POLISH ASSOCIATION FOR SPATIAL INFORMATION
ANNALS OF GEOMATICS 2009 ❍ VOLUME VII ❍ NUMBER 2(32)

QUANTIFICATION OF BARE SOIL AND ITS SPATIO-TEMPORAL DYNAMIC USING DIFFERENT IMAGE CLASSIFICATION METHODS

KWANTYFIKACJA ODSŁONIĘTEJ GLEBY ORAZ JEJ DYNAMIKA CZASOPRZESTRZENNA Z WYKORZYSTANIEM RÓŻNYCH METOD KLASYFIKACJI OBRAZU

Manuel Pulido Fernández, Joakín Francisco Lavado Contador, Álvaro Gómez Gutiérrez , Francisco Javier Lozano Parra

GeoEnvironmental Research Group, Philosophy and Letters Faculty
University of Extemadura, Cáceres, Spain

Keywords: bare soil, pasturelands, image classification, dynamic, algorithms
Słowa kluczowe: odsłonięta gleba, pastwiska, klasyfikacja obrazu, dynamika, algorytmy

Introduction

Soil is essential for the development of a wide range of vital functions. Its availability and productivity in the long-term is sometimes seriously threatened by human activities. In a general term, soil degradation is defined as the deterioration of the soil function, causing a reduction of its biological potential and productivity that include multiple processes affecting their physical, chemical and biological properties (Imeson, 1998).

In rangelands, physical, and also biological degradation, are mainly related to mismanagement. As an example, excessive trampling by livestock usually produces an increase in the bulk density of the soil and a decrease in porosity, significantly reducing the water holding capacity and infiltration rate, finally affecting the pasture productivity, degree of soil cover and soil vulnerability. (Gamougoun et al., 1984; Mulholland and Fullen, 1991). In this respect, the degree of ground cover is known to be one of the key elements in reflecting the soil condition and protecting soil from degradation.

The amount and type of vegetation cover strongly depends on climate, topographic factors, lithology and human factors such as the history of land use. The reduction of vegetative biomass and degree of soil cover in rangelands due to grazing causes a decrease in the interception of rainfall, an increase in the percentage of bare soil and, consequently, less

* Corresponding author

protection against the direct impact of raindrops on the soil surface. Indirectly the reduction of biomass, with the consequent reduction of soil organic matter, leads to the physical degradation of the soil, negatively affecting infiltration capacity and water retention properties, and increasing the amount of surface runoff (Imeson, 1998).

In the Mediterranean regions, the existence of long summer drought periods coupled with the intense autumn rains, are likely to cause considerable loss of soil material by rainfall erosion.

Therefore, rate of uncovered versus covered soil in rangelands is considered as an indicator of soil quality. Quantification of uncovered or bare soil can be done in different ways and at different spatiotemporal scales. Quantifying bare soil surface and its temporal dynamics by field work can be a highly demanding activity in terms of both time and economic costs for medium-to-large scale areas, such as rangeland farms. Using satellite images overestimates or underestimates bare soil surfaces at farm level. This leads to the need for exploring other methods for the quantification of bare soil surfaces.

Among the objectives of this study are: (1) to test a set of available techniques and methods for quantifying bare soil surfaces, (2) to analyse the possibilities of orthophotos for land cover analysis and the detection of land degradation phenomena, (3) to compare the efficiency of different methods of image classification and to improve them in order to monitor changes in vegetation cover and bare soil as a soil health indicator.

Study area

The study was carried out on the 1024 ha farm “Parapuńos de Doña María”, located in Monroy municipalty, Cáceres province, SW Spain (Fig. 1). Located about 24 kilometres northeast of the city of Cáceres, bounded to the south by the Almonte River, one of the main tributaries of the Tagus river. The farm can be considered as representative of the rangelands (dehesas) of the peneplain geomorphological domain of Cáceres province (Spain), characterized by the presence of thin soils, mainly developed on schist and subject to typical Mediterranean climate characteristics. Vegetation is characteristic of the dehesa wooded rangelands systems, with the presence of therophitic pasture communities with more or less dense shrubs and scattered trees *(Quercus rotundifolia)*. The predominant land use is livestock grazing of sheep, cows and pigs.

