

Uso di dati telerilevati multispettrali per la stima della massa legnosa tramite metodi di spazializzazione non-parametrici

Travaglini D., Bertini R., Chirici G.

geoLAB, Laboratorio di Geomatica, Dipartimento di Scienze e Tecnologie Ambientali Forestali, Università degli Studi di Firenze. Via S. Bonaventura, 13, I-50145 Firenze (Italy). Tel.: +39 055 30231221, Fax: +39 055 319179, e-mail: davide.travaglini@unifi.it

Riassunto

Il presente studio illustra l'utilizzo di immagini telerilevate Landsat 5 TM per la spazializzazione con metodo k -NN della provvigione legnosa rilevata nell'ambito dell'Inventario Forestale della Regione Toscana (IFT). Nelle condizioni esaminate le *performance* di stima mostrano un sensibile incremento assimilando le unità inventariali non a singoli punti ma a superfici d'intorno omogenee e spazialmente continue ottenute con tecniche di segmentazione delle immagini Landsat.

Abstract

The study illustrates the use of remote sensing data Landsat 5 TM for the spazialization of the growing stock by means of k -NN method within the Regional Forest Inventory of Tuscany. In the conditions here examined, the accuracy grows when field plots are compared not to single points but to homogeneous and spatially continuous area obtained with the use of Landsat image segmentation techniques.

Introduzione

Il crescente interesse verso le problematiche ambientali ha generato l'esigenza di ampliare e approfondire le attuali conoscenze sui sistemi naturali.

Per quanto riguarda il settore forestale si registra una crescita del numero e del tipo di attributi comunemente rilevati nell'ambito di indagini inventariali (Corona, Tabacchi, 2001), strumenti indispensabili per impostare una corretta pianificazione del territorio forestale.

Le statistiche desunte su base campionaria, però, non consentono una restituzione georeferenziata dell'informazione acquisita, ma solo aggregata per macro aree. Per superare questo limite sono stati sviluppati differenti metodi finalizzati a stimare sulla intera superficie di indagine il valore di attributi biofisici accuratamente rilevati su un congruo numero di aree campione, operazione indicata con il termine "spazializzazione". Tra i vari metodi a oggi disponibili in genere si preferisce ricorrere a quelli basati sul rilievo a terra di attributi forestali, integrato con informazioni multispettrali derivate da immagini telerilevate da satellite, in quanto capaci di fornire stime più accurate soprattutto in ambiti forestali complessi (Maselli et al., 2003).

La procedura prevede che l'intera area d'indagine sia coperta da una scena satellitare multispettrale. La superficie d'interesse viene quindi assimilata a una matrice di celle elementari (pixel), di dimensione pari alla risoluzione delle immagini telerilevate. Su un numero sufficientemente ampio di pixel bosco (unità di campionamento a terra) si esegue il rilievo a terra dell'attributo forestale da stimare (a esempio il volume o l'area basimetrica a ettaro) in modo tale che per ciascuna unità di campionamento a terra sia noto sia il valore dell'attributo d'interesse, sia il valore di una o più variabili spettrali derivate dalle bande del sensore satellitare prescelto o di indici ottenuti per combinazione tra le bande. L'operazione finale di stima dell'attributo per tutti i pixel bosco che

compongono la superficie d'indagine può essere operata secondo un approccio correlativo di tipo parametrico o non-parametrico.

In questi casi il grado di accuratezza di stima dipende dal livello di correlazione tra i valori dell'attributo rilevato a terra e i valori della/e variabile/i rilevata/e sulle immagini, che a sua volta è condizionato da una serie di fattori principali quali: *a)* la dimensione e la scelta del campione; *b)* l'attendibilità del rilievo a terra dell'attributo; *c)* la qualità della risposta spettrale telerilevata, a esempio in relazione agli effetti dovuti alla topografia o alla presenza di pixel misti (Chirici et al., 2003a); *d)* l'errore commesso nella georeferenziazione delle unità di rilevamento a terra, anche se oggi questo problema può essere contenuto entro limiti accettabili utilizzando sistemi di posizionamento satellitare adeguati in relazione alla risoluzione geometrica delle immagini telerilevate; *e)* il periodo di tempo intercorso tra il rilievo a terra e l'acquisizione del dato satellitare, visto che la risposta spettrale dei popolamenti forestali subisce modifiche anche per cause naturali o antropiche; *f)* l'errore commesso nella georeferenziazione delle immagini satellitari, un problema particolarmente sentito in aree montane a morfologia complessa e accidentata.

