

# Uso dell'informazione spaziale per il miglioramento della classificazione

C. Fiorentino (\*), C. Tarantino (\*\*), A. Castrignanò (\*), G. Pasquariello (\*\*)

(\*) CRA, Via Celso Ulpiani 5, Bari, Italia

(\*\*) ISSIA – CNR, Via Amendola 122/ D-O, Bari, Italia

## Abstract

L'uso dei dati telerilevati nel monitoraggio del territorio gioca un ruolo importante in molte applicazioni: prevenzione, pianificazione, controllo, ecc.. In questo contesto diventa importante sviluppare metodologie che consentano di produrre mappe tematiche sempre più accurate. A tale scopo, in questo lavoro, è stato realizzato un approccio combinato che integrasse informazione spaziale e spettrale per caratterizzare meglio ciascuna classe tematica. In particolare, l'informazione spaziale, elaborata con l'algoritmo geostatistico dell'*Indicator Kriging*, è stata inglobata in un classificatore statistico spettrale tradizionale di tipo *Maximum Likelihood* applicato a dati telerilevati nell'ottico a media risoluzione.

Questo approccio combinato ha consentito una migliore discriminazione della copertura del suolo rispetto alla metodologia tradizionale.

## Abstract

The use of remote sensed data for monitoring land use has an important role in many applications: prevention, planning, controlling etc.; in this context, in monitoring land cover is it very important to develop methodologies to obtain reliable maps. In order to achieve this objective a possible approach is to combine both "spectral" and "spatial" features to characterize each ground class. In this paper we propose the integration of a spectral classifier for remote sensed data at medium resolution in the optical spectrum, based on a traditional statistical supervised classifier as "Maximum Likelihood", with the spatial information provided by a geostatistical tool, as "Indicator Kriging" algorithm.

Using this combined approach, better results in land cover class discrimination have been obtained and the resulting maps look more homogenous than in the case with the spectral information only.

## Introduzione

Le immagini telerilevate costituiscono un insostituibile supporto nell'ambito del monitoraggio del territorio, quindi, mettere a punto metodologie di analisi sempre più complete è fondamentale se si vogliono produrre mappe accurate. A tale scopo applicare il tradizionale classificatore *Maximum Likelihood* che valuta la prossimità di un pixel ad una classe esclusivamente nello spazio multispettrale, significa trascurare le coordinate spaziali del pixel stesso. In tal modo l'immagine classificata può apparire 'disturbata' dalla presenza di pixel isolati mal classificati, perchè attribuiti ad una classe differente rispetto a quella dei pixels limitrofi.

Una soluzione efficace a questo problema è fornita dalla geostatistica che consente di integrare l'informazione spaziale nella classificazione. In particolare, applicando l'*Indicator Kriging* e interpolando i dati di verità a terra disponibili opportunamente trasformati, è possibile stimare la probabilità locale associata a ciascuna classe tematica. In tal caso si parla di probabilità “a posteriori” perché stimate anche nei punti non campionati dalle verità a terra dell'intorno. Integrare l'informazione spaziale nel classificatore *Maximum Likelihood* è possibile inserendovi valori puntuali di probabilità a priori che sono appunto quelli calcolati con l'*Indicator Kriging*. L'obiettivo di questo lavoro è la messa a punto e validazione di un algoritmo di classificazione supervisionata che integri l'informazione spettrale con quella spaziale.

## Materiali e Metodi

L'area di interesse è l'azienda sperimentale del CRA-Istituto di Zootecnia sita a Foggia, di cui sono note le destinazioni d'uso del suolo. Il data set adoperato è relativo alle sei bande spettrali, nel visibile, vicino e medio infrarosso, alla risoluzione spaziale di 30 metri, del sensore Landsat TM 5 riferito ad un'immagine del 14 maggio 1997. L'uso del suolo è caratterizzato prevalentemente da colture di graminacee e prati adibiti a pascolo, in particolare sono state selezionate 9 classi tematiche: avena granella, avena pascolo, grano duro, orzo, prato naturale, prato pascolo, vecchia-avena, incendio, suolo nudo.

Il classificatore *Maximum Likelihood* affronta in modo puramente statistico il problema della classificazione secondo la teoria Bayesiana che pone il problema della ‘decisione’ in termini probabilistici, assumendo i valori di probabilità noti.