Overview and references to related work

Many authors have emphasized the role of livestock activities, particularly in overgrazed areas of semiarid climates, with some physical and biological soil degradation indicators such as bare soil occurrence, (Gamougoun et al., 1984; Lavado, 2002; Mulholland, Fullen, 1991; Schnabel et al., 1996; Imeson, 1998; Kosmas et al., 1999; Shakesby et al., 2002; Schnabel, 2003; De Bello, 2006; Taboada, 2007).

In several works, land cover classifications were carried out (Chen et al., 1996; Hansen et al., 2000; Stuckens et al., 2000), and also monitoring of land cover changes (Wegmüller, Werner, 1997). Others tested different classification techniques and methods using diverse material at various scales. (Foody, 2002).

Oruc et al. (2004) compare pixel-based and object-oriented classification methods using Landsat-7 ETM images with 6 spectral bands to classify land cover types in Zonguldak (Turkey). They applied an unsupervised classification method based on the ISODATA algorithm to provide a priori knowledge, completing the final classification by supervised methods.

Carmel and Ronen (1998) classified Mediterranean vegetation in Mt. Meron (Israel) using two panchromatic aerial photographs (dated in 1964 and 1992) by supervised classification methods of three classes (trees, shrubs and herbaceous vegetation (including bare soil)) with a resulting accuracy close to 90% in many cases. Lecerf (2008) determined winter bare soil variability of intensive farming areas with MODIS data (250 metres spatial resolution).Some more recent work, Marpu et al. (2006), Gao et al. (2007), Nussbaum et al. (2008), started to use object-based image analysis methods.

Material and Methods

Unsupervised classification

Unsupervised classification uses only data from the image pixel spectral characteristics, performed by assigning classes to series of homogenous pixel groups called clusters. The first step in an unsupervised classification is the creation of groups or clusters using a multivariate technique for grouping or clustering, similar to factor analysis used to classify multiband images, based on ISODATA algorithm (Iterative Self-Organizing Data Analysis Technique). The selection of centroids and the allocation of pixels to them are made by k-means clustering algorithm (Mac Queen, 1967). The resulting classified image is computed using the maximum likelihood of each pixel of belonging to the created classes. Due to its spectral characteristics (high values of RGB and brightness digital levels), bare soil pixels were always classified according to the statistical properties of the last class in all the classifications performed. 42 classifications were performed, using different band combinations for each image (RGB, brightness and RGB+brightness) and different settings of the k-means algorithm from 2 to 15 predefined classes. The better classifications for each band combination were afterward selected on the premise of maximizing the minimum statistical distance among classes and reclassified into two classes, covered and uncovered (bare soil) surfaces.

Supervised classification

Supervised classification adds some additional information provided by the user concerning pixels or group of pixels of known certainty about its correspondence with the cover classes. Those pixels are used as training areas in the process of classification. The remaining pixels are assigned to the most similar class through different algorithms, such as minimum distance, maximum likelihood and Mahalanobis distance.

Training field sites were created by using 300 random points with known values of soil cover (250 and 50 sites for covered and uncovered areas respectively) used to generate the spectral signatures of each class with ERDAS® software.

Object-based classification

In this method of classification the focus of analysis are objects created on the image and not the pixel values. Object-oriented classifications take into account important aspects of image objects such as shape, texture and spectral information, amongst others (Arroyo et al., 2005). One of the advantages of this type of analysis is the close relationship between real objects and the created image objects, which enhances the value of the final classification (Benz et al., 2004).

Objects are created by a process of image segmentation done with the multi-resolution algorithm implemented in Definiens® software (Marpu et al., 2006).

125 training segments or image objects were selected at random places to train the classification algorithm based on the minimum distance or nearest neighbour (applied to objects), resulting in the final object-oriented classification

The multi-resolution segmentation algorithm for delimiting image objects is a key aspect in the method (Marpu et al., 2006).