Un modo per limitare la componente di errore causata da una imprecisa coregistrazione spaziale tra la posizione delle unità di rilevamento a terra e i corrispondenti pixel delle immagini da satellite prevede l'uso di filtri passa basso, così che la risposta spettrale delle unità campione viene calcolata come media dei valori spettrali degli $n \times n$ pixel circostanti (Trotter, Dymond, 1997; Tokola, 2000). Altri metodi più complessi prevedono l'impiego di tecniche di segmentazione delle immagini per associare ai punti campione le proprietà spettrali medie rilevate su regioni omogenee e spazialmente continue (Pekkarinen, 2002) o di procedure basate su un approccio multicriteriale (Halme, Tompoo, 2001).

Il presente studio illustra l'utilizzo di immagini telerilevate Landsat 5 TM per la spazializzazione con metodo non-parametrico della provvigione a ettaro rilevata nell'ambito dell'Inventario Forestale della Regione Toscana (IFT). In particolare, vengono confrontati in termini di accuratezza differenti configurazioni dell'algoritmo di stima, basati sull'impiego di dati spettrali derivati per pixel, con metodi che associano alle unità inventariali proprietà spettrali calcolate su un intorno di pixel definito con filtri di differente ampiezza o con tecniche di segmentazione delle immagini.

Area di studio

La sperimentazione è stata condotta nel settore orientale della Toscana e ha interessato le province di Pistoia, Prato, Firenze, Arezzo, Siena e in minima parte quelle di Pisa e Grosseto, per una superficie complessiva di circa 12000 km². Le principali formazioni forestali della zona sono rappresentate da castagneti, faggete, querceti di cerro e di roverella, pinete mediterranee e di pino nero e abetine di abete bianco, alle quali si aggiungono cespuglieti, arbusteti e zone a macchia mediterranea (Hofmann et al., 1998).

Materiali

L'IFT è impostato su un sistema di campionamento a due fasi che ha previsto: in prima fase la fotointerpretazione di un numero di punti campione distribuiti sistematicamente sulle foto aeree; in seconda fase la selezione di un campione ristretto di aree di saggio sul quale effettuare i rilievi a terra operando una stratificazione sulla base dei risultati ottenuti in prima fase. La procedura per la localizzazione e il posizionamento sul terreno dei punti inventariali e il metodo di rilievo degli attributi dendroauxometrici non è stato omogeneo per tutta la superficie d'indagine. Nel primo periodo dell'inventario forestale, che ha interessato la Toscana orientale, l'operazione di localizzazione dei punti è stata eseguita mediante poligoni aperte e il rilievo è stato condotto su quattro aree di saggio a raggio variabile con tecnica relascopica, disposte secondo una conformazione esagonale che prevede l'identificazione di sei aree satelliti virtuali a distanza di 50 m dall'area centrale. Oltre all'area centrale sono state rilevate le aree periferiche che in ordine azimutale occupano le posizioni pari e al punto inventariale è stato assegnato la media dei risultati ottenuti nelle quattro aree di saggio. Nel secondo periodo di attuazione dell'inventario, che ha

interessato la Toscana occidentale, il posizionamento a terra dei punti di rilevamento è stato eseguito con tecniche GPS e i rilievi si sono basati sull'esecuzione di aree di saggio a raggio fisso di ampiezza variabile secondo la struttura del soprassuolo (Hofmann et al., 1998).

Ai fini di questo studio sono state utilizzate le sole unità inventariali della Toscana orientale sottoposte a rilievo a terra e coperte dalla scena satellitare Landsat 5 TM *path* 192, *rows* 030, acquisita nell'estate del 1997.

Metodi

Pre-elaborazioni

La scena satellitare Landsat 5 TM è stata ortorettificata utilizzando un DEM (*Digital Elevation Model*) con passo di 75 m e oltre 60 GCPs (*Ground Control Points*) individuati per confronto su ortofoto digitali AIMA a 8 bit. L'errore commesso in questa fase è risultato inferiore alla dimensione del pixel. Per fotointerpretazione in ambiente GIS sono state individuate e mascherate le aree coperte da nuvole e da loro ombre.

Definizione della popolazione sperimentale

A partire dall'intero *set* di punti inventariali della Toscana orientale è stata eseguita una stratificazione dei dati nel tentativo di ridurre la variabilità della popolazione e ottimizzare le *performance* di calcolo dello stimatore. A questo scopo sono state selezionate le sole unità inventariali che rispondono ai seguenti requisiti: (i) rilievo eseguito a terra sulla base di cluster e non di singoli punti, considerato che in questo secondo caso è stata riscontrata una minore relazione tra il valore dell'attributo d'interesse e i valori spettrali; (ii) area basimetrica ≥ 20 m²/ha, si sono in questo modo escluse le formazioni particolarmente rade; (iii) l'unità inventariale è coperta dalla immagine satellitare libera da nuvole; (iv) la distribuzione delle unità inventariali tra le classi di uso/copertura del suolo deve essere pari ad almeno il 5% circa del totale di punti che costituiscono la popolazione. In definitiva sono stati scelti 253 punti inventariali, rappresentativi di 8 differenti classi di uso/copertura del suolo (Tabella 1).

