

Alma Mater Studiorum - Università di Bologna

**Dottorato di Ricerca in
SCIENZE GEODETICHE E TOPOGRAFICHE**

Ciclo XX

Settore scientifico disciplinare di afferenza: ICAR06

**MONITORAGGIO DI COLTURE AGRICOLE
MEDIANTE DATI SATELLITARI ASTER:
PROBLEMATICHE E APPLICAZIONI**

**Candidato:
Michaela De Giglio**

**Relatore:
Prof. Maurizio Barbarella**

**Coordinatore Dottorato:
Prof. Gabriele Bitelli**

Esame finale anno 2008

SOMMARIO

Premessa	1
Introduzione alla parte I.....	4
CAPITOLO 1: TELERILEVAMENTO MULTISPETTRALE	5
1.1 Breve introduzione al telerilevamento	5
1.2 Lo spettro elettromagnetico	7
1.3 Le tecniche di analisi dei dati multispettrali.....	11
1.4 Le caratteristiche spettrali della vegetazione	12
CAPITOLO 2: IL DATO MULTISPETTRALE ASTER	15
2.1 Cenni sulla missione TERRA.....	15
2.2 Il sensore ASTER.....	17
2.2.1 VNIR (Visible and Near-Infrared).....	20
2.2.2 SWIR (Short Wave Infrared)	22
2.2.3 TIR (Thermal Infrared)	23
2.2.4 APPLICAZIONI ALL'USO AGRICOLO DEL SUOLO	25
2.3 Il dato ASTER.....	29
2.3.1 Dati ASTER <i>Level-1A</i>	29
2.3.2 Dati ASTER <i>Level-1B</i>	30
2.4 Dipendenza tra risoluzione spettrale e geometrica	34
2.5 Il futuro dell'analisi multispettrale	35
CAPITOLO 3: TECNICHE DI CLASSIFICAZIONE	38
3.1 Introduzione alle tecniche di classificazione.....	38
3.2 Classificazione per pixel	39
3.2.1 Classificazione supervisionata	41
3.2.1.1 Logica generale	42
3.3 Classificazione Object-Oriented	49
3.3.1 Segmentazione	49
3.3.1.1 Procedura	50
3.3.2 Classificazione	57
3.3.2.1 Logica generale	58
Introduzione alla parte II.....	61

CAPITOLO 4: VALUTAZIONE DELLE INFORMAZIONI NECESSARIE.....	62
4.1 Area di studio	62
4.2 Tipi di dati utilizzati.....	64
4.2.1 Dati ASTER	65
4.2.2 Carta dell'uso del suolo.....	67
4.2.3 Carta della vegetazione	69
4.2.5 Mappe catastali	74
4.3 Principali problematiche relative ai dati	76
4.3.1 Armonizzazione dei dati	76
4.3.2 Contemporaneità dei dati	79
4.3.2.1 Programmazione dell'acquisizione di un dato Aster	79
4.3.2.2 Risultato della programmazione di acquisizione del dato Aster.....	87
4.3.2.3 Aster Data Available at No Cost.....	89
4.3.3 Analisi del fenomeno studiato.....	90
4.3.3.1 Cicli di maturazione delle colture erbacee.....	90
4.3.3.2 Indice di vegetazione: NDVI	99
Introduzione alla parte III	104
CAPITOLO 5: CLASSIFICAZIONI.....	105
5.1 La fase di addestramento	105
5.1.1 Classi individuate.....	105
5.1.2 Determinazioni dei campioni	108
5.1.2.1 Utilizzo della Carta dell'Uso del Suolo di Ravenna	108
5.1.2.2 Utilizzo della Carta della Vegetazione di Ravenna	110
5.1.2.3 Utilizzo delle dichiarazioni PAC e delle mappe catastali	111
5.1.3 Difficoltà della fase di addestramento	113
5.2 Classificazioni	117
5.2.1 Confronto multitemporale su base mensile.....	118
5.2.2 Confronto multitemporale su base annuale.....	125
5.2.3 Confronto qualitativo tra i metodi di classificazione	127
CAPITOLO 6: ACCURATEZZA DELLE CLASSIFICAZIONI.....	131
6.1 Introduzione	131
6.2 Valutazione dell'accuratezza	131
6.3 Matrice di confusione.....	133
6.3.1 Teoria	133
6.3.2 Misure di accuratezza.....	135
6.4 Coefficiente K	138
6.4.1 Introduzione	138
6.4.2 La stima di k.....	140
6.4.2.1 La distribuzione della variabile statistica Kappa	142
6.4.2.2 Conditional kappa	143
6.5 Z Test.....	144

6.6 Applicazione ai casi reali	146
6.6.1 Significatività dei singoli k globali	146
6.6.2 Z Test tra le coppie di classificazioni.....	147
6.6.3 Confronto tra le singole classi.....	148
6.6.3.1 Considerazioni	153
<i>CAPITOLO 7: SCHEMATIZZAZIONE DELLA PROCEDURA GENERALE DI LAVORO... 182</i>	
7.1 Flusso di lavoro.....	183
7.2 Studio di un'applicazione reale.....	187
7.2.1 Metodologia	188
A) Riorganizzazione dei dati e progettazione del Database	188
B) Classificazione delle immagini satellitari Aster.....	189
C) Georeferenziazione dei dati	189
D) Vettorializzazione delle immagini QuickBird	190
E) Aggiornamento dei poligoni delle colture.....	191
F) Creazione Geodatabase	192
Considerazioni finali	193
Conclusioni	194
Bibliografia	198

Premessa

L'utilizzo di immagini satellitari negli studi di carattere ambientale è diffuso ormai dagli inizi degli anni '80: indagini di tale natura sono state eseguite, ad esempio, grazie ai dati forniti dai sensori multispettrali delle missioni Landsat, operative da un trentennio. L'introduzione di sensori multispettrali ha permesso di approfondire le conoscenze in diversi ambiti scientifici, soprattutto in quello geologico, consentendo di estrarre un gran numero di informazioni da tutte le bande disponibili. Nonostante la buona qualità dei dati, però, alcune applicazioni su scala ridotta non hanno visto un grande sviluppo a causa delle risoluzioni geometrica e spettrale relativamente modeste.

ASTER ha rappresentato, agli inizi del 2000, un passo avanti nel compromesso tra tali risoluzioni: le 14 bande attraverso le quali acquisisce i suoi dati, comprese tra il visibile e l'infrarosso termico, si sono rivelate molto utili infatti all'interno di diversi studi, relativi ai tipi di suolo, al loro uso, alle coperture vegetali, ecc., soprattutto per le nuove dimensioni del pixel rispettivamente di 15, 30 e 90 m, previste dai tre sottosistemi che lo compongono.

Questo elaborato si colloca al centro di un progetto, cominciato in sede di tesi di Laurea, dedicato allo studio delle potenzialità del dato ASTER per il monitoraggio del territorio, rivolto in particolare al riconoscimento dell'uso del suolo, delle coperture vegetali e alla caratterizzazione delle acque interne e costiere. Le caratteristiche paesaggistiche dell'area di interesse, la fascia costiera emiliano-romagnola, hanno subito mostrato un'elevata variabilità che spazia da elementi naturali altamente sensibili, e quindi soggetti a vincoli di tutela, ad aree intensamente sfruttate da un punto di vista agricolo. E' stata proprio la difficile reperibilità dei dati a terra relativi alle colture agricole, le cosiddette verità al suolo, che ha messo in luce i problemi circa la gestione delle coltivazioni, soprattutto di quelle erbacee. Non esiste infatti un archivio ufficiale, obbligatorio e di libero accesso, nel quale sono registrate tutte le informazioni circa la collocazione, l'estensione e la produzione delle varie colture non arboree. Esistono però delle associazioni private che tutelano gli interessi degli agricoltori, specialmente all'interno di quelle politiche europee che erogano aiuti economici, in seguito alla presentazione e approvazione di dichiarazioni annuali riguardanti la propria attività (ex domande PAC – Politica Agricola Comune): le aziende agricole però devono esservi iscritte.

L'agenzia di riferimento che ha supportato l'attività di ricerca, fornendo un prezioso contributo sia in termini di dati forniti sia di consulenza tecnica, è stata l'Unione Agricoltori di Ravenna. La metodologia sperimentata, ovvero il riconoscimento delle varie coperture presenti a terra da immagini telerilevate, ha suscitato l'interesse dell'Associazione, suggerendo la restrizione

dell'applicazione del lavoro al solo utilizzo agricolo del suolo. Infatti la possibilità di identificare le diverse colture mediante dati satellitari può diventare un utile strumento a favore degli agricoltori durante le fasi di verifica delle dichiarazioni.

Attualmente i controlli dell'ex domanda PAC (ora Domanda di Premio Unica) prevedono alcuni stadi operativi: le liste delle aziende controllate arrivano agli studi che si occupano della verifica, i quali provvedono al recupero del materiale cartaceo necessario al lavoro dei tecnici in campagna (stampa della foto aerea della particella, dei dati dichiarativi e modulistica varia). Dopodichè comincia la rilevazione in campo al fine di verificare la corrispondenza fra quanto dichiarato e quanto effettivamente coltivato. Alcuni controlli vengono effettuati tramite fotointerpretazione di immagini aeree, quando sono disponibili, con tutti i problemi di costi e tempi che queste comportano. Gli esiti dei controlli di campo prevedono la convocazione dell'agricoltore solo nel momento in cui la situazione necessita di chiarimenti (es. difformità fra quanto dichiarato nella DPU e quanto riscontrato in campo dal tecnico o presenza di anomalie catastali). E' in questa fase che l'Unione Agricoltori necessita di un supporto valido che le permetta di far fronte all'eventuale contestazione: la classificazione dell'immagine diventerebbe così la prova della reale situazione a terra.

Naturalmente le esigenze dell'Associazione si traducono nella disponibilità di una procedura veloce ed economica da utilizzare in qualunque situazione di emergenza. Quindi contemporaneamente all'obiettivo principale della Tesi, ovvero testare le possibilità di riconoscimento delle colture mediante dati ASTER valutando tutte le difficoltà connesse, si è cercato di studiare una metodologia di lavoro, il più semplice possibile, destinata ad un utente generico il cui unico interesse è quello di dimostrare cosa c'è sul terreno in un preciso momento.

L'interesse verso questa tecnica è cresciuto notevolmente negli ultimi mesi, in seguito all'avvicinamento anche da parte degli organi di controllo alle classificazioni di immagini satellitari telerilevate.

PARTE I

Introduzione alla parte I

La prima parte della Tesi, articolata in tre capitoli, è dedicata alla descrizione degli strumenti necessari al raggiungimento dello scopo del lavoro, ovvero il riconoscimento delle colture agricole mediante immagini satellitari. Tali strumenti sono la tecnica mediante la quale l'oggetto di interesse viene studiato, il tipo di immagine satellitare, dal quale estrarre il maggior numero di informazioni possibile e, infine, gli algoritmi da applicare al dato.

Viene quindi introdotto il *Telerilevamento Multispettrale*, quale scienza che permette di esplorare contemporaneamente in varie lunghezze d'onda l'energia emessa o riflessa da una superficie, in posizione remota rispetto al sensore. Si è scelto di non richiamare in modo approfondito i principi generali del Telerilevamento, ormai noti in letteratura da diverso tempo, per dedicare maggiore attenzione a quegli aspetti del Multispettrale che sono alla base delle tecniche utilizzate nella sperimentazione. Disponendo infatti di una statistica di comportamento spettrale sufficientemente vasta, è possibile riconoscere, con un grado di precisione variabile, la natura dell'oggetto investigato esaminando la sua risposta spettrale: l'obiettivo del capitolo è quindi quello di introdurre le tecniche di classificazione basate su informazioni multispettrali.

Il dato multispettrale studiato in questo elaborato è acquisito dal sensore ASTER, uno dei cinque presenti sul satellite Terra, nato da una collaborazione americano-giapponese. Le 14 bande disponibili per ciascuna immagine, identificate in funzione della porzione dello spettro elettromagnetico a cui appartengono, ampliano le prospettive delle indagini multibanda tradizionali, presentano risoluzioni geometriche quasi doppie rispetto a quelle dei sensori Landsat, operativi in questo campo da oltre un trentennio, e consentendo un'osservazione maggiormente dettagliata della superficie terrestre. Purtroppo però non è previsto un seguito per la missione Terra, ufficialmente in fase di chiusura, ma solo la ripresa di quella Landsat, che comunque avrà rispetto ad ASTER un minor numero di bande e una minore risoluzione spaziale.

