CONTINUOUS CLASSIFICATION OF NATURAL VEGETATION IN A MEDITERRANEAN ECOSYSTEM BY LINEAR UNMIXING METHODS APPLIED TO MULTITEMPORAL TM DATA.

A. Canesi, G. Pennati
DSM / SMA - CISE SpA, 20090 Segrate (MI) - (Italy)
E-Mail 0987penn -*AT*- s1.cise.it
A. Galli
University of Ancona - Dipartimento di biotecnologie agrarie ed ambientali
M. D'Angelo
University of Sassari
R. Ambrogi
ENEL SpA - CRAM - Milan

Commission VII

Keywords: *Pixel* spectral unmixing, continuous *classification*, Mediterranean *vegetation*, *multitemporal* satellite data.

Abstract

The aim of this paper is to describe the classification methodology used to analyze remote sensing multispectral data of an area located in north-west Sardinia (Italy) characterized by various natural vegetation species association continuously distributed with different canopy closure and to present the results obtained by the linear unmixing approach in processing the satellite data.
A careful analysis of available information and data suggested to define two levels of classification: first a traditional statistical technique was used to determine the main cover classes, then a refinement of vegetation cover classes was obtained by applying a linear mixing model suited for the specific case.
Results obtained in vegetation classification by linear unmixing are presented and for each vegetation class the estimated proportions are shown. The whole map production process was performed using the EarthView(*) remote sensing image analysis package.

Introduction

Traditional classification techniques often fail in identifying land cover classes when complex landscapes are considered. In such cases the choice of an appropriate classification methodology, based on new methods (fuzzy sets, mixture models) becomes essential to obtain a precise land cover map of the study area. This approach oriented to vegetation continuum classification was necessary both for the pixel mixing situation and since this land cover classification is being used as an input parameter in a model for fire hazard assessment in Sardinia district.

1. MATERIALS

1.1 Study area and nomenclature definition

The study area represents a significant part of N-W Sardinia, extended in N-S direction between the town of Sassari (SS) and the town of Macomer (NU), with a surface of about 3,200 square kilometers.
It is characterized by a complex landscape,where agricultural fields and natural grazed or wooded lands are mixed together. His typical climate is temperate-warm, with annual average temperatures between 15 and 16.9 degrees centigrade and a level of precipitation from a minimum of 500 mm/year to a maximum of 800 mm/year. The geomorphology is characterized by the hilly landscape, while the alluvial plains are restricted to comparatively narrow surfaces. The lithology embraces old alluvial sediments, calcareous-arenaceous formations, volcanic formations.

The first step of the work was dedicated to analyze and to define the main natural and human elements of this complex ecosystem. We noted that human activities (above all deforestation, farming and grazing) have determined notable alterations of the natural landscape. Moreover the fire, a kind of endemic event in Sardinia, becomes more and more a powerful agent of landscape transformation.
Bibliographical references and aerial photographs interpretation, with some coverage parameters collected by field survey, were used to define the land cover legend of N-W Sardinia.
This legend has a hierarchical structure and it includes the main extra-agricultural classes, the most important agricultural classes, with particular regard to their biological cycles, the physiognomic groups of the natural vegetation.
At the first level it includes: 1. *artificial surfaces*, 2. *open spaces with little or no vegetation*, 3. *water bodies* , 4. *agricultural areas*, 5. *pasture and shrubs*, 6. *forested areas*.
At the second level the legend includes: 1.1*continuous urbanized areas* (density of buildings greater than 80%), 1.2 *discontinuous urbanized areas* (low density of buildings), 2.1 *bare rock and quarries*, 2.2 *burnt areas*, 3.1 *water bodies* (rivers, streams and channels), 3.2 *inland and coastal wetlands*, 4.1 *non-irrigated crops*, 4.2 *irrigated crops*, 4.3 *olives groves/vineyards/fruit trees*, 5.1 *herbaceous pastures* (density of shrubs less than 35%), 5.2 *low density shrubs* (density of shrubs between 35% and 70% with main presence *of Pistaccia lentiscus,Cistus spp.* et al.)), 5.3 *high density shrubs* (density of *Quercus* spp. shrubs greater than 7O%, density of trees less than 30%), 6.1 *sclerophyllous low density forests* (*Quercus ilex and Quercus suber*, density between 35% and 70%), 6.2 *sclerophyllous high density forests* (density greater than 70%), 6.3 *broad-leaved low density forests* (*Quercus*