<i>Classi di uso/copertura del suolo</i>	<i>Punti inventariali</i>	
	<i>n</i>	<i>%</i>
Boschi a prevalenza di faggio	34	13.4
Boschi a prevalenza di castagno	49	19.4
Boschi a prevalenza di roverella	16	6.3
Boschi a prevalenza di cerro	44	17.4
Boschi a prevalenza di leccio	12	4.7
Boschi a prevalenza di abete bianco	30	11.9
Boschi a prevalenza di pino nero	46	18.2
Boschi a prevalenza di pini mediterranei	22	8.7
Tot.	253	100.0

Tabella 1 – La popolazione sperimentale.

Scelta delle variabili spettrali

In base a una preliminare analisi di correlazione tra i valori di massa a ettaro rilevati a terra e le statistiche spettrali calcolate sulle immagini Landsat, assieme al valore di indici ottenuti per combinazione delle bande multispettrali (NDVI, IRI) e informazioni ausiliarie descrittive della stazione, sono stati identificati gli strati informativi che mostrano una migliore relazione tra i valori di provvigione rilevati a terra e i caratteri multispettrali delle unità inventariali. In questa fase sono stati scartati i fattori stagionali la cui determinazione richiederebbe lunghi e costosi tempi di rilievo, limitando la scelta alle sole variabili deducibili da comuni banche dati geografiche o da fonti

telerilevate. Tutte le bande Landsat con esclusione della sesta sono state utilizzate assieme a un DEM.

Scelta del campione

Al fine di valutare le *performance* di stima, dalla popolazione di riferimento è stato estratto un campione casuale del 50% scelto in modo proporzionale alla ripartizione delle unità inventariali nelle differenti classi di uso/copertura del suolo. Il campione così ottenuto è stato assunto come *set* di unità di rilevamento a terra, di seguito denominate unità *training*, con il quale individuare la migliore configurazione dell’algoritmo di stima da utilizzare per la previsione della provvigione a ettaro sul rimanente 50% di unità della popolazione, di seguito denominate unità *test*, per le quali si è ipotizzato di non conoscere l’entità dell’attributo ricercato.

Configurazione dell’algoritmo di stima e valutazione dell’accuratezza

Il metodo di stima non-parametrico testato in questo studio è quello basato sulle distanze spettrali, più conosciuto con il nome di *k Nearest Neighbors (k-NN)*. Il metodo, sviluppato secondo un approccio raster, prevede che per ogni pixel appartenente alla popolazione di riferimento per il quale è sconosciuto il valore dell’attributo d’interesse, vengano individuati i *k* pixel più vicini (simili) tra quelli corrispondenti alle unità di rilevamento a terra, scelti all’interno di uno spazio multidimensionale definito dal numero di variabili spettrali (bande del satellite e/o bande accessorie) utilizzate nella procedura di stima. I valori dell’attributo rilevati in corrispondenza dei *k* pixel più vicini sono quindi utilizzati per la stima finale, ottenuta come media dei *k* valori pesata sull’inverso della distanza spettrale, in modo da assegnare ai pixel più simili un peso maggiore rispetto agli altri (Tomppo, 1996):

$$M_{p_o} = \frac{\sum_{i=1}^k w_{p_i p_o} m_{p_i}}{\sum_{i=1}^k w_{p_i p_o}} \quad [1]$$

dove $w_{p_i p_o}$ = peso attribuito al pixel campione p_i (con valori noti dell’attributo considerato) rispetto al pixel incognito p_o ; m_{p_i} = valore dell’attributo misurato a terra in corrispondenza del pixel p_i .