Le tecniche utilizzate nell'elaborazione dei dati satellitari infine, delle quali è riportata una descrizione teorica generale, sono la *classificazione per pixel* e la *classificazione object-oriented*. Gli algoritmi appartenenti al primo metodo classificano separatamente ciascun pixel dell'immagine, mentre nell'altro approccio il maggior contenuto informativo del prodotto finale è associabile a interi poligoni individuati automaticamente. Lo scopo dell'utilizzo di entrambi i metodi consiste nel verificare l'utilità dei due all'interno di una procedura di lavoro, semplice e veloce, per la gestione delle colture, rivolta ad un generico utente, non esperto della tecnica, come può essere l'agricoltore o l'associazione che lo rappresenta.

CAPITOLO 1: TELERILEVAMENTO MULTISPETTRALE

1.1 Breve introduzione al telerilevamento

Il termine telerilevamento indica l'acquisizione e la misura di dati relativi alle proprietà di un fenomeno o di una realtà materiale, attraverso uno strumento di registrazione non in stretto contatto fisico con l'oggetto di analisi: tale tecnica permette la conoscenza delle diverse caratteristiche mediante l'osservazione di campi di forza, di radiazione elettromagnetica o di energia acustica, per mezzo di macchine fotografiche, sistemi laser e radar, ricevitori a radio frequenza, sonar, dispositivi termici, magnetometri ed altri strumenti. E' comunque prassi normale limitare il campo del telerilevamento alle informazioni derivanti dallo spettro delle onde elettromagnetiche: il principio di base consiste nella capacità di ogni corpo di riflettere, assorbire e trasmettere una radiazione incidente in percentuali differenti, a seconda delle proprie qualità strutturali, chimiche, cromatiche. Ogni massa, avente una temperatura superiore allo zero assoluto, è in grado, inoltre, di emettere radiazioni di ampiezza e lunghezza d'onda dipendenti dalle proprietà termiche dell'oggetto stesso. Gli strumenti da ripresa, che forniscono misure o immagini, vengono definiti sensori. Il rilievo remoto satellitare prevede alcune fasi distinte come riassunto schematicamente nella figura 1.1.

- ☞ raccolta dei dati dalle numerose stazioni di rilevamento terrestri, aerei e satellitari;
- ☞ elaborazione delle informazioni registrate;
- ☞ interpretazione dei dati, cui può seguire la restituzione, su supporto cartaceo o informatico, di materiale cartografico tematico.

E' importante ricordare la distinzione tra sensori *passivi* e *attivi*: i primi si limitano a registrare l'intensità con cui è riflessa o trasmessa l'energia emessa da una fonte esterna (sole, terra) e incidente sulla superficie investigata (tra questi, le fotocamere, gli strumenti a scansione, le termocamere e le telecamere). I secondi, invece, emettono loro stessi radiazioni captandone il segnale di ritorno (radar e le camere fotografiche con flash). In entrambi i casi lo scopo è quello di differenziare il maggior numero possibile di elementi o oggetti sul territorio (suolo, vegetazione,

acqua, urbanizzato, ecc.) cercando di descriverne le caratteristiche spettrali, o di ricostruire la superficie investigata e le eventuali deformazioni (caso specifico del rilievo radar).

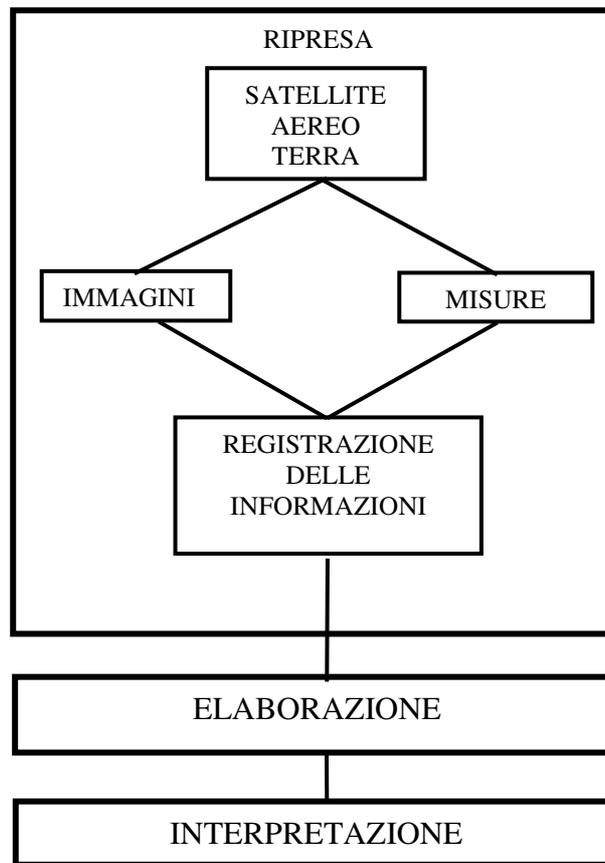


Fig. 1.1 – Schema delle tre fasi in cui si sviluppa un completo processo di Telerilevamento

Gli strumenti si suddividono, poi, in base alle modalità di funzionamento, nelle seguenti categoria:

- ☞ ottico: l'intervallo spettrale è compreso tra 0.3 e 0.9 μm , associato al telerilevamento passivo, identificato da sensori;
- ☞ pancromatico: possiede una sola banda che, normalmente, comprende tutto il visibile ed eventualmente parte dell'infrarosso vicino;
- ☞ multispettrale (include sistemi ottici): utilizza un numero limitato di bande spettrali;
- ☞ iper-spettrali: utilizza anche centinaia di canali;
- ☞ radar: tipico strumento attivo che può operare in singole frequenze o multi-frequenza, nell'intervallo delle microonde tra 1 mm e 100 m (caso tipico è la banda C con $\lambda = 5$ cm).

Gli aspetti peculiari e fondamentali del telerilevamento sono: la *sinotticità*, la possibilità cioè di una visione di porzioni di superficie diverse del territorio altrimenti non conseguibile con tecniche tradizionali; la *ripetibilità*, ottenibile con osservazioni cicliche o periodiche di determinate situazioni, utili per i confronti temporali; l'*aggiornabilità* dei dati, trasferibile alla redazione cartografica. In questo ambito, il risultato di un'operazione è, in genere, la produzione di tematismi che evidenziano aree dal comportamento omogeneo.

1.2 Lo spettro elettromagnetico

In accordo con la teoria ondulatoria l'energia radiante può essere descritta come un'onda armonica che si propaga nello spazio e consiste di due campi di forza ortogonali tra loro: quello elettrico **E**, la cui intensità varia continuamente secondo la direzione di propagazione **x**, e quello magnetico **M**. L'onda è un modello che ben rappresenta i meccanismi di interazione energia-materia su scala macroscopica ed è definita da tre parametri:

- ☞ la lunghezza d'onda λ , ovvero la distanza che separa due creste consecutive misurata con unità molto piccole, quali il micrometro (μm) o il nanometro (nm), a seconda degli intervalli presi in considerazione;
- ☞ la frequenza ν , definita come il numero di cicli compiuti dall'onda in un certo intervallo di tempo e che si solito si esprime in herz (1 Hz= 1 ciclo al secondo);
- ☞ l'ampiezza **A**, che equivale all'altezza di ogni picco e quindi all'intensità massima di **E**. E' spesso misurata in termini di energia, come l'irradianza **E** [$\text{Wm}^{-2} \mu\text{m}$].

Il prodotto della lunghezza d'onda λ e della frequenza ν è costante e rappresenta la velocità di propagazione dell'onda

$$c = \lambda\nu \quad [\text{m s}^{-1}]$$

Lo spettro elettromagnetico (figura 1.2) è una distribuzione monodimensionale continua dell'energia elettromagnetica ordinata normalmente per lunghezze d'onda λ crescenti: per motivi pratici però risulta più comodo suddividerlo in diversi *intervalli* o *bande spettrali*, definiti internazionalmente secondo termini convenzionali.

Nonostante nomi come *infrarosso* o *microonde* individuino regioni dello spettro ben precise e definite per convenienza, non esiste di fatto un confine netto fra una banda e l'altra. Inoltre l'energia

elettromagnetica non è distribuita in maniera uniforme lungo lo spettro ma, nell'ipotesi di Planck, essa si propaga per quantità finite dette *fotoni* o “quanti” di energia, la cui intensità è legata alla lunghezza d'onda: maggiore è la lunghezza d'onda, minore è l'energia coinvolta

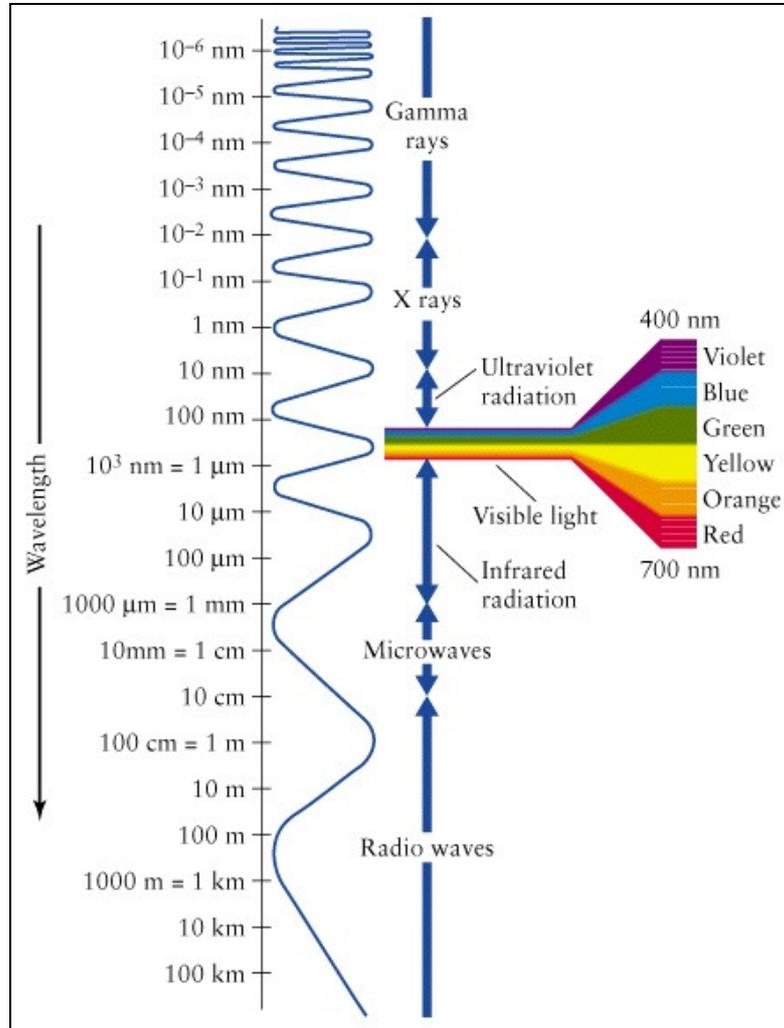


Fig. 1.2 – Rappresentazione dello spettro elettromagnetico in funzione della lunghezza d'onda dei segnali. La porzione del visibile è riportata in dettaglio

Questa è la causa, insieme ad altri fattori descritti nel capitolo successivo, della minore risoluzione dei sensori operanti nella banda dell'infrarosso termico rispetto a quella dei sensori funzionanti nel visibile.

Il telerilevamento impiega le bande Ultravioletto, Visibile, Infrarosso (si veda la tabella 1.1), mentre la tecnologia radar si sviluppa nelle bande del Microonde e tutto il mondo della comunicazioni in banda Radio. Ogni regione dello spettro gioca un ruolo importante in qualche aspetto del Telerilevamento, ma per la caratterizzazione migliore delle proprietà delle superfici terrestri alcune bande sono più adatte di altre:

- ☞ l'ultravioletto, il visibile e il vicino infrarosso, che si appoggiano per lo più sulla porzione di energia riflessa dal Sole;
- ☞ l'infrarosso termico, che si appoggia sull'energia emessa dalla Terra;
- ☞ le microonde, che si appoggiano sia sull'energia naturalmente emessa dalla superficie terrestre sia su quella riflessa e in origine prodotta da uno strumento come può essere un generatore di microonde.

