

# Influenza della frammentazione del territorio sull'accuratezza di classificazioni bosco/non bosco da dati satellitari.

(1) Caterina Gagliano, (2) Alessandro Canavesio (1) Flora De Natale, (2) Fabio Giannetti

- (1) I.S.A.F.A. – P.za Nicolini, 6 – 38050, Villazzano, Trento; Tel: 0461-381139, Fax: 0461-381131, e-mail: caterina.gagliano@entecra.it  
(2) I.P.L.A. S.p.A., C.so Casale 476, 10132, Torino; tel. 011/89.98.933, fax 011/89.89.333, e-mail: giannetti@ipla.org

## Riassunto

La valutazione dell'accuratezza delle classificazioni derivanti da immagini satellitari appare di fondamentale importanza per quantificare gli errori sia di tipo tematico sia di tipo spaziale. In particolare gli errori di tipo tematico possono essere anche direttamente correlati alla complessità strutturale del territorio ed alla sua eterogeneità. Con il presente lavoro si intende i) valutare la relazione tra l'eterogeneità territoriale e l'accuratezza tematica, ii) verificare la possibilità di valutare la complessità del paesaggio a partire da dati raccolti per l'analisi dell'accuratezza. A tal fine sono stati confrontati i valori di accuratezza ottenuti a partire da test set selezionati in base ad un'analisi della frammentazione.

## Abstract

The accuracy assessment of land cover maps obtained from satellite data is a crucial issue. Map accuracy depends on two different sources of errors: positional and thematic errors. In particular thematic errors are also related to land cover complexity and heterogeneity. The main topics of this work are to i) evaluate the relation between land heterogeneity and thematic accuracy, ii) verify the possibility of analysing the landscape complexity through accuracy assessment data.

## Introduzione

La disponibilità di informazioni relative alla composizione, estensione ed evoluzione della copertura forestale è considerata di fondamentale importanza per assicurare una corretta analisi e gestione delle risorse forestali

La classificazione delle immagini telerilevate permette di ottenere delle carte tematiche delle risorse agroforestali mediante le quali è possibile misurare le proprietà spaziali e fisiche degli elementi che compongono il paesaggio. Dal confronto di immagini derivate da periodi di ripresa diversi può essere inoltre analizzata la variazione, in termini quali-quantitativi, del paesaggio agroforestale (Frohn, 1998). Nel caso di confronti multitemporali l'accuratezza diventa fondamentale nella discriminazione tra reali cambiamenti di uso del suolo ed errori indotti dalla classificazione.

Per poter conoscere l'accuratezza di tali classificazioni risulta importante non solo valutare gli errori di classificazione, ma anche la loro distribuzione spaziale.

L'accuratezza delle mappe tematiche dipende principalmente da due tipi di errori distinti, anche se interdipendenti (Perdigao e Annoni, 1997; Latifovic e Olthof, 2004): di tipo tematico e di tipo spaziale. Gli errori di tipo tematico, in particolare, oltre a dipendere dalle caratteristiche spettrali dell'immagine e dall'efficacia del classificatore utilizzato, possono essere direttamente correlati alla complessità strutturale del territorio in termini di eterogeneità e frammentazione (Campbell, 1996). Queste caratteristiche spaziali possono determinare un aumento dei pixel misti il cui valore di riflettanza è dovuto alla presenza di più di un uso del suolo (Smith, 2003).

La frammentazione del paesaggio può determinare inoltre problemi di accuratezza spaziale, in particolare nelle zone di transizione tra un uso del suolo ed un altro (Congalton, 1988; Townshend et al., 2000). Per analizzare la diminuzione dell'affidabilità delle classificazioni nelle aree ad elevata

frammentazione alcuni autori hanno proposto di adottare diverse strategie di campionamento, confrontando i risultati ottenuti (Moisen et al., 1994).

L'obiettivo del presente lavoro è quello di valutare la relazione tra l'eterogeneità territoriale e l'accuratezza tematica verificando i) se la complessità del paesaggio può condizionare l'accuratezza tematica e ii) se è possibile valutare la complessità del paesaggio a partire da dati raccolti per l'analisi dell'accuratezza.

