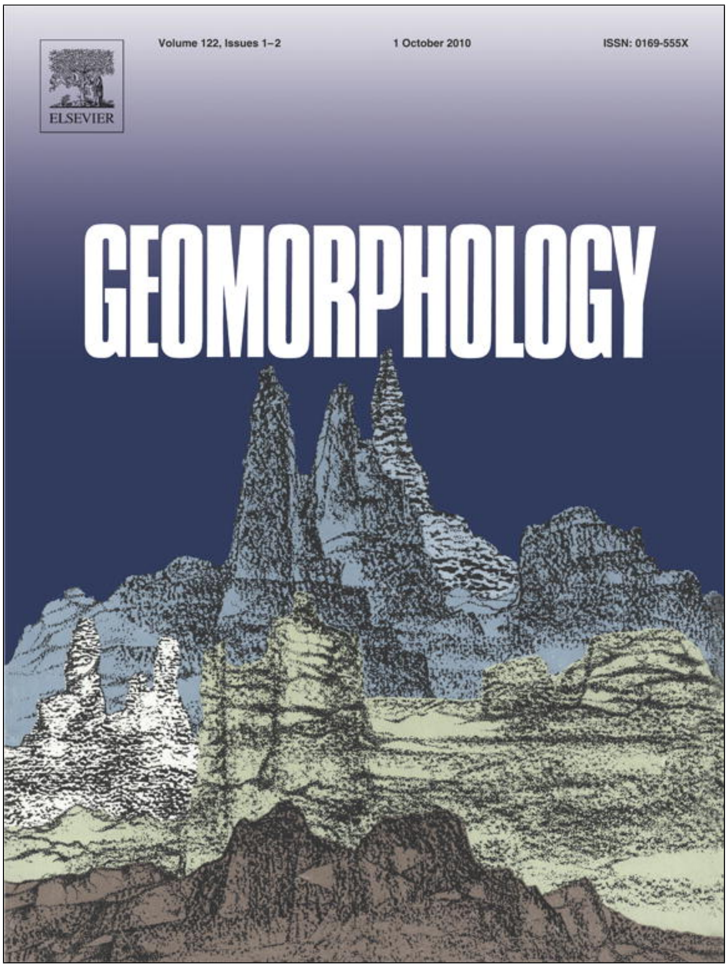
Volume 122, Issues 1–2
1 October 2010
ISSN: 0169-555X
ELSEVIER
GEOMORPHOLOGY

Geomorphology 122 (2010) 56–64

Contents lists available at ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

Evaluation of sedimentation vulnerability at small hillside reservoirs in the semi-arid region of Tunisia using the Self-Organizing Map

Achraf Hentati *, Akira Kawamura, Hideo Amaguchi, Yoshihiko Iseri

Department of Civil & Environmental Engineering, Tokyo Metropolitan University, 1-1 Minami-Ohsawa, Hachioji, Tokyo 192-0397, Japan

ARTICLE INFO

Article history:
Received 5 November 2009
Received in revised form 17 May 2010
Accepted 20 May 2010
Available online 2 June 2010

Keywords:
Erosion vulnerability
Hypsometric integral
Topographic index
Self-Organizing Map
Clustering
Semi-arid Tunisia

ABSTRACT

Small hillside reservoirs are fundamental for the sustainability of the water resources in the semi-arid region of Tunisia, where the government has constructed small dams since 1990. However, flash floods and severe erosion are jeopardizing the life of those small hydraulic structures prompting an assessment of the vulnerability of semi-arid catchments to erosion. In this study, we investigated the degree of erosion in 23 catchments by considering six variables related to erosion: the specific sediment yield, the hypsometric integral, the topographic index, the rainfall-related attribute as well as lithological and vegetation cover. More importance was assigned to the hypsometric integral and how it controls erosion vulnerability. The Self-Organizing Map of Kohonen, a robust clustering method, was implemented to analyze the patterns of the six variables. The results revealed four groups of catchments with different vulnerabilities to erosion. In addition, there is a global trend of increasing erosion risk from the Southwest to the Northeast of the semi-arid region of Tunisia.

1. Introduction

In many parts of the world, soil erosion affects the stability of ecosystems, often causing irreversible land degradation. Both on-site and off-site problems can be related to the variety of soil erosion processes. On-site effects refer mainly to a loss of agricultural and economic productivity, and to land degradation as a whole, whereas one of the most important off-site effects is certainly the sedimentation of reservoirs and their corresponding loss in water storage capacity (Verstraeten et al., 2003). Particularly in the semi-arid and arid Mediterranean regions, soil erosion is one of the major threats to the conservation of soil and water resources. The potential risk of soil erosion in these regions is due to the high climatic instability to which the extreme intensive rainfall events, the poor vegetation cover and the existence of highly vulnerable soils on steep slopes all contribute (Vandekerckhove et al., 2000; Raclot and Albergel, 2006; Al Ali et al., 2008). The soil erosion in Tunisia is severer than that in other Mediterranean countries. Mammou and Louati (2007) estimated that in 2030 the loss of storage capacity of the Tunisian reservoirs might reach 43% of their initial storage capacities.

In the early 1990s, the Tunisian government launched an ambitious program for constructing small hillside reservoirs in the northern and central regions of the country (Albergel and Rejeb, 1997). However, because of the intensive irregular rainfall and severe soil erosion, the life of these hydraulic structures is in jeopardy. Flash floods with a high sediment load are threatening many of these reservoirs and some of them have already filled up with sediment in less than 10 years after construction. For example, the Sadine reservoir in central Tunisia with a volume of 44,500 m^3 was entirely filled with sediments from two floods in August and September 1995 (Nasri et al., 2004). In 1993, a network for studying and observing over thirty small hillside catchments in the semi-arid region of Tunisia was set up to form a hydrological database. In 1996, the HYDROMED research project (Albergel and Nasri, 2001), led by the RID (Research Institute for Development) and financed by the European Union, selected several sites for pilot schemes in countries near the Mediterranean: Lebanon, Morocco, Syria and Tunisia. Several research projects aimed to build a hydrological model suitable for semi-arid Mediterranean catchment areas to highlight the importance of small dams in water harvesting and to study the sustainability of those hydraulic structures within a semi-arid context. Importance was assigned to the erosion vulnerability and the silting of the hillside reservoirs. Camus et al. (1995) described the bathymetric measurement technique for the sediment that was used in eight hillside reservoirs. Albergel et al. (1998) estimated the erosion using the USLE (Universal Soil Loss Equation) model in 24 catchments, and showed that it was difficult to estimate the siltation in reservoirs because of the high variability of erosion intensity. Mahjoub et al. (2001) considered 23 catchments and classified them by means of principal components analysis (PCA). Two classes were found: one represents the catchment areas with a high runoff coefficient and specific sediment yield, and the other includes the catchment areas with a low runoff coefficient. Felfoul et al. (2003) determined a significant linear relationship between a catchment lithologic number and the dam efficiency for fifteen small dams, indicating that an accurate lithological survey during the planning

* Corresponding author. Tel./fax: +81 42 677 2787.
E-mail address: achraf.hentati@gmail.com (A. Hentati).