BANDA	BANDA	INTERVALLO λ	AMPIEZZA $\Delta\lambda$	INTERVALLO ν
Raggi UV	Ultravioletto	10 – 380 nm	370 nm	30000 – 789 THz
	Violetto	380 – 430 nm	50 nm	789 – 697 THz
Visibile VIS	Blu	430 – 475 nm	45 nm	697 – 631 THz
	Blu-Verde	475 – 490 nm	15 nm	631 – 612 THz
	Verde	490 – 550 nm	60 nm	612 – 545 THz
	Giallo	550 – 580 nm	30 nm	545 – 517 THz
	Arancio	580 – 620 nm	40 nm	517 – 484 THz
	Rosso	620 – 750 nm	130 nm	484 – 400 THz
Infrarosso IR	IR vicino NIR	0.75 – 0.9 μm	0.15 μm	400 – 333 THz
	IR onde corte SWIR	0.9 – 1.3 μm	0.40 μm	333 – 231 THz
	IR onde medie MWIR	1.5 – 2.5 μm	1.00 μm	200 – 120 THz
	IR termico TIR	3.5 – 5.2 μm	1.70 μm	86 – 58 THz
		7.0 – 20 μm	13 μm	43 – 15 THz

Tab. 1.1 – Principali bande dello spettro elettromagnetico utilizzate nel telerilevamento

La mancanza di contiguità delle bande dipende dal fatto che esse sono localizzate nelle *finestre atmosferiche*, zone cioè di particolare trasparenza dell'atmosfera, evitando opportunamente le bande ove l'atmosfera risulta invece opaca. Considerando la distribuzione dell'energia in funzione della lunghezza d'onda, i diversi meccanismi di interazione con la materia, le caratteristiche di trasparenza atmosferica e, infine, la progettazione strumentale, lo spettro utile per i sistemi di Telerilevamento può essere concettualmente diviso in due finestre principali:

- ☞ finestra ottica (da 100 nm a 20 μm circa): l'energia studiata è di tipo riflesso od emesso in forma naturale dalle superfici;
- ☞ finestra radar (da 0.2 a 100 cm circa): l'energia studiata viene solitamente attivata in modo artificiale.

Per finestra ottica si intende quella porzione di spettro elettromagnetico nella quale sono operanti sistemi di ripresa che hanno come collettore di energia sistemi ottici, sia a rifrazione, come le lenti, che a riflessione, come gli specchi. La sottobanda dell'ultravioletto è utilizzata soltanto in condizioni particolari, per l'enorme diffusione atmosferica che rende estremamente difficile ogni ripresa in ambiente terrestre, anche in presenza di uno spessore ottico ridotto. Di fatto, si possono

usare energie di lunghezze d'onda inferiore a $0.38 \mu\text{m}$ solo da distanze ravvicinate, oppure in condizioni extra-atmosferiche, come succede per il monitoraggio dell'ozono nella troposfera.

La prima parte utile della finestra ottica ($0.38 - 3 \mu\text{m}$) diventa quella più adatta alle indagini sul comportamento spettrale delle superfici nei confronti dell'energia solare riflessa. Il coefficiente più importante e critico è la riflettività spettrale ρ_λ seguito, per i corpi idrici e la stessa atmosfera, dalla trasmittività spettrale τ_λ . Il coefficiente di assorbimento spettrale α_λ , in questa zona della finestra, fa da tramite con i fenomeni termici dovuti allo scambio di energia, laddove essa viene assorbita a lunghezza d'onda corta e restituita, sotto forma di irraggiamento, a lunghezza d'onda lunga. Da un punto di vista strumentale, nella parte riflessa della finestra ottica, si prenderanno in considerazione i radiometri, gli spettroradiometri e i colorimetri, gli scanner multispettrali, i sistemi fotografici, le telecamere, i tubi convertitori di spettro, funzionanti nel visibile e nel vicino infrarosso. I prodotti tipici di indagini multispettrali condotte sulla superficie della Terra in tale ambito sono le mappe di copertura del suolo (dalle quali deriva l'uso del suolo) e di colore delle acque (da cui si possono ricavare concentrazioni di clorofilla, presenza di alghe, distribuzione di inquinanti fisici superficiali, ecc.).

Nella parte emessa (o termica) della finestra ottica ($3 - 20 \mu\text{m}$), sono preponderanti gli effetti dell'energia emessa spontaneamente dalle superfici poste a temperatura ambiente, cioè intorno ai 300 K . È il dominio del coefficiente di emissione spettrale ε_λ , che condiziona fortemente l'interpretazione dei dati in termini di temperatura superficiale. Il prodotto tipico nella parte termica della finestra, infatti, è la mappa delle radianze emesse dalle superfici, dalla quale si possono desumere, oltre alle temperature, parametri indiretti come l'inerzia termica (con rilievi multitemporali) e altri tematismi correlati, quali ad esempio, l'umidità del suolo.

La finestra dell'ottico è ulteriormente suddivisa in tre sottobande:

- ☞ nella banda del visibile (VIS) si trovano le sottobande relative ai colori dell'arcobaleno. Tale divisione, puramente convenzionale, risulta utile nelle notazioni scientifiche, perché quando si parla di un certo colore, si fa riferimento in realtà ad una precisa banda di lunghezza d'onda;
- ☞ nella banda dell'infrarosso IR si trovano quattro sottobande principali: infrarosso vicino (NIR = *Near InfraRed*) e a onde corte (SWIR = *Short-Wave InfraRed*), impiegate in Telerilevamento per lo studio di fenomeni esclusivamente di tipo riflesso; a onde medie (MWIR = *Mid-Wave InfraRed*) per fenomeni di tipo sia riflesso che emesso, lunghe o IR termico (TIR = *Thermal InfraRed*) per fenomeni essenzialmente di tipo emesso. Poco impiegata

è la sottobanda nell'IR lontano (FIR = *Far InfraRed*) oltre i 20 μm di lunghezza d'onda.

1.3 Le tecniche di analisi dei dati multispettrali

Uno dei principali aspetti del telerilevamento, è la possibilità dell'*analisi multispettrale*, per mezzo della quale l'indagine viene condotta simultaneamente su porzioni diverse dello spettro elettromagnetico, in genere comprese tra il visibile e l'infrarosso termico, al fine di migliorare la riconoscibilità fisica delle superfici oggetto di studio. Questa caratteristica è legata al concetto di *risoluzione spettrale*, che si riferisce al numero e all'ampiezza delle bande in cui è suddiviso l'intervallo di sensibilità spettrale del sistema. A sua volta, la risoluzione spettrale è influenzata dal *potere risolvente spettrale*, il quale determina il passo di lunghezza d'onda di campionamento ($d\lambda$) a cui vengono effettuate le misure con un determinato strumento, ed esprime la capacità di un sistema nel distinguere due lunghezze d'onda contigue (Brivio et al.,2006).

Dalla stessa scena osservata si hanno quindi più immagini, una per ogni banda spettrale: il loro insieme viene detto *immagine multispettrale*, che risulta essere una matrice numerica a più dimensioni, ognuna delle quali descrive la scena osservata nella banda spettrale di competenza. Il generico elemento, ovvero il pixel, di coordinate (i,j) non è più costituito da una singola misura radiometrica (DN), ma può essere considerato come un vettore di misure di tante componenti quante sono le bande spettrali dell'immagine. La struttura del dato multispettrale è riportata nella figura 1.3.

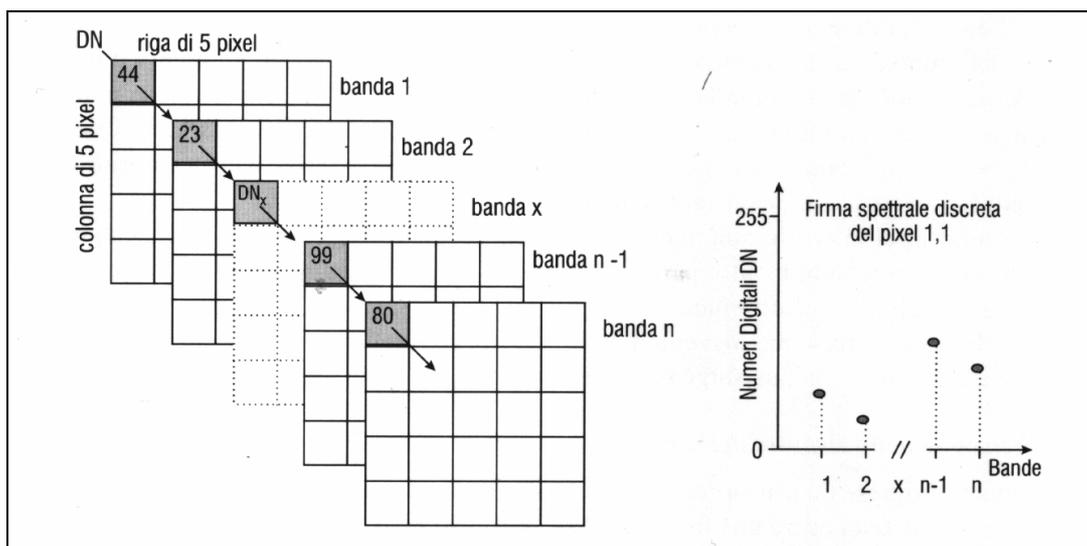


Fig. 1.3 – Struttura del dato multispettrale e comportamento spettrale del pixel 1,1: ogni pixel è caratterizzato da valori ($DN_1, DN_2, \dots, DN_{n-1}, DN_n$) tanti quante sono le bande spettrali, oltre alle sue coordinate-immagine (da Gomarasca M.A., 1997)

Il livello di energia registrati nelle varie bande in corrispondenza di un generico pixel consentono di ricostruire la cosiddetta *firma spettrale*, ovvero la curva caratteristica dell'oggetto indagato generata dalla variazione dei valori di riflettanza nelle diverse bande.

1.4 Le caratteristiche spettrali della vegetazione

L'equilibrio di un vegetale è regolato da un continuo scambio di energia con l'ambiente circostante: tale ponte energetico si attua su differenti intervalli spettrali. L'andamento della firma spettrale della vegetazione è regolato nelle varie lunghezze d'onda dai seguenti fattori:

- ☞ *Pigmenti fogliari*: la porzione di energia riflessa nel campo del visibile (VIS) è correlata alla presenza di pigmenti fogliari quali la clorofilla, la xantofilla e il carotene. La clorofilla in particolare (65% dei pigmenti) è responsabile delle due bande di assorbimento nel blu (0.45 μm) e nel rosso (0.66 μm) e del massimo relativo di riflettività nel verde (0.50 μm), secondo il quale ci appare il mondo vegetale; questo effetto è dovuto principalmente ai cloroplasti che contengono dei granuli dove sono localizzati i diversi tipi di clorofilla. I carotenoidi invece, impartiscono quei toni giallo-arancio alle foglie povere o prive di clorofilla.
- ☞ *Struttura fogliare*: essa è responsabile del comportamento spettrale nelle bande dell'infrarosso vicino (NIR), tra 0.70 e 1.35 μm . Per struttura fogliare, in questo contesto, si intende l'organizzazione spaziale delle cellule della foglia con riferimento alla sua evidenza morfologica che varia da una specie ad un'altra. La radiazione infrarossa non interagisce con i cloroplasti ma è influenzata notevolmente dalla struttura della foglia, che provoca una riflessione molto elevata, dell'ordine del 30-70%, dell'energia incidente. Questo fenomeno è collegato al meccanismo di difesa termica della pianta: per evitare danni irreparabili alle proprie cellule, la pianta cerca di mantenere la temperatura delle foglie al di sotto di 41°C. Oltre ai complicati processi di evapotraspirazione, uno dei sistemi di difesa dall'innalzamento di temperatura consiste nel "rifiutare" (cioè riflettere) l'energia alle piccole lunghezze d'onda dell'infrarosso vicino, assorbendo in parte dello spettro visibile l'energia indispensabile alla vita. La porzione di energia incidente residua viene soprattutto trasmessa, mentre quella assorbita è molto poca. Infatti, da un lato i pigmenti fogliari e la cellulosa che costituisce le pareti cellulari sono trasparenti, e dall'altro l'acqua contenuta nella foglia presenta deboli fenomeni di assorbimento ($\leq 10\%$). Si consideri, a tal proposito la seguente figura dove sono

rappresentate contemporaneamente tutte e tre le componenti di interazione dell'energia elettromagnetica con una foglia campione di oleandro (Fig. 1.4). Si possono fare le seguenti considerazioni:

