Monitoraggio della rinnovazione naturale post incendio tramite l'uso integrato di immagini telerilevate ad alta risoluzione: il caso della pineta di Castel Fusano

Gherardo Chirici¹, Marco Baisi², Roberta Bertini³, Nico Bonora⁴, Ugo Chiavetta¹, Marco Ottaviano¹, Piermaria Corona⁵, Andrea Lamonaca⁶, Diego Giuliarelli⁵, Alessandro Mastronardi⁷, Giovanni Nardinocchi², Valter Sambucini⁴, Daniela Tonti¹ e Marco Marchetti¹

¹ECOGEOFOR, Laboratorio di Ecologia e Geomatica Forestale, Dipartimento di Scienze e Tecnologie per l'Ambiente e il Territorio, Università del Molise. Contrada Fonte Lappone, I-86090 Pesche (IS).

E-mail: gherardo.chirici@unimol.it

²Dipartimento di Ingegneria Elettronica, Università "La Sapienza", via Eudossiana 18, I-00186, Roma

³geoLAB, Laboratorio di Geomatica, Dipartimento di Scienze e Tecnologie Ambientali Forestali, Università degli Studi di Firenze, Via S. Bonaventura 13, I-50145 Firenze

⁴Agenzia per la Protezione dell'Ambiente e per i Servizi Tecnici, Via Vitaliano Brancati, 48 - 00144 Roma
⁵sisFOR, Laboratorio di Inventari Forestali e Sistemi Informativi, Dipartimento di Scienze dell'Ambiente Forestale e delle sue Risorse, Università della Tuscia, Via S. Camillo de Lellis s.n.c., I-01100 Viterbo

⁶BIOFOR Italy srl, *spin off* accademico dell'Università degli Studi della Tuscia, Via Santa Maria di Galeria, 212 - 00123 Roma

⁷GM STUDIO, Via N. Tagliaferri 6 - 50127 Firenze

Riassunto

La pineta di Castel Fusano (Roma) è stata colpita il 4 luglio del 2000 da un importante incendio boschivo in seguito al quale si è avviata nell'area una intensa rinnovazione naturale sia per via gamica che agamica. Ai fini di monitoraggio della suddetta rinnovazione sono stati realizzati una serie di rilievi a terra in aree campione nel 2003 e nel 2006. Negli stessi anni è stata acquisita la copertura di immagini telerilevate multispettrali ad altissima risoluzione Ikonos e Quick Bird. Scopo del presente lavoro è la sperimentazione di diverse metodologie finalizzate alla modellizzazione delle relazioni esistenti tra i dati telerilevati acquisiti e le misure realizzate a terra per la stima e la mappatura dei fenomeni di rinnovazione gamica e agamica. Sono stati per questo sperimentati metodi sia tradizionali di analisi regressiva multivariata, sia di tipo non parametrico, con algoritmi basati su reti neurali (*Relevance Vector Machine* e *Multi-Layer Perceptron*) e *k-Nearest Neighbors*. Le attività si inquadrano nell'ambito del progetto GRINFOMED - MEDIFIRE per il quale è stato realizzato un apposito software denominato *Spatial Forest Modeller* (SFM) capace di analizzare le relazioni tra variabili telerilevate e misurate a terra e di individuare i modelli predittivi migliori in modo da derivare mappe tematiche delle variabili acquisite mediante campionamento a terra. Il contributo illustra i dati acquisiti, le metodologie di analisi e di modellizzazione e i risultati ottenuti. Viene inoltre illustrato il funzionamento del software SFM.

Parole chiave: incendi forestali, rinnovazione naturale, telerilevamento, spazializzazione, reti neurali, *k-Nearest Neighbors*.

Post fire natural regeneration monitoring with the integrated use of high resolution remotely sensed images: the case study of the Pineta di Castel Fusano

Abstract

*Stone pine stand of Castel Fusano (Rome) burnt on July the 4th 2000 during a huge wild-fire. As a consequence of the fire an intensive natural sexual and asexual regeneration began. In order to monitor such a regeneration field surveys were carried out in 2003 and 2006 in sample plots. Remotely sensed high resolution images from Ikonos and Quick Bird were acquired for the same years. The purpose of this work is to test different methodologies for modeling existing relationships between remotely sensed images and ground collected data in order to estimate and to map both sexual and asexual regeneration. For such a purpose different methodologies were tested: step-wise Multiple Linear Regression, Neural Networks (Relevance-Vector-Machine and the Multi-Layered-Perceptron) and the *k*-Nearest-Neighbors. These activities were carried out within the framework of the GRIN-FOMED-MEDIFIRE also developing a specific software named Spatial Forest Modeler (SFM) able to analyze existing relationships between remotely sensed variables and data collected in the field in order to identify the best available models to map and estimate the studied variables acquired on the basis of a field sampling design. The present paper presents data collected in the field, analysis and modeling methods and achieved results. The SFM software is also presented.*

Keywords: forest wildfires, natural re generation, remote sensing, spatialisation, neural networks, *k*-Nearest Neighbors.