Material

The whole work was carried out on the basis of three orthorrectified aerial photographs: The first one of 0.4 m pixel size, dated on July 2002 taken *ex profeso*. A second image of 0.5 metres pixel size (digitally resampled to 0.4 m) was taken on April 2006 and belongs to the SIGPAC project (Geographical Information System of Agricultural Plots from Spanish Ministry of Environment). A third 1 metre pixel (resampled to 0.4 metres) panchromatic image dated on February 1998 was taken from the Spanish Olive tree Geographical Information System *(SIG Oleícola Español)* carried out by the former Spanish Ministry of Agriculture, Fisheries and Food.

All the significant artificial elements as human constructions and water bodies were masked and removed from the original images in order to avoid interference with the representativeness of the results. For the two RGB images (2002 and 2006) the brightness component was calculated and for the panchromatic image only one layer was used for every classification method.

The software used were ArcGis®, Erdas® and Definiens® for unsupervised, supervised and object-based classifications, respectively.

Validation of the classification results

To validate the results, 1000 random points were generated over the farm and their cover characteristics thoroughly checked on the screen based on the prior knowledge of the study area characteristics. The statistical algorithms used for measuring the classification accuracy were contingency matrices, which show the percentage accuracy, and the Area Under the ROC Curve (Receiver Operating Characteristic, AUC), based on values of sensitivity and specificity that predict the chances of success in the decision to assign a pixel as uncovered or covered soil. AUC can be interpreted as an average sensitivity and specificity over all possible values of specificity and sensitivity (López, Pita, 1998), the values of the AUC ranges from 0.5 (complete randomness) to 1 (perfect discrimination).

Results and Discusion

Unsupervised classification

The best results were obtained from a combination of the RGB image + brightness component, classifying them to first obtain the optimum number of classes (k) and afterwards carrying out a reclassification process in order to obtain two final classes as described in methods. For spring 2006 the optimum number of classes was 6; 5 in the case of summer 2002 and 9 in winter 1998. The validation procedure showed high accuracy levels: 93.4% for the spring image, 93.1% for the summer one and 98.7% for the winter one (Table 1). The Area Under the Curve (AUC) from the ROC validation technique amounted to 0.94, 0.91 and 0.99 for spring, summer and winter respectively (Table 2). 9.10% of the study area was

Table 1. Percentage of accuracy of unsupervised classification using 1000 field work random validation points

k	1998	2002			2006		
	PAN	RGB	BR	RGB+BR	RGB	BR	RGB+BR
2	50.1	56.5	56.5	56.9	41.6	39.8	41.2
3	75.4	75.5	74.9	71.2	61.5	60.8	61.2
4	86.5	87.4	86.5	85.0	81.6	80.6	81.5
5	91.4	92.9	92.8	93.1	88.5	87.3	88.3
6	94.6	95.1	94.5	94.7	93.8	93.0	93.4
7	97.4	95.0	95.1	94.8	96.3	95.9	96.2
8	98.1	94.3	95.2	94.8	97.2	96.8	97.1
9	98.7	94.0	94.9	94.2	96.4	96.1	96.3
10	99.4	93.3	94.5	93.9	96.2	95.6	95.9
11	98.8	92.4	94.0	92.8	95.6	95.5	95.7
12	98.8	92.4	93.6	92.5	95.6	95.4	95.5
13	98.6	92.4	93.2	92.2	95.4	95.4	95.6
14	98.2	92.1	92.8	92.2	95.4	95.4	95.5
15	98.2	92.1	92.8	91.8	95.3	95.1	95.3

k: number of classes; PAN: Panchromatic image; RGB: RGB Image; BR: Brightness component of the RGB image; RGB + BR: Combination of RGB image with its brightness component

Table 2. Area Under the ROC Curve (AUC) of unsupervised classification

k	1998	2002			2006		
	PAN	RGB	BR	RGB+BR	RGB	BR	RGB+BR
2	0.705	0.734	0.734	0.737	0.676	0.666	0.673
3	0.850	0.847	0.843	0.821	0.782	0.782	0.780
4	0.915	0.906	0.905	0.900	0.893	0.888	0.893
5	0.944	0.912	0.923	0.913	0.927	0.921	0.926
6	0.963	0.903	0.901	0.900	0.939	0.939	0.941
7	0.979	0.873	0.887	0.885	0.931	0.928	0.930
8	0.983	0.848	0.876	0.865	0.891	0.889	0.891
9	0.987	0.831	0.859	0.843	0.838	0.841	0.837
10	0.982	0.811	0.847	0.828	0.819	0.798	0.808
11	0.962	0.785	0.831	0.797	0.789	0.793	0.794
12	0.962	0.785	0.819	0.788	0.789	0.783	0.784
13	0.955	0.785	0.808	0.799	0.774	0.779	0.780
14	0.941	0.777	0.797	0.799	0.770	0.774	0.775
15	0.941	0.777	0.797	0.768	0.765	0.755	0.765