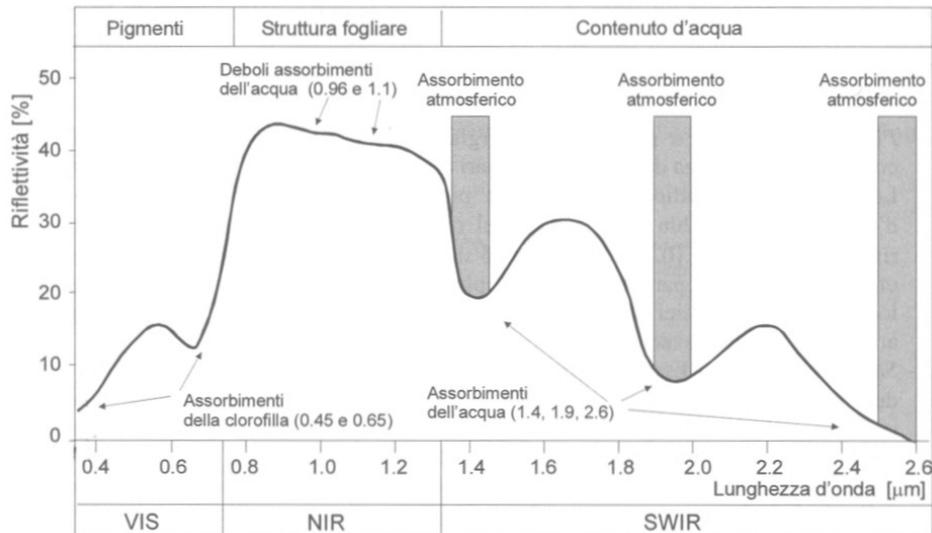


Fig. 1.4 – Tipico andamento della riflettività relativa alla vegetazione con principali fenomeni di assorbimento nella regione SWIR dello spettro elettromagnetico

All'inizio dell'intervallo NIR la riflessività passa da pochi punti percentuali a valori vicini al 50%, con un ripido gradiente attorno a 0.8 μm: il livello di riflettività è tanto più alto quanto più il tessuto è costituito da cellule dalle forme irregolari e dal contenuto eterogeneo.

Poiché il mesofillo, cioè lo strato fogliare situato tra l'epidermide superiore e inferiore della foglia, è più o meno sviluppato a seconda dello stato di salute della pianta, della varietà specifica e dello stato fenologico (e cioè dei rapporti tra clima e processi di fioritura o di produzione fogliare), lo studio di riflessione nel NIR diventa un potente strumento d'indagine, sia per discriminare specie diverse sia per analizzare gli stadi evolutivi e sia per determinare eventuali patologie all'interno di una stessa comunità. Infatti, si è osservato che uno stato di stress e una condizione patologica si manifestano come riduzione di riflettività delle foglie nell'intervallo dell'infrarosso vicino. Questo si spiega col fatto che le cellule dell'epitelio fogliare riflettono tanto più vigorosamente la radiazione infrarossa vicina, quanto più le loro pareti sono turgide e lisce; campioni di epitelio raccolti su piante in stato di stress mostrano o la parziale distribuzione dell'impalcatura delle cellule, o il loro afflosciamento. Entrambi sono fenomeni che riducono la riflettività a spese

dell'assorbimento di radianza. Si può innescare allora un processo a catena che porta alla distruzione delle cellule per superamento della soglia di 41°C. Riduzione della riflettività nell'infrarosso vicino e medio, e aumento dell'emissione di calore (cioè radianza nell'infrarosso termico), sono quindi indici di alterazione della vegetazione. Spesso, questi fenomeni sono accompagnati da un aumento di riflettività nel verde e nel rosso. Globalmente, la vegetazione alterata tende a comportarsi come un oggetto inerte che segue le vicissitudini microclimatiche senza alcuna possibilità di regolazione.

☞ *Contenuto d'acqua:* nel dominio dell'infrarosso a onde corte (SWIR), tra 1.35 e 2.70 μm , le proprietà spettrali della vegetazione dipendono strettamente dal tenore in acqua nella foglia dove, come si vede in figura 1.4, a 1,4 μm , 1.9 μm e 2.6 μm si hanno intensi fenomeni di assorbimento della radiazione da parte dell'acqua, e anche i massimi relativi di riflessione sono parzialmente influenzati dal tenore in acqua delle foglie stesse. In tal modo, le lunghezze d'onda di assorbimento, come per i suoli, possono diventare diagnostiche rispetto allo stress idrico delle piante.

Ugualmente interessante è la verifica della dipendenza reciproca tra suoli e vegetazione in funzione della loro diversa copertura areale (Fig. 1.5). Come si può vedere, l'andamento tipico dei suoli è ottenuto direttamente attraverso i diversi contributi nelle tre porzioni spettrali considerate, in funzione della diversa copertura della porzione vegetata.

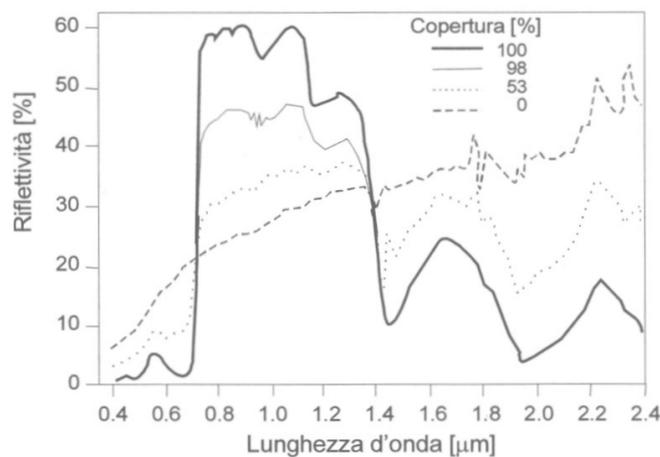


Fig. 1.5 – Tipico andamento della riflettività relativa alla vegetazione in relazione alla percentuale di copertura rispetto al suolo

CAPITOLO 2: IL DATO MULTISPETTRALE ASTER

2.1 Cenni sulla missione TERRA

In orbita a 705 chilometri di quota, il satellite chiamato *TERRA* sta conducendo un esame dettagliato di vari aspetti biotici ed abiotici del pianeta.

TERRA è uno dei satelliti facente parte dell'Earth Observing System (EOS), un programma della NASA che ha lo scopo di supportare gli scienziati nella decifrazione dei numerosi fattori climatici, indispensabili per la comprensione dei mutamenti in atto. Le informazioni raccolte dai sensori del satellite stanno fornendo utili informazioni circa i vari cicli naturali, le alterazioni dovute all'attività antropiche, ma anche sulle diverse coperture al suolo o sui fenomeni in atto. Lungo la sua orbita polare, il satellite *TERRA* attraversa la linea equatoriale alle ore 10:30 UTC, mentre ripete una completa rivisitazione dell'orbita ogni sedici giorni. Dato che la traccia a terra prodotta dall'orbita del satellite è quasi perpendicolare al verso di rotazione della Terra, le strisciate (*swath*) prodotte dai successivi passaggi comporranno un mosaico dell'intera superficie terrestre.

La piattaforma *TERRA*, caratterizzata dalla presenza di cinque sensori, nasce dalla collaborazione tra le Agenzie spaziali di diversi paesi ed in particolare Stati Uniti, Giappone e Canada. L'Agenzia statunitense si è occupata del veicolo spaziale e dei seguenti sensori:

CERES: Clouds and the Earth's Radiant Energy System;

MISR: Multi-Angle Imaging Spectroradiometer;

MODIS: Moderate-resolution Imaging Spectroradiometer.

La Canadian Space Agency e University of Toronto hanno fornito il sensore

MOPITT: Measurements Of Pollution In The Troposphere

Invece il *Japanese Ministry of International Trade and Industry* ha realizzato il sensore da cui provengono la immagini oggetto della tesi:

ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer

La durata inizialmente prevista per la missione *TERRA* era di sei anni, ma al momento in cui si scrive questo elaborato il satellite è ancora pienamente operativo. I responsabili del progetto EOS hanno identificato ventiquattro fattori che nell'insieme hanno un ruolo fondamentale nella

determinazione del clima. Fra essi vi sono il flusso radiante proveniente dal sole e quello di altri corpi radianti, la concentrazione dei gas-serra, la copertura di neve e ghiaccio, le nubi e gli aerosol, le coperture vegetate e altre caratteristiche della superficie. Nella tabella 3.1 compaiono alcuni dei principali comparti investigati tramite la piattaforma *TERRA*. Viene sottolineato l'uso del suolo, principale interesse di questo lavoro di ricerca.

PARAMETRI MISURATI DA *TERRA*

AEROSOL
TEMPERATURA DELL'ARIA
NUBI
INCENDI
GHIACCIAI TERRESTRI
TEMPERATURA DELLE TERRE EMERSE
USO DEL SUOLO
DISASTRI NATURALI
PRODUTTIVITA' DEGLI OCEANI
TEMPERATURA DEGLI OCEANI
INQUINAMENTO
IRRAGGIAMENTO
GHIACCI MARINI
COPERTURA NEVOSA
VEGETAZIONE
VAPORE ACQUEO

Tab. 2.1 – I 16 parametri studiati dal satellite TERRA.

Anche la materia SUOLO presenta molti aspetti rilevanti per l'osservazione satellitare: le modifiche antropiche al paesaggio, i fenomeni di urbanizzazione, disboscamento tropicale, grandi aree adibite ad agricoltura, ma anche l'indagine della ciclicità di tutti questi fenomeni. I GAS-SERRA sono un altro oggetto di grande interesse per la piattaforma *TERRA*. Tutte le indagini climatiche, infatti, non possono prescindere dalle considerazioni relative all'aumento dei livelli di gas-serra nell'atmosfera, soprattutto di anidride carbonica, come risultato dell'uso di combustibili fossili e degli incendi delle foreste tropicali.

Una delle applicazioni principali dei dati del satellite *TERRA* riguarda le indagini sulla VEGETAZIONE, intesa come biomassa complessiva o prodotta per coltivazione. Questo soprattutto nei paesi in via di sviluppo, dove è fondamentale il controllo delle risorse agricole e forestali. Lo studio della vegetazione entra poi in contatto con altri aspetti relativi al suolo, in un delicato bilancio di calore, luce, nutrienti e acqua. Il tutto a comporre vari fattori climatici. La vegetazione influisce soprattutto nella regolazione del *ciclo di carbonio*. Con la fotosintesi, le piante assorbono anidride carbonica dall'aria e, sfruttando l'energia solare, la combinano con l'acqua per generare molecole complesse chiamate carboidrati. Materia viva per l'esistenza della biosfera, la vegetazione, seguita poi dagli altri livelli della catena alimentare, usa i glucidi sia come nutrimento

sia per costruire la propria struttura. Piante e animali, attraverso la respirazione, bruciano carboidrati come combustibile per il metabolismo e li riseparano in molecole di acqua e anidride carbonica, successivamente rilasciata. Quindi, attraverso la respirazione e la decomposizione dei tessuti, il carbonio ritorna in atmosfera concludendo il suo ciclo.

Negli ultimi cento anni, come già detto, si è assistito ad un aumento della concentrazione di CO₂ pari al 25%, con conseguente innalzamento di 0.5°C della temperatura globale: è stato dimostrato che, da circa venti anni, questi due effetti hanno alterato alcune stagionalità nello sviluppo di varie specie nell'emisfero settentrionale, ottenendo una maggiore produttività annuale delle piante. Da un punto di vista ambientale, tale conseguenza corrisponde ad un eccesso di anidride carbonica immagazzinata nelle piante, con il rischio di trasformare la biosfera da dispersore di carbonio ad una sua fonte netta. Comunque, le osservazioni suggeriscono che, parallelamente a questi fenomeni, la vegetazione ha cominciato a ridurre l'evapotraspirazione (perdita di acqua per evaporazione) per conservare l'acqua e per non compromettere la fotosintesi. Le informazioni trasmesse da *TERRA* permettono di mettere in relazione la produttività delle piante con l'anidride carbonica e con la temperatura globale. In generale gli strumenti operanti in questo senso sono:

MODIS e MISR, che forniscono una copertura globale continua, consentendo uno studio completo stagionale e inter-annuale della vegetazione terrestre;

ASTER, che fornisce informazioni dettagliate su zone di speciale interesse, come ad esempio quelle disboscate.