Indichiamo con  $\mathbf{u}$  la posizione spaziale del generico pixel dell'immagine e siano  $s$  i punti  $\mathbf{u}_\alpha$  di cui si conosce con precisione la classe di appartenenza; l'informazione spettrale associata a ciascuno di tali punti sarà allora  $X_b(\mathbf{u}_\alpha)$  con  $b=1,2,\dots,B$  (numero di bande dell'immagine multispettrale) e  $\alpha=1,2,\dots,s$ .

Sia  $\mathbf{X}=\mathbf{X}(\mathbf{u}_\alpha)$  un vettore *random* a  $b$  (elementi, ove  $\mathbf{u}_\alpha$  rappresenta la posizione dei punti di verità a terra campionati e sia  $\{s_k; k=1,\dots,n\}$  una variabile categoriale con  $n$  classi che si escludono vicendevolmente. L'algoritmo *Maximum Likelihood* assegna ogni pixel alla classe a cui è associata la massima probabilità a posteriori:

Si assegna la posizione  $\mathbf{u}$  alla classe  $s_k \Leftrightarrow P(s_k(\mathbf{u})|\mathbf{X}(\mathbf{u}_\alpha)) = \max_s P(s|\mathbf{X}(\mathbf{u}_\alpha))$

ove  $P(s_k|\mathbf{X}(\mathbf{u}_\alpha))$  è la probabilità a posteriori associate alla classe  $s_k$  nella posizione  $\mathbf{u}$ .

Applicando il teorema di Bayes:

$$P(s_k(\mathbf{u}) | \mathbf{X}(\mathbf{u}_\alpha)) = \frac{p(\mathbf{X}(\mathbf{u}_\alpha)|s_k(\mathbf{u}))p(s_k(\mathbf{u}))}{p(\mathbf{X}(\mathbf{u}_\alpha))} \quad [1]$$

ove  $p(\mathbf{X}(\mathbf{u}_\alpha)|s_k(\mathbf{u}))$  è la funzione di distribuzione di probabilità ‘spettrale’ che un *pixel* appartenente alla classe  $s_k$  assuma il valore  $\mathbf{X}$ ;  $p(s_k(\mathbf{u}))$  rappresenta la probabilità a priori associata alla classe  $s_k$  nella posizione  $\mathbf{u}$ , mentre  $p(\mathbf{X}(\mathbf{u}_\alpha))$  è un fattore di normalizzazione

In generale si assume che i pixel di ciascuna classe seguano una distribuzione multivariata normale (Gaussiana); in tal caso è possibile esprimere la funzione di densità di probabilità come segue:

$$p(\mathbf{X}(\mathbf{u}_\alpha)|s_k(\mathbf{u})) = \frac{1}{(2\pi)^{\frac{B}{2}} |\Sigma_k|^{-\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{X}(\mathbf{u}) - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\mathbf{X}(\mathbf{u}) - \boldsymbol{\mu}_k)\right) \quad [2]$$

ove  $\mathbf{X} = \mathbf{X}(\mathbf{u}) = [X_1, X_2, \dots, X_B]$  è il vettore colonna associato ai valori del generico pixel in tutte le B bande dell'immagine, mentre  $\boldsymbol{\mu}_k = [\mu_k^1, \mu_k^2, \dots, \mu_k^B]$  e  $\Sigma_k$  sono rispettivamente il vettore media delle B bande spettrali e la matrice di covarianza (di dimensione B x B) associati alla k-esima classe;  $\Sigma_k^{-1}$  rappresenta la matrice di covarianza inversa,  $|\Sigma_k|$  è il determinante di tale matrice.

La letteratura tradizionale definisce tale probabilità "a priori" indipendente dalla posizione spaziale. L'approccio adottato in questo lavoro si basa sull'idea che la probabilità "a priori" associata a ciascuna classe non sia uguale ovunque, ma dipenda dalla posizione del *pixel* (Castrignanò A. e al., 2000), perciò le probabilità "a priori" sono state sostituite delle probabilità "a posteriori" calcolate utilizzando l'algoritmo geostatistico dell' *Indicator Kriging*

Ciascun pixel dell'immagine corrisponderà ordinatamente al nodo di una griglia, e su cui, applicando l'*Indicator Kriging*, è possibile stimare la probabilità associata a ciascuna classe.

Nei punti campionati  $\mathbf{u}_\alpha$  non vi è incertezza sulla classe di appartenenza, quindi la probabilità a priori vale 1 per la classe osservata, e 0 per le altre n-1 classi.