(*) *EarthView is a Copyright by CISE Spa and ENEL Spa* *pubescens*, density between 35% and 70%), 6.4 *broad-leaved high density forests* (density greater than 70%), 6.5 *coniferous reforestation* , 6.6 *broad-leaved reforestation..*

The definition of the endmembers and vegetation continuum legend has been obtained starting from the nomenclature described above; this nomenclature has been further refined and integrated by iterative spectral analysis of the multitemporal dataset, evaluation of the classification results and assessment of the errors occurred in the mixing model analysis.

1.2 Image data

Two Landsat 5 - Thematic Mapper data quarter scenes (track 193, frame 92) were acquired to obtain spectral and multitemporal discrimination of the vegetation canopies; in order to get discrimination capability in evergreen and deciduous quercus species, as well as for pasture classes and shrubs, a mid spring date and a deep summer date were chosen: 31st March and 22nd August 1994.

About half of the quarter scene (80Km X 40Km) was processed to obtain the land use and vegetation map. About 20% of this area resulted recently burnt in the August image, reporting highest fire events occurred at the beginning of that month.

2. METHODS

2.1 Geometric and radiometric preprocessing

A first geometric rectification process was made in order to correlate the two Landsat scenes each other; due to the almost identical flight direction of the satellite overpasses the registration was performed with a simple 1st order polynomial transformation with a rotational coefficient very close to 0.

The projection on the UTM 32 coordinate system was made after the classification process in order to limit the resampling effects on radiometry and on the multitemporal registration.

During this last geometric rectification process no digital elevation model correction was applied due to the smooth landscape and low elevation range occurring in the study area; the estimated maximum planimetric error introduced by this approximation for such an elevation range is less than 20 mt, which is comparable to the planimetric location error obtained in the control points and test points used for the UTM projection transformation. This error is compatible with 1:100000 scale products geometric specifications.

Despite the low elevation range, the typical Mediterranean landscape is often characterized by steep slopes and discontinuities that create small but serious topographic effects that cannot be recovered easily without an accurate DEM (generally unavailable or extremely expensive) and a comparable spatial resolution in the imagery. To recover these effects spectral normalization has been applied to the two Landsat scenes by means of chromatic components computed starting from each TM channel reflectance.

The chromatic transformation, while giving generally gøod results in topographic effects normalization, maintains the linear relation originally existing among the total pixel reflectance and the sum of area-weighted reflectances for each endmember class, thus conserving the necessary hypothesis to allow the application of linear unmixing methods.

2.2 Training sets, endmembers and channel selection

Within the study area, a square of 500 x 500 pixels was selected. This area is representative of the most important cover classes we put in the described legend. With the help of aerial phothograps, showing the same zone, a reasonable set of polygons were selected for each classes on a RGB 432 composite of the Thematic Mapper august scene. In this way, a overall number of 73 polygons was defined. The standard statistical parameters, referred to each training-set, were then extracted from all original and all chromatic bands. Using different kinds of digital tools, e.g. scatterograms and plots, it was analyzed the spectral behavior of each cover classes within the available spectral space. This spectral analysis allowed to estimate the discrimination between the cover classes considered in the legend and it was very useful for the choice of the better set of bands to use in statistical classification. It was observed a good spectral separation between the land use-land cover classes hereinafter indicated.

The definition of endmembers, which represents a crucial point in the unmixing method, has been driven starting from the analysis of typical vegetation components, the ground truth plot location on the imagery and the subsequent spectral distribution analysis using tools available with EarthView. While running this activity various interesting aspects arose among which the presence of further endmembers and the usefulness of visualization of spectral spaces in endmembers choice.