2.2 Il sensore ASTER

ASTER è il sottoinsieme con risoluzione geometrica maggiore tra quelli presenti sulla piattaforma *TERRA*: l'area ricoperta dai pixel va da 15 a 90 metri quadrati passando dal sensore visibile a quello infrarosso termico sulle complessive quattordici bande dello spettro elettromagnetico (si veda la tabella 2.2). Diversamente dagli altri sensori, che raccolgono dati con continuità, ASTER opera per un tempo limitato, durante le porzioni diurne e notturne del ciclo orbitale: l'intera registrazione delle informazioni avviene con una media di otto minuti per orbita. Limitando il numero di bande, o differenziando i *gains* o con operazioni simili, è possibile ridurre le configurazioni e implementarle per rispondere alle esigenze degli utenti.

CARATTERISTICHE PECULIARI	SENSORI	NUMERO BANDA	BANDE SPETTRALI (μm)	RISOLUZIONE SPAZIALE (m)	DIMENSIONE SWATH (Km)
Massima risoluzione spaziale tra tutti i sensori di <i>TERRA</i> e capacità di puntare verso specifici bersagli	VNIR (visibile e vicino infrarosso)	1	0.52-0.60	15	60
		2	0.63-0.69		
		3N	0.78-0.86		
		3B	0.78-0.86		
	SWIR (infrarosso a breve lunghezza d'onda)	4	1.60-1.70	30	
		5	2.145-2.185		
		6	2.185-2.225		
		7	2.235-2.285		
		8	2.295-2.365		
		9	2.360-2.430		
	TIR (infrarosso termico)	10	8.125-8.475	90	
		11	8.475-8.825		
		12	8.925-9.275		
		13	10.25-10.95		
14		10.95-11.65			

Tab. 2.2 - Caratteristiche tecniche del sensore ASTER

Occorre descrivere i tre sottosistemi cross-track di telescopi del sensore ASTER: VNIR, SWIR e TIR (figura 2.1). Come visto anche nella tabella 2.2 ognuno di essi lavora in una diversa regione spettrale, ottenuta tramite separazione dello spettro tramite filtri. La relativa alta risoluzione spaziale-radiometrica e la capacità di cambiare gli angoli di osservazione, consentono la produzione di immagini stereoscopiche, prodotte in particolare nelle bande 3N e 3B, che consentono anche l'elaborazione di modelli particolareggiati di elevazione del terreno.

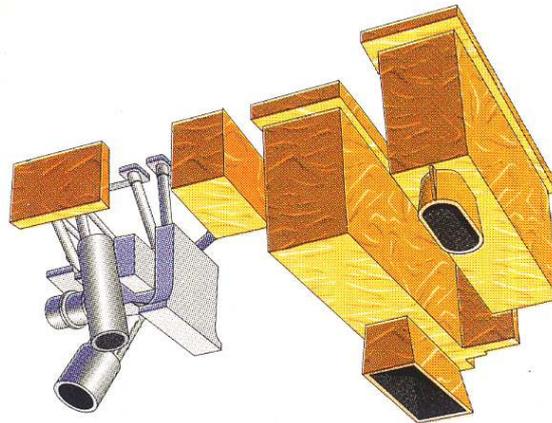


Fig. - 2.1 - Sensore ASTER con i tre sottosistemi, da sinistra verso destra, VNIR, SWIR, TIR

Ogni sensore, quindi, fornirà immagini (o meglio dati) dal diverso contenuto informativo, come mostrato nella figura 2.2, attraverso la soluzione falsi colori.

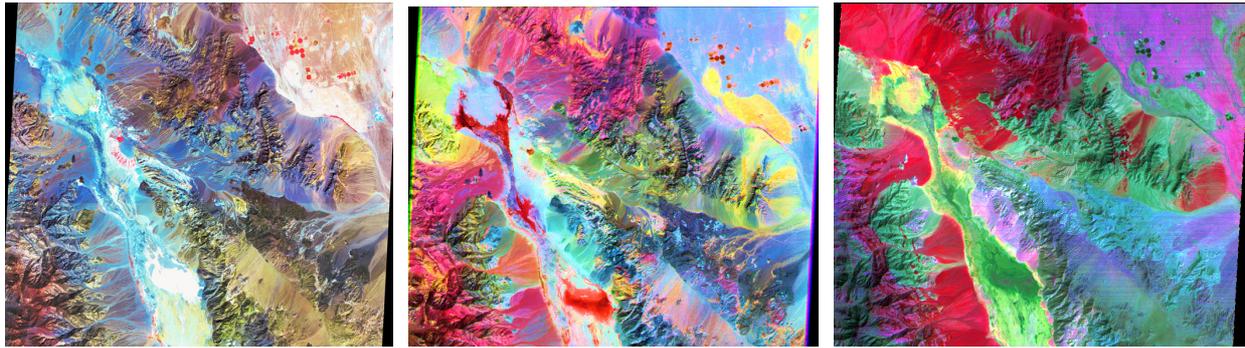


Fig. 2.2 – Immagini in falso colore della medesima porzione di territorio prodotte dai sottoinsiemi VNIR, SWIR, TIR

I valori di energia e calore rilevati sono fondamentali per stimare il bilancio radiativo terrestre e sono particolarmente utili per identificare litologie, suoli e vegetazione. Nello specifico, vengono acquisite informazioni sulla composizione e distribuzione dei materiali sulla superficie, necessarie per individuare nuovi giacimenti di minerali o di combustibili fossili in regioni scarsamente esplorate e per controllare l'uso del territorio. In agricoltura, le immagini multispettrali aiutano a capire come i cambiamenti di temperatura, di pendenza del terreno e di natura del suolo influiscono sulla resa delle colture. ASTER può anche verificare altri tipi di variazioni in corso - come il ritiro di ghiacciai e l'evoluzione delle calotte polari, l'avanzamento dei deserti, la deforestazione, le inondazioni e gli incendi – per i quali gli studiosi devono distinguere le cause naturali da quelle artificiali.

Inoltre, il sensore ASTER può essere considerato in continuità scientifica con la serie di dati acquisiti già nel 1972 dai satelliti LANDSAT. In figura 2.3 compare il confronto tra le caratteristiche dei sensori ASTER e LANDSAT-7. I due hanno in comune molte caratteristiche tecniche: si muovono con un groundspeed (velocità rispetto al suolo) di circa 7 Km/s sulla stessa orbita, a 705Km di quota, e con un ritardo di ASTER di quindici minuti rispetto a LANDSAT-7. La risoluzione minima del primo, pari a 15m, è rispettata dal secondo solo in un canale pancromatico, avendo in generale una risoluzione di 30m.

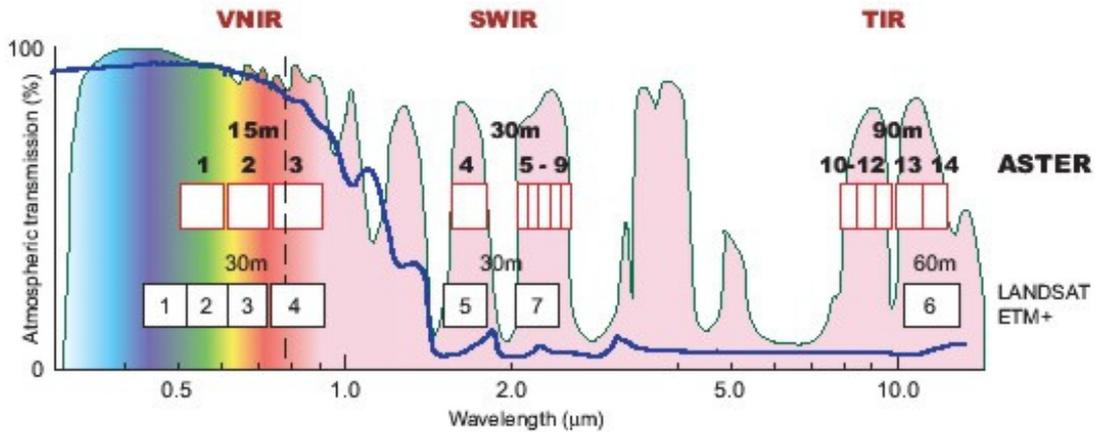


Fig. 2.3 – Confronto tra le bande dei sensori ASTER e LANDSAT 7

2.2.1 VNIR (Visible and Near-Infrared)

Il sottosistema VNIR comprende due telescopi indipendenti a riflessione-rifrazione, uno *nadirale* e uno *backward-looking*, in modo da realizzare immagini stereo sulla stessa orbita, minimizzando le distorsioni generate singolarmente (Figura 2.4).

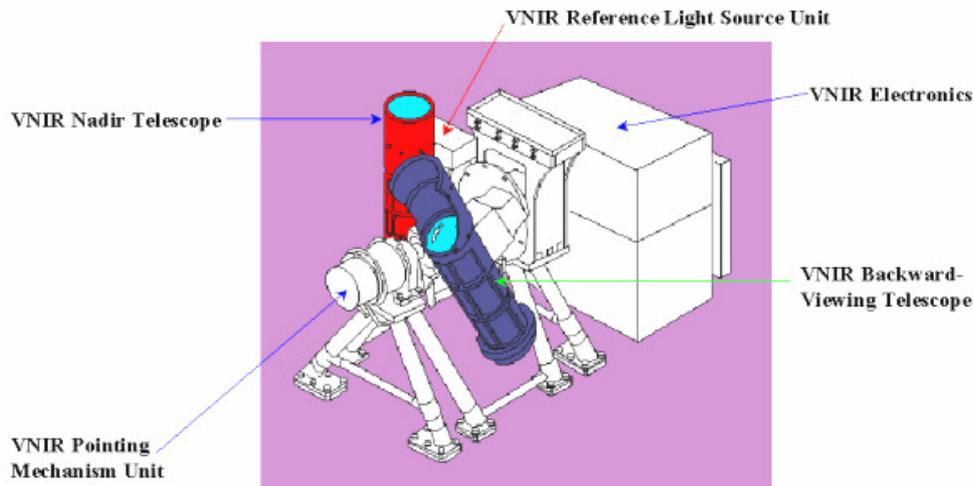


Fig. 2.4 – Sottosistema VNIR, (<http://asterweb.jpl.nasa.gov/instrument/subsystem.htm>)

Questo componente di ASTER, opera nelle bande visibili verde e rossa e nel vicino infrarosso con risoluzione spaziale di 15m e risoluzione radiometrica di 8 bits (256 livelli). La coppia di telescopi (Schmidt design) può essere ruotata di $\pm 24^\circ$, fornendo l'estensione massima della capacità di puntamento cross-track. La risposta spettrale nelle bande VNIR è riportata nella successiva figura 2.5.

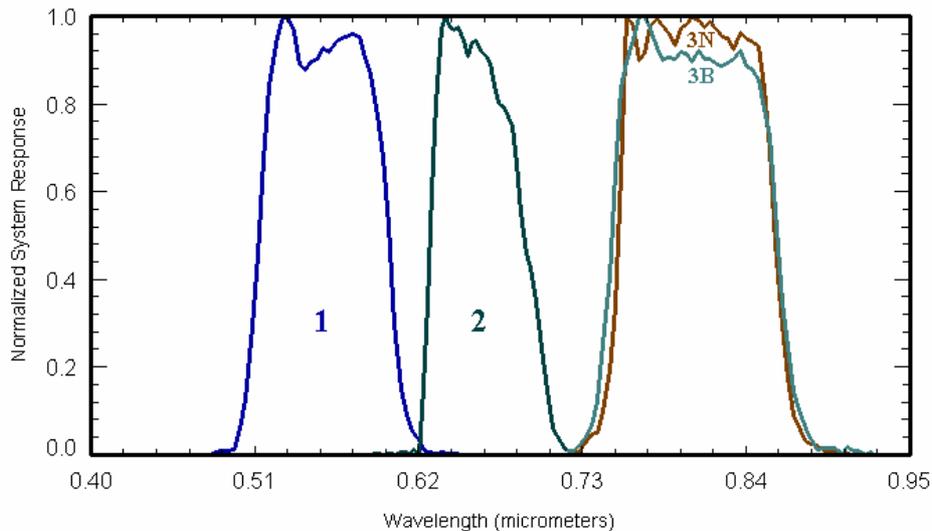


Fig. 2.5 – Risposta spettrale nelle bande VNIR, (<http://asterweb.jpl.nasa.gov/instrument/vnirchart.htm>)

Il piano focale dello strumento ottico nadirale contiene tre array lineari (banda 1, 2, 3), formati da 5000 elementi CCD al silicio, costituenti i rivelatori per ogni banda spettrale: solo 4000 di essi sono impiegati a tempo pieno. Tra l'acquisizione dell'immagine nadirale e quella backward intercorre un intervallo di tempo durante il quale la rotazione terrestre causa lo spostamento del centro dell'immagine. VNIR estrae automaticamente i 4000 pixel corretti sulla base delle informazioni orbitali provenienti dalla piattaforma EOS. Per permettere alle tre bande di vedere simultaneamente la stessa area, sono utilizzati un prisma dicroico e filtri di interferenza per la separazione spettrale. Il piano focale del telescopio backward-looking contiene un solo array di rivelatori (banda 3) e usa un filtro di interferenza per distinguere le lunghezze d'onda.