A ciascun punto  $\mathbf{u}_\alpha$  verrà quindi associato un vettore i cui elementi saranno n variabili indicatrici tali che:

$$\begin{aligned} i(\mathbf{u}_\alpha; s_k) &= 1 & \text{if } \mathbf{u}_\alpha \in s_k & \quad k=1, 2, \dots, n \\ i(\mathbf{u}_\alpha; s_k) &= 0 & \text{altrimenti} \end{aligned} \quad [3]$$

Se  $\mathbf{u}$  è una locazione non campionata, la probabilità ad esso associata per una data classe  $s_{k0}$  può essere stimata mediante il *kriging* sfruttando le variabili indicatrici note in un intorno della locazione considerata. In tal modo è possibile stimare la probabilità  $i^*(\mathbf{u}; s_{k0})$  che ciascuna classe ha di prevalere in un particolare punto:

$$i^*(\mathbf{u}; s_{k0}) = \text{Prob} \{ \mathbf{u} \in s_{k0} | (n) \} \quad [4]$$

dove la notazione (n) rappresenta l'informazione condizionale contenuta nell'intorno di  $\mathbf{u}$ . Un algoritmo adatto a tale stima è il *kriging* ordinario applicato alle variabili indicatrici (Goovaerts P. 2002).

Una misura della diversità delle assegnazioni, al crescere della distanza  $\mathbf{h}$  tra le coppie di punti, è data dal semivariogramma che è calcolato, per ciascuna variabile indicatrice, come segue:

$$\gamma(\mathbf{h}; s_k) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [i(\mathbf{u}_\alpha; s_k) - i(\mathbf{u}_\alpha + \mathbf{h}; s_k)]^2 \quad [5]$$

Ove  $N(\mathbf{h})$  è il numero di coppie di dati separati dal vettore  $\mathbf{h}$ .

Una volta calcolati gli n modelli di semivariogramma associati a ciascuna classe tematica, si è applicato l'*Indicator Kriging* per stimare le probabilità condizionale  $p(s_k | x_{geo})$  che il generico nodo  $\mathbf{u}$  appartenga alla classe  $s_k$  condizionale all'informazione spettrale nei punti  $\mathbf{u}_\alpha$  o di verità a terra. In tal modo si calcolano le probabilità associate al generico punto di coordinate  $\mathbf{u}$ :

$$p(s_k | x_{geo}) = \sum_{\alpha=1}^s \lambda(\mathbf{u}_\alpha; s_k) i(\mathbf{u}_\alpha; s_k) \quad [6]$$

Per ciascuna classe i pesi  $\lambda(\mathbf{u}_\alpha; s_k)$  si ottengono risolvendo il sistema di s+1 equazioni del Kriging Ordinario:

$$\begin{cases} \sum_{\alpha=1}^s \lambda(\mathbf{u}_\beta; s_k) \gamma(\mathbf{u}_\alpha - \mathbf{u}_\beta; s_k) - \mu_i = \gamma(\mathbf{u}_\alpha - \mathbf{u}; s_k) \\ \forall \alpha = 1, \dots, s \\ \sum_{\beta=1}^s \lambda(\mathbf{u}_\beta; s_k) = 1 \end{cases} \quad [7]$$

L'ultima equazione deriva dalla condizione di non distorsione, ossia di uguaglianza tra il valore vero e il valore atteso. Il sistema è ottenuto dalla minimizzazione della varianza dell'errore calcolato come differenza tra il valore stimato ed il valore vero.

In ogni punto  $\mathbf{u}$  le  $n$  probabilità associate a ciascuna classe devono soddisfare le seguenti condizioni:

$$0 \leq p(s_k | x_{geo}) \leq 1 \quad [8]$$

$$\sum_{k=1}^n p(s_k | x_{geo}) = 1. \quad [9]$$

La prima condizione potrebbe non sempre essere soddisfatta poiché alcuni pesi del *kriging* potrebbero essere negativi, pertanto i valori di probabilità negativi sono posti a 0 e quelli maggiori di 1 ad 1. La seconda condizione è realizzata mediante una standardizzazione a posteriori delle probabilità individuali stimate.