k, PAN, RGB, BR, RGB+BR – see Table 1

estimated as bare soil in summer, 3.45 % in spring and 3.38 % in winter. 77.8% of the farm surface was permanently covered, and a 1.58% was always bare. Approximately 20% of the surface changed from covered to bare or vice versa (Fig. 2).

Supervised classification

For spring 2006 the best results were obtained using a minimum distance or nearest neighbour algorithm, in the case of summer 2002 by maximum likelihood and in winter 1998 by Mahalanobis distance. The validation procedure showed high accuracy levels: 94.67% for the spring image, 96.43% for the summer one and 97.10% for the winter one and AUC values of 0.905, 0.880 and 0.952 for spring, summer and winter respectively (Table 3). 12.9 % of the study area was estimated as bare soil in summer, 1.05% in spring and 9.99% in winter. 66.3% of the farm surface was permanently covered, and a 1.99 % was always bare. Approximately 30% of the surface changed from covered to bare or vice versa (Fig. 3).

Table 3. Percentage of accuracy (% acc.) and Area Under the ROC Curve (AUC) of supervised classification

	1998		2002		2006	
	% acc.	AUC	% acc.	AUC	% acc.	AUC
MLK	96.30	0.950	96.43	0.880	93.70	0.898
MD	82.50	0.886	81.86	0.874	94.67	0.905
MHD	97.10	0.952	91.71	0.874	85.00	0.837

MLK: Maximum likelihood; MD: Minimum distance; MHD: Mahalanobis distance

Object-based classification

The validation procedure also showed high accuracy levels: 95.0% for the spring image, 90.9% for the summer one and 94.8% for the winter one. The Area Under the Curve (AUC) from the ROC validation technique amounted to 0.834, 0.900 and 0.853 for spring, summer and winter respectively. 9.50% of the study area was estimated as bare soil in summer, 1.42% in spring and 2.90% in winter. 81.43% of the farm surface was permanently covered, and 1.19% was always bare. Approximately 20% of the surface changed from covered to bare or vice versa (Fig. 4).

Discussion and conclusions

We compare classification methods based on pixel and objects as well as other authors (Oruc et al., 2004; Gao et al., 2007), who used classification algorithms as the Mahalanobis distance. Other methods could also be considered, such as fuzzy logic or SeaTH algorithms (Benz et al., 2004; Nussbaum et al., 2008). All the classification methods tested proved to be appropriate for bare soil surface quantification in the study area with more than 90% of accuracy in many cases, in agreement with results obtained by other authors (Oruc et al., 2004; Nussbaum et al., 2008).

Supervised classification methods do not improve the results obtained by the unsupervised one; this is particularly relevant by avoiding time-consuming labor for selecting supervised training areas. Object-based methods improved the "realism" of the resulting image by significantly removing the "salt and pepper" effect.

Comparing images from different dates, lead to a better description of the dynamics of covered and uncovered areas as well as to the appropriate identification of changing surfaces as spring or winter cultivated areas, places where the shrubs has been cleared or overgrazed areas in summer, particularly relevant in terms of soil degradation.

Acknowledgements

The authors wish to thank to Junta de Extremadura for financially supporting this work by the Research Project PRI06A281 Soil Degradation Indicators in Rangelands.

References

Arroyo Méndez L.A., Cocero Matesanz D., Manzanera de la Vega J.A., García Montero L.G., Pascual Castaño C., 2005: El empleo de clasificadores de contexto para la obtención de cartografía en la interfase urbano forestal. *Geofocus*, 5: 115-128.

Benz U.C., Hofmann P., Willhauck G., Lingenfelder I., Heynen M., 2004: Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *Photogrammetry & Remote Sensing*, 58: 239-258.