Entrambi gli strumenti sono mantenuti a 296±3K e la sorgente della radiazione è una lampada alogena: un raggio divergente dal filamento della lampada irradia uniformemente gli array. La calibrazione a bordo dei due telescopi è ottenuta da dispositivi indipendenti, ciascuno costituito da due fotodiodi al silicio: uno controlla direttamente il filamento della lampada, il secondo verifica la calibrazione del raggio proprio di fronte al primo elemento ottico del telescopio. VNIR, costruito da NEC Corporation, produce, di gran lunga, la quantità di dati più consistente tra i tre sottosistemi ASTER.

2.2.2 SWIR (Short Wave Infrared)

Il sottosistema SWIR usa un solo telescopio a rifrazione asferico (Fig. 2.6), operando in sei bande dell'infrarosso ad onda corta (da $1.6\mu\text{m}$ a $2.5\mu\text{m}$, come si vede nella figura 2.7 relativa alla risposta spettrale delle stesse bande), con risoluzione spaziale pari a 30m e risoluzione radiometrica di 8 bits (256 livelli).

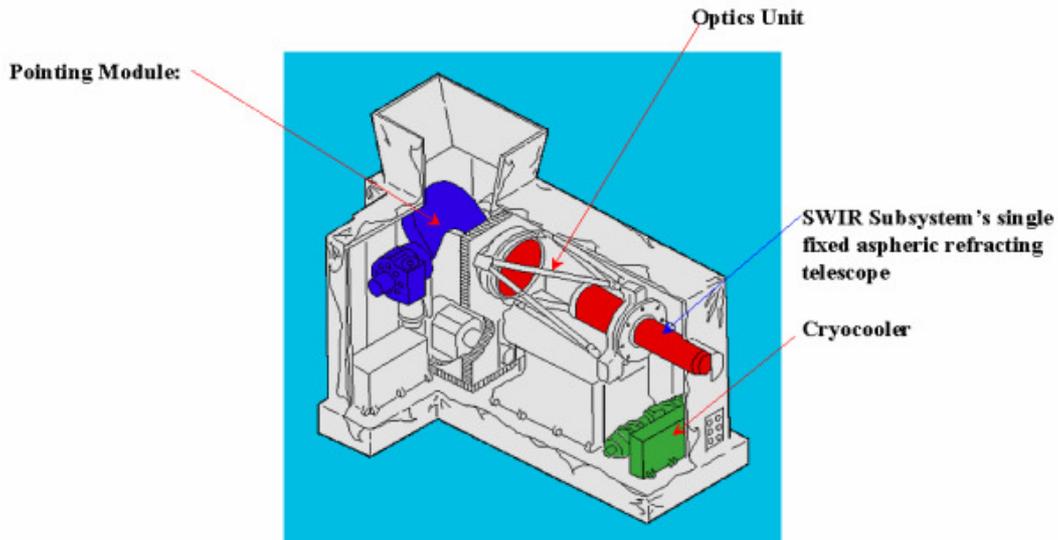


Fig. - 2.6 – Sottosistema SWIR, (<http://asterweb.jpl.nasa.gov/instrument/subsystem.htm>).

Il rivelatore, in ogni banda, è dato da un array lineare di tipo Schottky al platino-silicio raffreddato ad 80K. Il raffreddamento è fornito da un crioraffreddatore a ciclo Stirling, con compressori opposti ed un bilanciere attivo per compensare gli spostamenti dovuti all'espansione. Il calore viene rimosso e irradiato nello spazio attraverso un radiatore locale, collegato al compressore: non è utilizzata una classica piattaforma, provvista di piatto freddo, per lasciare la massima libertà di movimento al telescopio, necessaria per il puntamento cross-track. Sebbene ASTER registri ad intervalli abbastanza brevi, il crioraffreddatore lavora continuamente, perché i tempi di raffreddamento e stabilizzazione sono lunghi.

Per la separazione spettrale vengono usati sei filtri ottici passabanda, al posto dei prismi o elementi dicroici visti in precedenza. La calibrazione è affidata ad un solo dispositivo, simile a quelli del sottosistema VNIR.

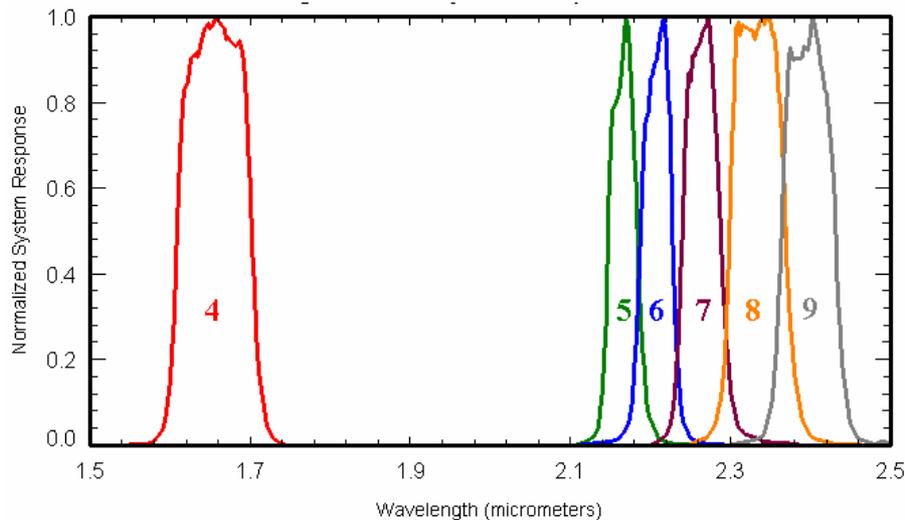


Fig. - 2.7 - Risposta spettrale nelle bande SWIR, (<http://asterweb.jpl.nasa.gov/instrument/vnirchart.htm>).

E' presente uno specchio che può ruotare di $\pm 8.54^\circ$ dalla direzione nadirale, per permettere la copertura di tutti i punti sulla superficie terrestre, durante i sedici giorni di ciclo del satellite. Periodicamente, serve anche per indirizzare la luce, da due lampade di calibrazione, al telescopio. Il sottosistema SWIR è stato realizzato da Mitsubishi Electric Company (MELCO).

2.2.3 TIR (Thermal Infrared)

Il sottosistema TIR usa un sistema catadiottico newtoniano, con uno specchio primario asferico e lenti per la correzione delle aberrazioni, (Fig. 2.8). Opera nell'infrarosso termico (da $8\mu\text{m}$ a $12\mu\text{m}$, come si vede in figura 2.9 relativa alla risposta spettrale delle stesse bande), con risoluzione spaziale di 90m e risoluzione radiometrica di 12 bits.

Diversamente da VNIR e SWIR, il telescopio del sottosistema è fisso: la scansione e il puntamento sono eseguiti da uno specchio. Ogni banda utilizza dieci rivelatori al mercurio-cadmio-tellurio con filtri ottici passabanda su ogni elemento: l'insieme delle prestazioni aumenta se sono usati rivelatori fotovoltaici. Anche TIR ricorre ad un raffreddamento meccanico a ciclo Stirling, per garantire la temperatura di 80K: in questo caso però, il calore disperso viene rimosso per mezzo di una piattaforma con un piatto freddo.

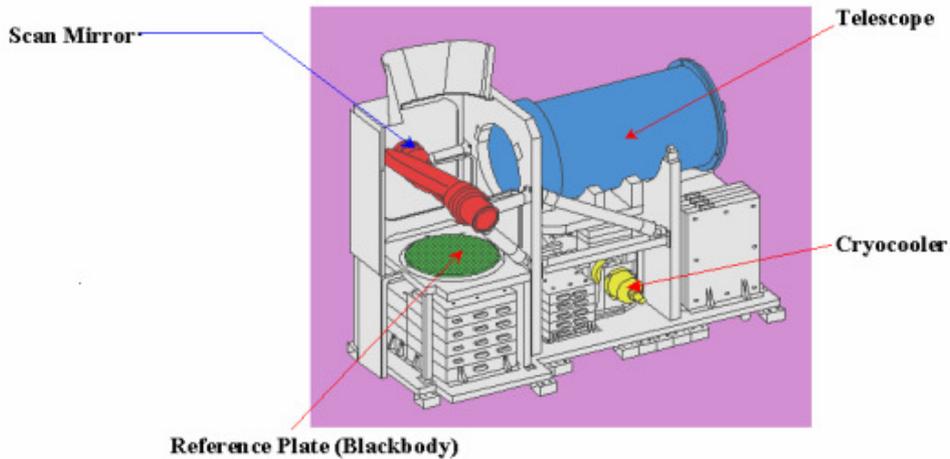


Fig. 2.8 – Sottosistema SWIR, (<http://asterweb.jpl.nasa.gov/instrument/subsystem.htm>).

Potendo lavorare sia di giorno che di notte, la durata del ciclo corrisponde al 16% dell'intera registrazione del satellite (invece dell'8% dei casi precedenti): il crioraffreddatore opera con continuità. Per la calibrazione, lo specchio ruota di 180° dalla posizione nadirale per mostrare all'interno un corpo nero che può essere caldo o freddo: essendo impedita la vista dello spazio freddo, ogni volta viene individuata una singola temperatura di calibrazione (ricalcolabile più volte, in volo, se necessario). Un chopper di controllo e monitoraggio della temperatura rimuove le frequenze più basse. Ad intervalli brevi, il corpo nero fornisce i punti termici di calibrazione, variando la propria quantità di calore. TIR è stato realizzato da Jujitsu Ltd.

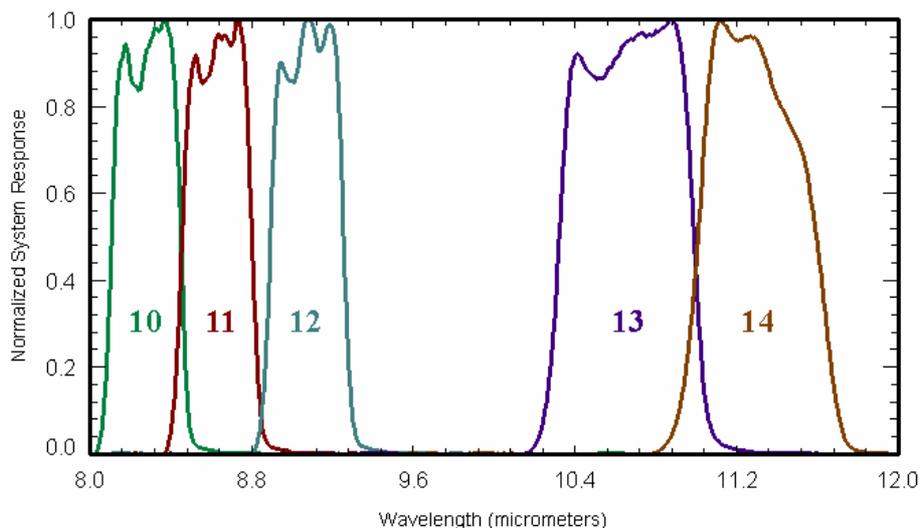


Fig. 2.9 - Risposta spettrale nelle bande TIR, (<http://asterweb.jpl.nasa.gov/instrument/vnirchart.htm>)

2.2.4 APPLICAZIONI ALL'USO AGRICOLO DEL SUOLO

In ogni parte del mondo, l'attività agricola si è sempre sviluppata in funzione della topografia del terreno, del tipo di suolo, di coltura, del regime annuale di pioggia, e delle usanze tradizionali. In figura 2.10 sono riportati i dettagli di sei immagini Aster relative a differenti pratiche agricole, sia dal punto di vista geometrico che da quello dimensionale:

- ☞ Minnesota – USA (riquadro alto a sinistra): la griglia regolare riflette l'agrimensura del XIX secolo, la dimensione degli appezzamenti è legata all'introduzione di macchinari e dettata da una notevole efficienza organizzativa;
- ☞ Kansas – USA (riquadro alto al centro): il sistema di irrigazione è responsabile della forma circolare dei campi;
- ☞ Nord-Ovest Germania – UE (riquadro alto a destra): la dimensione e il pattern random delle colture hanno origine medievale;
- ☞ Santa Cruz – Bolivia (riquadro o a sinistra): la forma singolare degli appezzamenti è decisa secondo uno schema comunitario;
- ☞ Bangkok – Thailandia (riquadro basso al centro): le piantagioni di riso sono alimentate da una fitta rete di canali centenari;
- ☞ Cerrado – Brasile: (riquadro basso a destra): il basso costo del terreno e la morfologia pianeggiante hanno permesso la creazione di coltivazioni molto estese.