Le relazioni tra errori di classificazione ed eterogeneità del paesaggio sono state analizzate confrontando le stime di accuratezza ottenute a partire da test set selezionati in base ad un'analisi della frammentazione.

### **Area di studio**

Il presente lavoro ha riguardato un'area delle Marche, estesa circa 72.000 ha, compresa tra i territori delle comunità montane di Catria-Nerone e Catria-Cesano e caratterizzata da un paesaggio vario ed articolato sia dal punto di vista dell'uso del suolo che dell'orografia, ad altitudini comprese tra 150 e 1700 m s.l.m. Alle basse quote prevalgono aree agricole a forte pressione antropica mentre alle quote più elevate si concentrano le aree naturali più compatte. Il 45% circa del territorio analizzato è coperto da foreste, con una netta prevalenza di specie decidue, se si escludono alcuni impianti di conifere distribuiti in modo eterogeneo sul territorio e costituiti prevalentemente da pino nero.

### **Materiali e metodi**

#### 1) Data set ed elaborazione delle immagini

Per il presente lavoro è stato utilizzato un data-set di tre immagini Landsat TM, riprese nella tarda primavera (3 giugno), in estate (6 agosto) e all'inizio dell'autunno (23 settembre) del 1999 al fine di evidenziare le diverse risposte fenologiche della vegetazione (Oetter et al., 2000).

Le immagini Landsat TM sono state ortorettificate usando un DEM con una risoluzione di 10 m ed applicando un modello di correzione geometrica basato su un'equazione di collinearità. In seguito le tre immagini sono state co-registrate al fine di garantirne una perfetta sovrapposizione (De Natale et al., 2005).

Ogni singola immagine è stata classificata con il metodo *Maximum Likelihood* mediante il software ENVI 4.1. Le aree di training utilizzate nella classificazione sono state selezionate mediante una campagna a terra (2003) e tramite fotointerpretazione di ortofoto digitali in bianco e nero del 1998 a risoluzione di 1 metro (Agea ex-AIMA).

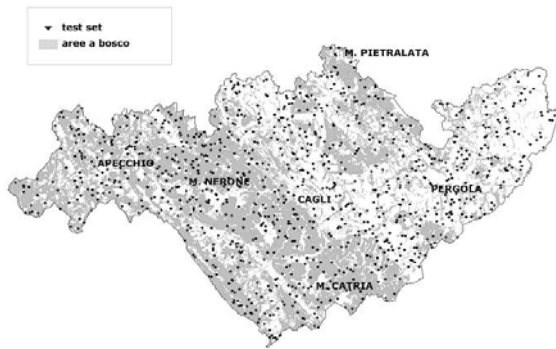
Nella fase di training sono state individuate 3 classi di uso del suolo non forestali (agricolo, pascoli e urbano) e 3 di tipo forestale (impianti di conifere, faggete e popolamenti misti). Infine, per migliorare l'affidabilità del test set per l'analisi dell'accuratezza, da ogni mappa di uso del suolo è stata derivata una classificazione semplificata bosco/non bosco.

Oltre alle immagini satellitari è stata utilizzata la Carta forestale delle Marche, in scala 1:25.000 (Regione Marche, 2000) per misurare il grado di frammentazione del territorio attraverso strumenti di analisi spaziale. A partire dal dato vettoriale, rasterizzato secondo una griglia con origine e risoluzione corrispondenti a quelle della matrice delle immagini satellitari, è stata eseguita una analisi dell'indice di CVN (numero di celle differenti dalla cella centrale, su un'area di analisi di 3x3 pixel), calcolato mediante l'estensione del software ArcView 3.2 Texture Analysis and Neighbourhood Statistics (Behrens, 2003).

#### 2) Selezione del campione

L'accuratezza della mappa bosco/non bosco è stata valutata su un campione di 1000 punti distribuiti in maniera casuale sul territorio studiato (fig. 1) mediante l'estensione *Random Point Generator* (Jeness, 2001) del software ArcView.