In our experiment, after this analysis, five classes were indicated as endmembers: three species of woods (quercus ilex, quercus pubescens, quercus suber), a typical shrub association of secondary species (*Cistus spp, Pinaccia Lentiscus*) and an open pasture class were chosen as endmembers, since they represented the main vegetation species in this study area; other typical vegetation species, even present in the Sardinia district, (e.g. coastline maquis), were excluded for their extremely scarce presence in the study area.

2.2 Land Use classification

The general land -use classification has been driven by the following steps: a first raw classification allowed to separate water bodies, bare rock outcrops and permanently bare soils by multitemporal NDVI thresholding. Secondly, manual identification of urban areas, recently burnt areas and horticultural areas characterized by small-parcels was done. Finally the Mahalanobis classifier was applied to classify the remaining areas in the following other classes: quercus woods, shrubs, pasture, 3 crop types and coniferous reforestation.

2.3 Linear unmixing classification

The classification of natural vegetation continuum has been obtained using Linear Mixing Models. Linear Mixing Models allow to classify multispectral satellite images at subpixel level. For each pixel proportions between a number of different materials are determined recalling that the spectral radiance recorded by a satellite results from a mixture of the energy reflected by all surface types within the instantaneous field of view of the satellite sensor. The main problem with these methods is a correct identification and characterization of a reasonably small and representative number of surfaces whose mixture can reproduce the spectral distribution observed. It's

easy to understand how crucial the choice of endmembers is to ensure the accuracy of the classification.
The interesting issues in this experiment are both the multitemporal processing of Thematic Mapper data in the linear unmixing, the spectral bands used and the relatively wide range of vegetation cover classes discriminated within the continuum.
The multitemporal dataset was necessary both for discrimination purposes and for the mathematical constraint given by the linear unmixing method which requires a number of channels (equations) greater than the number of reflectances of endmember classes (unknowns). The multitemporal dataset can be in principle considered a way to extend the number of statistically uncorrelated bands available within a single date image while maintaining the linear weighted reflectance hypothesis on the pixel reflectance, unless major changes occur in the pixel land cover composition (for instance due to clear cuts or fires) between the two dates.

3. RESULTS

3.1 Ground truth set

The ground truth dataset has been organized in several plots identified by aerial photography and existing forest maps; the uncertainty arose in this first part of the ground set definition was removed after a field survey in uncertain areas.
To assess the potential of the proposed technique in estimating the percentages of vegetation canopy cover, for each of the endmembers a number of plots (about 40) were identified with 10 different mixing degrees ranging from almost pure (>70% canopy cover) to bare pasture in various species compositions (See Table 2). The overall size of the test areas was about 2000 TM pixels wide.

3.2 Vegetation continuum

The result summary of results in mixing component estimation is given in Table 1 while in Table 2 the description of the plots' composition is given: as the name of the vegetation elements and the proportion they were estimated in the mixed pixel (between brackets): in each of the classes, except for the "pure pasture" one, the remaining percentage is intended to be occupied by pasture itself. For each ground plot type representing the 10 categories (the five concerning "pure" species and the 5 concerning mixed areas) the average percentage is computed for all the pixels. In Table 1 the average cover percentage for each endmember (rows) is given for each of the 10 classes (columns).
Linear unmixing residuals have been given for each ground truth class together with percentages of unexplained pixel composition as a report of the accuracy of the continuous classification. Overall unexplained composition percentage was 8.6 (See Table 1); this means that a residual unexplained percentage mainly for the mixed classes remains.
For the 5 endmembers only the cover percentage of the quercus suber was underestimated (52.6%) with respect to the expected result (greater than 70%), giving higher percentage canopy cover to the ilex class; this is probably due to the low spectral separability shown by the quercus suber and quercus ilex classes and to the common presence of a vegetation composition of the two quercus species in the whole Sardinia district.

