

Sezione Speciale: Atti 5° Congresso SISEF: Foreste e Società - Cambiamenti, Conflitti, Sinergie
(a cura di: E. Lingua, R. Marzano, G. Minotta, R. Motta, A. Nosenzo, G. Bovio)

Metodi semplificati di spazializzazione di dati campionari: un'applicazione ai dati di prima fase dell'INFC in Sicilia

Cullotta S^{(1)*}, La Mela Veca DS⁽¹⁾, Maccari F⁽²⁾, Cibella R⁽³⁾, Marchetti M⁽⁴⁾

(1) Dipartimento di Colture Arboree, Università degli Studi di Palermo, v.le delle Scienze 11, 90128 Palermo; (2) v. Dante 145 - 92100 Agrigento; (3) Dipartimento Foreste, Regione Siciliana, v.le della Regione Siciliana 2246 - 90135 Palermo; (4) Lab. EcoGeoFor - Dipartimento di Scienze e Tecnologie per l'Ambiente e il Territorio (STAT), Università del Molise, v. Mazzini 8 - 86170 Isernia. - *Corresponding author: Sebastiano Cullotta - cullotta@unipa.it

Abstract: Simplified methods for spatial sampling: application to first-phase data of Italian National Forest Inventory (INFC) in Sicily. Methodological approaches able to integrate data from sample plots with cartographic processes are widely applied. Based on mathematic-statistical techniques, the spatial analysis allows the exploration and spatialization of geographic data. Starting from the punctual information on land use types obtained from the dataset of the first phase of the ongoing new Italian NFI (INFC), a spatialization of land cover classes was carried out using the Inverse Distance Weighting (IDW) method. In order to validate the obtained results, an overlay with other vectorial land use data was carried out. In particular, the overlay compared data at different scales, evaluating differences in terms of degree of correspondence between the interpolated and reference land cover.

Keywords: Classification, Corine Land Cover, Geographic Information Systems - GIS, Inverse Distance Weighting - IDW, Land cover class, Forest Inventory.

Received: Jan 10, 2006 - Accepted: Jun 11, 2006

Citation: Cullotta S, La Mela Veca DS, Maccari F, Cibella R, Marchetti M, 2006. Metodi semplificati di spazializzazione di dati campionari: un'applicazione ai dati di prima fase dell'INFC in Sicilia. Forest@ 3 (3): 407-419. [online] URL: <http://www.sisef.it/>

Introduzione

La necessità di ottenere informazioni sulle caratteristiche dei popolamenti arborei e su analoghi usi del suolo, in campo ambientale e forestale, limitando il rilievo degli attributi in campo perché molto oneroso, ha spinto ormai da tempo la ricerca verso l'adozione di nuovi approcci e metodologie capaci di integrare il processo cartografico con i rilevamenti di tipo campionario, per aumentare le sinergie e i reciproci benefici.

Negli ultimi anni la procedura più usata per derivare informazioni cartografiche tematiche è stata l'interpretazione di fotogrammi aerei e telerilevati con l'uso di supporti GIS (Chirici et al. 2003a). L'ultima famiglia di classificatori in ordine di tempo e svi-

luppo è basata su algoritmi *object-oriented* (Chirici et al. 2003b): essa permette in molti casi di superare le difficoltà e i limiti mostrati dai precedenti classificatori *pixel-oriented* (con approcci *supervised* o *unsupervised*), attraverso una classificazione di poligoni vettoriali generati tramite processi di segmentazione dell'immagine grezza, dimostrando di essere promettente nei *targets* complessi come i sistemi naturali e seminaturali (Chirici et al. 2003c). In particolare, per la classificazione delle immagini con alta complessità di relazioni tra attributi e risposte spettrali delle superfici a copertura forestale, è necessario l'utilizzo di sistemi flessibili ed efficienti, come la categoria dei classificatori non parametrici, tra cui il sistema "*k-Nearest Neighbors*" (k-NN, Maselli et al.

2003). L'uso di tali metodi ha avuto applicazioni di successo dirette all'acquisizione di una migliore conoscenza sulle caratteristiche del patrimonio forestale, da tempo richiesta, e ai fini della pianificazione dell'uso delle risorse (Marchetti 1990, Ambrosini et al. 1994, Castagnoli et al. 1997).

Oggi l'uso di metodologie classiche d'inventariazione delle risorse forestali si scontra con la necessità di ottenere sia una maggiore quantità e un miglioramento nella qualità dei dati sia un dettaglio d'informazione elevato; ciò comporta una naturale ibridazione metodologica con nuove tecniche di rilievo. Il ricorso a tecniche di telerilevamento con l'uso di foto aeree per l'individuazione dei punti bosco è da sempre consuetudine in molti inventari regionali (Preto 1983, Corona & Ambrosini 1995, Tosi & Marchetti 1998). I sistemi GIS, le tecniche di telerilevamento e di rilievo in campo, i diversi sistemi di classificazione dell'immagine, concorrono e cooperano per lo stesso obiettivo, ibridandosi a diversi livelli per la corretta quantificazione delle risorse ambientali (La Mela Veca & Cullotta 2005; Clementi et al. 2006). E' dunque necessario non solo quantificare i fenomeni ma anche localizzarli con buona accuratezza su base spaziale, il che comporta costi elevati.

Lo scopo del presente lavoro è stato quello di sperimentare l'applicazione di processi semplici di spazializzazione su base geografica (Naoum & Tsanis 2004) al fine di ottenere dati aggiornati sull'uso e la copertura del suolo per la Sicilia tramite l'utilizzo e l'adattamento di alcune metodologie e algoritmi comunemente presenti nei più comuni *software GIS* senza ricorrere all'uso di sistemi di *image processing*. Nello specifico è stata utilizzata una procedura di analisi spaziale *IDW* (*Inverse Distance Weighting* - Watson & Philip 1985, Ware et al. 1991, Watson 1992, Bartier & Keller 1996) per estrapolare da singole unità campionarie informazioni quantitative di superficie sulla copertura forestale; l'obiettivo è quello di ottenere una stima delle tipologie di uso del suolo per ogni *pixel* dell'immagine *raster* ottenuta quale risultato dell'operazione. Il dato puntuale di partenza, originato dai risultati di prima fase dell'INFC (Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio - INFC 2005, Tosi & Monteccone 2004), rappresenta il *set* di dati campione (*dataset*). L'analisi di tutti i dati è stata realizzata con l'utilizzo di un semplice protocollo che fa uso di metodi di analisi spaziale implementati in ambiente *GIS*.

Nello specifico si è proceduto attraverso le seguenti fasi:

- analisi dei dati preliminari;

- adattamento della metodologia di spazializzazione e individuazione dei parametri d'impostazione;
- creazione di tematismi *raster* di uso del suolo secondo i parametri INFC;
- validazione dei dati elaborati tramite il confronto con altre fonti informative cartografiche a scala e dettaglio differenti.

Materiali e metodi

Protocollo inventariale e dati preliminari

Il nuovo Inventario Nazionale per le Foreste e i Serbatoi di Carbonio - INFC, si propone l'obiettivo di aggiornare e approfondire la conoscenza del patrimonio forestale del territorio Italiano, secondo uno schema di analisi che si articola in tre fasi distinte. La prima fase è consistita nella individuazione del tipo di copertura del suolo delle unità di campionamento di circa 300000 punti (in Sicilia 25709 punti) della maglia inventariale derivanti da campionamento sistematico, con una densità di un punto per km² e ogni punto localizzato in modo casuale all'interno di ogni cella (fig. 1), in linea con quanto previsto dallo schema inventariale (INFC 2003, 2005).

I punti d'interesse inventariale sono stati individuati secondo un sistema di classificazione di copertura del suolo con due livelli di approfondimento (classe e sottoclasse), ottenuta per quanto possibile tramite fotointerpretazione a video della copertura forestale su ortofoto digitali b/n (AGEA 2000-2002). L'attribuzione delle classi e sottoclassi, secondo lo schema INFC, prevede in primo luogo l'individuazione, per ogni punto, del poligono omogeneo (per copertura del suolo) posto all'interno di un intorno di analisi quadrangolare di circa 2.2 ha; in questa prima fase il dettaglio al suolo fa quindi riferimento a questo intorno di analisi prestabilito.