Carmel Y., Ronen K., 1998: Computerized classification of Mediterranean vegetation using panchromatic aerial photographs. *Journal of Vegetation Science*, 9: 445-454.

Chen K.S., Huang W.P., Tsay D.H., Amar F., 1996: Classification of multifrequency polarimetric SAR imagery using a dynamic learning neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 34 (3): 814-820.

De Bello F., 2006: Consecuencias de cambios de presión ganadera sobre la estructura de la vegetación a lo largo de gradientes climáticos. *Ecosistemas*, 15 (2): 106-112.

Foody G.M., 2002: Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80:. 185-201.

Gamougoun N.D., Smith R.P., Wood K., Pieper R.D., 1984: Soil vegetation and hydrologic response to grazing management at Fort Stanton, New Mexico. *Journal of Range Management*, 37: 538-541.

Gao Y., Mas J.F., Niemeyer I., Marpu P.R., Palacio J.L., 2007: Object-based image analysis for mapping land-cover in a forest area. Spatial Data Quality 2007.

Hansen M.C., DeFries R.S., Townshend J.R.G., Sohlberg R., 2000: Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21 (6-7): 1331-1364.

Imeson A.C., 1998: Una vía de ataque eco-geomorfológica al problema de la degradación y erosión del suelo. [In:] Desertificación en Europa, MOPU, Madrid, Spain, pp. 161-181.

Kosmas C., Kirkby M.J., Geeson N., 1999: Manual on key indicators of desertification and mapping environmentally sensitive areas to desertification. The Medalus project Mediterranean desertification and land use. European Commission, Brussels, Belgium.

Lavado Contador J.F., Schnabel S., Trenado Ordóñez R., 2002: Relación entre cubierta vegetal y humedad edáfica bajo diferentes usos de suelo del suroeste peninsular. [In:] Pérez-González A., Vegas J., Machado M.J. (eds) Aportaciones a la Geomorfología de España en el inicio del tercer milenio, pp. 283-291.

Lecerf R., 2008: Detection and analysis of winter bare soils variability in intensive farming areas with medium resolution images. Symposium Spatial landscape modelling: from dynamic approaches to functional evaluations. Toulouse. France.

López de Ullibarri Galparsoro I., Pita Fernández S., 1998: Curvas ROC. *Cadernos de Atención Primaria*, 5 (4): 229-235.

Mac Queen J., 1967: Some methods for classification and analysis of multivariate observations. [In:] Le Cam L.M., Neyman J. (eds.) Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. University of California Press, Berkeley, California, United States, pp. 281-297.

Marpu P.R., Gloaguen R., Niemeyer I., 2006: Evaluation of the efficiency of object-based classification in the identification of geological structures. Case study: Extraction of the morphology of the normal faults. Proceedings IEEE International Geosciences and Remote Sensing. IGARSS'06. Denver. United States.

Mulholland B., Fullen M.A., 1991: Cattle trampling and soil compaction on loamy sands. Soil Use and Management, 7: 189-193.

Nussbaum S., Niemeyer I., Canty M.J., 2008: SeaTH – A new tool for automated feature extraction in the context of object-based image analysis. [In:] Nussbaum S., Menz G. (eds.) Object-based image analysis and treaty verification. New approaches in Remote Sensing – Applied to nuclear facilities in Iran. Springer, Netherlands.

Oruc M., Marangoz A.M., Buyuksalih G., 2004: Comparison of pixel-based and object-oriented classification approaches using Landsat-7 ETM spectral bands. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 35 (4): 1118-1122.

Schnabel S., Gómez Amelia D., Bernet Herguijuela R., 1996: La pérdida de suelo y su relación con la cubierta vegetal en una zona de dehesa. [In:] Campesino A.J., Velasco C. (eds) Portugal-Espańa: Ordenación territorial del suroeste comunitario, Universidad de Extremadura, Cáceres, Spain, pp. 195-206.

Schnabel S., 2003: Variabilidad espacio-temporal de la pérdida de suelo en áreas con aprovechamiento silvopastoril. [In:] Bienes R., Márquez M.J. (eds.) Control de la erosión y degradación del suelo, Instituto Madrileńo de Investigación Agraria y Alimentaria, Alcalá de Henares, Madrid, Spain, pp. 475-478.