Fig. 2.10 - Dettagli di dati Aster relativi all'uso agricolo del suolo. La superficie delle singole aree riportate è pari a $10.5 \times 12 \text{ Km}^2$

Finney County, Kansas



Fig. 2.11 – Finney County: immagine acquisita il 24/06/2001. L'area ricoperta è di $37.2 \times 38.8 \text{ Km}^2$

Situata nella parte sud-occidentale del Kansas, questa zona (figura 2.11) è stata trasformata da immensa prateria ad area ad intenso sfruttamento agricolo (in particolare sono coltivati mais, sorgo

e frumento). I sistemi di irrigazione utilizzati realizzano degli appezzamenti circolari con un diametro variabile tra 800 e 1600m. Il verde scuro indica un'intensa attività vegetativa mentre i colori più chiari individuano aree incolte o sulle quali è appena stato effettuato il raccolto.

Santa Cruz de la Sierra, Bolivia

Nella figura seguente (Fig. 2.12) sono confrontate un'immagine Landsat del 1986 e una Aster del 2001 di un'area di foresta tropicale secca, situata nei pressi di Santa Cruz de la Sierra, in Bolivia. Da metà degli anni Ottanta, l'insediamento umano e il largo sviluppo dell'attività agricola sono stati alla base del disboscamento di questa zona. La forma radiale o "a torta" del pattern degli appezzamenti sono previsti dallo schema San Javier: al centro di ogni unità si trova una piccola comunità che include la chiesa, la scuola, i negozi e il campo da calcio. Le aree rettangolari più chiare sono coltivazioni di soia destinata all'esportazione, in genere gestiti da finanziatori stranieri; le strisce nere invece sono paraventi che permettono di difendere il terreno dall'erosione provocata dal vento.

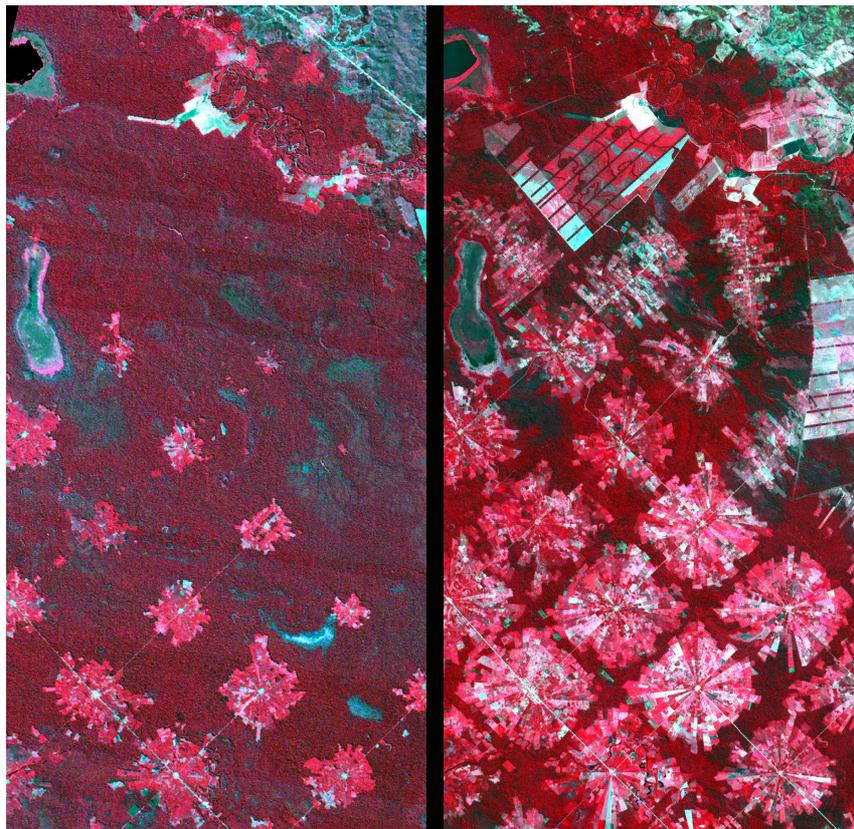


Fig. 2.12 – Santa Cruz de la Sierra, Bolivia: confronto tra l'immagine Landsat acquisita il 04/05/1986 e l'immagine Aster acquisita l'11/05/2001. L'area coperta da ciascuna immagine è 17.6×35 Km².

Coachella Valley, California

L'intera immagine di figura 2.13 copre la Coachella Valley, in California. Il riquadro giallo comprende la zona a nord-ovest del lago Salton, detto Salton Sea: è la regione della California maggiormente destinata ad uso agricolo, soprattutto a coltivazioni da frutta produttive durante tutto l'anno. Le differenti combinazioni in RGB delle bande ASTER servono per migliorare l'identificazione delle varie colture: in alto a sinistra sono state usate le bande 3, 2, 1, in alto a destra le bande 4, 2, 1, e in basso a sinistra le bande 4, 3, 2.

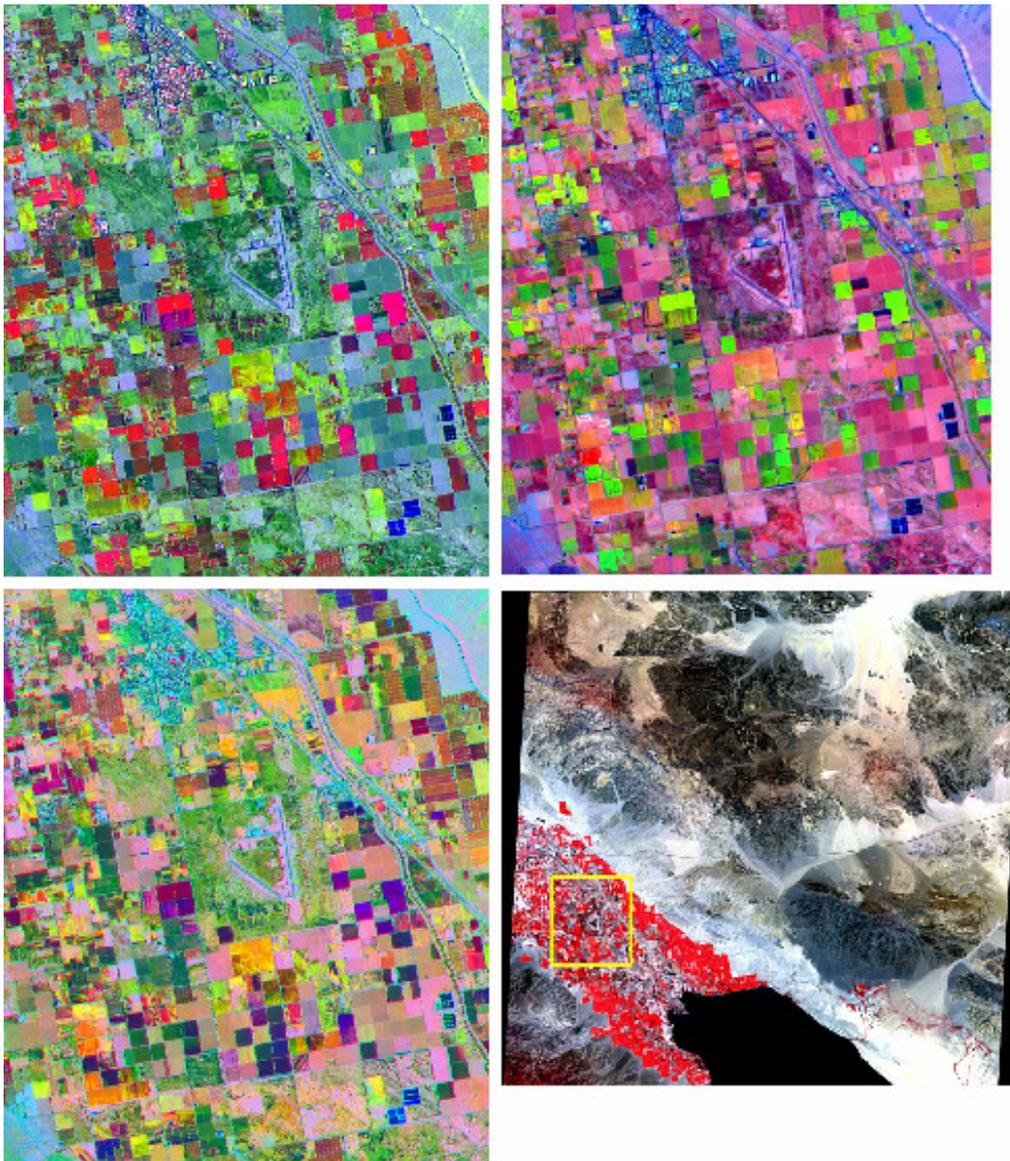


Fig. 2.13 – Coachella Valley (California): l'immagine è stata acquisita il 4 giugno 2000 e copre un'area di 11x13.5 Km, (<http://asterweb.jpl.nasa.gov/gallery/gallery.htm?name=Salton>).

2.3 Il dato ASTER

ASTER, il principale dei sensori di *TERRA*, può registrare dati relativi all'intero globo con cicli di durata pari all'8% per orbita. Tale acquisizione comporta 650 scene al giorno, processate al *Level-1A*: di queste circa 150 (al massimo 310, in funzione della copertura delle nuvole) vengono processate al *Level-1B*. Tutte le scene *1A* e *1B* sono trasferite dall'archivio di EOSDIS a EROS Data Center's (EDC) Land Processes Distributed Active Archive Center (LP-DAAC), per l'immagazzinamento, la distribuzione e l'elaborazione di dati ad alto livello (*livello 2*, come Brightness Temperature, e *livello 3*, come DEM). Tutte le informazioni ASTER prodotte vengono raccolte con una specifica implementazione di Hierarchical Data Format chiamata HDF-EOS.

2.3.1 Dati ASTER Level-1A

Gli ASTER *Level-1A* (*L1A*) grezzi sono dati ricostruiti dal *Level-0* e non sono processati dagli strumenti digitali di calcolo. Tale produzione contiene immagini non sottoposte ad impacchettamento e multiplex, con allegati i coefficienti non applicati di correzione geometrica e di calibrazione radiometrica. I dati VNIR e SWIR hanno profondità di 8 bits e una regolazione variabile di gain, mentre i TIR sono di 12 bits e hanno un solo gain: per tutti è richiesto il formato HDF-EOS.

Nell'ASTER *Level-1A* è prevista anche, per ciascuno dei tre sensori, una versione ricampionata e compressa delle immagini con correzioni simili a quelle effettuate ai dati *Level-1B*. Tutte le immagini (VNIR, SWIR, TIR) sono compresse a 24-bit JPEG e immagazzinate in HDF file. La tabella 2.3 fornisce le principali caratteristiche del *Level-1A* browse-image.

TELESCOPIO	DIMENSIONE (PIXEL×LINE)	METODO DI COMPRESSIONE	BLU	VERDE	ROSSO
VNIR	224×208	JPEG	Band 1	Band 2	Band 3N
SWIR	224×208	JPEG	Band 4	Band 5	Band 9
TIR	224×208	JPEG	Band 10	Band 12	Band 14

Tab. - 2.3 - Caratteristiche del browse-images dei dati Level-1A.

I valori di radianza sono determinati tramite la conversione dei DN (Digital Number), ottenuta con la seguente formula:

$$L = \frac{A \cdot V}{G} + D \quad (\text{VNIR e SWIR})$$
$$L = A \cdot V + C \cdot V^2 + D \quad (\text{TIR})$$

dove:

L è la radianza $\left(\frac{W}{m^2 \cdot sr \cdot \mu m} \right)$;

A è un coefficiente lineare;

C è un coefficiente non lineare;

D è l'offset;

V è il valore DN;

G è il gain.

I dati sono prodotti da Ground Data System (GDS) a Tokyo (Giappone) e inviati a EROS Data Center's (USA) per l'archiviazione e l'elaborazione ad alti livelli. Attualmente vengono forniti unicamente i dati al Livello 1° mentre, a partire dal marzo 2006, i dati al livello 1B devono essere espressamente richiesti dall'utente.