Introduzione

La pineta di Castel Fusano costituisce un territorio boscato di circa 1100 ha a ridosso del litorale compreso tra Ostia e la Tenuta Presidenziale di Castel Porziano e rappresenta uno degli elementi di maggiore rilievo ecologico e paesaggistico della Riserva Naturale Statale del Litorale Romano. La vegetazione forestale naturale è costituita da macchie e boschi mesoigrofilo [Mantero, 1995]. Passando dalle dune più prossime al mare alle aree più interne si incontra: la macchia mediterranea a dominanza di leccio, poi i querceti igrofilo con farnia e carpino bianco e quindi i querceti caducifogli con cerro, roverella, farnetto, leccio e sughera [Agrimi et al., 2002]. Tale vegetazione è stata quasi completamente stravolta in conseguenza sia della bonifica idraulica attuata tra il 1884 e il 1891 sia della piantagione di pino domestico iniziata fin dal 1700.

Il nucleo più antico della pineta, coltivato fino circa al 1970, è caratterizzato da un soprassuolo pluristratificato con uno strato superiore di pino domestico e uno inferiore di leccio consociato con altre specie della macchia mediterranea [Agrimi et al., 2002].

A seguito del devastante incendio che ha colpito il 4 luglio del 2000 circa 270 ha di questo nucleo più antico è stata avviata nel 2003 una campagna di rilievo della rinnovazione naturale post incendio, successivamente ripetuta nell'anno 2006 [Corona e Marchetti, 2002]. Scopo della presente nota è la illustrazione dei metodi sperimentati per la stima di indici di rinnovazione naturale post incendio tramite immagini telerilevate multi spettrali ad alta risoluzione. Il contributo descrive i dati raccolti a terra, le immagini telerilevate acquisite, le metodologie di stima basate sull'algoritmo *k*-Nearest Neighbors e su reti neurali e i risultati ottenuti.

Materiali

Rilievi a terra

Lo sviluppo della rinnovazione gamica e agamica all'interno dei 270 ha di pineta distrutti dall'incendio del 2000 è stato realizzato mediante due campagne di rilievi a terra. Nella

Figura 1 - Dislocazione dei transetti utilizzati per la raccolta delle informazioni nella campagna del 2003 sulla base dell'immagine Quick Bird utilizzata nello studio (qui l'indice NDVI in toni di grigio). L'area incendiata nel 2000 è in grigio più scuro, le aree più chiare interne al perimetro dell'incendio sono le nuvole presenti nell'immagine Quick Bird.

campagna condotta nel mese di agosto 2003 sono stati realizzati 178 transetti di forma rettangolare e di superficie di 20 m², georeferenziati mediante uso di GPS (Fig. 1). Su un sottocampione di 49 transetti degli originari 178 è stato ripetuto il rilievo nell'agosto del 2006.

Per ciascun transetto sono stati rilevati i seguenti parametri:

- posizione topografica (coordinate UTM) rilevate con GPS in correzione differenziale in post processing con precisione metrica;
- altezza dei ricacci vegetativi di ciascuna ceppaia di latifoglie arboree, differenziata per specie e copertura della chioma ottenuta con una misurazione parallela e una perpendicolare alla direzione del transetto;
- altezza di tutti i semenzali e plantule da seme differenziata per specie arborea;
- altezza e percentuale della copertura dello strato erbaceo eventualmente presente;
- altezza e percentuale di copertura della lettiera eventualmente presente.

Per ogni singolo transetto sono stati calcolati i seguenti indici di rinnovazione:

- indice di rinnovazione gamica delle latifoglie (*IRGI*):

$$IRGI = \sum_{i=1}^n h_i \quad [1]$$

dove h_i è l'altezza della i -esima pianta (latifolia arborea) nata da seme;

- l'indice di rinnovazione gamica delle conifere ($IRGc$) calcolato come:

$$IRGc = \sum_{i=1}^n h_i \quad [2]$$

dove h_i è l'altezza della i -esima pianta (conifera) nata da seme;

- l'indice di rinnovazione agamica 1 ($IRAl$) calcolato come:

$$IRAl = \sum_{i=1}^n d1_i \cdot d2_i \cdot h_i \quad [3]$$

dove $d1_i$, $d2_i$ e h_i sono, rispettivamente, i due diametri ortogonali della proiezione della chioma e l'altezza della i -esima pianta nata per via vegetativa la cui base ricade all'interno transetto;

- l'indice di rinnovazione agamica 2 ($IRAal$) calcolato come:

$$IRAal = \sum_{i=1}^n d1_i \cdot d2_i \cdot h_i \cdot f_i \quad [4]$$

dove, rispetto all'indice precedente, il fattore correttivo f_i , compreso tra 0 e 1, diminuisce la stima del volume apparente della chioma sottraendo la porzione che ricade al di fuori del margine del transetto.

Immagini telerilevate

Lo studio è basato su due immagini telerilevate multispettrali: una scena Quick Bird acquisita nell'estate del 2003 e già ortorettificata nel sistema UTM WGS84 e una immagine Ikonos acquisita nell'estate del 2006 e coregistrata con l'immagine Quick Bird del 2003. La coregistrazione è stata realizzata mediante 87 punti di controllo omologhi individuati manualmente sulla base delle bande pancromatiche delle due immagini. La correzione dell'immagine del 2006 è stata realizzata utilizzando *Rational Polynomial Functions* (RPF) che hanno permesso di contenere l'errore posizionale medio entro 1 metro. Le funzioni sono poi state applicate alle bande multispettrali che sono state quindi ricampionate con algoritmo *nearest neighbors*.