0169-555X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.geomorph.2010.05.013

of new reservoirs is necessary to reduce dam siltation. Ayadi et al. (2008) made use of PCA and hierarchical clustering for regionalizing erosion vulnerability. Findings revealed three groups of catchments on the basis of their runoff coefficient, global slope index and vegetation cover.

Although catchment geomorphology is believed to be a major factor controlling the catchment erosion vulnerability (Strahler, 1952; Verstraeten and Poesen, 2001; Singh et al., 2008), previous studies paid more attention to factors like rainfall intensity, the runoff coefficient and lithological cover to evaluate erosion risk in the semi-arid region of Tunisia. There were no detailed studies about the geomorphological aspect. In a previous work (Hentati et al., 2009), we used the hypsometric curve (Strahler, 1952) and the topographic index (Beven and Kirkby, 1979) to describe the geomorphological characteristics of 30 catchments in the semi-arid region of Tunisia. Then we classified these catchments depending on the shape of the hypsometric curve, and found a correlation between the shape and the distribution of the topographic index. In this paper we consider these two geomorphological indices as well as the above-mentioned erosion-related factors to investigate the hillside reservoir sedimentation problems in the semi-arid region of Tunisia.

Integrated analyses of reservoir sediment data with respect to geomorphological properties, soil cover, vegetation and rainfall-related attributes help understanding the dominant factors governing the erosion and sedimentation vulnerability of reservoirs. It is necessary to use a multivariate data mining technique to detect and interpret any pattern in the raw data. Although most previous studies on soil erosion and catchment characteristics employed classical statistical methods such as the Clustering Analysis (CA) and PCA, a new method called the Self-Organizing Map (SOM) has been used recently as a multivariate method in various research fields such as electric engineering chemistry (Chen and Gasteiger, 1997; Kohonen, 2001). The SOM has a strong capability of extracting complicated patterns and visualizing their relationships using low dimensional arrays. Several studies (Ultsch and Vetter, 1994; Liu et al., 2006; Annas et al., 2007; Astel et al., 2007; Iseri et al., 2009) present the superior performance of the SOM over the conventional methods such as PCA. However, only a few studies in geomorphology (Ehsani and Quiel, 2008) utilized the SOM.

In this study, the SOM has been applied to classify 23 small hillside reservoirs in the semi-arid region of Tunisia in order to evaluate their sedimentation vulnerability by considering six variables related to catchment erosion. In contrast to the previous studies in this region, more importance was assigned to two geomorphological indices (the hypsometric integral and the topographic index) and how they control erosion. Specific sediment yield, the annual average occurrence of erosive rainfall events, lithology and vegetation were also considered.

2. Study area and data

2.1. Study area

Tunisia is a country situated on the Mediterranean coast of North Africa. Despite its small size, Tunisia has a relatively large climatic diversity. The Tunisian climate is influenced by the Mediterranean climatic perturbation from the North and the arid desertic climate from the South, with a semi-arid climate in the central region. It is generally characterized by hot dry summers, mild to cool rainy winters, but coastal areas are warm and temperate. Strong solar radiation in summer leads to low humidity and high evapotranspiration rates (Berndtsson, 1987; Fiorentino et al., 2006). The vegetation cover is sparse and unevenly distributed. Intense rainstorms over the sparsely vegetated surface create pronounced erosion.

Since 1993, the Direction of Soil and Water Conservation, Tunisia (DSWC) and the Research Institute for Development, France (RID) have collaborated on a research program for small hillside reservoirs. In the semi-arid Dorsal Mountains of central Tunisia that extend from the Northeast of the country to the Algerian border on the West (dashed line in Fig. 1), 30 hillside catchments were chosen to make up a network of hydrological observations within the HYDROMED (hydrological observation network) project (Hentati et al., 2009). We selected 23 catchments smaller than 10 km^2 for this study because the RID did not provide data for larger catchments (Fig. 1). The 23 catchments have an elevation range from 70 to 1352 m (Table 1). They are located along the rainfall gradient within the semi-arid region of Tunisia, with 250 to 550 mm of annual rainfall.

2.2. Data

Digital Elevation Models (DEMs) for the 23 catchments were derived from the Shuttle Radar Topographic Mission (SRTM) data with a 3-arc second resolution (http://srtm.csi.cgiar.org/). The DEM made it possible to obtain the hypsometric curve and the topographic index distribution for each catchment (Hentati et al., 2009). The rainfall data came from 23 rain gauges of the HYDROMED, over a period from 1993 to 2003. The rainfall was measured by automated tipping-bucket rain gauges. The lithological and vegetation cover data were taken from Temple-Boyer (2002). They describe the ratio of the catchment area covered by four lithological types and three vegetation types (Table 1). Siltation measurements started in 1993 with up to five observations for each reservoir during about ten years. The bathymetric surveys of reservoirs were made by echo-sounding along pre-defined transects (Camus et al., 1995). The change in reservoir volume was deduced from the difference between one observation time and another (Table 1).

3. Methodology

3.1. Self-Organizing Map

The SOM method was invented by Teuvo Kohonen to provide a way of representing multidimensional data in a much lower dimensional space, usually one or two dimensions (lattice). This process of reducing vector dimensionality is essentially a data compression technique known as “vector quantization” (Kohonen, 2001). In addition, the technique creates a network of information in which any topological relationships within the training set are maintained.

The SOM does not need a target output to be specified, unlike other types of neural networks. From the initial distribution of random weights, and in many different iterations, the SOM eventually settles into a map of stable zones. Training to obtain an SOM occurs in several steps (Table 2).

The number of the output neurons in an SOM (the map size) is important to detect the deviation of the data. If the map size is small, it might not explain some important differences that should be detected. Conversely, if the map size is too big, the differences are too small. The number of output neurons in an SOM can be selected using the heuristic rule suggested by Vesanto et al. (2000):

$$m = 5\sqrt{n}, \tag{1}$$

where m is the map size and n is the number of the training samples. Using this formula, the map size can be efficiently determined without trial and error. It has been observed that a rectangle SOM is better than a square SOM, since the former can more easily align with the training data, which may distributed along a dominant axis (May et al., 2010). In case of the rectangular SOM, the two largest eigenvalues of the training data are first calculated, then the ratio between side lengths of the map grid is set to the ratio between the two maximum eigenvalues (Vesanto et al., 2000). The actual side lengths are finally set so that the product of vertical and horizontal length is close to the number m given by Eq. (1). In this work, the SOM tool box, which calculates the map size by the rule given by Vesanto et al. (2000), was used to determine m.