Attraverso fotointerpretazione, i punti sono stati assegnati al bosco o al non bosco secondo la definizione di bosco della FAO-FRA2000 (Global Forest Resources Assessment, 2000) che include nelle superfici boscate i territori con una copertura arborea maggiore del 10%, un'estensione minima di 5000 m<sup>2</sup>, una larghezza minima di 20 m ed un'altezza minima di 5 m.



Quest'ultimo criterio non è stato adottato in quanto non è valutabile da ortofoto. Alle ortofoto è stata sovrapposta una griglia a maglia quadratica 30x30 m, coincidente con la matrice delle immagini Landsat TM, al fine di localizzare con precisione ogni pixel dell'immagine da satellite sulle ortofoto e misurare la percentuale di copertura forestale di ogni pixel contenente un punto di campionamento. Tali pixel sono stati classificati come bosco quando la copertura superava il 50% e come non bosco nel caso contrario.

*Figura 1 – Distribuzione dei punti di campionamento nell'area di studio*

Sono stati così individuati tre test set: i) punti di campionamento, ii) pixel contenenti i punti di campionamento ed iii) finestra di analisi di 3x3 pixel intorno ad ogni pixel contenente i punti di campionamento.

Ad ogni punto di campionamento è stata associata una serie di informazioni sulle caratteristiche spaziali delle aree definite dal pixel e dalla finestra di analisi di 3x3 pixel, segnalando la presenza di chiarie<sup>1</sup>, strade<sup>2</sup> e alberi fuori foresta<sup>3</sup> per entrambe le aree, mentre per il solo pixel è stato riportato il grado di frammentazione calcolato applicando il CVN<sup>4</sup>. Mediante un sistema di 4 cerchi concentrici con raggio rispettivamente di 25 m, 50 m, 75 m e 100 m è stato possibile inoltre attribuire ogni punto ad una classe di distanza dal margine del bosco.

Tutti questi dati sono stati rilevati da fotointerpretazione o derivati da dati ausiliari e hanno permesso di selezionare sottocampioni significativi per l'analisi dell'accuratezza e della frammentazione.

### 3) Analisi dell'accuratezza e delle caratteristiche strutturali del paesaggio

La valutazione dell'accuratezza delle mappe bosco/non bosco è stata fatta sulla base delle matrici di confusione (Congalton, 1991) calcolate per l'intero campione di punti di verifica con riferimento alle tre unità campionarie punto, pixel ed intorno, nonché per alcuni sottocampioni scelti sulla base delle informazioni raccolte. In particolare sono stati utilizzati:

- il sottocampione dei pixel puri (ovvero i pixel la cui superficie è completamente occupata da una classe)
- il sottocampione dei pixel puri ricaduti in un intorno puro (ovvero i pixel del sottocampione precedente che non presentano pixel di classe diversa nel kernel 3x3)

Successivamente è stata fatta un'ulteriore distinzione dei sottocampioni in base alla presenza o assenza di chiarie, strade e alberi fuori foresta. L'accuratezza è stata poi calcolata sui punti distinti per classe di distanza dal margine e sui pixel distinti per classi di CVN.

<sup>1</sup> Aperture nel bosco che non superano i limiti dimensionali della definizione di bosco adottata.

<sup>2</sup> Soluzioni di continuità che non pregiudicano l'attribuzione di un punto alla classe bosco secondo la definizione utilizzata ma che, avendo caratteristiche spettrali molto diverse, possono rappresentare un fattore di disturbo nella classificazione automatica dell'immagine da satellite.

<sup>3</sup> Barriere frangivento e fasce boscate di larghezza inferiore a 20 m, superfici boscate inferiori a 5000 m<sup>2</sup>, che possono influire sulla classificazione.

<sup>4</sup> La mappa del CVN distingue il numero di celle differenti dalla cella centrale, su un'area di analisi di 3x3 pixel, mediante 9 classi. La mappa è stata riclassificata in 4 classi (0, 1, 2 e >2) al fine di ottenere dei campioni sufficientemente numerosi in ciascuna classe per il calcolo dell'accuratezza.

## Risultati

### 1) Elaborazione delle immagini

Le classificazioni bosco/non bosco ottenute per ogni singola immagine mostrano un'alta frammentazione nella parte orientale dell'area di studio, caratterizzata da morfologia collinare con versanti poco acclivi ed uso agricolo intensivo, ed una maggiore compattezza della copertura forestale nella parte occidentale, caratterizzata da versanti a maggiore pendenza.