Per assegnare una classe di copertura ai poligoni individuati, secondo le definizioni *FRA 2000* (UN-FAO 2001), sono stati considerati struttura (forma, dimensioni > 5000 m²), tessitura (tipo di copertura) e tono (densità e tipo di copertura - UN-FAO 2001, INFC 2003 e INFC 2005).

La prima fase dell'INFC rappresenta il dato preliminare di partenza per l'elaborazione della spazializzazione: si propone una trasposizione dell'informazione puntuale dell'uso del suolo ad una superficie. Il campionamento di prima fase INFC ha un errore generale contenuto, le stime provvisorie delle formazioni forestali, comprensive di aree temporaneamente prive di soprassuolo e degli impianti di arboricoltura da legno, al netto delle superfici incluse

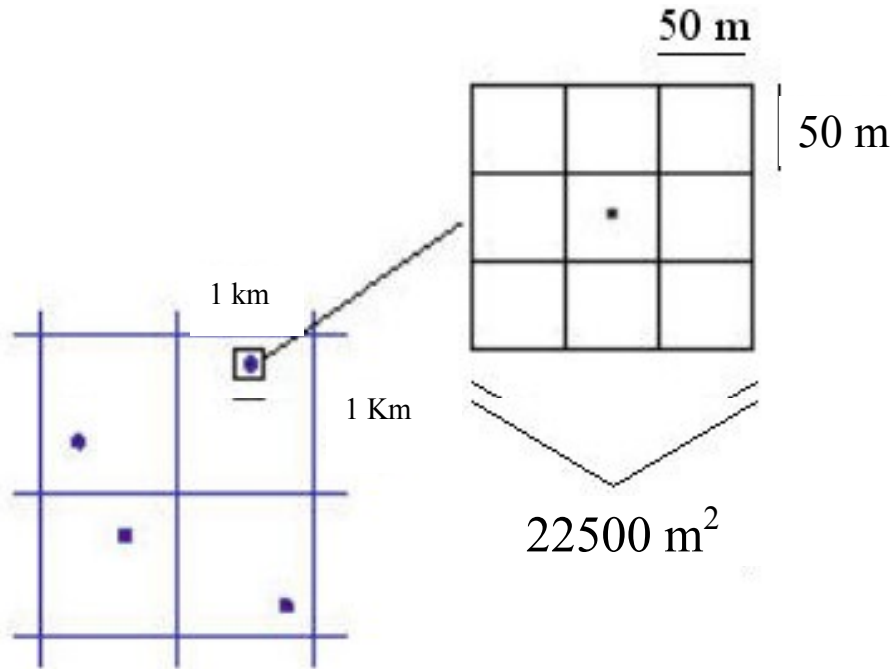


Fig. 1 - Disegno di campionamento sistematico. Dislocazione casuale dei punti all'interno di maglie di un km² (in basso a sinistra) ed intorno di analisi di superficie complessiva di 22500 m² ripartito in 9 quadranti ciascuno di 2500 m².

di varia natura entro le aree boscate, risultano corredate di incertezze campionarie decisamente contenute; per la Sicilia l'errore standard calcolato è 1.6 (INFC 2005).

Analisi spaziale

Per la procedura di analisi spaziale, si è fatto ricorso ad un metodo di interpolazione di dati per estendere l'informazione puntuale dei dati INFC all'intera superficie analizzata, utilizzando una procedura di "data-mining" (Berry & Linoff 1997) che consente di ottenere dal dato grezzo dati più o diversamente leggibili secondo altre chiavi di analisi.

L'interpolazione spaziale, attraverso un operatore software, fornisce una stima del valore di un attributo (in tal caso uso del suolo) in un sito (dato puntuale) in cui tale valore non è conosciuto. L'interpolazione spaziale si basa sulla legge di Tobler (1979): "tutti i siti sono correlabili e gli attributi dei siti vicino hanno valori maggiormente correlabili; questo significa che la stima di un valore di un attributo in un determinato punto è funzione dal valore degli attributi misurati nei punti più vicini.

I dati necessari per l'interpolazione possono essere dati puntuali provenienti da diverse fonti. Nel presente lavoro, l'obiettivo è stato quello di ottenere una stima di *land cover class* (di *pixel*), tramite l'interpolazione dei fotopunti disponibili della maglia inventariale INFC, ed ottenere da essi una superficie di stima (*raster*) della classe di uso e copertura del suolo laddove l'informazione mancava tramite l'uso della

metodologia *IDW*. L'interpolazione *IDW* permette la stima di valori incogniti di un attributo a partire da un *set* di dati di tipo puntuale, come media ponderata degli attributi relativi a *n* punti prossimi, assegnando pesi maggiori ai punti più vicini (Bartier & Keller 1996, Lopez et al. 2001).

L'equazione generica è (eqn. 1):

$$\hat{Z}(x, y) = \frac{\sum_{i=1}^n Z_i W_i}{\sum_{i=1}^n W_i}$$

dove $Z(x,y)$ rappresenta l'attributo del punto da stimare, Z_i il valore dell'attributo nel punto noto, W_i il peso attribuito al punto i . L'equazione più esplicitamente può essere scritta come (eqn. 2):

$$\hat{Z}(x_j) = \frac{\sum_{i=1}^n Z(x_i) \cdot d_{ij}^{-2}}{\sum_{i=1}^n d_{ij}^{-2}}$$

dove x_j è il valore dell'attributo nel punto da stimare, n il numero di punti utilizzati, x_i l'attributo del punto a valore noto e d_{ij} è la distanza fra x_j e x_i .

Il modello matematico usato è un interpolatore esatto: restituisce in fase di stima, per i punti che coincidono a quelli esistenti di valore noto, lo stesso valore misurato o osservato attribuendo ad essi peso

pari a 1 ed ai restanti peso pari a 0 (Goodchild & Lam 1980). Ciò significa che si conserva il dettaglio dell'informazione originaria. Tramite questo metodo, l'attributo di ogni cella di valore sconosciuto sulla superficie *raster* è interpolato mediando i valori relativi al numero scelto di punti più vicini o ad un insieme di punti compresi all'interno di un raggio di ricerca definito. Il risultato è una superficie *raster* i cui *pixel* avranno un attributo di valore pari a quello del punto noto, nel caso di sovrapposizione ad esso, o un valore pesato in funzione dell'inverso del quadrato della distanza tra la cella (centroide) in esame ed i punti vicini.

Il numero di punti utilizzati in questa spazializzazione per la Sicilia è di 25435, dato che l'elaborazione non ha riguardato i punti localizzati nelle isole minori (274).

Adattamento metodologico della spazializzazione

La metodologia *IDW* così come si presenta non permette di ottenere stime adeguate alla determinazione delle categorie di uso e copertura del suolo. Partendo da questo presupposto, la suddivisione in categorie del territorio siciliano secondo le specifiche INFC (per classe e sottoclasse) è stata ottenuta mediante l'utilizzo dell'indice *IDW* opportunamente adattato.

L'indice permette la stima di valori incogniti di una variabile continua, creando un vettore di risposte concentriche intorno ai punti noti (Curtis & Clyde 1999); questa caratteristica ne permette un uso nella stima di variabili che generano modelli di ri-

sposta continua e che non sono influenzate da fattori geografici legati al suolo e ad attività antropiche (Tomczak 1998), come ad esempio le precipitazioni dove tale metodologia trova largo impiego (Jolly et al. 2005).

La copertura del suolo è un attributo di tipo qualitativo e non esprime un valore numerico specifico per classe di copertura. L'impiego dell'interpolazione *IDW* implica, nei confronti delle variabili di *land cover* utilizzate, delle limitazioni all'utilizzo di questa tecnica (Burrough & McDonnell 1998); attribuire valori numerici (ad esempio una sequenza numerica 1, 2, ..., n) alle unità di una classificazione di copertura del suolo, quale quella INFC, consente di stimare il valore incognito, della copertura del suolo, nelle aree non coperte da valori e quindi non analizzate. Si è operato quindi attribuendo alle unità di copertura una semplice sequenza numerica di valori interi, legando al dato qualitativo un valore numerico di comodo.