Per l'immagine Quick Bird del 2003 è stato necessario creare una maschera delle nuvole e delle relative ombre per fotointerpretazione e digitalizzazione manuale.

A partire dalle bande originarie delle due immagini sono stati calcolati gli indici di attività fotosintetica NDVI (*Normalised Difference Vegetation Index*): (banda infrarossa - banda rossa) / (banda infrarossa + banda rossa) e SAVI (*Soil Adjusted Vegetation Index*), calcolato a partire dall'NDVI con un *adjustment factor* L pari a 0,5 [Huete, 1988].

Sia le originali quattro bande multispettrali che ognuno dei due indici di attività fotosintetica sono stati filtrati con operatori a finestra mobile quadrati di 3x3, 5x5 e 7x7 pixel. Per ogni finestra mobile è stata calcolata la media, la deviazione standard e il massimo dei valori dei pixel rientranti nella finestra di calcolo.

Entrambe le immagini sono state segmentate con algoritmo *object oriented* [Baatz et al., 2004], per ogni segmento ottenuto è stato calcolato il valore medio, la deviazione standard e il massimo dei valori dei pixel delle bande originali e dei due indici di attività foto sintetica NDVI e SAVI. L'area di studio della sperimentazione è stata definita sulla base della cartografia dell'area percorsa dal fuoco realizzata a partire da un'immagine Ikonos acquisita nell'agosto del 2000 [Chirici et al., 2001].

Stima degli indici di rinnovazione mediante immagini telerilevate

Al fine di stimare la relazione tra gli indici di rinnovazione misurati a terra e i dati ancillari ottenuti mediante telerilevamento per ottenere una mappatura sull'area percorsa dall'in-

condio sono stati applicati due approcci non parametrici: il primo basato sull'algoritmo *k-Nearest Neighbors*, il secondo basato su reti neurali. In un precedente studio, limitatamente ai dati del 2003 [Chirici e Corona, 2005], lo stesso risultato era stato ottenuto attraverso l'applicazione di tradizionali metodi parametrici multiregressivi.

k-Nearest Neighbors

Il metodo *k*-NN permette di stimare in modo non parametrico il valore di una variabile *Y* per gli *N* elementi di una popolazione per la quale sia noto il valore vero di variabili ausiliarie correlate con *Y*, posto che per un campione di *n* elementi della popolazione sia noto anche il valore vero di *Y*.

Nella sperimentazione la popolazione è costituita da tutti gli *N* pixel delle immagini telerilevate non coperti da nuvole (4237000 pixel per l'immagine Ikonos del 2003 e di 3564000 pixel dell'immagine Quick Bird del 2006), la variabile *Y* è di volta in volta uno degli indici di rinnovazione naturale misurati a terra in corrispondenza delle *n* unità del campione (detto *reference set*) (15678 pixel per il 2003; 12356 pixel per il 2006). Per tutti gli *N* pixel sono noti i valori delle variabili ausiliarie rappresentate dai valori spettrali delle immagini telerilevate utilizzate o da indici ottenuti dalla combinazione di queste ultime.

Il valore incognito della variabile *Y* per ciascuna *j*-esima unità dell'insieme *N-n* (detto *target set*) può essere stimato come media pesata dei valori di *Y* misurati in corrispondenza delle *k* unità del *reference set* più vicine alla *j*-esima unità nello spazio multidimensionale definito dalle variabili ausiliarie:

$$\tilde{y}_j = \frac{\sum_{i=1}^k w_{ij} y_i}{\sum_{i=1}^k w_{ij}} \quad [5]$$

dove il peso *w* può essere posto pari a $1/k$ (in questo caso il valore $\tilde{y}_j$ è pari alla media aritmetica dei valori di *Y* misurati nelle *k* unità del *reference set* più vicini alla *j*-esima unità) o, come avviene più frequentemente, può essere calcolato in modo inversamente proporzionale alla distanza multidimensionale tra la *j*-esima unità e ciascuna delle *k* unità del *reference set* a essa più vicine.

Come accennato, la distanza multidimensionale è misurata nello spazio definito dalle variabili ausiliarie considerate e può essere calcolata attraverso diversi tipi di misure: nella presente sperimentazione sono state utilizzate la distanza euclidea [De Maesschalck et al., 2000], quella di Mahalanobis [Holmstrom et al., 2001; Maselli et al., 2005; Bertini et al., 2007] e una distanza *fuzzy* modificata [Maselli, 2001].

All'aumentare di *k* tende in genere ad aumentare l'accuratezza della stima $\tilde{y}_j$. Con valori alti di *k* la variabilità nei valori stimati per le singole unità tende però a essere minore di quella reale a causa dell'effetto di livellamento prodotto dalla ponderazione. In genere, per la maggior parte di applicazioni di questa procedura su immagini telerilevate ad alta risoluzione possono essere orientativamente consigliati valori di *k* compresi tra 1 e 15.