### 2) Selezione del campione

La tabella 1 mostra i diversi test set usati per il calcolo dell'accuratezza.

Mentre il campione (a) e quello (b) sono composti dalle 1000 unità di campionamento, gli altri rappresentano dei sottocampioni.

Test set	Numerosità campionaria
<b>(a) Punti</b>	<b>1000</b>
<b>(b) Pixel</b>	<b>1000</b>
(b.1) pixel - presenza di chiarie, alberi fuori foresta e strade	228
(b.2) pixel - assenza di chiarie, alberi fuori foresta e strade	772
<b>(c) pixel puri</b>	<b>776</b>
(c.1) pixel puri - presenza di chiarie, alberi fuori foresta e strade	131
(c.2) pixel puri - assenza di chiarie, alberi fuori foresta e strade	645
<b>(d) pixel puri in blocchi puri</b>	<b>498</b>
(d.1) pixel puri in blocchi puri - presenza di chiarie, alberi fuori foresta e strade	76
(d.2) pixel puri in blocchi puri - assenza di chiarie, alberi fuori foresta e strade	422

Tabella 1 - Test set usati per il calcolo dell'accuratezza

### 3) Analisi dell'accuratezza e delle caratteristiche strutturali del paesaggio

Nella tabella 2 si può osservare che i valori di accuratezza ottenuti per le tre classificazioni si differenziano a seconda del periodo di ripresa, con un'accuratezza crescente da giugno a settembre. I motivi principali di questa differenza sono da imputare principalmente alla fase fenologica per quanto riguarda giugno, mentre si può ipotizzare un disturbo radiometrico dato da effetti derivati dall'elevata umidità atmosferica al suolo nell'immagine di agosto.

		Test set											
		a			b			c			d		
		PA	UA	k	PA	UA	k	PA	UA	k	PA	UA	k
03/06/99	B	53,86	88,79		54,89	88,48		62,44	91,10		74,06	89,55	
	NB	91,80	62,26		91,79	63,91		93,09	68,64		91,42	78,03	
	<b>OA</b>	<b>70,70</b>		0,44	<b>72,06</b>		0,45	<b>76,81</b>		0,54	<b>82,77</b>		0,66
06/08/99	B	72,06	86,92		73,50	86,70		79,02	90,50		84,59	88,24	
	NB	86,92	72,06		87,04	74,08		90,61	79,23		88,81	85,30	
	<b>OA</b>	<b>78,79</b>		0,55	<b>79,80</b>		0,60	<b>84,46</b>		0,69	<b>86,70</b>		0,73
23/09/99	B	68,75	93,27		69,55	92,27		78,54	95,27		85,34	91,53	
	NB	94,01	71,38		93,30	72,73		95,58	79,72		92,16	86,36	
	<b>OA</b>	<b>80,20</b>		0,61	<b>80,60</b>		0,62	<b>86,53</b>		0,73	<b>88,76</b>		0,78

Tabella 2 – Producer's accuracy, User's accuracy, Overall accuracy e coefficiente k ottenuti per i test set punti (a), pixel (b), pixel puri (c) e pixel puri in intorno puro (d).

La stima dell'accuratezza tende ad aumentare all'aumentare dell'omogeneità del campione considerato, come dimostra l'accuratezza raggiunta con il campione dei pixel puri e con quello dei pixel puri in kernel 3x3 puri.

Considerando il sotto-campione di pixel con presenza di strade, alberi fuori foresta o chiarie (b.1 - tab. 3), si stimano valori di accuratezza più bassi: per le immagini di giugno e agosto questi sono inferiori di circa il 9% rispetto a quelli ottenuti con l'intero test set (tab. 2) e del 12% circa rispetto a

Test set	TM 030699	TM 080699	TM 230999
	OA (%)	OA (%)	OA (%)
(b.1)	65,92	72,65	71,30
(b.2)	74,05	82,14	83,44
(c.1)	74,22	78,91	78,91
(c.2)	77,54	85,65	87,99
(d.1)	78,47	84,21	85,65
(d.2)	87,76	93,01	95,10

quelli ottenuti con il campione di pixel nel cui intorno non è stata segnalata la presenza di questi elementi.