Another source of error, unavoidable in practice, is given by the need of geometric correction of multitemporal imagery, which causes approximations in pixel alignment since for different satellite overpasses the location of the sensor's sampling grid on the terrain is independent and in general non-overlapping. Resampling methods could in part overcome this problem if they didn't introduce radiometric distortion in the original data. In any case, this fact adversely affects the linear mixing hypothesis whenever extended on multitemporal imagery, while its impact on fraction cover estimation is related to the autocorrelation of the cover class fractions for each training plots and of the spatial distribution of the classes at the sensor's resolution.

4. CONCLUSIONS

A 1:100000 scale land-use map and a continuous map of natural vegetation in a Mediterranean ecosystem (NW Sardinia district - Italy) has been produced using multitemporal analysis of Thematic Mapper data. The linear unmixing method, suitable for mixed pixels, was employed for the analysis of vegetation continuum which mainly characterizes the Mediterranean landscape, ranging from high percentage of canopy closure (>70%) up to open pasture areas.
For each of the TM pixels which were classified in the first step as "vegetated" or abandoned to "pasture" regrowth, the canopy closure percentages for four vegetation types plus the open-pasture one, were obtained.
To enhance the encouraging present results, more effort will be given to consider the influence of shadow and soils in the endmember characterization and to enhance the endmember selection process.
Since the proposed strategy proved to be effective even with small datasets, further analisys will be driven to improve the accuracy of the classification, the estimation of canopy closure percentage and, possibly, to enhance again the endmember selection, by means of new aerial CIR datasets recently acquired.
Within the current year an integration of this methodology and results with AVHRR data will be achieved in order to assess the potential of multiresolution and multitemporal data integration for vegetation classification in the mediterranean ecosystem.

5. REFERENCES

- Arrigoni, P. V., 1968. Fitoclimatologia della Sardegna. Webbia 23, Firenze, Italy.
- Corrias B., Diana-Corrias S, Valsecchi F, 1982. Carta della vegetazione della Nurra di Alghero, CNR PF promozione della qualita' dell'ambiente, Roma Italy.
- Cross, A.M., Settle, J.J., Drake, N., Paivinen, R.T. M. 1991. Subpixel measurement of tropical forest cover using AVHRR data. Int J. of Rem. Sens. Vol 12 No 5, pp 1119-1129.
- Gong, P., Miller, J. R., Spanner, M., 1994. Forest canopy closure from classification and spectral unmixing of scene components - Multisensor evaluation of an open canopy. IEEE on GE vol 32 No 5, pp 1067-1078.
- Idda, L., 1989. Marghine-Planargia. CNR IPRA Centro di sperimentazione agricola e forestale. Roma, Italy.
- Lawson, C. L., Hanson R. J., 1974. Solving Least squares Problems. Englewood Cliffs, NJ.

– Regione Autonoma della Sardegna, 1988. Carta forestale della Sardegna. Istituto Geografico De Agostini, Italy.
– Shimabukuro, Y. E., Smith, J. A., 1991. The least squares mixing models to generate fraction images derived from remote sensing multispectral data. IEEE on GE Vol 29 No 1.pp 16- 20.

	A	B	C	D	E	F	G	H	I	L	
A	**96.70**	33.10	4.98	2.90	4.58	45.37	11.13	47.45	3.76	60.94	**Pure Quercus ilex** (>70%)
B	0.74	**52.66**	0.62	0.66	0.04	2.15	3.13	15.68	0.27	11.37	**Pure Q. suber** (>70%)
C	1.64	10.82	**79.45**	0.62	0	0	1.92	2	1.02	0.27	**Pure Q. pubescens** (>70%)
D	0	0.98	3.33	**77.30**	6.78	0	15.84	9.02	9.92	2.43	**Shrubs** (>70%)
E	0.08	1.84	0.15	14.23	**78.98**	52.04	65.05	17.25	59.60	3.56	**Pure Pasture**
	1.61	0.47	1.09	1.88	1.80	2.82	1.76	1.45	1.53	1.21	Linear Unmixing Residual
	99.1	99.4	88.5	95.7	90.3	99.5	97.0	91.4	74.5	78.5	Sum of components (%)
	0.9	0.6	11.5	4.3	9.7	0.5	3.0	8.6	25.5	21.5	Unexplained (%)
											Average Unexplained: 8.6 (%)