La scelta delle variabili ausiliarie, del tipo di distanza multidimensionale e di *k* è in genere condotta empiricamente attraverso una procedura *leave-one-out* (LOO) di valutazione dell'accuratezza delle stime prodotte. Questo tipo di procedura prevede la stima mediante *k*-NN del valore della variabile *Y* per ciascuna *j*-esima unità del *reference set* con l'accortezza di escludere, ai fini della stima stessa, il valore vero y_j corrispondente a quell'unità: si ottengono così *n* valori stimati $\tilde{y}_j$ che confrontati con i corrispondenti valori veri y_j permettono di valutare l'accuratezza delle stime prodotte.

Sulla base dei risultati della procedura LOO viene definita la configurazione localmente ottimale dell'algoritmo k -NN (in termini di variabili ausiliarie considerate, tipo di distanza multidimensionale, valore di k) capace di fornire le stime più accurate sulle unità del *target set*.

Reti neurali

Le reti neurali sono state applicate a un vasto campo di problemi di approssimazione, regressione e classificazione. Il modello più utilizzato è il *Multi-Layer Perceptron* (MLP) [Hertz et al., 1991], per il quale è stata sviluppata un ampio bagaglio di conoscenze teoriche (in particolare la dimostrazione che esso costituisce un approssimatore universale di funzioni) e di esperienze pratiche.

Il MLP stima una generica i -esima unità dell'insieme N - n (detto *target set*) attraverso una funzione nonlineare del vettore delle variabili ausiliarie usate in ingresso:

$$\tilde{y}_i = b^{(2)} + \sum_{j=1}^{n_{hid}} w_j^{(2)} \tanh \left[b_j^{(1)} + \sum_{k=1}^{n_{in}} w_{jk}^{(1)} x_{ik} \right] \quad [6]$$

In questa formula x_{ik} indica la k -esima variabile ausiliare delle n_{in} utilizzate, presa in corrispondenza del punto in cui si desidera stimare la $\tilde{y}_i$, n_{hid} è il numero dei j -esimi neuroni nascosti, che determina la complessità del modello, e le quantità w e b sono parametri del modello che vengono ottimizzati attraverso una procedura di apprendimento che utilizza i dati del *reference set*.

Il problema principale nell'uso del MLP è legato alla selezione del modello, che consiste essenzialmente, facendo riferimento a reti con uno strato nascosto, alla selezione del numero di neuroni appartenenti a tale strato. L'obiettivo di tale ottimizzazione è il miglior compromesso fra la capacità di generalizzazione (migliore quando i neuroni nascosti sono relativamente pochi) e l'errore commesso sull'insieme degli esempi di apprendimento (che migliora con il numero di neuroni nascosti). Tale compromesso può essere ottimizzato attraverso una procedura di validazione incrociata, in particolare (specialmente in casi come quello considerato, in cui l'insieme di apprendimento è piccolo) di LOO.

Tuttavia, il metodo LOO è molto costoso dal punto di vista computazionale, e per questo motivo è stata sviluppata una tecnica che consente di ottenere essenzialmente lo stesso tipo di risultato senza eseguire il LOO. Si tratta del *Virtual Leave-One-Out* (VLOO) [Monari e Dreyfus, 2000], applicato in questo lavoro.

Il VLOO funziona valutando una quantità chiamata "leva" (*leverage*) per ciascun esempio di apprendimento. La leva fornisce una valutazione di quanto l'esempio corrispondente è stato importante nell'ambito dell'apprendimento, attraverso il calcolo di uno jacobiano dei parametri stimati rispetto agli esempi, e in questo modo stima (con un procedimento esatto al primo ordine) l'errore che verrebbe commesso se tale esempio venisse omissso, senza bisogno di rifare l'apprendimento tante volte quante sono gli esempi stessi.

Il secondo modello neurale utilizzato è basato su unità che calcolano funzioni di tipo "kernel", ovvero "*radial basis functions*". Si tratta in particolare della *Relevance Vector Machine* (RVM) [Tipping, 2001], che determina i parametri della rete attraverso una ottimizzazione di tipo probabilistico, basata sulla teoria di Bayes. In questo caso la funzione approssimante assume il seguente aspetto:

$$\tilde{y}_i = w_{(1)} + \sum_{j=1}^{d_{RV}} w_{j+1} \exp \left[-\frac{1}{r^2} \sum_{k=1}^{d_X} \left(\frac{x_i^{(k)} - v_j^{(k)}}{\sigma^{(k)}} \right)^2 \right] \quad [7]$$

in cui i j -esimi vettori v sono alcuni dei vettori x appartenenti al *reference set*, scelti opportunamente dall'algoritmo di apprendimento, e d_{rv} è il numero di questi vettori, detti rilevanti, e determina la complessità del modello. Il parametro r determina l'ampiezza del dominio significativo della funzione gaussiana utilizzata come base dell'approssimazione (nucleo). Anche in questo caso i parametri w vengono ottimizzati dalla procedura di apprendimento, e si è scelto di normalizzare le variabili di ingresso alla loro deviazione standard $\sigma(k)$ per migliorare l'efficacia del metodo.

Il vantaggio di tale approccio consiste nell'automatizzare la decisione sulla base di funzioni approssimanti, che vengono a costituire un insieme che ottimizza il compromesso fra accuratezza e generalizzazione nel senso della massima verosimiglianza. In questo modo, l'unico parametro che deve essere scelto attraverso una opportuna euristica è l'ampiezza r delle funzioni "*radial-basis*". Le funzioni approssimanti sono scelte dall'algoritmo fra tutte quelle possibili centrate sui singoli esempi di apprendimento. La scelta dell'ampiezza delle funzioni di base può essere basata sulla valutazione di una funzione di verosimiglianza come indicato dallo stesso Tipping, oppure su una preliminare procedura LOO.