Il fenomeno è ancora più accentuato per l'immagine di settembre (tab. 3).

Nel caso dei campioni di pixel puri e di pixel puri in kernel puri si può osservare lo stesso trend, ma con variazioni meno marcate.

Tabella 3 – Valori di overall accuracy ottenuti per i test sets (b.1), (b.2), (c.1), (c.2), (d.1) e (d.2)

Infine i valori di accuratezza calcolati utilizzando sottocampioni di punti scelti in base alla loro distanza dal margine mostrano che, per distanze superiori ai 100 m, l'accuratezza può sfiorare il valore del 90% mentre per distanze inferiori i valori si abbassano considerevolmente quasi fino a dimezzarsi (fig. 2).

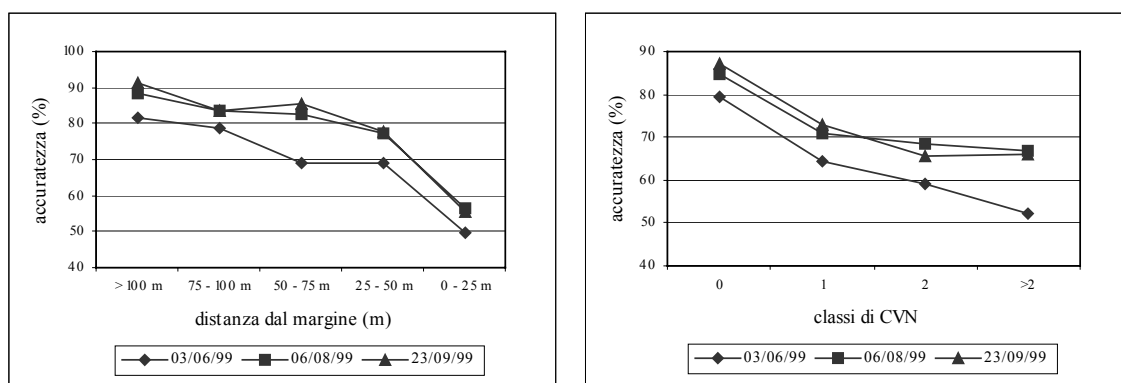


Figura 2 - Effetto della eterogeneità del paesaggio sull'accuratezza delle classificazioni.

La mappa del CVN estratta dalla Carta Forestale delle Marche mostra che le accuratezze maggiori sono localizzate in corrispondenza delle aree più compatte ed omogenee e, viceversa, l'errore di classificazione maggiore si concentra nelle aree più eterogenee (fig. 2).

È da notare inoltre che, per valori di CVN elevati, l'accuratezza di giugno risulta decisamente più bassa rispetto alle altre, che si attestano attorno ad un valore del 70%. Questo potrebbe essere dovuto ad una minore capacità di riflessione del bosco, considerando la sua risposta fenologica in questo periodo, e, contemporaneamente, ad una più forte riflettanza delle aree agricole legata in particolare alla cerealicoltura, in piena fase di sviluppo.

## Conclusioni

Nel presente studio sono state adottate diverse strategie di campionamento al fine di osservare l'effetto dell'eterogeneità e della frammentazione del paesaggio sul calcolo dell'accuratezza. Si osserva, infatti, che in presenza di elementi in grado di interrompere la continuità dell'uso del suolo (chiarie, strade, alberi fuori foresta), l'accuratezza tende a diminuire. Viceversa, se si escludono le aree di margine e quelle più eterogenee si osserva una sovrastima dell'accuratezza di circa il 10%. Altri studi riportano che l'esclusione delle aree di margine può determinare degli incrementi di accuratezza anche del 25% (Fuller *et al.*, citato in Foody, 2002).

I risultati dell'analisi dell'accuratezza sui sotto-campioni distinti in base alla distanza dal margine concordano con quelli riferiti alle classi di CVN e confermano che gli errori di classificazione sono concentrati nelle aree di transizione tra un uso del suolo ed un altro (Czaplewski, 2003).