## Risultati

In primo luogo sono stati scelti 479 punti secondo una distribuzione casuale che ricoprisse abbastanza uniformemente l'area di interesse. Ciascuno di tali punti è stato attribuito, quindi, alla rispettiva classe tematica, sulla base delle verità a terra fornite dagli operatori dell'azienda.

Tali dati campionati sono stati processati utilizzando l'*Indicator Kriging*, in tal modo si è ottenuta una mappa di probabilità a posteriori relativa a ciascun tematismo.

A titolo di esempio di seguito (fig. 1) sono riportate le mappe relative alle classi: grano duro e prato pascolo.

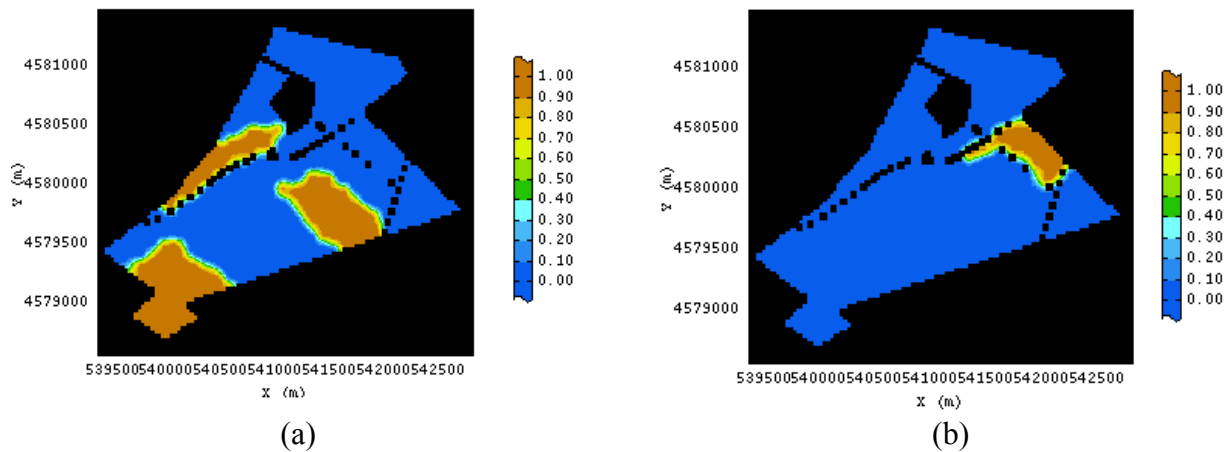


Figura 1. Mappe di probabilità a posteriori ottenute applicando l'“Indicator Kriging” e relative alle classi grano duro (a) e prato pascolo (b).

Tali mappe sono state inviate in input al classificatore *Maximum Likelihood* ed opportunamente integrate nel programma in modo che diventassero probabilità “a posteriori”, puntuali, nell’algoritmo.

Di seguito sono riportate (Fig. 2, 3 ) le mappe, di classificazione supervisionata, ottenute rispettivamente con l’ipotesi di equiprobabilità tra le classi e con le probabilità puntuali elaborate con l’ *Indicator Kriging*.

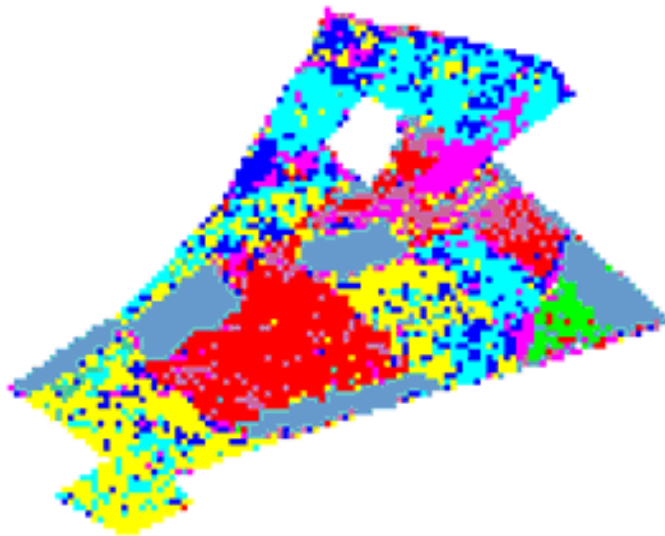


Figura 2. Mappa di classificazione ottenuta nell’ipotesi di equiprobabilità a priori.