Messa a punto di un software per la spazializzazione mediante tecniche non-parametriche

A causa delle mancanza di un *software* specifico per è stato implementato. È stato sviluppato uno specifico pacchetto software dedicato all'applicazione di algoritmi k -NN e di reti neurali *Spatial Forest Modeller* (SFM) che permette l'applicazione di queste tecniche attraverso apposita interfaccia grafica.

SFM è sviluppato in ambiente Windows e utilizza il motore di gestione di informazioni geografiche messo a disposizione da IDRISI. La versione del *software* SFM realizzata specificatamente per il progetto prende il nome di "*MEDIFIRE edition*" in quanto permette l'elaborazione esclusivamente del set di dati (rilievi a terra e immagini telerilevate) acquisito nella pineta di Castel Fusano.

SFM si presenta con un'interfaccia utente che presenta nella schermata iniziale il *database* di informazioni acquisite nei diversi transect rilevati a terra. La schermata permette la selezione di un campo contenente la variabile *target* oggetto di stima e n campi contenenti le informazioni in *input* derivate dalle immagini telerilevate.

Una volta definito il set di dati su cui lavorare è possibile selezionare i due moduli presenti: quello destinato alla k -NN e quello basato sulla creazione di una rete neurale.

Il modulo di k -NN è basato sul motore del *software forest k-nn* sviluppato da BIOFOR ITALY srl, *spin off* accademico dell'Università degli Studi della Toscana, su una procedura in due step. Nella prima fase la configurazione dell'algoritmo viene ottimizzata tramite *leave-one-out* (LOO) permettendo la scelta del valore k e del tipo di distanza spettrale. Una volta definito l'algoritmo localmente ottimizzato viene applicato a tutti i pixel dell'immagine in analisi. Il risultato è costituito da una matrice raster contenente per ogni pixel la stima della variabile *target*.

Il modulo di rete neurale che implementa il MLP esegue una preliminare valutazione su un opportuno campo di possibili valori del numero di neuroni nascosti, valutando mediante VLOO due parametri che stimano da una parte l'accuratezza (errore di VLOO, che costituisce una stima attendibile dell'errore di LOO) e dall'altra la capacità di generalizzazione (attraverso un parametro che misura la relativa uniformità delle leve).

Per quanto riguarda l'algoritmo RVM, il modulo implementato valuta preliminarmente il parametro di verosimiglianza associato ad un opportuno campo di possibili scelte dell'ampiezza delle funzioni di base, in modo da permetterne una rapida scelta euristica, per poi eseguire la regressione. Parallelamente, mostra la cardinalità dell'insieme di funzioni di base, che è indice dell'adeguatezza della capacità di generalizzazione (quando il numero delle funzioni è significativamente minore del numero di esempi).

Protocollo sperimentale

k-Nearest Neighbors

Attraverso il metodo LOO sono state sperimentalmente valutate le migliori configurazioni *k*-NN per la stima delle quattro variabili *target* (*IRGl*, *IRGc*, *IRAl* e *IRAal*), per valori di *k* compresi tra 1 e 10, per i due anni di studio, per i quattro metodi di estrazione dei valori spettrali (singolo pixel, finestra 3x3, 5x5, 7x7 e per segmentazione) sulla base di un set standard di variabili in *input* costituito dai valori medi delle quattro bande spettrali disponibili. Complessivamente sono state testate 80 configurazioni *k*-NN per ognuna delle quattro variabili *target*, per ogni configurazione tramite LOO è stata calcolata l'accuratezza della stima sulla base dell'indice di correlazione *r* di Pearson e il valore di Errore Quadratico Medio (RMSE) tra valori veri e valori stimati.

Per ogni variabile *target* sono state selezionate le migliori configurazioni *k*-NN in termini di *k* e metodo di estrazione dei valori spettrali. Queste sono state ulteriormente ottimizzate in modo reiterato tramite LOO in modo da definire le migliori variabili in *input*. Le migliori configurazioni *k*-NN sono state quindi applicate per derivare una stima per ognuna delle variabili *target*. L'accuratezza delle stime è stata valutata confrontando i valori veri con i valori stimati.

Reti neurali

Per la scelta delle variabili di ingresso da utilizzare per i modelli neurali, è stata effettuata una preliminare valutazione euristica esplorando in modo non esaustivo le possibili scelte, iniziando da singole variabili e aggiungendone progressivamente alle migliori trovate (nel senso dell'errore di LOO). Da questo tipo di esperienze, si è evidenziato che normalmente non è opportuno utilizzare più di due o tre variabili di ingresso. Considerato il ridotto numero di esempi, e la loro elevata variabilità, questo tipo di osservazione è del tutto coerente con gli abituali criteri euristici.

Si è quindi proceduto valutando le prestazioni dell'algoritmo su singole variabili di ingresso selezionando arbitrariamente il criterio di estrazione di finestre 3x3 e scegliendo la migliore, esplorando poi su quella variabile il criterio di estrazione migliore. Successivamente si è tentata l'aggiunta di un'altra variabile procedendo in modo simile, finché le soluzioni trovate mostravano un miglioramento apprezzabile.