L'interpolazione delle classi di copertura del suolo non può essere eseguita per tutte le categorie contemporaneamente. Infatti, eseguendo un'elaborazione che avesse tenuto contemporaneamente in conto tutte le classi di copertura del suolo si sarebbe commesso un errore: l'indice stima i dati incogniti creando una scala di valori ognuno dei quali fa riferimento ad una copertura del suolo specifica. Ciò significa che si sarebbe associato ad ogni attributo un valore intero tra 1 e *n* ed il risultato sarebbe stato una superficie di *trend* costituita da *pixel* con valori intermedi tra 1 e l'ennesima classe di copertura (cioè tra due

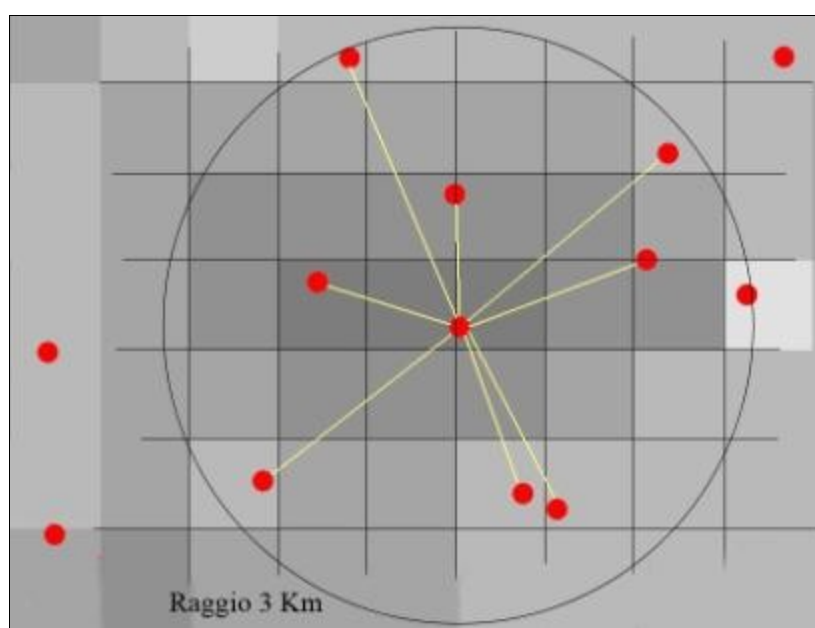
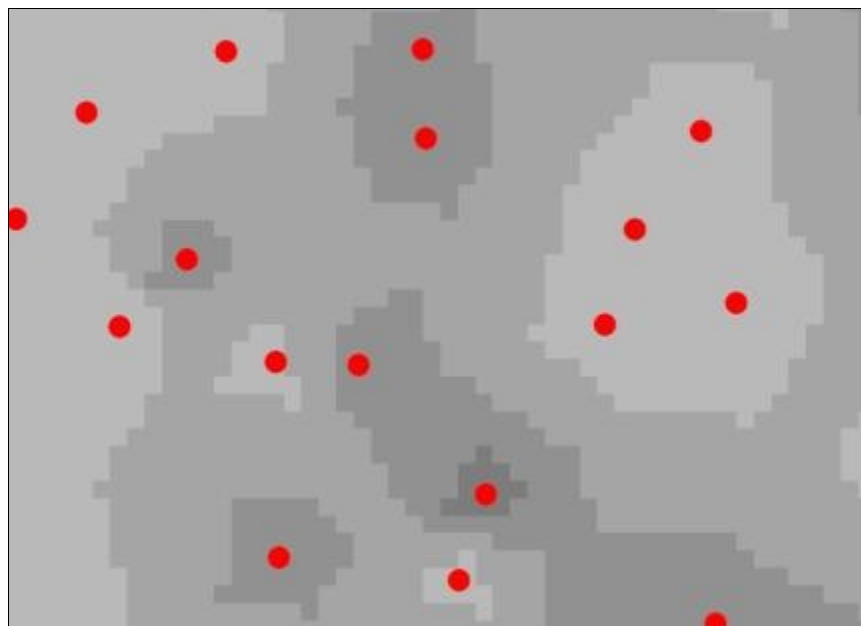


Fig. 2 - Esempio di attribuzione dei valori: i 9 punti più vicini a quello con i pixel classificati, i valori di grigio presenti si riferiscono ad una scala compresa tra 0 e 1.

Fig. 3 - Esempio di calcolo di superficie *raster*. La figura rappresenta parte della superficie globale calcolata come unione dei singoli raster interpolati per ogni unità di copertura del suolo; in particolare si evidenzia la differenza di attribuzioni di valori ai pixel (100 m) in funzione della distribuzione dei punti e dei loro differenti valori (in questo caso i valori di grigio discriminano ognuno una unità di copertura specifica).



punti vicini di valore 1 ed n l'interpolazione stimebbe valori di attributo - unità di copertura - che in realtà non esistono).

Per non incorrere in errori di stima, la procedura di spazializzazione è stata perciò eseguita valutando un solo tipo di copertura del suolo alla volta, accomunando tutti gli altri punti, pur appartenenti ad unità diverse, ad un solo tipo "fittizio" di convenienza. Quindi, per ogni classe di copertura è stata interpolata un'unità reale ed una fittizia, attribuendo loro valori di 1 (unità di copertura reale) e 0 (unità fittizia). Questo ha permesso di stimare, per la classe considerata, la superficie interessata dall'interpolazione. La superficie residua del *raster* è occupata dalle altre classi di copertura e rappresenta il *no-data*. In questo modo la superficie occupata dalla singola classe è esattamente quella che si intendeva stimare tra i punti. Ogni *raster* creato, infatti, avrà una scala di valori di *pixel* compresi tra 0 e 1, riclassificati in due valori unici (0 e 1) sulla base del valore soglia determinato dalla media aritmetica di tutti i valori calcolati.

Usando per ogni elaborazione (una per ogni classe di copertura del suolo) tutti i punti disponibili è stata mantenuta inalterata la superficie d'indagine ed il dettaglio, imposto dalla densità delle unità campionarie (un punto ogni km²).

Individuate le singole unità di superficie di uso del suolo è stata generata un'unica superficie complessiva attraverso la sovrapposizione di tutte le unità, interpolate ognuna singolarmente, ed in un secondo momento unite in un unico strato. Tale operazione non rischia di generare errori di sovrapposizione tra

superfici proprio perché sono state opportunamente utilizzati tutti i punti disponibili per la classificazione, ciò consente di generare sempre lo stesso numero di *pixel* per riga e colonna sempre nella medesima posizione.

La sovrapposizione, quindi, è stata condotta semplicemente tramite l'unione di superfici *raster* considerando solo il dato di unità (valore 1) di uso del suolo ed eliminando il resto del dato.

Per l'elaborazione dei dati è stato utilizzato l'estensione *Spatial Analyst* del software *ArcGis* 8.3 che consente di generare superfici *raster* attraverso l'indice *IDW*. Nello specifico è richiesto di settare il numero dei punti da considerare, il raggio di ricerca (che può anche essere variabile) e l'esponente.

L'algoritmo matematico di spazializzazione è stato impostato per tenere in considerazione al massimo 9 punti per volta (fig. 2, fig. 3) all'interno di un raggio di 3 km per le classi e sottoclassi forestali e agricole (maggior numero di punti) e con un raggio di 1 o 0.5 km per le altre. La valutazione del raggio è stata effettuata singolarmente per ogni interpolazione in funzione dell'entità numerica dei punti e della loro distribuzione e dispersione, diminuendo il raggio s'impone la condizione d'investigare una superficie più piccola, ciò si rende necessario per evitare una sovrastima areale dovuta alla bassa entità numerica di punti inerenti ad unità di copertura che occupano piccole superfici. In tal caso, tali unità sono meno rappresentate e riducendo il raggio s'intende attribuire loro un minor contributo relativo alla copertura complessiva.

