

7. Toomanian, N., Esfandiarpour Boroujeni, I. (2017). Outcomes of applying a geopedologic approach to soil survey in Iran. *Desert*, 22(2), 239-247.

8. Zinck, J.A. (1989). *Physiography and soils*. Lecture notes. Enschede, The Netherlands: International Institute for Aerospace Survey and Earth Sciences (ITC)

9. Mohamed, M.A., 2020. Classification of Landforms for Digital Soil Mapping in Urban Areas Using LiDAR Data Derived Terrain Attributes: A Case Study from Berlin, Germany. *Land*, 9: 319.

10. Hartemink, A.E., Hempel, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., MacMillan, R.A. (2010). GlobalSoilMap.net–A new digital soil map of the world. In: Digital Soil Mapping: Bridging Research, *Environmental Application and Operation* (eds. J. L. Boettinger, D.W. Howell, A.C. Moore, A.E. Hartemink & S. Kienast-Brown), pp. 423–428. Springer Science, Dordrecht, The Netherlands.

11. Bagheri Bodaghabadi, M., Toomanian, N. (2019). Investigation of the Relationship between Geopedology and Soil Taxonomy in the Soil-Landscape Models using Conditional Probabilities. *Water Soil Research Iran*, 50(5), 1152-1168 (Abstract in English).

12. Jafari, A., Ayoubi, S., Khademi, H., Finke, P.A., Toomanian, N. (2013). Selection of a taxonomic level for soil mapping using diversity and map purity indices: A case study from an Iranian arid region. *Geomorphology*, 201, 86-97.

13. McBratney, A.B., Minasny, B., (2007). On measuring pedodiversity. *Geoderma*, 141, 149−154.

14. Ibáñez, J.J., De Alba, S., Lobo, A., Zucarello, V. (1998). Pedodiversity and global soil patterns at coarse scales (with discussion). *Geoderma*, 83, 171–214.

15. Ibáñez, J.J., De-Alba, S., Bermúdez, F.F., García-Alvarez, A. (1995). Pedodiversity: concepts and tools. *Catena*, 24, 215–232.

16. Toomanian, N. (2013). Pedodiversity and landforms. in: Ibáñez, J.J., Bockheim, J. (Eds.), *Pedodiversity*. CRC Press, Boca Ratón, CA, pp. 133–152.

17. Toomanian, N., Jalalian, A., Khademi, H., KarimianEghbal, M., Papritz, A. (2006). Pedodiversity and pedogenesis in Zayandeh-rud Valley, Central Iran. *Geomorphology*, 81, 376–393.

18. van der Meij, W.M., Temme, A.J.A.M., Wallinga, J., Sommer, M. (2020). Modelling soil and landscape evolution– the effect of rainfall and land use change on soil and landscape patterns. *SOIL* 6: 337–358. https://doi.org/10.5194/soil-6-337-2020

19. Hawinkel, P., De Pauw, E., Deckers, J. (2016). Probabilistic soil mapping by Bayesian inference to assess suitability for derocking in northwest Syria. *Soil Use Management*, 32, 137–149.

20. Mikhailova, E.A.; Zurqani, H.A.; Post, C.J.; Schlautman, M.A.; Post, G.C. (2021) Soil Diversity (Pedodiversity) and Ecosystem Services. *Land* , *10*, 288.

21. Fu T., Liu J., Jiang G., Gao H., Qi F. and Wang F. (2022) Influences of pedodiversity on ecosystem services in a mountainous are. Catena 217(11):106505

22. Soil Survey Staff.(2014). *Keys to soil taxonomy* (12th ed.). NRCS, USDA, USA

23. National Cartographic Center of Iran (NCC), 2007. Aerial photo, 1:55000 scale, https://www.ncc.gov.ir/en/.

24. Schoeneberger, P.J., Wysocki, D.A., Benham, E.C., Broderson, W.D., (Eds.). (2002*). Field book for describing and sampling soils*, Version 2/0. Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE.

25. Gee, G.W., Bauder, J.W., (1986). Particle-size analysis in: Klute, A. (Ed.), *Methods of soil analysis,* Part 1. 2nd ed. Agronomy Monograph, vol. 9. American Society of Agronomy, Madison, WI, pp. 383–409.

26. Kalra, Y.P. (1995). Determination of pH of soils by diff erent methods: Collaborative study. *Journal of AOAC Internationa,* 78:310–324.

27. Rhoades, J.D. (1996). Salinity: Electrical conductivity and total dissolved solids. p. 417–435. In D.L. Sparks (ed.) *Methods of soil analysis*. Part 3. Chemical methods. SSSA Book Ser. 5. SSSA, Madison, WI.

28. Schnitzer, M. (1982). Total carbon, organic matter, and carbon, in: Page, A.L., Miller, R.H., Keeney, D.R. (Eds.), *Methods of Soil Analysis*, 2nd ed. Agronomy Monograph, vol. 9. American Society of Agronomy, Madison, WI, pp. 539–577 (Part 2).

29. Nelson, R.E. (1982). Carbonate and gypsum. In: Page, A.L. (Ed.), *Methods of Soil Analysis.* Part 2. American Society of Agronomy, Madison, pp. 181–197

30. Phillips, J.D. (2005). Weathering instability and landscape evolution. *Geomorphology*, 67, 255-272.

31. Taalab, K., Corstanje, R., Zawadzka, J., Mayr, T., Whelan, M.J., Hannam, J.A., Creamer. R., (2015). On the application of Bayesian Networks in Digital Soil Mapping. *Geoderma*, (259–260), 134–148.

32. Saldana, A., Ibáñez, J.J. (2004). Pedodiversity analysis at large scales: an example of three fluvial terraces of the Henares River (central Spain). *Geomorphology* 62, 123–138.

33. Gue, Y., Gong, P., Amundson, R. (2003). Pedodiversity in the United States of America. *Geoderma*, 117, 99–115.

34. Ibáñez, J.J., Caniego, J., San Jose, F., Carrera, C. (2005). Pedodiversity-area relationships for islands. *Ecological Modeling*, 182, 257-269.

35. Bagheri-Bodaghabadi, M., Salehi, M.H., Martínez-Casasnovas, José A., Mohammadi, J., Toomanian, N., Esfandiarpoor Borujeni, I., 2010, Using Canonical Correspondence Analysis

(CCA) to identify the most important DEM attributes for digital soil mapping applications. *Catena*, 86, 66-74.

36. Esfandiarpoor Borujeni, E., Toomanian, N., Salehi, M.H., Mohammadi, J. and Poch R.M. (2010). Assessing geopedological soil mapping approach by statistical and geostatistical methods: A case study in the Borujen region, Central Iran. *Catena* 82 (15): 1-14.
37. Maleki, S., Khormali, F., Bagheri Bodaghabadi, M., Mohammadi, J., Hoffmeister, D., Kehl, M. (2020). Role of geomorphic surface on the above-ground biomass and soil organic carbon storage in a semi-arid region of Iranian loess plateau. *Quaternary Internatioinal* 552, 111–121.
38. Keshtkar, S., Jafari, A., Farpoor, M.H. (2018). The effect of environmental and pedogenic factors on soil diversity in Kerman and Lalehzar regions. *Journal Soil Management and Sustanability*, 8(1), 89-106.
39. Nazari, N., Mahmoodi, S., Masihabadi, M.H. (2016). Employing Diversity and Similarity Indices to Evaluate Geopedological Soil Mapping in Miyaneh, East Azerbaijan Province, Iran. *Open Journal Geology*, 6, 1221-1239.