Parallelamente, si è tentata l'utilizzazione delle variabili che presentano la massima correlazione con la variabile da stimare nei punti noti. Si è osservato euristicamente che, nonostante questo tipo di scelta non sia sempre la migliore possibile, quando è possibile trovare insiemi di variabili di ingresso migliori la differenza non è tuttavia significativa. D'altra parte, quest'ultimo criterio di scelta è completamente automatizzabile e comporta un solo tentativo in luogo di un numero molto elevato e comunque mai esaustivo (a meno di esplorare una quantità di tentativi proibitiva).

Risultati

Per quanto riguarda l'applicazione dell'algoritmo *k-Nearest Neighbors* la procedura di ottimizzazione dell'algoritmo tramite LOO ha permesso di escludere a priori la stima di alcune delle variabili *target*: l'indice *IRGc* per la limitata disponibilità di dati e *IRAl* in quanto la sua stima si è sempre dimostrata peggiore di quella *IRAal*.

La stima è stata quindi realizzata per *IRGl* e *IRAal*. Sulla base della procedura di LOO le configurazioni *k*-NN localmente dimostratesi migliori per la stima delle due variabili *target* per i due anni di studio sono riportate in Tabella 1. La relazione tra l'indice di correlazione di Pearson e i valori di *k* per i diversi tipi di distanze multidimensionali testate sono riportate nelle Figure da 2 a 5.

Tabella 1 - Configurazione localmente ottimali dell'algoritmo di stima k -NN determinati tramite tecnica di *leave-one-out* per la stima degli indici di rinnovazione IRAal e IRG, per i due anni di studio 2003 e 2006. La colonna "metodo" si riferisce alla metodologia di estrazione dei dati spettrali: "segm" si riferisce alla tecnica di segmentazione automatica delle immagini, 7x7 e 5x5 all'estrazione tramite tradizionale finestra mobile. La colonna " k " indica il valore di k scelto per l'implementazione dell'algoritmo k -NN. Le variabili di *input* sono denominate componendo l'identificazione della banda (o dell'indice di attività fotosintetica) con l'operatore di aggregazione dei valori spettrali calcolati nelle finestre mobili o nei poligonati originati per segmentazione (per esempio nel caso dell'IRG del 2003 NDVIstd indica che all'interno della finestra di 7x7 pixel è stato calcolato il valore di deviazione standard dell'indice NDVI). R e RMSE riportano rispettivamente l'indice di correlazione di Pearson e l'errore quadratico medio calcolati tramite LOO tra valori veri e valori stimati.

Indice	Anno	Metodo	k	Variabili input	R	RMSE
IRAal	2003	segm	3	b1med, b2med, b2std, b3med, b4std, NDVImax, NDVImed	0,448	0,108
IRG	2003	7x7	4	b2med, b3std, b3med, b4med, b4std, NDVIstd, NDVImax, NDVImed	0,3523	0,671
IRAal	2006	7x7	4	b1med, b1std, b2med, b4med, NDVImed, SAVImed	0,5104	0,109
IRG	2006	segm	4	b1med, b3med, b4med, NDVImed, SAVImed	0,5259	0,3557

Figura 2 - Valori di R calcolati tramite LOO con metodo k -NN con diversi valori di k per l'indice IRAal al 2003 stimato secondo i parametri riportati in Tabella 1.

Figura 3 - Valori di R calcolati tramite LOO con metodo k -NN con diversi valori di k per l'indice IRG al 2003 stimato secondo i parametri riportati in Tabella 1.

Figura 4 - Valori di R calcolati tramite LOO con metodo k -NN con diversi valori di k per l'indice IRAal al 2006 stimato secondo i parametri riportati in Tabella 1.

Figura 5 - Valori di R calcolati tramite LOO con metodo k -NN con diversi valori di k per l'indice IRGI al 2006 stimato secondo i parametri riportati in Tabella 1.

Le migliori configurazioni k -NN (Tab. 1) sono state applicate per la stima dei due indici per i due anni di studio ottenendo risultati con valori di RMSE percentuale nel reference set del 42% per IRAal al 2003 (Fig. 6) e del 38% al 2006 (Fig. 7) e del 13% per IRGI al 2003 (Fig. 8) e del 27% al 2006 (Fig. 9).

Per quanto riguarda l'applicazione delle reti neurali la rete di tipo RVM ha permesso la stima solo per l'indice IRGI (Tab. 2) ottenendo valori di RMSE percentuali dell'63% per il 2003 (Fig. 10) e del 39% per il 2006 (Fig. 11). La rete di tipo MLP ha invece permesso la stima solo di IRGI (Tab. 3) al 2006, con un RMSE del 41%.

Tabella 2 - Risultati ottenuti con la RVM (stesso significato di simboli e notazioni della Tab. 1). Per IRG al 2006 sono mostrati due risultati, il primo ottenuto con una ricerca empirica della variabile di ingresso migliore, il secondo scegliendo gli ingressi in base alla massima correlazione con la variabile da stimare, come fatto per i dati 2003.