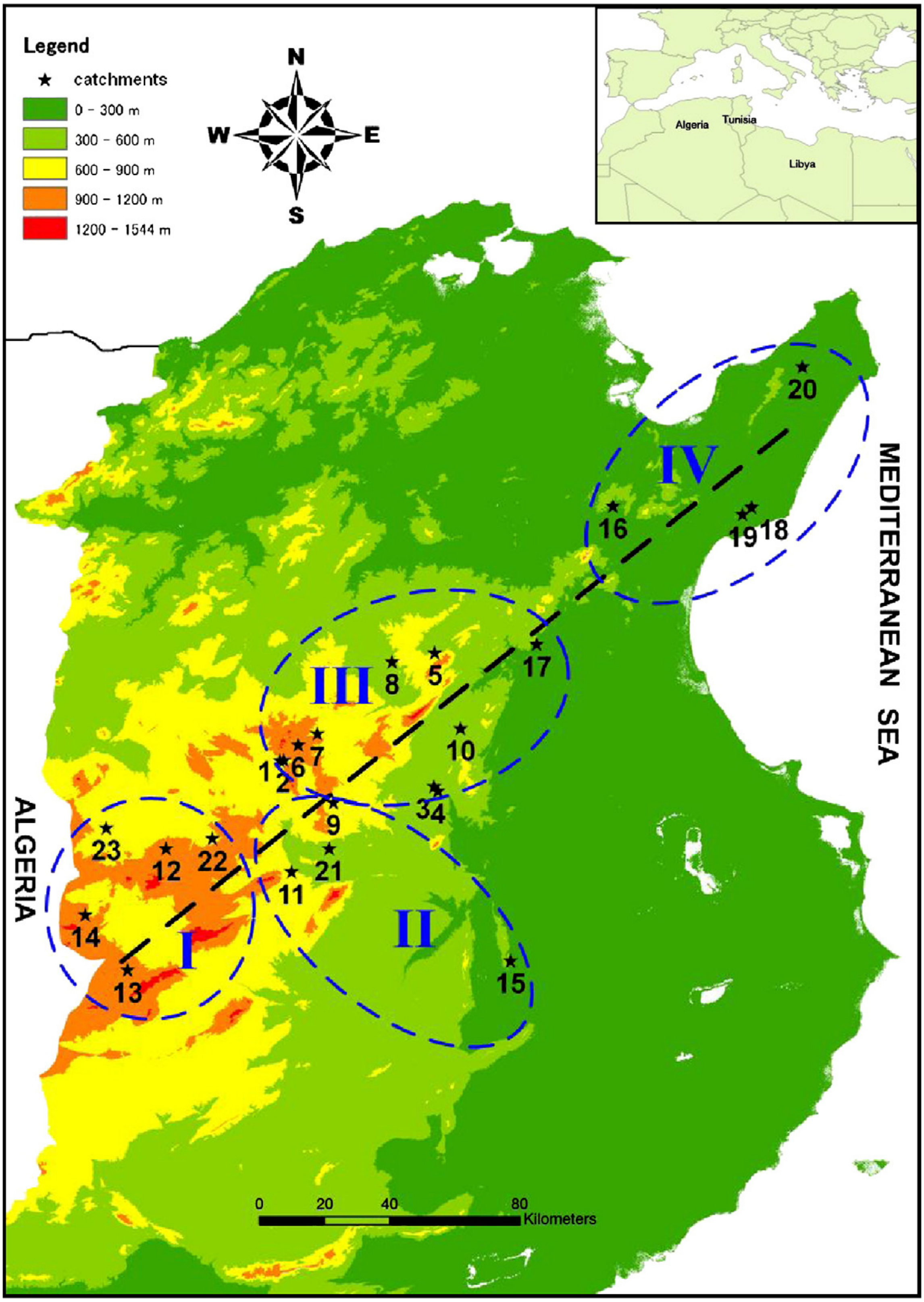

Fig. 1. Location of the 23 studied catchments.

3.2. Ward's method

In this paper, the clustering procedure is based on a two-level approach, where the data set is firstly clustered using an SOM, and then the SOM is clustered by Ward's method. The method is a hierarchical agglomerative clustering that uses the Ward criterion, and aims at obtaining a classification such that each element of a cluster is as close as possible to the elements of this cluster, and as far as possible from elements belonging to any other clusters (Leloup et al., 2007). The method computes a hierarchical clustering of the

Table 1
Properties of 23 studied catchments.

Catchment		Main catchments characteristics			Main reservoirs characteristics				Geom. attributes		Rainfall			Lithology						Vegetation			
No.	Name	*Area* (km^2)	*Alt.* Min (m)	*Alt.* Max (m)	*Storage* (m^3)	*Sedim.* (m^3)	*Sed. Obs.* (yr)	*SSY* (m^{-3} ha^{-1} yr^{-1})	*HI*	*PFR* (%)	*Start* Obs.	*End* Obs.	PI_{15} (events yr^{-1})	*MRL* (%)	*CLA* (%)	*SND* (%)	*LIM* (%)	*SST* (%)	*REI*	*F* (%)	*CL* (%)	*BL* (%)	*VEI*
1	Sadine1	3.84	842	1250	44500	41085	12.51	8.55	0.44	19.87	Feb-92	Sep-00	0.89	90	0	0	10	0	0.90	0	68	32	0.66
2	Sadine2	6.53	825	1267	108800	106585	10.50	15.54	0.55	22.54	Dec-95	Jan-00	0.80	90	0	0	10	0	0.90	38	62	0	0.31
3	Fidh Ben naceur	1.69	350	462	47110	14720	9.41	9.26	0.43	21.00	Apr-94	Sep-01	0.63	80	0	0	20	0	0.80	0	57	43	0.72
4	Fidh Ali	4.12	335	444	134710	49840	8.67	13.95	0.41	27.24	Jan-93	Sep-04	0.42	80	0	0	20	0	0.80	0	12	88	0.94
5	Mrichet	1.58	590	730	41780	8990	8.73	6.51	0.50	26.46	Nov-93	Oct-01	0.63	90	0	0	10	0	0.90	0	92	8	0.54
6	Hadada	4.69	900	1246	84970	14060	4.38	6.85	0.45	21.53	Sep-93	Aug-05	0.18	30	0	0	70	0	0.30	0	76	24	0.62
7	Jannet	5.21	820	1191	94280	55910	6.36	16.86	0.38	24.64	Sep-93	Oct-03	0.45	70	0	0	30	0	0.70	0	62	38	0.69
8	Hanach	3.95	447	834	77220	18590	4.41	10.66	0.41	26.30	Oct-93	Oct-02	0.20	90	0	0	10	0	0.90	1	43	56	0.78
9	Abdessadok	3.07	815	1189	92530	25960	8.72	9.70	0.31	21.38	Oct-93	Dec-00	0.63	80	0	0	0	20	0.80	0	57	43	0.72
10	Dekikira	3.07	380	479	219100	21300	5.45	12.73	0.44	21.38	Oct-93	Oct-03	0.73	60	0	0	30	10	0.60	33	25	42	0.55
11	Es Senaga	3.63	618	883	86420	27780	7.43	10.30	0.33	20.30	Nov-93	Mar-02	0.80	40	0	0	10	50	0.40	0	34	66	0.83
12	Echar	9.17	970	1190	186840	5300	3.47	1.67	0.45	21.61	Nov-93	Jul-99	0.29	90	0	0	10	0	0.90	0	81	19	0.60
13	Abdeladim	6.42	1030	1224	174870	6895	7.58	1.42	0.33	17.53	Nov-93	Sep-05	0.23	100	0	0	0	0	1.00	20	76	4	0.42
14	Arara	7.08	910	1352	91150	58940	5.59	14.89	0.40	17.23	Nov-93	Oct-01	0.11	90	0	0	10	0	0.90	59	41	0	0.21
15	Mouidhi	2.66	235	363	142770	26980	10.09	10.05	0.42	18.46	Nov-93	Sep-05	0.64	50	10	0	40	0	0.60	0	10	90	0.95
16	Sbaihia	3.24	300	473	135100	10600	3.75	8.72	0.47	28.07	Dec-93	Oct-03	0.36	90	0	0	10	0	0.90	80	20	0	0.10
17	Saadine	2.72	245	552	35620	27610	6.62	15.32	0.35	21.64	Jan-94	Oct-98	1.60	70	0	0	30	0	0.70	30	70	0	0.35
18	Es Segir	4.31	70	232	192450	2020	4.76	0.98	0.44	24.31	Jan-94	Sep-03	0.80	0	20	80	0	0	0.36	20	80	0	0.40
19	Maleh	0.85	90	144	19875	4836	8.62	6.60	0.50	30.00	Jan-94	Aug-99	1.00	0	90	0	0	10	0.90	40	30	30	0.45
20	Kamech	2.45	95	203	142560	29900	6.96	17.53	0.53	23.79	Mar-94	Sep-05	0.64	0	80	0	0	20	0.80	0	75	25	0.63
21	Brahim	4.64	570	1015	86540	30160	7.71	8.43	0.40	17.62	Mar-94	Sep-99	0.83	60	0	0	35	5	0.60	30	27	43	0.57
22	Baouejer	4.86	987	1118	66030	6510	7.73	1.73	0.34	22.15	May-93	May-00	0.63	90	0	0	10	0	0.90	80	20	0	0.10
23	Mrira	6.13	770	940	126350	11760	5.47	3.51	0.35	20.47	Apr-93	Sep-05	0.55	90	0	0	10	0	0.90	90	10	0	0.05