Table 1 - Percentage Composition (**bold typed**=endmembers)

A	**Pure Quercus Ilex (>70%)**
B	**Pure Q. Suber (>70%)**
C	**Pure Quercus Pubescens (>70%)**
D	**Shrubs (Cistus spp, Pinaccia Lentiscus) (>70%)**
E	**Pure Pasture (>70%)**
F	Quercus Ilex (35-70%)
G	Q. Ilex - Shrubs (35-70%)
H	Mix Q. shrubs and maquis (A,B,C,D) (>70%)
I	Q. Ilex - Shrubs (<35%)
L	Mix Q. woods (ilex + suber (>70%)

Table 2 - Ground truth vegetation compositions

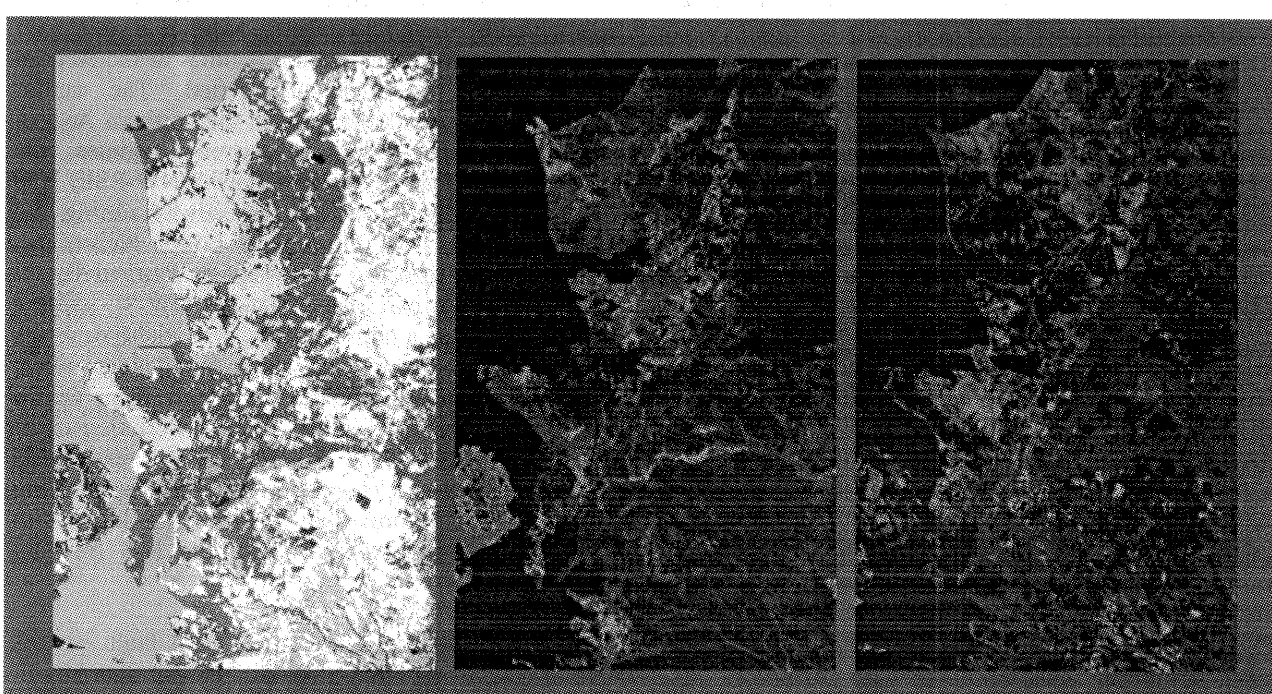

Figure 1 - Portion of 10 X 18 Km within the study area. Left: land use map. Center : fraction image of Quercus spp. Right: fraction image of pasture. Left image obtained by statistical method, center image and right image obtained by linear spectral unmixing.