Tab. 1 - Esempio di comparazione tra le definizioni di copertura e uso del suolo di interesse forestale tra INFC (INFC 2003) e CLC2000 (APAT 2005).

Classe INFC	Sottoclasse INFC	Categoria CLC di 3° livello
Superfici boscate e ambienti seminaturali	Formazioni forestali e Formazioni forestali rade	311 Boschi di latifoglie
		312 Boschi di conifere
		313 Boschi misti di conifere e latifoglie
		322 Brughiere e cespuglieti
		323 Aree a vegetazione sclerofilla
		324 Aree a vegetazione boschiva ed arbustiva in evoluzione
	Praterie pascoli e incolti	321 Aree a pascolo naturale e praterie
	Aree temporaneamente prive di soprassuolo	-
	Aree con vegetazione rada o assente	331 Spiagge, dune e sabbie
		331 Rocce nude, falesie, rupi, affioramenti
		332 Aree con vegetazione rada

Altro parametro richiesto per l'elaborazione è l'esponente, il quale regola il grado di "smoothing"¹: accrescendolo se ne aumenta la funzione di peso e quindi l'effetto inversamente proporzionale alla distanza. La necessità di mantenere i valori sulla superficie più uniformi possibile, in maniera che i valori stimati dei *pixel* siano più simili ai punti più vicini piuttosto che a quelli lontani, ha determinato la scelta dell'esponente 2.

Per ogni unità di copertura del suolo e per i due livelli di dettaglio (classe e sottoclasse) sono state calcolate tante superfici *raster* una per ogni unità di copertura con *pixel* di 100² (fig. 3).

I *raster* d'interpolazione relativi a ciascuna classe sono stati successivamente uniti a formare due cartografie tematiche, una relativa al livello di classe e un'altra relativa al livello di sottoclasse. Ciò ha permesso una verifica eseguendo un'operazione di "dis-

solve"³ gerarchico dal livello di sottoclasse al livello di classe, confrontando il risultato con quello ottenuto dalla spazializzazione e verificando la congruità reciproca delle superfici derivate.

Validazione con altre fonti informative cartografiche

Per verificare l'errore di stima commesso dalla procedura di spazializzazione sono stati effettuati dei controlli quantitativi. Un primo *test* è basato sul confronto percentuale, per il livello di sottoclasse (livello più dettagliato), tra la percentuale dei punti inventariati di prima fase appartenenti ad una determinata sottoclasse (dato puntuale) e quella della superficie occupata dai *pixel* cui l'interpolazione ha assegnato la medesima sottoclasse di copertura (dato *raster*).

Non potendo effettuare rilevamenti a terra (*ground control point*), il secondo gruppo di test ha invece previsto la validazione delle previsioni attraverso confronti diretti tra la spazializzazione dei punti di prima fase INFC e altri dati vettoriali di natura indipendente (previa trasformazione *raster* con *pixel* 100 m), disponibili per la Sicilia a diverse scale di analisi. In tali casi non si è considerato il dettaglio delle altre classificazioni in quanto l'indagine è stata condotta per tre diverse scale di dettaglio cartografico e nomenclaturale, secondo le classificazioni su basi tipologiche forestali disponibili, corrispondenti a:

- **scala regionale:** INFC - Categorie di copertura del

³ *Dissolve:* operazione che riclassifica poligoni in funzione di un parametro in comune; permette il passaggio da un livello dettagliato ad uno più generale.

1 *Smoothing:* letteralmente smussare, appianare. In questo caso assume il significato di calcolare un valore medio dei punti in un intorno simmetrico al punto considerato, assegnando ai *pixel* derivati dall'interpolazione compresi tra i punti prossimi a quello in esame, lo stesso peso.

2 La scelta dell'ampiezza del *pixel* di 100 m è dettata dal fatto che tale dimensione determina un contorno delle superfici stimate più uniforme; più piccolo è il *pixel* maggiore sarà la scala di valori compresi tra il punto considerato e i prossimi più vicini aumentando la precisione di stima, in ogni caso tale dimensione non rappresenta un valore di dettaglio al suolo dell'immagine risultante che rimane quello derivante dalla maglia inventariale puntuale originaria.



Fig. 4 - Esempio di confronto tra le classificazioni nell'area del Bosco della Ficuzza: l'immagine a sinistra rappresenta il dato spazializzato, al centro il dato relativo a CLC ed a destra il risultato della sovrapposizione tra le classificazioni (in rosso la differenza e in verde l'uguaglianza).

suolo di CLC (*Corine Land Cover 2000* - APAT 2005);

- **scala comprensoriale:** INFC - Categorie forestali del comprensorio dei Monti Sicani (Sicilia Centro-occidentale - La Mela Veca et al. 2004);
- **scala aziendale:** INFC - Tipi forestali della R.N.O. "Bosco della Ficuzza" (Sicilia Nord-occidentale - AA VV 2004).

In alcuni casi le differenze delle classificazioni messe a confronto hanno reso necessario un adeguamento dei diversi sistemi di nomenclatura e quindi delle corrispondenti legende derivabili (tab. 1) per renderle congrue ad un confronto. Inoltre, è da evidenziare come confrontando dati di diversa natura (per esempio CORINE LC e INFC) emergono limiti legati alle diverse definizioni di "Uso-Copertura del suolo" (UN-FAO 2001, Mollicone & Federici 2005a, Mollicone & Federici 2005b).

Per *Corine Land Cover* la gerarchizzazione ha interessato le unità del terzo livello della classificazione, mentre per il comprensorio dei Monti Sicani, il livello è quello ben più dettagliato delle categorie forestali (La Mantia et al. 2000, La Mantia et al. 2001, Cullotta 2003).

La comparazione per la scala comprensoriale (Monti Sicani) ha previsto una riclassificazione delle aree per renderle conformi al sistema INFC; questo è stato ottenuto attraverso una semplificazione nomenclaturale, mantenendo il dettaglio cartografico-territoriale legato al dato iniziale. Nella scala aziendale, invece, il dettaglio cartografico è quello delle tipologie forestali (Marchetti & Cullotta 2003), ma la gerarchizzazione per il confronto è stata riportata al livello di classi principali, per esempio boschi e foreste, secondo le definizioni *FRA 2000* usate nell'INFC.

La procedura di validazione attraverso le comparazioni è stata ottenuta mediante l'*overlay* di strati raster (due alla volta), analizzando e quantificando in

termini percentuali le aree di sovrapposizione con procedure software di selezione ed estrapolazione. Le superfici di disuguaglianza risultanti (fig. 4) non rappresentano l'errore ma solo una differenza di classificazione per quelle aree, dove si presume sia possibile rilevare un errore. Il dato di verifica (e ciò ne definisce anche un limite) si riferisce quindi solo alle aree di esatta sovrapposizione, aree dove nelle diverse classificazioni si è giunti al medesimo risultato e la classificazione dei pixel corrisponde con le classi del *dataset* di validazione. L'errore riscontrato non è attribuibile univocamente alla stima della superficie generata dall'interpolazione, ma bensì a tutte le classificazioni affette anch'esse da incertezze di stima. Eccezione fa il confronto con la scala aziendale dove la tipologia di costruzione del dato originario realizzato ad una scala di 1:10000 ne determina un elevato dettaglio.

Risultati

Spazializzazione

Dalla procedura di spazializzazione è stata ottenuta una superficie omogenea per tutto il territorio regionale, che riporta le aree in cui si localizzano i tipi di uso e copertura del suolo classificati per i due livelli, classe (fig. 5) e sottoclasse (fig. 6).