I metodi sviluppati per calcolare la distanza spettrale sono molteplici. Quelli utilizzati in questo studio si basano sulla distanza spettrale euclidea, di Mahalanobis, pesata con pesi fuzzy (Franco-Lopez, 2001, Maselli et al., 2003). La migliore configurazione dell’algoritmo *k-NN* è stata scelta preventivamente con metodo *leave one out* applicato alle sole unità *training*, secondo una comune procedura di *cross-validation* (Katila e Tomppo, 2001; Maselli et al., 2003). In particolare sono stati testati valori di *k* compresi tra 1 e 20, tre tipi di distanza spettrale e, per confronto, differenti approcci per l’estrazione dei caratteri spettrali della popolazione d’indagine: (i) per singolo pixel, (ii) su immagini filtrate con filtro passa basso di 3x3, 5x5, 7x7, (iii) su aree ottenute per segmentazione di immagini Landsat tracciate secondo due differenti fattori di scala (2 e 3), senza tenere conto del fattore di forma secondo la procedura descritta in Chirici et al. (2003b). La migliore configurazione trovata è stata usata per la stima dei valori di volume a ettaro sulle unità *test*. L’accuratezza di stima è stata valutata sulle unità *test* confrontando i valori di volume rilevati a terra con quelli ottenuti con metodo *k-NN*.

Risultati

La distanza spettrale di Mahalanobis ha fornito i migliori risultati nelle prove di *cross-validation* eseguite con metodo *leave one out*. Come riportato in Figura 1, l’andamento delle *performance* di stima evidenzia un sensibile incremento quando i caratteri spettrali sono determinati assimilando i punti inventariali non a singoli pixel ma a un intorno di pixel.

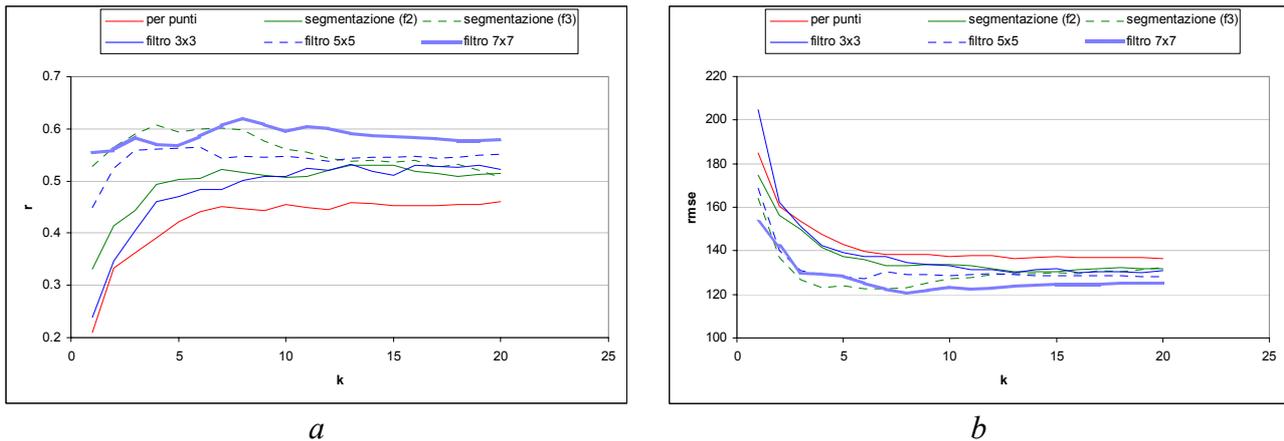


Figura 1 – Risultato del leave one out eseguito sulle unità training utilizzando la distanza spettrale di Mahalanobis, valori crescenti di k e differenti metodi per l'estrazione delle statistiche spettrali delle unità inventariali (a. coefficiente di correlazione tra valori di massa rilevati a terra e valori stimati con metodo k -NN, b. RMSE).

Nelle condizioni esaminate, l'uso di superfici d'intorno di maggiore ampiezza (Tabella 2) è risultato preferibile.

Metodo	Superficie media	Superficie minima	Superficie massima	Deviazione standard
	ha	ha	ha	ha
Per pixel	0.09	0.09	0.09	0
Filtro (3x3)	0.81	0.81	0.81	0
Filtro (5x5)	2.25	2.25	2.25	0
Filtro (7x7)	4.41	4.41	4.41	0
Segmentazione (f=2)	1.07	0.09	5.76	1.05
Segmentazione (f=3)	2.66	0.09	15.03	2.17

Tabella 2 – Dimensione media, minima, massima e deviazione standard della superficie d'intorno abbinata alle unità inventariali per l'estrazione delle caratteristiche spettrali.