The catchment numbers correspond to those in Figs. 1, 4 and 5. *Alt.*: altitude, *Sedim.*: sediment, *Sed. Obs.*: sediment observation period, *SSY*: specific sediment yield, Geom. Attributes: geomorphological attributes, *HI*: hypsometric integral, *PFR*: peak frequency of the topographic index distribution, PI_{15}: annual average occurrence of 15-min rainfall events with peak intensity higher than 80 mm h^{-1}, *MRL*: marl, *CLA*: clay, *SND*: sand, *LIM*: limestone, *SST*: sandstone, *REI*: Rock cover erodibility index, *F*: forest, *CL*: cultivated lands, *BL*: bare lands, *VEI*: vegetation cover erodibility index.

Table 2
Training steps of SOM.

Step	Description	Formulas
1	Each neuron's weight W_i is randomly initialized.	
2	A vector V_i is chosen randomly from the set of training data and presented to the lattice.	
3	The weighting of every neuron is compared to the input vector and the neuron with the best match (BMU: best matching unit). The Euclidian distance D is calculated.	$D = \sqrt{\sum_{i=1}^{n} (V_i - W_i)^2}$ $i = 1, 2, 3, \ldots\ldots$
4	The radius σ of the neighborhood of the BMU is now calculated. This is a value that starts large, typically set to the 'radius' of the lattice, but diminishes at each time-step. This is accomplished by making use of the exponential decay function.	$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\lambda}\right)$ where t is time, σ_0 is the width of the lattice at time t_0, and λ is a decay constant.
5	Weight of each neighboring neuron found in step 4 is adjusted to provide the input vector. If a neuron is closer to the BMU, its weight is more altered. Basically the new adjustment weight for the neuron is equal to the old weight W plus a fraction of the difference in θL between the old weight and the input vector V.	$W(t+1) = W(t) + \theta(t)L(t)[(V(t) - W(t)]$ $t = 1, 2, 3\ldots\ldots$ where L is the learning rate and θ is the neighborhood function.
6	Repeat from step 2 for N iterations.	

reference vectors of the SOM to form only one cluster containing all the neurons in the end.

3.3. Preparation of input data

3.3.1. Hypsometric integral

The hypsometric curve represents the relative proportion of a basin below a given height. For a selected basin, its relief (i.e. elevation difference between the summit and the outlet) is divided into equal elevation intervals, and for each interval, the proportion of the basin area is calculated. Then, elevations and areas for each interval are normalized by the relief and total area of the catchment area, respectively. Harlin (1978) quantitatively describe the hypsometric curve as a cumulative probability distribution:

$$F(Z) = \int_{Z}^{1} f(Z)dZ \qquad (2)$$
$$0 \le F(Z) \le 1; \quad 0 \le Z \le 1$$

where F is the cumulative probability distribution of a normalized catchment area at or above a normalized altitude Z, and f is the probability density function of area change with altitude.

The Hypsometric Integral (HI) is defined as the area under the hypsometric curve (Strahler, 1952):

$$HI = \int_{0}^{1} f(Z)dZ \qquad (3)$$

This value can be used to estimate catchment erosion. Strahler (1952) interpreted the low values of HI (concave hypsometric curves) as indicating stabilized catchments and the high values (convex hypsometric curves) as symptoms of catchments with a greater susceptibility to erosion.

The hypsometric curves of the catchments studied were obtained by using the DEMs to determine the point-pairs (area, height). Then we calculated HI according to Eq. (3), as an input variable for creating the SOM (Table 1).

3.3.2. Topographic Index

The Topographic Index (TI) represents the propensity of a point within a catchment to generate saturation excess overland flow, or a topographic control of surface and subsurface flow. Beven and Kirkby (1979) defined TI at a point/pixel of a catchment as:

$$TI = \ln\left(\frac{a}{\tan\beta}\right) \qquad (4)$$

where a is the specific upslope area per unit contour length draining through the point, and β is the local topographic slope at the point (Fig. 2A). Flat areas with large specific upslope areas will have a higher TI values than steep areas with small upslope areas.

Our previous study (Hentati et al., 2009) revealed a correlation between HI and the peak frequency (PFR) of TI distribution (Fig. 2B). We selected both HI and PFR as input for the SOM, first to confirm their correlation by the SOM and second to study the control of the geomorphological indices on erosion.