La classe più rappresentata è quella relativa alle "Superfici agricole" con il 63% del territorio regionale. Le "Superfici boscate e gli ambienti seminaturali" incidono per circa il 33% (847902 ha) del territorio, di cui circa il 14% (345131 ha) sono costituite da "Formazioni forestali" che, secondo le definizioni INFC, sono rappresentate da alberi o arbusti che presentano contemporaneamente una superficie maggiore di 5000 m², un grado di copertura maggiore del 10% ed una larghezza maggiore di 20 m. A questa sottoclasse appartengono i boschi di latifoglie, di conifere e mi-

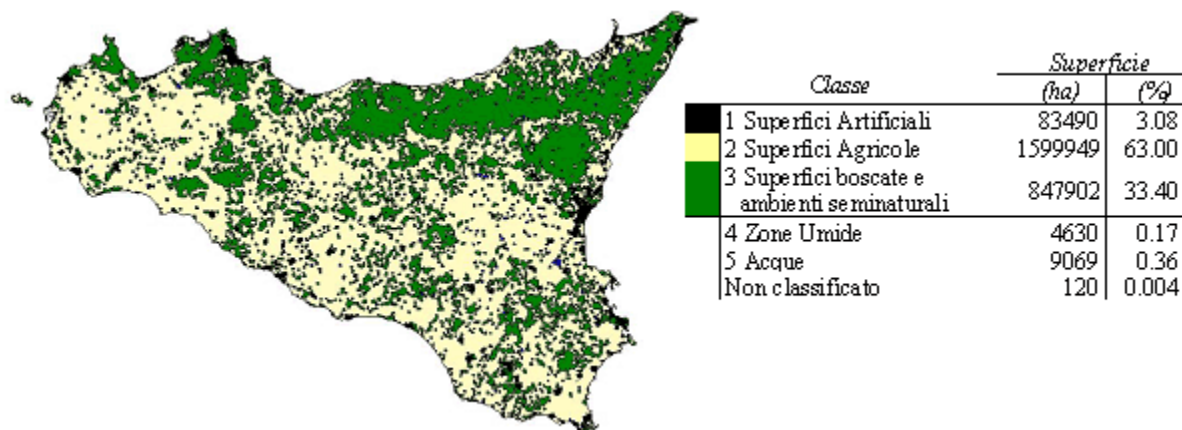


Fig. 5 - Distribuzione della copertura del suolo per il livello di classe dopo la spazializzazione dei punti di I fase INFC (si omette di riportare alcuni colori delle sottoclassi in quanto non apprezzabili visivamente nella mappa visto il loro basso valore di copertura di superficie).

sti, i rimboschimenti, i castagneti da frutto, le sugherete, gli arbusteti e la vegetazione a macchia mediterranea.

La sottoclasse "Praterie, pascoli e incolti" (18 % in termini di superficie) rappresenta il secondo tipo di copertura del suolo della Sicilia (della classe sopra citata) e comprende tutte le formazioni vegetali occupate da vegetazione erbacea spontanea (con copertura superiore al 40%), aree a pascolo naturale, praterie, aree incolte, formazioni di arbusti bassi con un'altezza < di 50 cm e con una copertura arborea inferiore al 5%.

Dal confronto tra i dati puntuali (maglia inventariale INFC) ed i dati spazializzati emerge una corrispondenza che esaminata con il test del chi-quadro

di Pearson è pari a 0.43; le differenze in termini percentuali oscillano tra 0.1 e 1.2 (per esempio, sottoclasse "praterie, pascoli e incolti" - fig. 7) ed in media l'errore stimato è dello 0.31% (valore molto basso). Tutte le sottoclassi possiedono quindi una superficie che, rispetto al dato complessivo, presentano una percentuale di copertura e di distribuzione sulla superficie interpolata del tutto analoga e convergente alla percentuale delle unità di uso del suolo riscontrate nei punti campione.

Validazione

Scala regionale

L'area messa a confronto è quella della superficie regionale con esclusione delle isole. Il confronto ri-

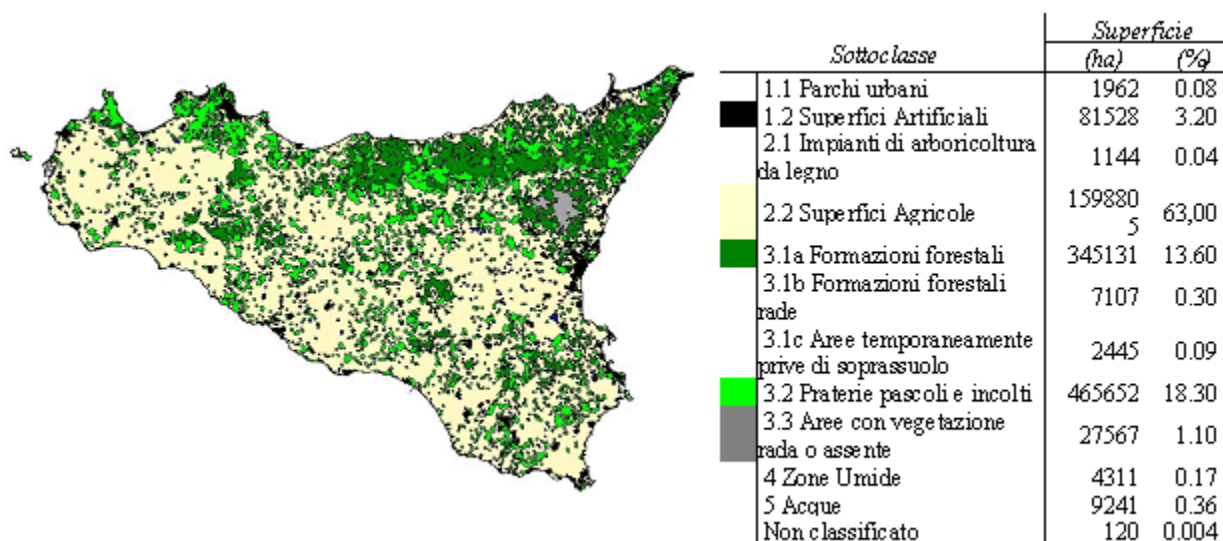


Fig. 6 - Distribuzione della copertura del suolo per il livello di sottoclasse dopo la spazializzazione dei punti di I fase INFC (si omette di riportare alcuni colori delle sottoclassi in quanto non apprezzabili visivamente nella mappa visto il loro basso valore di copertura di superficie).

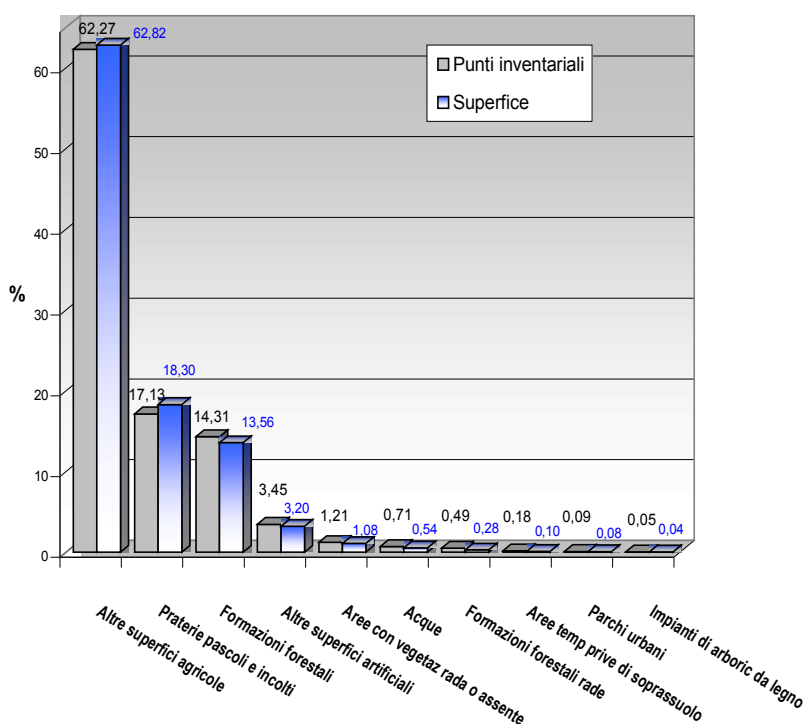


Fig. 7 - Confronto tra i punti inventariali e le superfici (%) dopo la spazializzazione.

guarda il dato INFC spazializzato ed il dato CLC 2000 al 3° livello (fig. 8 e fig. 9).