Indice	Anno	Metodo	Variabili input	R	RMSE
IRG	2003	segm	b1max, b4max, NDVI _{max}	0,439	0,654
IRG	2006	3×3	SAVI _{max}	0,538	0,419
IRG	2006	7×7	b1 _{sd} , NDVI _{med}	0,580	0,418

Figura 6 - Risultato della stima dell'indice IRAal per l'anno 2003 ottenuto tramite algoritmo k -NN. Per la configurazione dell'algoritmo si veda la Tabella 1. Sullo sfondo in toni di grigio l'indice NDVI calcolato dall'immagine Quick Bird.

Figura 7 - Risultato della stima dell'indice IRAal per l'anno 2006 ottenuto tramite algoritmo k -NN. Per la configurazione dell'algoritmo si veda la Tabella 1. Sullo sfondo in toni di grigio l'indice NDVI calcolato dall'immagine Ikonos.

Figura 8 - Risultato della stima dell'indice IRGI per l'anno 2003 ottenuto tramite algoritmo k -NN. Per la configurazione dell'algoritmo si veda la Tabella 1. Sullo sfondo in toni di grigio l'indice NDVI calcolato dall'immagine Quick Bird.

Figura 9 - Risultato della stima dell'indice IRGI per l'anno 2006 ottenuto tramite algoritmo k -NN. Per la configurazione dell'algoritmo si veda la Tabella 1. Sullo sfondo in toni di grigio l'indice NDVI calcolato dall'immagine Ikonos.

Figura 10 - Risultato della stima dell'indice IRGI per l'anno 2003 ottenuto tramite algoritmo RVM. Sullo sfondo in toni di grigio l'indice NDVI calcolato dall'immagine Quick Bird.

Figura 11 - Risultato della stima dell'indice IRGI per l'anno 2006 ottenuto tramite algoritmo RVM. Sullo sfondo in toni di grigio l'indice NDVI calcolato dall'immagine Ikonos.

Tabella 3 - Risultati ottenuti con MLP (stesso significato di simboli e notazioni della Tab. 1).

Indice	Anno	Metodo	Variabili input	R	RMSE
IRG	2006	3x3	NDVImax, SAVImax	0,394	0,440

Discussione e conclusioni

I rilievi di campagna realizzati nella pineta di Castel Fusano hanno permesso di valutare la dinamica temporale del fenomeno della rinnovazione naturale innescatosi in seguito all'incendio del 2000 sulla base delle variazioni degli indici di rinnovazione gamica e agamica dei transetti su cui è stato effettuato il rilievo sia nella campagna del 2003 che in quella del 2006.

Complessivamente a livello aggregato si evidenzia un trend di crescita della rinnovazione agamica generatasi successivamente all'incendio. Sulla base dei rilievi (e quindi sul campione di valori veri del *reference set* del 2006 per i quali esistono anche le misure del 2003) mediamente i valori dell'indice di rinnovazione agamico delle latifoglie (*IRGI*) è aumentato del 60% rispetto a quelli del 2003 e del 79% considerando *IRGal*. L'indice di rinnovazione gamico delle latifoglie è invece diminuito, in media, del 22%. Tendenzialmente i transetti che dimostrano una più elevata rinnovazione gamica delle latifoglie dimostrano anche una più marcata rinnovazione agamica: come atteso i due indici di rinnovazione agamica *IRGI* e *IRGal* tendono a essere correlati.

La relazione tra i dati telerilevati multispettrali e gli indici di rinnovazione si è dimostrata relativamente debole. A titolo esemplificativo, la correlazione tra l'indice di attività fotosintetica NDVI e i diversi indici di rinnovazione adottati non ha mai superato valori di r di Pearson pari a 0,22 per le immagini del 2003 e a 0,14 per quelle del 2006 (in entrambi gli anni tali valori sono riferiti all'indice *IRGI*). La debole relazione tra misure multispettrali telerilevate e dati misurati a terra è probabilmente dovuta a due fattori: (i) per l'attività fotosintetica dello strato erbaceo che ha velocemente ricolonizzato l'area incendiata e che tende a confondersi con la risposta spettrale dovuta all'attività fotosintetica della rinnovazione e (ii) per le ridotte dimensioni dei transetti che non permettono l'estrazione di valori spettrali esclusivamente riferiti ai soli pixel afferenti alle aree misurate rilevate a terra (questo effetto è particolarmente vero per l'anno 2006 per il quale è stata utilizzata un'immagine Ikonos che ha risoluzione geometrica più ridotta di quella Quick Bird usata per il 2003).

Nonostante tali limitazioni i metodi non parametrici hanno comunque permesso la derivazione di cartografie tematiche dei diversi tipi di rinnovazione nell'area incendiata con accuratezze comparabili con quelle riportate per altri attributi tradizionalmente applicati in campo forestale [Chirici et al., 2007].

I risultati ottenuti nel 2003 sono in genere migliori rispetto a quelli del 2006, indipendentemente dalla tecnica utilizzata e dall'indice di rinnovazione testato. Tale risultato è probabilmente dovuto alla più ridotta risoluzione geometrica delle immagini telerilevate impiegate e, in secondo luogo, alla più elevata numerosità campionaria dei rilievi.