The TI distribution was obtained using the Spatial Analyst extension of ArcGIS software with six steps: 1) preparation of the DEM; 2) DEM pre-processing by the sink removal algorithm; 3) implementation of a

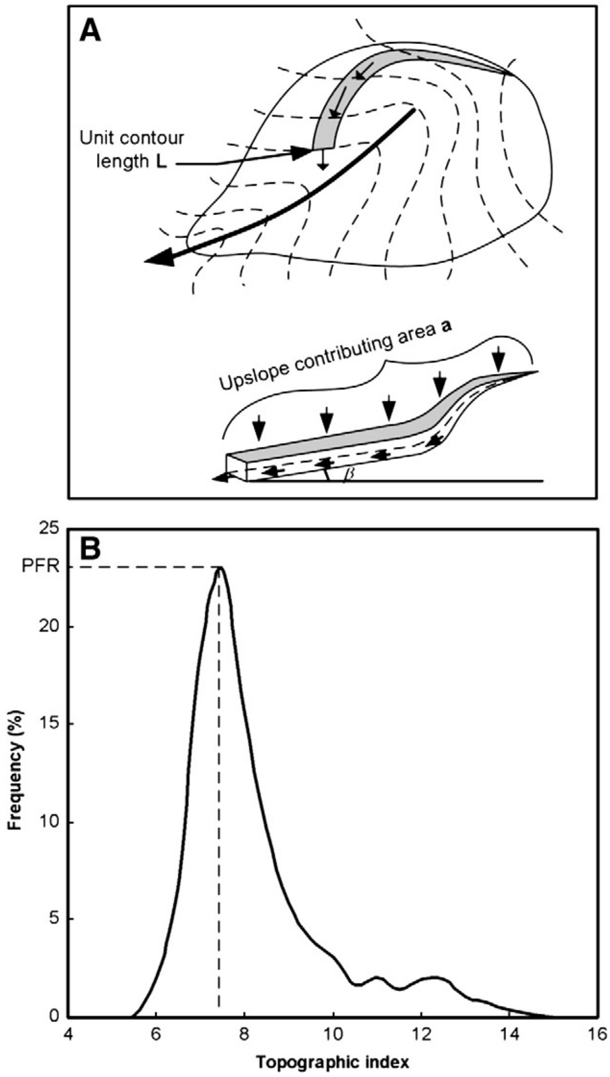

Fig. 2. Definitions of (A) the topographic index (Tarboton, 2001) and (B) PFR (peak frequency of the topographic index distribution).

single flow direction algorithm on the processed DEM; 4) determination of flow accumulation; 5) computation of the slope by applying a slope algorithm on the original DEM; and, 6) calculation of *TI* values to obtain *TI* distributions and their *PFR* (Table 1).

3.3.3. Rainfall

Jebari et al. (2008) have shown that the 15-minute duration is the most representative erosive rainfall in the semi-arid region of Tunisia. We calculated the annual average occurrence of exceptional rainfalls having a peak intensity >20, 40, 60 and 80 mm h^{-1}, and found that the frequency of 15-minute duration events with a peak intensity >80 mm h^{-1} (PI_{15}) correlated the greatest with the observed specific sediment yields in the 23 catchments (Table 1).

3.3.4. Lithology cover

Previous studies on sediment yield in Mediterranean areas have revealed that catchment lithology exerts a dominant influence. Lahlou (1988) and Verstraeten et al. (2003) showed that catchments underlain by marls or clays have a significantly higher specific sediment yield (*SSY*) than catchments underlain by limestone or sandstone. We gave the following erodibility scores to lithology types: 0 for limestone and sandstone, 0.2 for sand and 1 for marl and clay. The rock erodibility index (*REI*) was calculated as follows (see Table 1):

$$REI = \sum_{i}^{n_L} a_{Li} sc_{Li} \quad (5)$$

where a_{Li} is the ratio of the catchment area covered by lithology i,sc_{Li} is its score and n_L is the number of different rock components.

3.3.5. Vegetation cover

The vegetation cover of sloping land also plays a very important role in sediment production. Vegetation in semi-arid environments is characterized by heterogeneous patterns of bare and vegetated lands. Bare lands consist mainly of bare rock and crusted soils with poor structure and low infiltration rates, whereas vegetated lands have better soils with higher organic matter contents, stronger aggregation, and higher infiltration rates. We gave a score of 1 for bare lands, 0.5 for cultivated lands and 0 for forests. The vegetation erodibility index (*VEI*) is defined in the same way as *REI*:

$$VEI = \sum_{i}^{n_V} a_{Vi} sc_{Vi} \quad (6)$$

where a_{Vi} is the ratio of the catchment area covered by vegetation i, sc_{Vi} is its score and n_V is the number of the different vegetation covers.

3.3.6. Sediment yield

Sediment yield was measured based on bathymetric observations in each catchment reservoir during 10 years from 1993, with up to five observations for each reservoir. The estimated sediment yield was standardized by the total catchment area to obtain *SSY* (Table 1).

4. Results

The six variables we used to obtain an SOM for the 23 reservoirs are: *SSY*, *HI*, PI_{15}, *PFR*, *REI* and *VEI*. The size of the SOM is determined as 6×4, and then the map was trained with the input variables to self-organize the 23 input vectors.

In Fig. 3, six component planes (CPs) and their color bars in gray scale. On the right were used to visualize the distribution of the 23 input vectors for the six variables via the 6×4 neuron map. Black corresponds to high values, gray moderate values, and white low values. To accurately determine the catchment groups, the data set was firstly clustered using the SOM; then the SOM reference vectors were clustered by Ward's method. At the first step, we had 24 clusters or reference vectors. Then at

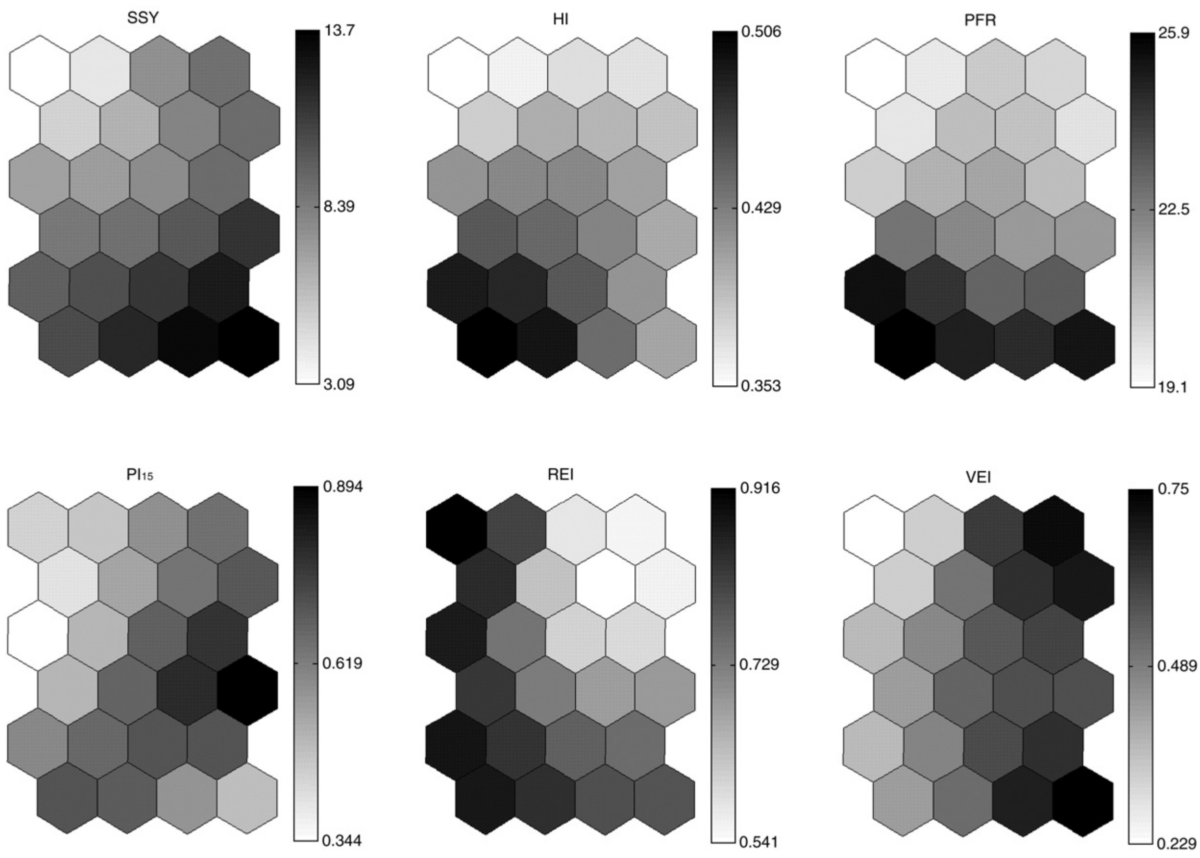

Fig. 3. Component planes representing the denormalized values of the six variables. Correlation between *HI* and *PFR* can be confirmed from this SOM visualization.