Il valore di corrispondenza più elevato è quello relativo alle classi INFC (fig. 8): le superfici coincidono per il 74 % del territorio; aumentando il dettaglio e scendendo al livello di sottoclasse si riscontra una sensibile diminuzione statisticamente significativa di

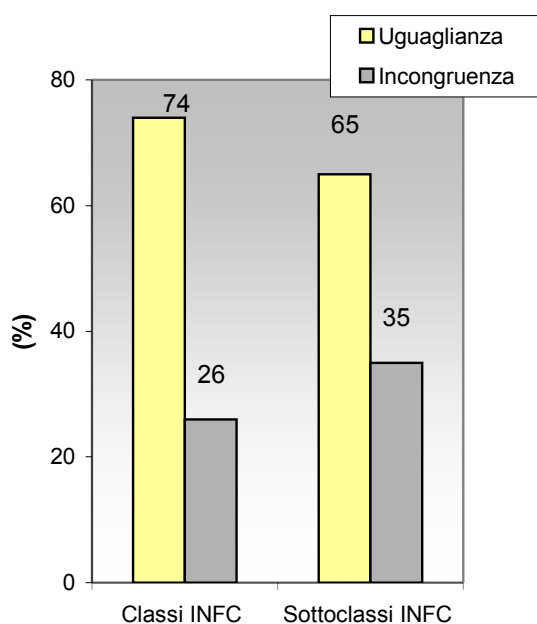


Fig. 8 - Confronto tra INFC e CLC tra le classi e tutte le sottoclassi.

uguaglianza. Le sottoclassi appartenenti alla classe "Superfici boscate e ambienti seminaturali", che nel complesso hanno un dato di uguaglianza del 66%, confrontate con le relative categorie di CLC mostrano valori contrastanti, dal 50% (circa) delle sottoclassi "Formazioni forestali e Formazioni forestali rade" al 27% delle "Praterie, pascoli e incolti" (in questo caso la comparazione, per la difficoltà oggettiva di ricondurre la classificazione CLC ad una che ne permettesse la perfetta comparazione, è stata condotta solo per le sottoclassi citate per le quali una definizione più univoca ne ha permesso una effettiva comparazione - fig. 9). Le motivazioni di queste oscillazioni nel dato di corrispondenza sono da ricercare nella difficoltà di confrontare dati riferiti alla stessa classe di uso, essendo i dati di origine diversa e calcolati con metodologie che considerano parametri diversi. Questo è più evidente quando si esaminano classi di uso del suolo definite in modo differente dalle fonti utilizzate, come nel caso dalla sottoclasse "Praterie, pascoli e incolti".

Scala comprensoriale

L'area confrontata è quella del comprensorio dei Monti Sicani (33404 ha); in questo caso è stato realizzato un doppio confronto INFC-CLC e INFC - Categorie forestali Monti Sicani, verificando il grado di corrispondenza tra i dati (fig. 10 - La Mela Veca et al. 2004).

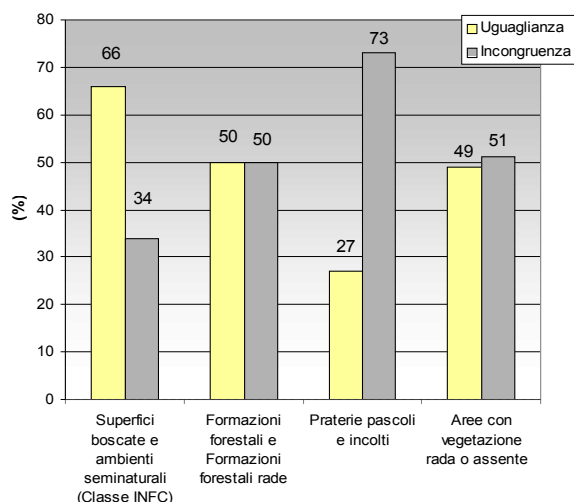


Fig. 9 - Confronto tra INFC e CLC: per la classe “Superfici boscate e ambienti seminaturali” e le sottoclassi relative.

Il dato più evidente è l’elevata corrispondenza in entrambi i casi della sottoclasse “Formazioni forestali e Formazioni forestali rade” (77 e 95% rispettivamente nei confronti di CLC e Sicani) e dell’unità di copertura “Praterie pascoli e incolti” (70 e 74%). Di contro è altrettanto elevata l’incongruenza delle unità “Aree con vegetazione rada o assente” (0,3 e 5% per CLC e Sicani). Le aree a “Superfici boscate e ambienti seminaturali” in questo caso presentano un modesto valore di con-

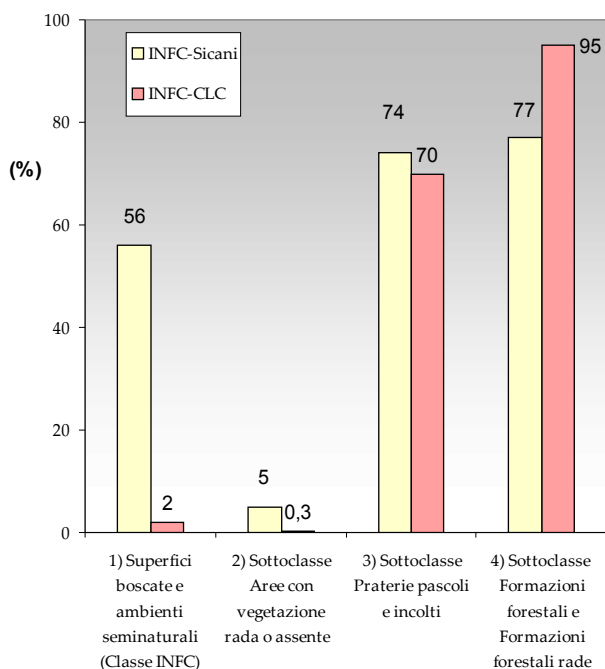


Fig. 10 - Confronto tra INFC, CLC e Sicani: valori in % di congruenza; in tale analisi l’area coinvolta è quella relativa ai “Sicani” di circa 33404 ha.

gruenza solo per il confronto tra INFC e Monti Sicani (56%).

Scala aziendale

Per la scala aziendale, in Sicilia generalmente coincidente con i demani forestali regionali, l’area di confronto è quella del “Bosco della Ficuzza” (7345 ha); il dettaglio di indagine cartografico è in scala 1:10.000 (Maccari 2002, AA. VV. 2004), con una classificazione in tipi forestali.

Nello specifico la classificazione è stata semplificata a due categorie, paragonabili alla classe “Superfici boscate e ambienti seminaturali” e “Praterie pascoli e incolti” di INFC. Anche in questo caso è stato realizzato un doppio confronto. La classificazione tipologica per “Ficuzza” (per tipologia d’indagine cartografica e di classificazione) presenta il dettaglio di copertura

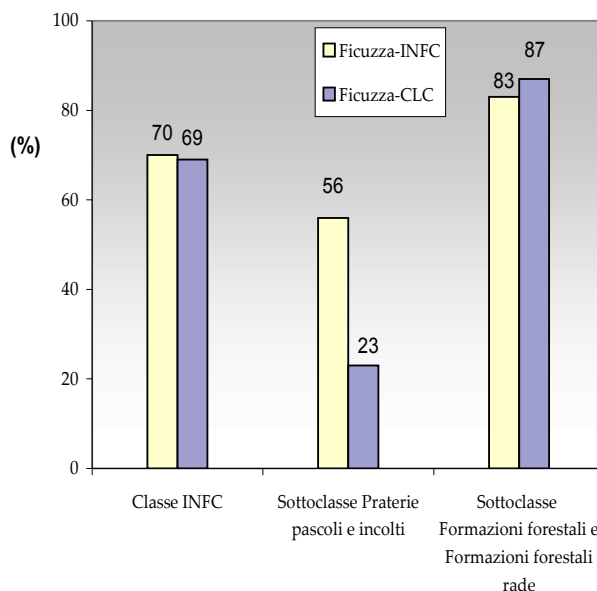


Fig. 11 - Confronto tra “Ficuzza”, INFC e CORINE: valori in % di congruenza; in tale analisi l’area coinvolta è quella relativa a “Ficuzza” di circa 7345 ha.

del suolo più elevato, perciò la verifica è stata condotta tra “Ficuzza”-INFC e “Ficuzza”-CLC (fig. 10). Il dato risultante conferma l’andamento dei precedenti confronti, con elevati livelli di corrispondenza. Il valore di corrispondenza dell’unità “Formazioni forestali e formazioni forestali rade” arriva all’83 % con il dato dell’INFC e all’87 % con il dato CLC (fig. 10). La sottoclasse “Praterie pascoli e incolti” mostra invece valori più bassi (fig. 11).