Shakesby R.A., Coelho C.O.A., Schnabel S., Keizer J.J., Clarke M.A., Lavado Contador J.F., Walsh R.P.D., Ferreira A.J.D., Doerr S.H., 2002: A ranking methodology for assessing relative erosion risk and its application to dehesas and montados in Spain and Portugal. *Land Degradation & Development*, 13: 129-140.

Stuckens J., Coppin P.R., Bauer M.E., 2000: Integrating contextual information with per-pixel classification for improved land cover class. *Remote Sensing of Environment*, 71(3): 282-296.

Taboada M.A., 2007: Efectos del pisoteo y pastoreo animal sobre suelos en siembra directa. 4ş Simposio de Ganadería en Siembra Directa, Aapresid, Potrero de los Funes, San Luis, Argentina, pp. 71-83.

Wegmüller U., Werner C., 1997: Retrieval of vegetation parameters with SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1): 18-24.

Streszczenie

Pomiar obszarów odsłoniętej gleby na pastwiskach jest bardzo ważny, gdyż silny wpływ na stopień erozji gleby mają jej podatność na degradację i stan zdrowia. Szczególnie w ubogich w roślinność obszarach śródziemnomorskich, na początku sezonu deszczowego, duża intensywność wypasu zwierzyny hodowlanej może prowadzić do ogołocenia i degradacji gleby, zwłaszcza w miejscach stale odsłoniętych i tych o mniejszym udziale roślinności trawiastej. Zatem ważnym wydaje się opracowanie metody umożliwiającej identyfikację oraz kwantyfikację obszarów odsłoniętej gleby oraz opis ich dynamiki czasowej.

W artykule przedstawiono metodę służącą do kwantyfikacji obszarów odsłoniętej gleby i analizy jej czasoprzestrzennej dynamiki, przy użyciu ortofotomap lotniczych. Ortofotomapy z okresów letniego, wiosennego i zimowego przedstawiają farmę położoną na mało zalesionym obszarze, z dużą ilością terenów wypasu (dehes), typowym dla południowo-zachodniej części Półwyspu Iberyjskiego. Do analizy tych ortofotomap zastosowano algorytmy klasyfikacji obrazu.

W badaniu wykorzystano obraz panchromatyczny z zimy 1998, dwa trójkanałowe obrazy (RGB) zarejestrowane latem 2002 i wiosną 2006 wraz z komponentem jasności, wszystkie z pikselem o wymiarze 0.4 m. W celu określenia obszarów odsłoniętej gleby zdjęcia zostały sklasyfikowane przy użyciu klasyfikacji nienadzorowanej, nadzorowanej oraz obiektowej.

Porównano wyniki uzyskane różnymi metodami. Stopień dokładności klasyfikacji był bardzo zbliżony dla wszystkich porównywanych metod (powyżej 90%). Wykorzystując najlepszą metodę klasyfikacji, określono dynamikę w obszarach o glebie odsłoniętej i porośniętej jako funkcję czasu, analizując zmiany powstałe na obszarach zarówno porośniętych, jak i nieporośniętych.