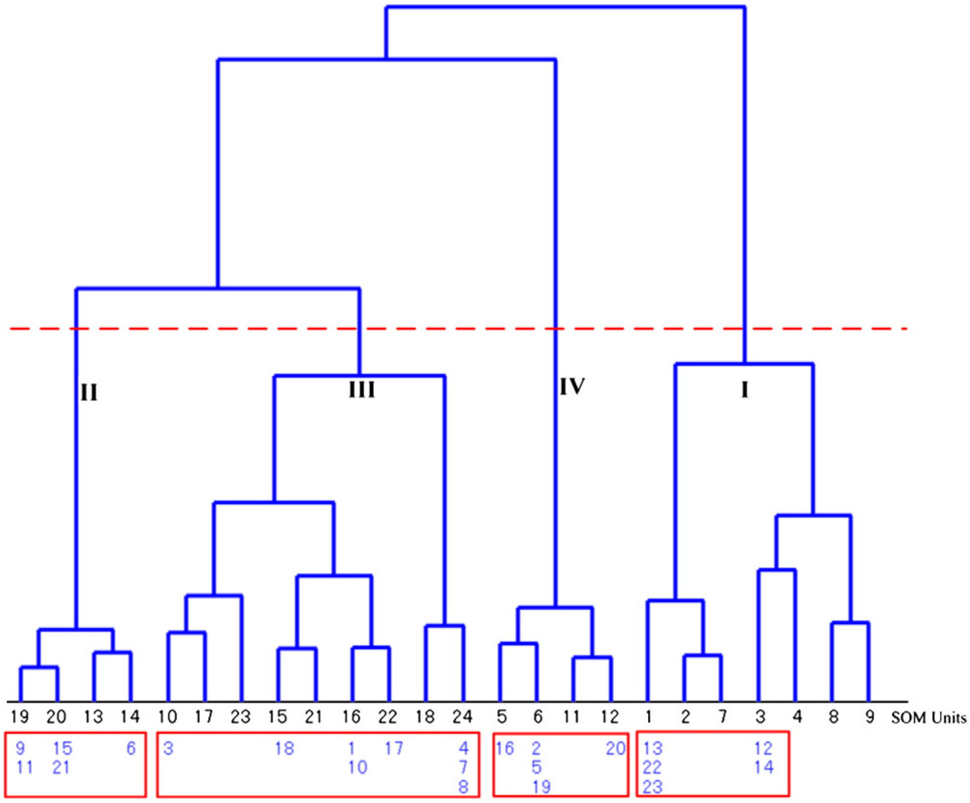

Fig. 4. Classification tree. Latin cardinal numbers show the four groups of catchments. Arabic numbers in the rectangles are the catchment numbers.

each step, the two closest clusters were aggregated according to the Ward criterion. The resultant dendrogram (Fig. 4) represents the partitioning of the 24 neurons of the SOM. Each leaf of the tree is a subset of neurons; each node represents the conjunction of two clusters; and the size of a branch represents the distance between the two clusters.

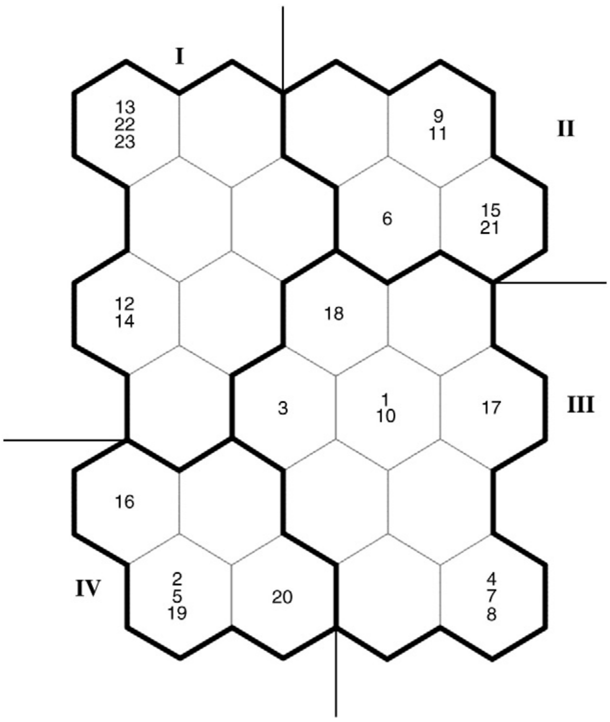

Fig. 5. Two-level clustered SOM. Borders between clusters are obtained by cutting the dendrogram at the level of four clusters.

Fig. 5 shows the result of the two-level clustering of the SOM. The 23 catchments were classified into four groups (I, II, III and IV) each of which represents a union of SOM neurons. Group I includes catchments 12, 13, 14, 22 and 23; Group II catchments 6, 9, 11, 15 and 21; Group III catchments 1, 3, 4, 7, 8, 10, 17 and 18; and Group IV catchments 2, 5, 16, 19 and 20. Fig. 6 shows hexagrams representing the mean standardized values of reference vectors for the four groups. The figure allows the interpretation of the global characteristic of each group. Fig. 7 shows the averaged hypsometric curves for each group and the corresponding averaged *TI* distributions. The curve of Group IV is located highest, followed by that of Group III, II and I (Fig. 7A). *PFR* of the averaged *TI* distributions follows the same order (Fig. 7B).

5. Discussion and conclusions

The component planes of Fig. 3 can be used for correlation hunting. For example, the color shades of SOM neurons for *SSY* and *VEI* exhibit a decrease from the right to the left, because soil protected by vegetation is less eroded. Fig. 3 also shows the correlation between *HI* and *PFR* as was already noted in Hentati et al (2009). Fig. 7 confirms this correlation; *HI* is positively correlated with *PFR* of the four groups. This is because *HI* is negatively correlated with catchment drainage density (Strahler, 1952). When *HI* is low, the river network is well developed and the global slopes are gentle. Consequently, the high values of *TI* are more frequent and hence *PFR* decreases.

As shown in Fig. 6, Group I is representative of potentially erodible catchments because *REI* is high. However, their distribution in the westernmost region of the Dorsal Mountains (Fig. 1) is characterized by infrequent erosive rainfall events and a dense vegetation cover, and thus the erosion intensity in Group I is low (Fig. 6). Furthermore, the Group II catchments are mainly located along the East of the Dorsal Mountains (Fig. 1), characterized by low *REI* values but the relatively frequent occurrence of erosive rainfall events and sparse vegetation cover, resulting in moderate erosion.