Come suggerisce Foody (op cit.) è preferibile fornire più misure dell'accuratezza per migliorare l'informazione sull'affidabilità delle mappe considerate. Poiché l'accuratezza tematica varia in funzione dell'eterogeneità e della dimensione delle tessere del paesaggio (Smith *et al.*, 2003) si è ritenuto utile associare un'informazione sulla frammentazione del paesaggio alle stime di

accuratezza. I risultati ottenuti confermano che la complessità del paesaggio determina una riduzione dell'accuratezza tematica e che è possibile valutare la complessità del paesaggio a partire da dati raccolti per l'analisi dell'accuratezza. Viceversa si può presumere che passando da un'analisi puntuale ad un'analisi di tipo continuo del grado di eterogeneità nel territorio si possa fornire un'indicazione sulla possibile distribuzione dell'errore di classificazione.

## **Bibliografia**

- Behrens T. (2005), *Texture Analysis and Neighbourhood Statistics*. <http://arcscrips.esri.com/details.asp?dbid=12646>
- Campbell J.B. (1996), *Introduction to remote sensing (2nd ed.)*. New York, NY: Guilford Press, 622 pp.
- Congalton R. S. (1988), "A practical look at the sources of confusion in error matrix generation", *Photogrammetric Engineering & Remote Sensing*, 54(5): 587-592
- Congalton R.G. (1991), "A review of assessing the accuracy of classifications of remotely sensed data", *Remote Sensing of Environment*, 37, 35-46
- Czaplewski, R.L. (2003), *Accuracy assessment of forest condition: statistical design and methodological considerations*. In M. Wulder and S. Franklin (Ed.), *Remote Sensing of Forest Environments: Concepts and Case Studies* (pp.115-140). Kluwer Academic Publishers, London.
- De Natale F., Puzzolo V., Giannetti F., Canavesio A. (2005), *Assessment of Landsat TM forest-non forest classification: how the selection of the test set affects the classification accuracy evaluation*. In FORESTSAT2005: Operational tools in forestry using remote sensing techniques. Boras, 31 maggio-3 giugno 2005.
- Frohn R.C., (1998), *Remote sensing for landscape ecology - new metric indicators for monitoring, modelling and assessment of ecosystems*. Lewis publishers/CRC Press, Florida. 99 pp.
- Foody, G.M., (2002), "Status of land cover classification accuracy assessment", *Remote Sensing of Environment*, 80: 185-201.
- Global Forest Resources Assessment (2000), *Main report. FAO Forestry Paper*, 140. [www.fao.org/forestry/fo/fra/main/index.jsp](http://www.fao.org/forestry/fo/fra/main/index.jsp)
- Jeness J. (2001), Random point generation. [http://www.jenessent.com/arcview/random\\_point.htm](http://www.jenessent.com/arcview/random_point.htm)
- Latifovic R., Olthof I. (2004), "Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data", *Remote Sensing of Environment*, 90: 153-165
- Moisen G.G., Edwards T.C., Cutler D.R. (1994), "Spatial sampling to assess classification accuracy of remotely sensed data", *Environmental Information Management and Analysis: Ecosystem to Global Scales* (Michener W.K., Brunt J.W., Stafford S.G., editors), Taylor and Francis, London, England, pp. 159-176
- Oetter D.R., Cohen W.B., Berterretche M., Maiersperger T.K., Kennedy R.E. (2000), "Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data", *Remote Sensing of Environment*, 76: 139-155.
- Perdigao, V., Annoni, A. (1997), *Technical and methodological guide for updating CORINE Land Cover Data Base*. JRC/EEA.
- Regione Marche – IPLA S.p.a. (2000), *I tipi forestali delle Marche – Inventario e carta Forestale della Regione Marche*. SELCA, Florence
- Smith J. H., Stehman S.V., Wickham J.D., & Yang L. (2003), "Effects of landscape characteristics on land-cover class accuracy", *Remote Sensing of Environment*, 84: 342-349.
- Townshend J.R.G, Huang C., Kalluri S.N.V., DeFries R.S., Liang S. (2000), "Beware of per-pixel characterization of land cover", *International Journal of Remote Sensing*, 21(4): 839-843