Per quanto riguarda il confronto tra il metodo k -NN e le due reti neurali utilizzate, il primo ha permesso di ottenere sempre un risultato cartografico, anche se in alcuni casi il modesto livello di accuratezza finale ottenuto non permette un impiego operativo del prodotto finito. Le reti, nel caso di dati poco significativamente correlati, tendono a non portare ad alcun risultato. Quando riescono invece a completare il loro iter di apprendimento, le reti neurali conducono a risultati di accuratezza paragonabile, in termini di RMSE nella prova LOO, a quelli ottenuti con k -NN.

L'accuratezza dei risultati prodotti dipende da condizioni locali e per questo le procedure di ottimizzazione sono scarsamente generalizzabili. La miglior tecnica consigliata è dunque

quella euristica testando tutte le possibili configurazioni degli algoritmi in modo da individuare il metodo localmente ottimizzato, ovvero quello capace di ottenere la stima con il più elevato livello di accuratezza. In tal senso la possibilità di disporre di un software come *Spatial Forest Modeller* che permette di testare automaticamente tutte le diverse configurazioni di stima risulta essenziale per rendere operativamente utilizzabili questi strumenti di stima di tipo non-parametrico.

In conclusione le stime prodotte nella sperimentazione per i diversi indici di rinnovazione non sono sufficientemente accurate per poter essere utilizzate a livello di singolo pixel. Le cartografie tematiche realizzate possono comunque essere di supporto generale delle attività di recupero in corso nella zona della pineta di Castel Fusano percorsa dal fuoco.

Bibliografia

- Agrimi M., Bollati S., Giordano E., Portoghesi L. (2002) - *Struttura dei popolamenti e proposte di gestione per le pinete del litorale romano*. L'Italia Forestale e Montana 3: 244-260.
- Baatz M., Benz U., Dehghani S., Heynen M., Höltje A., Hofmann P., Lingenfelder I., Mimler M., Sohlbach M., Weber M., Willhauck G. (2004) - *eCognition user guide*. Definies imaging GmbH, München, Germany.
- Bertini R., Chirici G., Corona P., Travaglini D. (2007) - *Confronto di metodi parametrici e non-parametrici per la spazializzazione della provvigione legnosa tramite integrazione di misure a terra, dati telerilevati e informazioni ancillari*. Forest@ 4: 110-117.
- Chirici G., Barbati A., Corona P., Marchetti M., Travaglini D., Maselli F., Bertini R. (2007) - *Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and mediterranean forest ecosystems*. Remote Sens. Environ., in press.
- Chirici G., Corona P. (2005) - *An overview of passive remote sensing for post-fire monitoring*. Forest@ 3: 282-289.
- Chirici G., Corona P., Dellisanti R.M., Di Giovine M., Marchetti M., Rossini P., Travaglini D. (2001) - *Confronto e integrazione di dati telerilevati Ikonos e Landsat 7 ETM+ nella valutazione dei danni da incendio in ambiente Mediterraneo: il caso della Pineta di Castel Fusano*. Atti V Conferenza ASITA, La qualità nell'informazione geografica, Rimini, pp. 529-537.
- Corona P., Marchetti M. (2002) - *Indagini condotte dall'Accademia Italiana di Scienze Forestali nel comprensorio boschivo della pineta di Castel Fusano, a seguito dell'incendio del 4 luglio 2000*. In: Blasi C., Cignini B., Dellisanti R.M., Montagna P. (Eds.), Il recupero ambientale della pineta di Castel Fusano, studi e monitoraggi. Palombi Editori, Roma, pp. 18-23.
- De Maesschalck R., Jouan-Rimbaud D., Massart D.L. (2000) - *The Mahalanobis distance*. Chemometrics and Intelligent Laboratory System 50: 1-18.
- Hertz J.A., Palmer R.G., Krogh A (1991). *Introduction to the Theory of Neural Computation*. Santa Fe Institute Studies in the Sciences of Complexity.
- Holmstrom H., Nilsson M., Ståhl G. (2001) - *Simultaneous estimations of forest parameters using aerial photograph-interpreted data and the k nearest neighbor method*. Scandinavian Journal of Forest Research 16: 67-78.
- Huete A.R. (1988) - *A Soil-Adjusted Vegetation Index (SAVI)*. Remote Sens. Environ., 25: 295-309.
- Mantero F.M. (1995). *Dalla selva al bosco. Un ambiente, la sua storia e i suoi abitanti*. In: Comune di Anzio - WWF Italia, P.d.R. (Ed.), Tor Caldara. Dalla selva al bosco, un ambiente, la sua storia, i suoi abitanti. Casa editrice Viella, pp. 7-111.

- Maselli F. (2001) - *Extension of environmental parameters over the land surface by improved fuzzy classification of remotely sensed data*. Int. J. Remote Sens., 17: 3597-3610.
- Maselli F., Chirici G., Bottai L., Corona P., Marchetti M. (2005) - *Estimation of Mediterranean forest attributes by the application of k-nn classifiers to multitemporal Landsat ETM+ images*. Int. J. Remote Sens., 26: 3781-3796.
- Monari G., Dreyfus G. (2000) - *Local Overfitting Control via Leverages*. Neural Computation, 14: 1481-1506.
- Tipping M.E. (2001) - *Sparse Bayesian Learning and the Relevance Vector Machine*. Journal of Machine Learning Research, 1: 211-244

Manoscritto ricevuto il 21/11/2007, accettato il 9/01/2008.