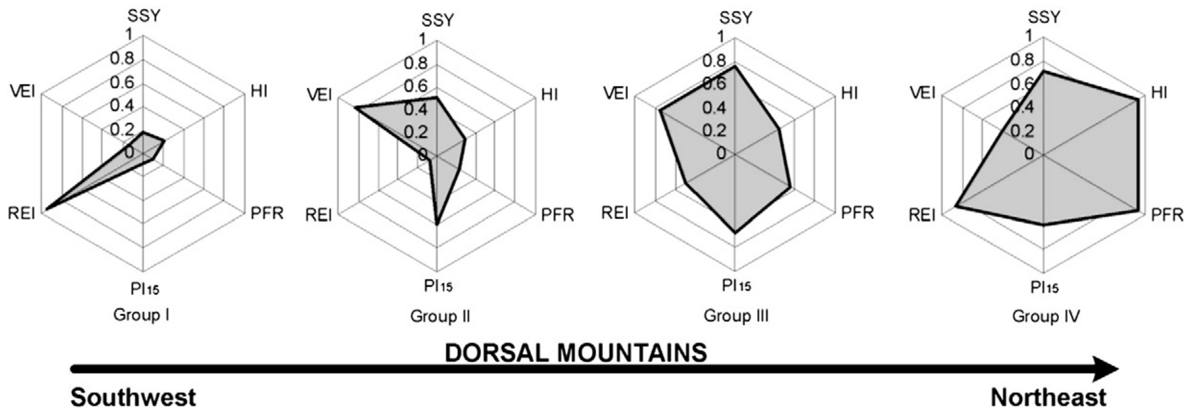

Fig. 6. Hexagrams of the four catchment groups.

Group III in the eastern part of the study area has a bigger hexagram than Groups I and II (Fig. 6). All the factors have relatively high values particularly PI_{15}, *VEI*, and *REI*. The geomorphological indices *HI* and *PFR* consistently increase from Group I to Group III. The higher *HI* values for Group III (0.35 to 0.44) indicate that they are more susceptible to erosion than the catchments in Groups I and II. This condition coupled with higher PI_{15}, *VEI*, and *REI* leads to rapid erosion in the Group III catchments. Group IV in the northeastern end of the Dorsal Mountains (Fig. 1) have the highest *HI* and *PFR* values (Fig. 6), with high *REI* and PI_{15}, leading to a highly erodible condition despite dense vegetation cover represented by low *VEI*.

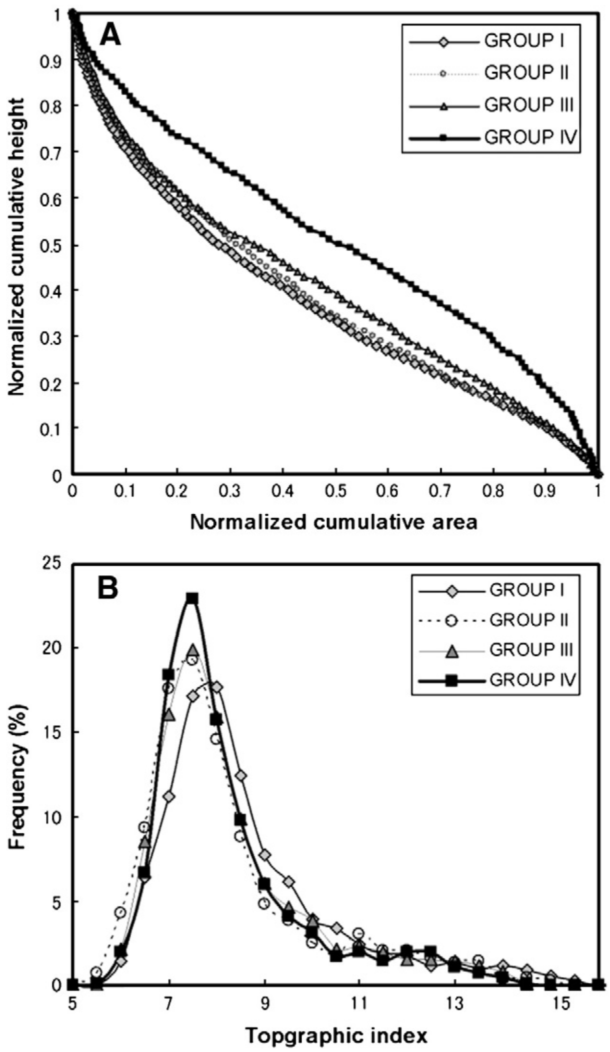

Fig. 7. Averaged hypsometric curves (A) and the averaged topographic index distributions (B) of the four groups.

The catchments of Groups III and IV need both vegetative and mechanical soil/water conservation measures to reduce erosion and sediment load. Although the catchments of both Groups I and II only need minimum mechanical and vegetative measures, they may require structures to stock water at appropriate locations for conjunctive water use.

The integrated analysis of the six variables by the SOM revealed a global trend of the erosion risk which increases from the Southwest to Northeast of the semi-arid region of Tunisia (Fig. 1). However, some catchments do not simply follow this trend. For instance, the catchments 14 and 18 have *SSY* different from the general values of Groups I and IV, respectively. *SSY* for the catchment 14 is 14.89 $m^3\ ha^{-1}\ year^{-1}$, which is higher than that of the other catchments of the Group I. This value is due to the exceptional 5-year return period erosive rainfall events (Jebari et al., 2010) and to the fact that 90% of the catchment is occupied by highly erodible marl (Table 1). *SSY* for the catchment 18 is 0.98 $m^3\ ha^{-1}\ year^{-1}$, which is much lower than that in the other catchments of Group IV. This low sediment yield is due to soil conservation practices (Jebari et al., 2010).

We compare our results with previous studies about sedimentation at hillside reservoirs in the semi-arid region of Tunisia. Mahjoub et al. (2001) divided the 23 small catchments of the HYDROMED project into two groups based on PCA for their areas and runoff coefficients. One group contains two catchments having large areas and high runoff coefficient, and the remaining catchments belong to the other. Ayadi et al. (2008) applied a hierarchical clustering technique and PCA and classified 24 catchments into three groups based on the overall catchment slope and the runoff coefficient. In both cases, the catchments classified are distributed almost randomly. In contrast, our catchment classification is more detailed (into four groups) and consistent with the general environmental gradient. Consequently, our grouping seems to be superior to the previous classifications.

Acknowledgments

This research was carried out within the "Support Project for the Advanced Education and Research" funded by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the research project "Solutions for the water related problems in Asian Metropolitan areas" supported by the Tokyo Metropolitan Government, Japan.

References

Al Ali, Y., Touma, J., Zante, P., Nasri, A., Albergel, J., 2008. Water and sediment balances of a contour bench terracing system in a semi-arid cultivated zone (El Gouazine, central Tunisia). Hydrological Sciences Journal 53, 883–892.