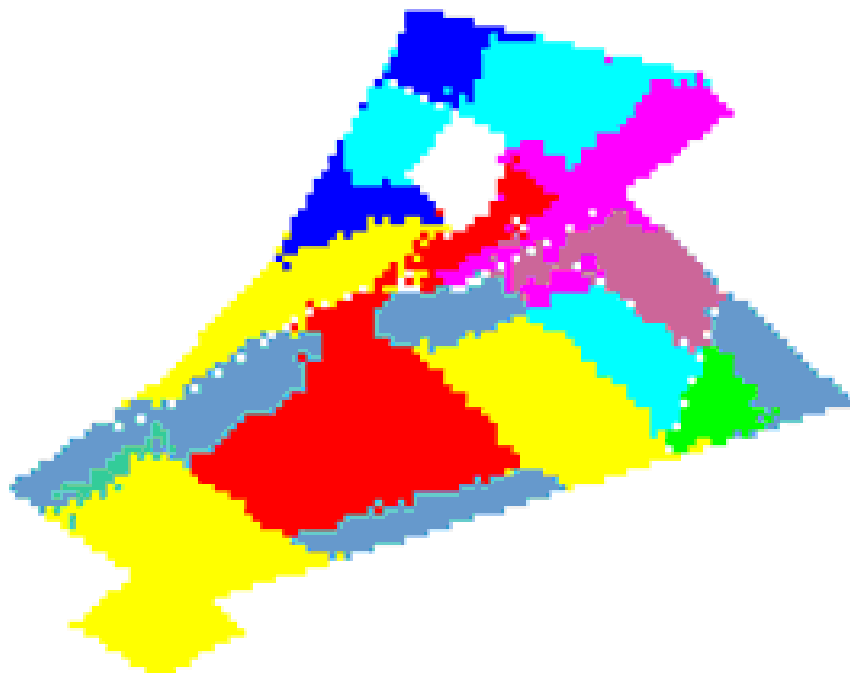


Figura 3. Mappa di classificazione ottenuta con le probabilità “a posteriori” elaborate con l’”Indicator Kriging”.

Per la validazione dei risultati (in entrambi i casi in analisi) è stato utilizzato un *training set* di 300 punti ed un set di test di 200 punti differenti da quelli del *data set* iniziale.

Dal confronto tra le due mappe nelle figure precedenti è evidente come i risultati più affidabili sono stati ottenuti inserendo nell'algoritmo di classificazione le probabilità "a posteriori" variabili localmente per le classi individuate nella scena, piuttosto che probabilità "a priori" uniformi spazialmente per ciascuna classe.

Integrando informazione spaziale e spettrale della scena si nota, come si evince dalla Figura 3, una maggiore omogeneità o compattezza dei pixel di ciascuna classe.

Il risultato qualitativo è confermato anche da quello quantitativo delle matrici di confusione e dalla statistica Kappa:

- **Accuratezza totale = 63.0 % e  $K = 0.5875$**  nel caso di equiprobabilità a priori.
- **Accuratezza totale = 91.5 % e  $K = 0.9076$**  nel caso di probabilità elaborate con l'*IK*.

Nella classificazione mostrata nella Figura 2 sono presenti pixel isolati che potrebbero essere anche realistici da un punto di vista spettrale, ma vanno ritenuti mal classificati nel processo di riconoscimento di aree omogenee contigue. Le mappe di probabilità a posteriori in Figura 1, invece, tengono conto esclusivamente dell'informazione spaziale e mostrano notevole incertezza nelle zone di transizione. Integrando, invece, informazione spaziale e spettrale (Figura 3) l'accuratezza nella classificazione risulta sensibilmente migliorata.

## Conclusioni

Si presenta un nuovo approccio combinato tra classificatore un tradizionale classificatore *Maximum Likelihood* e un interpolatore geostatistico al fine di ottenere un classificatore supervisionato più accurato. La metodologia proposta risulta particolarmente efficiente e potrebbe essere applicata vantaggiosamente in vari studi di caratterizzazione ambientale.

## Bibliografia

Castrignanò A., Goovaerts P., Lulli L., Bragato G (2000) A geostatistical approach to estimate probability of occurrence of *Tuber melanosporum* in relation to some soil properties, in: *Geoderma*, 98, 95 – 113.

Goovaerts P. (2002) Geostatistical incorporation of spatial coordinates into supervised classification of hyperspectral data, in: *Geographical Systems*, 4, 99– 111.