Manuel Pulido Fernández
mapuliod@unex.es

Dr. J.F. Lavado Contador, Prof. UEX
frlavado@unex.es

Dr. Álvaro Gómez Gutiérrez, Prof. UEX
alvgo@unex.es

Francisco Javier Lozano Parra,
fjavierlp@hotmail.com

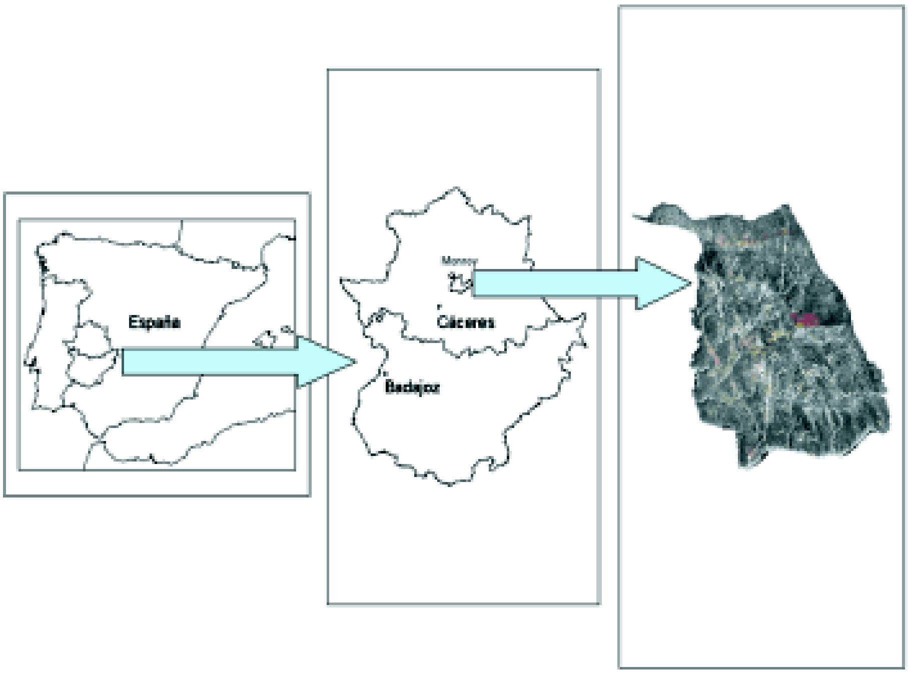

Figure 1. Location of study area

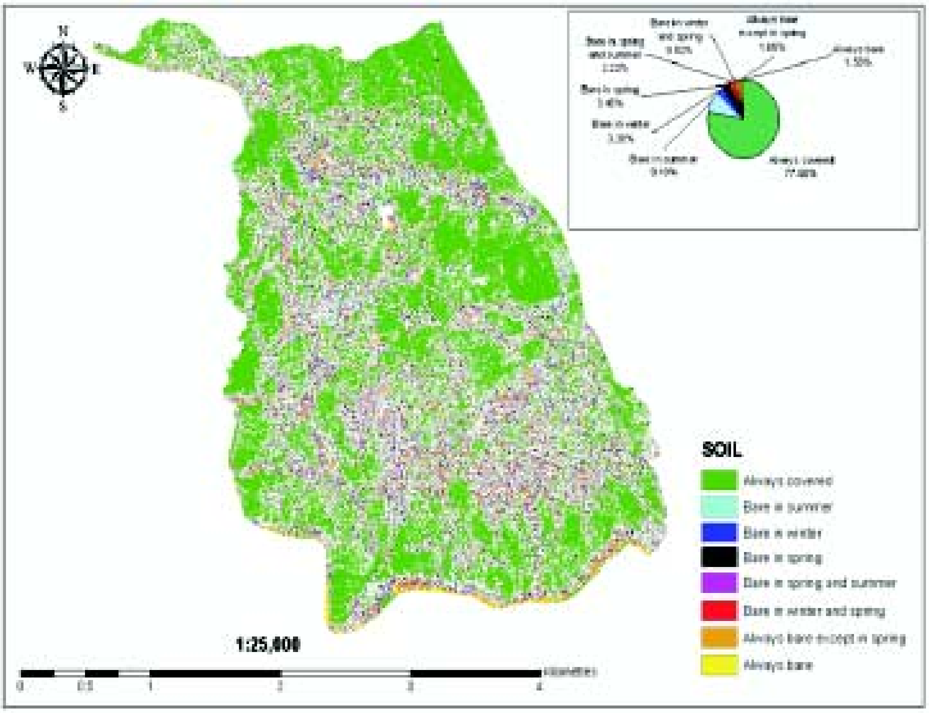

Figure 2. Bare soil and its seasonal variation using the unsupervised classification method