In particolare, con una configurazione dell'algoritmo impostata su un valore di $k=4$ e utilizzando la distanza spettrale di Mahalanobis calcolata sui valori medi di *digital number* dei pixel appartenenti a aree d'intorno individuate con tecniche di segmentazione delle immagini con fattore di scala 3, si registra una riduzione dell'RMSE calcolato con metodo *leave one out* del 17% rispetto a quanto ottenuto con uno stesso numero di k e con la stessa distanza spettrale calcolata sui valori estratti per singolo pixel. Per quanto riguarda il coefficiente di correlazione (r) tra valori rilevati e valori predetti sulle unità di *training* si registra un incremento del 56%. La scelta di un $k=4$ è stata effettuata in considerazione del fatto che all'aumentare di k la variabilità dei valori predetti si riduce rispetto a quella effettiva, a causa del livellamento operato dal metodo di calcolo, impostato su una media pesata [1] (Katila, Tompoo, 2001).

Questa configurazione del sistema k -NN è stata utilizzata per stimare i valori di volume a ettaro sulle unità *test*. Confrontando i valori così ottenuti con quelli rilevati a terra nell'ambito dell'IFT si ottiene un RMSE di 122.9 m³/ha e un coefficiente di correlazione di 0.62. Lo scarto medio e lo scarto medio assoluto tra valori stimati e valori rilevati sulle unità *test* è rispettivamente di 16.1 m³/ha e di 89.4 m³/ha.

Conclusioni

La sperimentazione dimostra come l'accuratezza di stima ottenuta nell'ambito di indagini inventariali con di tecniche di spazializzazione non-parametriche basate sulla distanza spettrale possa essere migliorata assimilando i punti campione a superfici derivate con tecniche di segmentazione delle immagini per contenere gli errori dovuti a una imprecisa coregistrazione tra le unità di campionamento a terra e le immagini telerilevate. Sebbene nelle condizioni esaminate l'errore di stima sia ancora elevato per proporre una applicazione operativa del metodo, si attendono migliori *performance* di calcolo riducendo ulteriormente la variabilità della popolazione d'indagine con opportune tecniche di stratificazione, a esempio lavorando per singola tipologia forestale.

Bibliografia

- Chirici G., Corona P., Marchetti M., Travaglini D. (2003a), *Testing Ikonos and Landsat 7 ETM+ potential for stand-level forest type mapping by soft supervised approaches*. Advances in Forest Inventory for Sustainable Forest Management and Biodiversity Monitoring, editors P. Corona, M. Köhl, M. Marchetti. Kluwer Academic Publishers: 71-85.
- Chirici G., Corona P., Travaglini D., Filiberti F. (2003b), *Segmentazione multirisoluzione e classificazione object oriented di immagini telerilevate*. Atti della 7° Conferenza Nazionale ASITA: "L'informazione territoriale e la dimensione tempo", Verona, 28-31 ottobre 2003, Volume I: 729-734.
- Corona P., Tabacchi G. (2001), *Inventariazione delle risorse forestali su ampi territori. Finalità, metodi e prospettive*. Monti e Boschi, 6: 27-38.
- Franco-Lopez H., Ek A.R., Bauer M.E. (2001), *Estimation and mapping of forest stand density, volume and cover type using k-nearest neighbours method*. Remote Sensing of Environment, 77: 251-274.
- Halme M., Tomppo E. (2001), *Improving the accuracy of multisource forest inventory estimates by reducing plot location error — a multicriteria approach*. Remote Sensing of Environment, 78: 321–327.
- Hofmann A., Goretti D., Merendi G. A., Tabacchi G., Vignoli M., Bernetti G. (1998), *L'inventario Forestale. Boschi e Macchie di Toscana*. Edizioni Regione Toscana, Firenze.
- Katila M., Tomppo E. (2001), *Selecting estimation parameters for the Finnish multisource National Forest Inventory*. Remote Sensing of Environment, 76: 16-32.
- Maselli F., Bottai L., Chirici G., Corona P., Marchetti M., Travaglini D. (2003), *Stima di attributi forestali in ambiente mediterraneo tramite integrazione di misure a terra e dati telerilevati*. L'Italia Forestale e Montana, 4: 250-263.
- Pekkarinen A. (2002), *Image segment-based spectral features in the estimation of timber volume*. Remote Sensing of Environment, 82: 349-359.
- Tokola T. (2000), *The influence of Field Sample Data Location on Growing Stock Volume Estimation in Landsat TM-Based Forest Inventory in Eastern Finland*. Remote Sensing of Environment, 74: 422-431.
- Tomppo E. (1996), *Multi-source National Forest Inventory of Finland*. In: New Thrusts in Forest Inventory. Päivinen R., Vanclay J. And Miina S. (eds.). EFI Proceedings, 7: 27-41.
- Trotter C., Dymond J. (1997), *Estimation of timber volume in a coniferous plantation forest using Landsat TM*. International Journal of Remote Sensing, 18 (10): 2209– 2223.