Discussioni e conclusioni

L’obiettivo preposto è stato quello di estrapolare, da un dato puntuale (INFC), un dato areale che di-

stribuisca l'informazione individuata del dato puntuale ad una superficie, attraverso l'utilizzo di semplici procedure implementate nei più comuni software GIS. La finalità è quella di ottenere velocemente e a basso costo un dato che esprima e traduca in una chiave di lettura diversa e più leggibile il rilievo campionario di un set di dati di origine puntuale.

La metodologia di spazializzazione adottata (IDW) per dati di natura qualitativa mostra molti limiti. Nel caso specifico l'estrapolazione di un dato quale l'uso del suolo, per sua natura molto complesso, ha comportato un'estremizzazione del suo utilizzo tramite un adattamento metodologico creato ad hoc. Partendo dai limiti dell'indice si è semplificato il problema analizzando due classi di uso alla volta per poi riunificare il tutto in un unico dato. Tale soluzione adottata costituisce un semplice accorgimento metodologico che ha consentito facilmente di escludere l'uso di altre metodologie certamente più precise ma più complesse.

La spazializzazione di un dato qualitativo senza far ricorso a procedure di *image processing* è in ogni caso una forzatura connessa e legata alla densità dei punti della maglia inventariale. La superficie generata è ovviamente una superficie matematica che stima un dato numerico al quale si attribuisce un valore qualitativo. L'accorgimento di usare per ogni elaborazione tutti i punti, permette di evitare che si sovrappongano superfici inerenti a unità di copertura diverse proprio perché ogni volta si stima tutta la superficie coperta dai punti campionari. Come già evidenziato l'indice stima solo una superficie compresa tra punti di valore noto attribuendo un peso in funzione della distanza. La semplificazione di elaborare un'unità di uso del suolo alla volta, usando solo due valori numerici legati al dato qualitativo, ha permesso di stimare la superficie interessata e il resto della superficie che è a sua volta legata ad altre tipologie di uso del suolo temporaneamente associate in unità fittizia.

In sintesi, le superfici derivate dalla spazializzazione dei dati inventariali hanno mostrato un contributo percentuale corrispondente a quello delle classi di copertura relative ai punti inventariali.

La comparazione dei dati INFC (prima fase) con altri dati spazializzati a diversa scala ha evidenziato i limiti del dettaglio e la necessità di approfondimento del rilievo campionario e della cartografia.

A livello di classe i dati mostrano, in tutti i casi, una buona congruenza con le altre classificazioni esistenti. Passando dalla scala regionale a quella com-

prensoriale le differenze dalla classe INFC aumentano; un dettaglio maggiore determina una maggiore precisione e quindi una più alta frammentazione (*pattern* disomogeneo) della copertura del suolo con conseguente aumento delle differenze tra i vari livelli.

Le incongruenze riscontrate si localizzano principalmente nelle aree di margine (bordi) tra diversi tipi di uso, e ciò è spiegabile con il fatto che le superfici ottenute sono determinate da un algoritmo matematico che considera solo dei punti.

Le diversità tra classificazioni evidenziano inoltre limiti di comparazione imputabili alla difficoltà di attribuire le aree a unità di copertura perfettamente coincidenti utilizzando sistemi di definizioni differenti.

Il dato CLC mostra, nelle comparazioni, corrispondenze maggiori al livello di classe INFC (scala regionale) rispetto al livello di sottoclasse. Le classificazioni alla scala comprensoriale e aziendale presentano complessivamente una migliore rispondenza rispetto al dato CLC, visto che le definizioni delle classi di copertura del suolo di CLC prendono in considerazione parametri diversi nell'individuazione delle tipologie di copertura.

La classificazione su basi tipologiche ha invece definizioni più simili agli *standard* FRA 2000.

Confrontando il numero di *pixel* classificati a livello di sottoclasse con il dato di riferimento in CLC, le corrispondenze maggiori si rilevano proprio laddove le definizioni sono più univoche e identificano unità quali il bosco e le praterie. Il resto delle unità di *land cover* sono di difficile comparazione e contribuiscono alla determinazione delle differenze riscontrate. Il dato CLC fa infatti riferimento ad una scala di 1:100000 ed individua superfici minime di 25 ha; determinando l'individuazione di aree a "Praterie, pascoli e incolti" molto grandi, non considerando aree di margine o superfici più piccole che vengono inglobate in aree più vaste da altre coperture del suolo. Questo giustifica anche il perché tale sottoclasse presenta valori di corrispondenza bassi; tutto ciò non accade ovviamente con la classificazione INFC applicata ad una fotointerpretazione per punti, dove la stima se pur affetta da possibili errori di valutazione fa riferimento al solo fotopunto valutato in un intorno quadrato di 22500 m². Come è risultato evidente dai confronti con le fonti informative di maggiore dettaglio cartografico si è notata una generale tendenza alla diminuzione delle differenze, ossia un aumento della corrispondenza fra dati confrontati. Per la sola sotto-

classe "Formazioni forestali e Formazioni forestali rade" si è passati da una corrispondenza del 50% con il dato CLC, al 77% con il dato relativo ai "Sicani", fino all'83% con il dato relativo alla cartografia tipologica di dettaglio di "Ficuzza" (fig. 9, fig. 10 e fig. 11).

Nonostante i limiti, è stato ottenuto un risultato che ha valenza su scala regionale. La discreta densità dei punti della maglia inventariale di prima fase ha consentito di stimare una superficie che permette di avere in modo semplice e veloce un dato generalizzato sulla tipologia di copertura del suolo. Tale modello certamente rappresenta un compromesso tra l'accuratezza/costi consentiti rispetto ad altre metodologie più complesse.

I dati relativi alla validazione consentono di stabilire che è possibile, seppure con incertezza, stimare facilmente superfici di unità di uso del suolo verosimili. Aumentando la densità dei punti campionari, tale modello potrebbe essere un valido supporto per la stima di unità di copertura superficiali, utili per gli scopi della pianificazione territoriale a scala regionale e comprensoriale. Dal confronto tra i dati di prima fase INFC ed i dati cartografici di maggiore dettaglio emerge la necessità, di un approfondimento del rilievo campionario, ottenibile attraverso il previsto Inventario Forestale della Regione Sicilia (attualmente in fase di realizzazione) che, mantenendo lo stesso impianto metodologico utilizzato nell'INFC, permetterà una conoscenza più dettagliata delle classi di copertura del suolo e tipologiche di interesse forestale, preforestale e pastorale. In quest'ottica il modello utilizzato nel presente lavoro potrà rappresentare un supporto per la restituzione cartografica delle unità di copertura del suolo con dettaglio maggiore.

Ringraziamenti

Si ringrazia il Prof. F. Maetzke per la revisione del testo ed i consigli forniti. Un particolare ringraziamento al Referee per le critiche ed i suggerimenti indicati. Il lavoro è stato realizzato nell'ambito della cooperazione esistente tra Dipartimento di Colture Arboree dell'Università di Palermo e il Dipartimento Foreste della Regione Siciliana e reso possibile dalla disponibilità dei dati dell'Inventario delle Foreste e dei Serbatoi di Carbonio con autorizzazione all'uso e pubblicazione gentilmente concessa dal Corpo Forestale dello Stato.