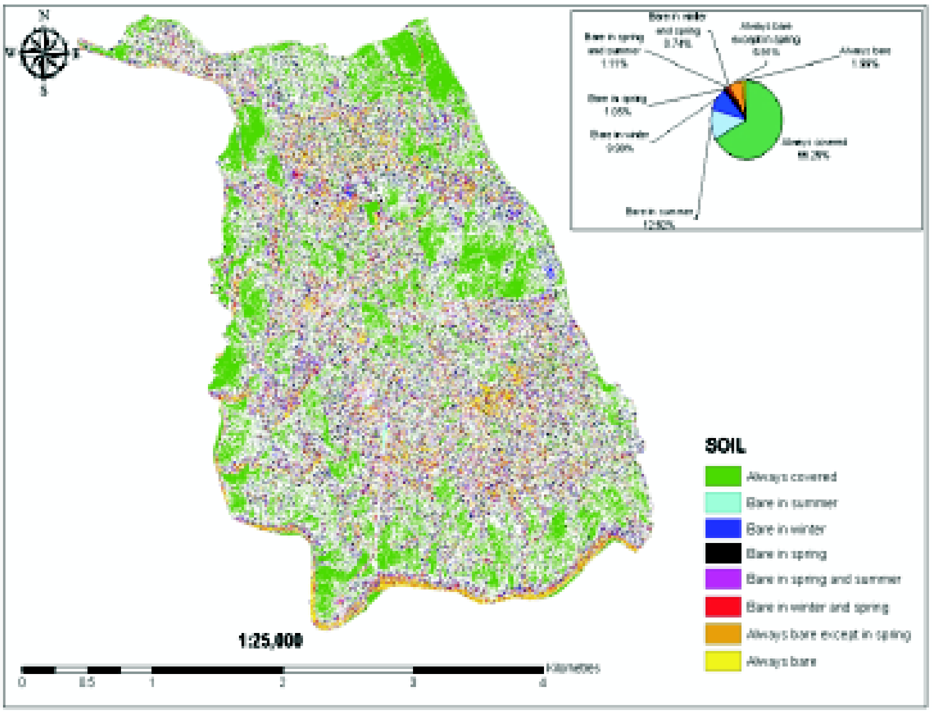

Figure 3. Bare soil and its seasonal variation using the supervised classification method

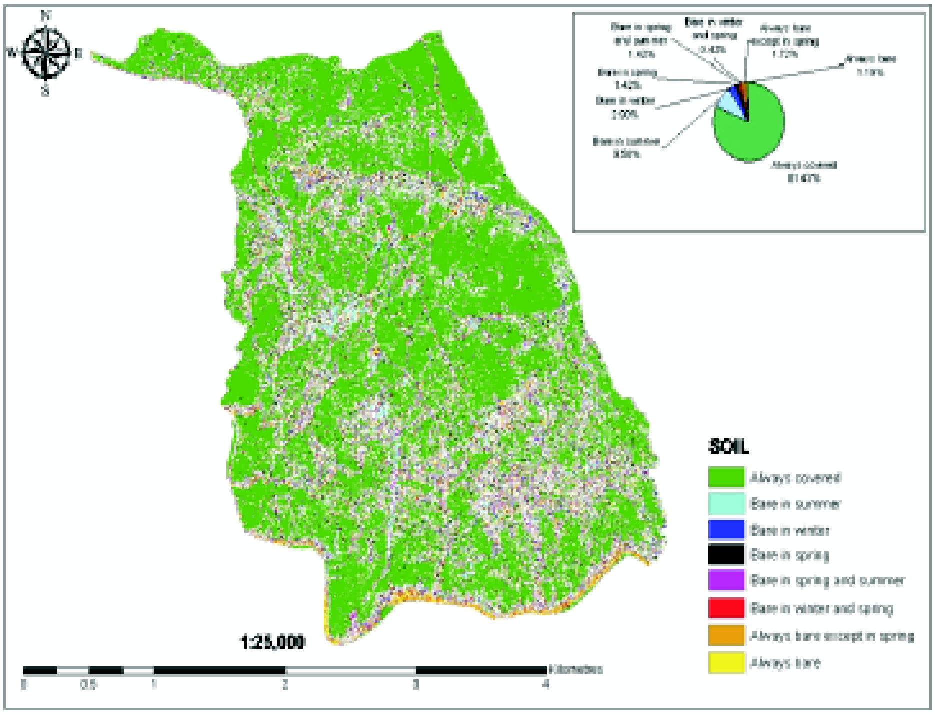

Figure 4. Bare soil and its seasonal variation using the object-based classification method