Albergel, J., Nasri, S., 2001. HYDROMED: Rapport final du programme de recherche sur les lacs collinaires dans les zones semi-arides du pourtour méditerranéen. Contrat européen INCO DC ERBIC 18 CT 90091-STD4. IRD/INGREF Tunis. .

Albergel, J., Rejeb, N., 1997. Les lacs collinaires en Tunisie: Enjeux, contraintes et perspectives. Comptes Rendus de l'Académie d'Agriculture de France 77–88.

Albergel, J., Boufaroua, M., Pepin, Y., 1998. Bilan de l'érosion sur les petits bassins versants des lacs collinaires en climat semi-aride tunisien. Bulletin de l'ORSTOM 18, 67–75.

Annas, S., Kanai, T., Koyama, S., 2007. Principal component analysis and self-organizing map for visualizing and classifying fire risks in forest regions. Agricultural Information Research 16 (2), 44–51.

Astel, A., Tsakouski, S., Barbieri, P., Simeonov, V., 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Research 41, 4566–4578.

Ayadi, I., Abida, H., Djebbar, Y., 2008. Caractérisation des lacs collinaires de la Tunisie centrale. International Water Resources Association, XIIIth World Water Congress, Montpellier, France.

Berndtsson, R., 1987. Spatial and temporal variability of rainfall and potential evaporation in Tunisia. IAHS Publication 168, 91–100.

Beven, K.J., Kirkby, M.J., 1979. A physically-based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin 24, 43–69.

Camus, H., Guiguen, N., Ben Younes, M., 1995. Note sur l'envasement des lacs collinaires en zone semi-aride tunisienne. Rapport publié par la Direction de la Conservation des Eaux et du Sol (D/CES) and l'Institut de Recherche pour le Développement en coopération (ORSTOM).

Chen, L., Gasteiger, J., 1997. Knowledge discovery in reaction databases: landscaping organic reactions by a self-organizing neural network. Journal of the American Chemical Society 119, 4033–4042.

Ehsani, A.H., Quiel, F., 2008. Geomorphometric feature analysis using morphometric parametrization and artificial neural networks. Geomorphology 99, 1–12.

Felfoul, M.S., Snane, M.H., ALbergel, J., Mechergui, M., 2003. Relationship between small dam efficiency and gully erodibility of the lithologic formations covering their watershed. Bulletin of Engineering Geology and the Environment 62, 315–322.

Fiorentino, M., Carriero, D., Iacobellis, V., Manfreda, S., Portoghese, I., 2006. Medclub-starting line and first activities. Predictions in Ungauged Bassins (PUB): Promises and Progress. IAHS Publication 303, 1–14.

Harlin, J., 1978. Statistical moments of the hypsometric curve and its density function. Mathematical Geology 10, 59–72.

Hentati, A., Kawamura, A., Amaguchi, H., 2009. Regional geomorphological characteristics of small hillside river basins in semi-arid region of Tunisia. Annual Journal of Hydraulic Engineering, JSCE 53, 43–48.

Iseri, Y., Matsuura, T., Iizuka, S., Nishiyama, K., Jinno, K., 2009. Comparison of pattern extraction capability between self-organizing maps and principal component analysis. Memoirs of the Faculty of Engineering. Kyushu University 69 (2), 37–47.

Jebari, S., Berndtsson, R., Bahri, A., Boufaroua, M., 2008. Exceptional rainfall characteristics related to erosion risk in semi-arid Tunisia. The Open Hydrology Journal 1, 1–9.

Jebari, S., Berndtsson, R., Bahri, A., Boufaroua, M., 2010. Spatial soil loss risk and reservoir siltation in semiarid Tunisia. Hydrological Sciences Journal 55, 121–137.

Kohonen, T., 2001. Self-organizing Maps, third edition. Springer-Verlag, Berlin.

Lahlou, A., 1988. The silting of Moroccan dams. Sediment Budgets, Proceedings of the Porto Alegre Symposium. IAHS Publication 174, 71–77.

Leloup, J.A., Lachkar, Z., Boulanger, J.P., Thiria, S., 2007. Detecting decadal changes in ENSO using neural networks. Climate Dynamics 28, 147–162.

Liu, Y., Weisberg, R.H., Mooers, C.N.K., 2006. Performance evaluation of the self-organizing map for feature extraction. Journal of Geophysical Research 111, C05018.

Mahjoub, M.R., Bergaoui, M., Souissi, A., Boufaroua, M., 2001. Régionalisation de l'envasement des lacs collinaires au niveau de la dorsale tunisienne. Sud Sciences & Technologies 7, 4–19.

Mammou, A.B., Louati, M.H., 2007. Evolution temporelle de l'envasement des retenues de barrages de Tunisie. Revue des sciences de l'eau 20, 201–210.

May, R.J., Maier, H.R., Dandy, G.C., 2010. Data splitting for artificial neural networks using SOM-based stratified sampling. Neural Networks 23, 283–294.

Nasri, S., Cudennec, C., Albergel, J., Berndtsson, R., 2004. Use of a geomorphological transfer function to model design floods in small hillside catchments in semi-arid Tunisia. Journal of Hydrology 287, 197–213.

Raclot, D., Albergel, J., 2006. Runoff and water erosion modelling using WEPP on a Mediterranean cultivated catchment. Physics and Chemistry of the Earth 31, 1038–1047.

Singh, O., Sarangi, A., Sharma, M.C., 2008. Hypsometric integral estimation methods and its relevance on erosion status of North-Western Lesser Himalayan catchments. Water Resources Management 22, 1545–1560.

Strahler, A.N., 1952. Hypsometric (area–altitude) analysis of erosional topography. Bulletin of the Geological Society of America 63, 1117–1142.

Tarboton, D.G., 2001. Terrain analysis using Digital Elevation Models (TauDEM), Presentation at GIS Hydro 2001. ESRI Users Conference Pre-Conference Seminar in San Diego, July 8, 2001.

Temple-Boyer, L., 2002. Note sur les données «relief»: Caractéristiques morphométriques, occupation du sol et aménagemets. IRD (ORSTOM), Tunis.

Ultsch, A., Vetter, C., 1994. Self-organizing-feature-maps versus statistical clustering methods: A benchmark. FG Neuroinformatik & Künstliche Intelligenz. Research Report, 0994 (14 pp.).

Vandekerckhove, L., Poesen, J., Oostwood Wijdenes, D., Nachtergaele, J., Kosmas, C., Roxo, M.J., De Figueiredo, T., 2000. Thresholds for gully initiation and sedimentation in Mediterranean Europe. Earth Surface Processes and Landforms 25, 1201–1220.

Verstraeten, G., Poesen, J., 2001. Factors controlling sediment yield from small intensively cultivated catchments in a temperate humid climate. Geomorphology 40, 123–144.

Verstraeten, G., Poesen, J., De Vente, J., Koninckx, X., 2003. Sediment yield variability in Spain: a quantitative and semiqualitative analysis using reservoir sedimentation rates. Geomorphology 50, 327–348.

Vesanto, J., Himberg, J., Alhoniemi, E., Parahankangas, J., 2000. SOM Toolbox for Matlab 5. Helsinki University Report A57.