Bibliografia

AA. VV. (2004). Ricerche per la redazione del Piano di Gestione della R. N. O. "Bosco della Ficuzza, Rocca Busambra, Bosco del Cappelliere e Gorgo del Drago" (PA). Rela-

zione Finale. Dip. Colture Arboree Palermo - Dip. Regionale Azienda Foreste Demaniali.

APAT (2005). La realizzazione in Italia del progetto *Corine Land Cover 2000*. APAT, Rapporti 36/2005, pp. 86.

Ambrosini I, Buffoni, Gherardi L, Viti ML (1994). Valutazione del contenuto informativo ai fini forestali di riprese aeree bianco/nero realizzate ad uso topografico. *L'Italia Forestale e Montana* 49 (3): 241-254.

Bartier PM, Keller P (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Computer & Geosciences* 22 (7): 795-799.

Berry MJA, Linoff G (1997). *Data Mining Techniques*. Wiley, New York, USA.

Burrough PA, McDonnell RA (1998). *Principles of Geographical Information Systems* (2nd edn). Oxford University Press, Oxford, UK, pp. 333.

Castagnoli A, Marchetti M, Bottai L, Chirici G, Orandini A (1997). Uso del sensore iperspettrale MIVIS per la realizzazione di cartografia forestale a media e grande scala. *Genio Rurale* 9: 54-60.

Chirici G, Corona P, Marchetti M, Maselli F, Bottai L (2003a). Spatial distribution modelling of forest attributes coupling remotely sensed imagery and GIS techniques. In: "Modelling forest systems" (Amaro A, Reed D, Soares P eds). CAB International, Wallingford, UK, pp. 41-50.

Chirici G, Corona P, Travaglini D (2003a). Sperimentazione di tecniche di classificazione *object-oriented* di immagini *Quickbird* a fini forestali. *L'Italia Forestale e Montana* 58 (4): 231-250.

Chirici G, Corona P, Marchetti M, Travaglini D (2003c). Rilievamento campionario e spazializzazione cartografica della necromassa legnosa in biocenosi forestali. *Monti e Boschi* 54 (6): 40-45.

Clementi G, Cullotta S, La Mela Veca DS (2006). Tecniche integrate di geomatica per la cartografia tipologica forestale nei siti NATURA 2000 dei Monti Sicani (Sicilia centro-occidentale). *Rivista Italiana di Telerilevamento* (in corso di stampa).

Corona P, Ambrosini I (1995). Inventari forestali in Italia. *L'Italia Forestale e Montana* 50 (2): 198-214.

Cullotta S (2003). *Forest and Pre-forest Types of Sicily (Italy): classification system, nomenclature, distribution, ecology and management*. Ph.D. Thesis, Mendel University of Agriculture and Forestry, Brno, Czech Republic, pp. 382.

Curtis DC, Clyde PDBS (1999). Comparing Spatial Distributions of Rainfall Derived from Rain Gages and Radar. *Journal of Floodplain Management*. NEXRAIN Corporation Folsom, CA.

Goodchild MF, Lam NS (1980). Areal interpolation: a variant on the traditional spatial problem. *Geoprocessing* 1

- (3): 207-312.
- Jolly WM, Graham JM, Michaelis A, Nemani R, Running SW (2005). A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales. *Environmental Modelling & Software* 20: 873-882.
- INFC (2003). Manuale di fotointerpretazione per la classificazione delle unità di campionamento di prima fase. Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio. MiPAF - Ispettorato Generale del Corpo Forestale dello Stato, CRA-ISAFA, pp. 37.
- INFC (2005). Linee generali del progetto per il secondo inventario forestale nazionale. Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio. MiPAF - Ispettorato Generale del Corpo Forestale dello Stato, CRA-ISAFA, Trento, pp. 57. [online] URL: http://www.isafa.it/scientifica/pubblicazioni/pu_infc/Pubblic.html
- La Mantia T, Marchetti M, Cullotta S, Pasta S (2000). Materiali conoscitivi per una classificazione dei tipi forestali e preforestali della Sicilia. I Parte: Metodologia ed inquadramento generale. *L'Italia Forestale e Montana* 55 (5): 307-326.
- La Mantia T, Marchetti M, Cullotta S, Pasta S (2001). Materiali conoscitivi per una classificazione dei tipi forestali e preforestali della Sicilia. Parte II: Descrizione delle categorie. *L'Italia Forestale e Montana* 56 (1): 24-47.
- La Mela Veca DS, Ciruolo G, Clementi G, Guzzardo E (2004). Spatial analysis of natural and seminatural habitats of Natura 2000 Network in the Sicani Mountains (W Sicily, Italy). In: "Monitoring and indicators of forest biodiversity in Europe - from ideas to operationality" (Marchetti M ed). *EFI Proceedings* No. 51: 343-357.
- La Mela Veca DS, Cullotta S (2005). I tipi forestali della Riserva Naturale Orientata "Monte Cammarata" (Monti Sicani - Sicilia Centro-occidentale). Regione Siciliana-Azienda Regionale Foreste Demaniali. Collana Sicilia Foreste, 29.
- Lopez HF, Ek AR, Bauer ME (2001). Estimation and mapping of forest stand density, volume, and cover type using the *k*-nearest neighbors method. *Remote Sensing of Environment* 77: 251-274.
- Maccari F (2002). Monitoraggio delle risorse forestali non legnose delle R. N. O. Bosco della Ficuzza, Rocca Busambra, Bosco del Cappelliere, Gorgo del drago. Tesi di Laurea. Facoltà di Agraria, Università degli Studi di Palermo.
- Marchetti M (1990). Telerilevamento e pianificazione forestale: nuovi strumenti e metodologie d'impiego. *Monti e Boschi* 41 (2): 4-11.
- Marchetti M, Cullotta S (2003). Classificazioni tipologiche forestali: relazioni e parallelismi per una valutazione comparata a livello europeo. *Monti e Boschi - Sezione Ricerca* 54 (3-4): 23-38.
- 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* 58 (4): 251-263.
- Mollicone D, Federici S (2005a). La stima delle superfici forestali per l'Italia: le incertezze del Nuovo Inventario delle Foreste e del Carbonio. *Forest@ 2* (2): 143-150. [online] URL: <http://www.sisef.it/>
- Mollicone D, Federici S (2005b). L'uso delle terre: il Protocollo di Kyoto, la definizione FAO di foresta e l'Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio. *Forest@ 2* (4): 321-330. [online] URL: <http://www.sisef.it/>
- Naoum S, Tsanis IK (2004). Ranking spatial interpolation techniques using a gis-based dss. *Global Nest: the Int. J.* 6 (1): 1-20.
- Preto G (1983). La fotointerpretazione per punti. Modalità di esecuzione. *Monti e Boschi* 34 (4): 55-60.
- Tosi V, Marchetti M (1998). I sistemi informativi forestali in Italia: uniformità e divergenze tra gli inventari delle risorse forestali. *L'Italia Forestale e Montana* 53 (5): 220-252.
- Tosi V, Monteccone M (2004). Standard per gli inventari forestali di area vasta. Uno studio comparativo per il territorio italiano. *Forest@ 1* (2): 148-164. [online] URL: <http://www.sisef.it/>
- Tobler W (1979). Cellular Geography. In: "Philosophy in Geography" (Gale S, Olsson G eds). D. Reidel Publishing Co., Dordrecht, Holland, pp. 379-386.
- Tomczak (1998). Spatial Interpolation and its Uncertainty Using Automated Anisotropic Inverse Distance Weighting (IDW) - Cross-Validation/Jackknife Approach. *Journal of Geographic Information and Decision Analysis* 2 (2): 18-30.
- UN-FAO (2001). Global Forest Resources Assessment - Main Report. [online] URL: <http://www.fao.org/forestry/site/fra2000report/en>.
- Ware C, Knight W, Wells D (1991). Memory intensive algorithms for multibeam bathymetric data. *Computer & Geosciences* 17 (7): 985-993.
- Watson DF (1992). Contouring: a guide to the analysis and display of spatial data. Pergamon press, Oxford, UK, pp. 340.
- Watson DF, Philip GM (1985). A refinement of inverse distance weighted interpolation. *Geo-Processing* 2: 315-327.