2.3.2 Dati ASTER Level-1B

Gli ASTER *Level-1B* (*LIB*) sono i dati rilevati dai differenti telescopi in *Level-1A*, radiometricamente calibrati e geometricamente co-registrati nelle diverse bande con i coefficienti sopra citati. L'immagine *LIB* è proiettata su una mappa rotante (che si orienta tramite un elenco di coordinate) a piena risoluzione dello strumento. I dati generati includono le registrazioni dei valori VNIR, SWIR e TIR: sono previsti lo stesso numero di bande e le stesse risoluzioni del precedente livello. L'informazione *Level-1B* chiama "SceneCenter" il centro geodetico della corrispondente immagine *L1A*, definito con coordinate *LIB*, cioè latitudine geocentrica e longitudine geodetica (quest'ultima coincide con la longitudine geocentrica perché riferita allo stesso meridiano). La radianza viene ottenuta dai valori DN con la seguente formula:

$$L = (V - 1) \cdot U$$

dove:

L è la radianza $\left(\frac{W}{m^2 \cdot sr \cdot \mu m} \right)$;

V è il valore DN;

U è il coefficiente di conversione dell'unità (Unit conversion coefficient)

Questo valore viene mantenuto costante durante l'intera missione (si veda la tabella 2.4 che riporta alcuni valori tipici per questo coefficiente).

Band No.	Coefficient ($W/(m^2 \cdot sr \cdot \mu m)/DN$)			
	High gain	Normal Gain	Low Gain 1	Low gain 2
1	0.676	1.688	2.25	N/A
2	0.708	1.415	1.89	
3N	0.423	0.862	1.15	
3B	0.423	0.862	1.15	
4	0.1087	0.2174	0.290	0.290
5	0.0348	0.0696	0.0925	0.409
6	0.0313	0.0625	0.0830	0.390
7	0.0299	0.0597	0.0795	0.332
8	0.0209	0.0417	0.0556	0.245
9	0.0159	0.0318	0.0424	0.265
10	N/A	6.822×10^{-3}	N/A	N/A
11		6.780×10^{-3}		
12		6.590×10^{-3}		
13		5.693×10^{-3}		
14		5.225×10^{-3}		

Tab. 2.4 – Coefficienti di conversione dell'unità, (http://asterweb.jpl.nasa.gov/documents/aster_user_guide_v2)

In mancanza di precisazioni, il dato *LIB* viene prodotto nella proiezione UTM (orientamento dello swath) e ricampionato col metodo Cubi Convolution (Convoluzione Cubica). Altre combinazioni possono essere richieste al GDS, secondo la tabella 2.5

METODO DI RICAMPIONAMENTO	PROIEZIONE
Nearest Neighbor (NN)	Geographic (EQRECT)
Cubic Convolution (CC)	Lambert Conformal Conic (LAMCC)
Bi-Linear (BL)	Space Oblique Mercator (SOM)
	Polar Stereographic (PS)
	Universal Transverse Mercator (UTM)

Tab. 2.5 - Metodi di ricampionamento e proiezioni disponibili per la produzioni di dati Level-1B

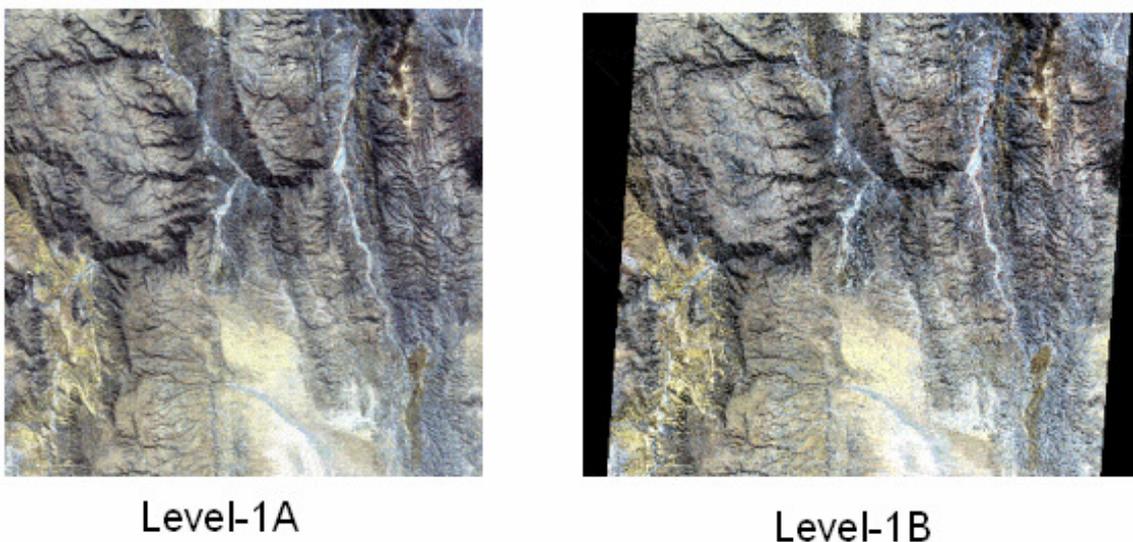
Ciascuna immagine contiene informazioni di geolocalizzazione immagazzinate in una serie di array, uno per ogni telescopio nadirale: questi sono costituiti da 11×11 elementi, di cui il primo in alto a sinistra è indicato con (0,0). L'immagine con vista backward risulta sfalsata rispetto alla precedente, pur avendo lo stesso array di latitudine e longitudine.

Le scene ASTER *Level-1* hanno una dimensione a terra di 60×60 Km²: sono leggermente più grandi le misure relative alle VNIR 3B usate nella generazione di immagini stereoscopiche per la produzione di DEM. Come prima, dati sono prodotti da Ground Data System (GDS) e inviati a EROS Data Center's per l'archiviazione e l'elaborazione ad alti livelli (Tabella 2.6).

SENSORE/TELESCOPIO	ASTER LIA	ASTER LIB
VNIR: Bands 1, 2, 3N	4200 rows × 4100 columns	4200 rows × 4980 columns
VNIR: Band 3B	4600 rows × 5400 columns	4600 rows × 4980 columns
SWIR: Bands 4 - 9	2100 rows × 2048 columns	2100 rows × 2490 columns
TIR: Bands 10 - 14	700 rows × 700 columns	700 rows × 700 columns

Tab. 2.6 - Dimensioni delle immagini dei dati ASTER LIA e LIB.

Nella figura 2.14 il confronto tra i due livelli di elaborazione per la stessa immagine



*Fig. 2.14 – Confronto tra Level 1A e Level 1B,
(Overview, Remote Sensing of the Earth's Environment from Terra,
La scuola per la società dell'informazione, August 25-30, 2002, L'Aquila, Italia)*

Il formato HDF (*Hierarchical Data Format*) è lo standard per tutti i dati prodotti dalla NASA (*Earth Observing System, EOS*). Comodo per gli utenti per trasferire e manipolare dati scientifici, nasce dall'incrocio di diversi sistemi operativi e dall'utilizzo di interfacce e utilità di FORTRAN e C. HDF supporta diversi tipi di informazioni: array n-dimensionali, tabelle, testi, immagini raster con associate le palette di colori e i metadati. La libreria HDF permette di immagazzinare e recuperare questi tipi di dati da altri formati, compressi e non. Per ciascun dato-oggetto nell'HDF, un'etichetta predefinita identifica la tipologia, il valore e la dimensione, oltre alla locazione

all'interno dello stesso file. La capacità autodescrittiva di un HDF file fornisce la possibilità di capire completamente la sua struttura e il metodo con cui trovare le informazioni ai diversi livelli: un programma specifico interpreta l'etichetta e processa il corrispondente dato. Il singolo file può accogliere numeri, simboli, grafici: tuttavia, le immagini raster e gli array multidimensionali sono spesso non georiferiti. Dato che la geolocalizzazione è necessaria per le misure alla base degli studi relativi al nostro pianeta, la NASA ha elaborato il formato HDF-EOS. Quest'ultimo contiene tre tipi di dati geospaziali (grid, point e swath) e la sua libreria facilita l'indagine sul contenuto del file, fornendo coordinate e informazioni temporali. Per comprendere le motivazioni che hanno spinto la NASA a scegliere questo formato, bisogna risalire al 1990, quando le ricerche scientifiche interdisciplinari hanno assunto un ruolo centrale nello studio della Terra in generale e, nel particolare, tra gli scopi di EOS: l'ente spaziale americano ha così cercato un formato che agevolasse lo scambio dei dati e la loro interpretazione.

Tra i software di processamento di immagini presenti in commercio, quelli in grado di manipolare i dati ASTER sono:

- ENVI version 3.5 (da RSI);
- Imagine version 8.5 (da ERDAS);
- PCI version 8.2 (PCI Geomatics);
- Matlab (da MathWorks);
- Noesys (RSI);
- IDRISI (da Clark University).

La versione IDRISI 32, realizzata da Clark University (Worcester, Massachusetts), è stata scelta come software per l'elaborazione dell'immagini prevista nella fase sperimentale della tesi. Il team responsabile di ASTER ha sviluppato degli strumenti per l'utilizzo di questi dati, che ne prevedono la conversione in altri formati: tra quelli disponibili vi sono WINVICAR, HEG -HDF to GeoTiff-Tool, Java HDF Viewer, MultiSpec, Webwinds.

Il software eCognition usato per la classificazione object oriented ha richiesto la trasformazione del formato originale in TIFF.

2.4 Dipendenza tra risoluzione spettrale e geometrica

Come già visto, le risoluzioni geometriche dei tre sottosistemi di Aster, che operano nel visibile-vicino infrarosso, nell'infrarosso a media lunghezza d'onda e nel termico, sono rispettivamente di 15, 30 e 90 metri, collocando tale tipo di dato tra quelli a media risoluzione spaziale.

Ma è evidente che la risoluzione geometrica non può essere valutata separatamente dalle altre risoluzioni in gioco:

- ✓ risoluzione radiometrica: intervallo di energia registrabile dal sensore
- ✓ risoluzione temporale: tempo che intercorre tra una ripresa e l'altra
- ✓ risoluzione spettrale

Il compromesso tra le risoluzioni deve essere preso in considerazione già nella fase di progettazione del sensore: per ottenere alte risoluzioni spaziali deve essere previsto un IFOV (Instantaneous Field of View) ridotto, il quale, però, comporta una riduzione della quantità di energia registrabile nella cella corrispondente all'area sul suolo, ovvero una bassa risoluzione radiometrica. Si sottolinea, inoltre, che se da un lato il flusso radiante Φ rilevato non dipende geometricamente dalla distanza di osservazione, ma, come detto, dalle caratteristiche strumentali di angolo di vista β (IFOV) e dall'area del rivelatore C , secondo la seguente formula

$$\Phi = L\beta^2 C, \quad \text{con } L = \text{radianza} \left(\frac{W}{m^2 \cdot sr \cdot \mu m} \right)$$

è, altresì, vero che la stessa quantità radiante dipende anche dall'ampiezza dell'intervallo spettrale considerato. Per aumentare l'energia registrabile senza ridurre la risoluzione spaziale, si dovrebbe ampliare, quindi, il range di lunghezza d'onda per un particolare canale o banda: ma questo significherebbe diminuire la risoluzione spettrale del sensore. Al contrario, una risoluzione spaziale bassa dovrebbe migliorare quella radiometrica e/o spettrale.

Perciò, nel dimensionamento spettrale di un sensore, si deve sempre tener conto della sua risoluzione geometrica al suolo, ma anche della risoluzione radiometrica, nel senso che bande spettrali troppo strette implicherebbero valori di flussi radianti troppo esigui da poter essere rilevati entro una dinamica opportuna di rappresentazione.

Le coesistenze tra le differenti risoluzioni rappresentano a volte un fattore limitante per l'utilizzo dei dati telerilevati nelle diverse applicazioni: devono essere bilanciate con la capacità desiderate e gli obiettivi del sensore. Due strade sono possibili: concentrare l'attenzione sulla risoluzione più

importante ai fini dello scopo perseguito, aspettandosi allo stesso tempo scarsi risultati per le altre risoluzioni, oppure trovare un accordo tra valori medi di ciascuna risoluzione.

In genere, comunque, al fine di mantenere un buon rendimento radiometrico nel sistema impiegato, si cerca un compromesso tra le caratteristiche di risoluzione spettrale e geometrica (si veda la figura 2.15 a tal proposito)

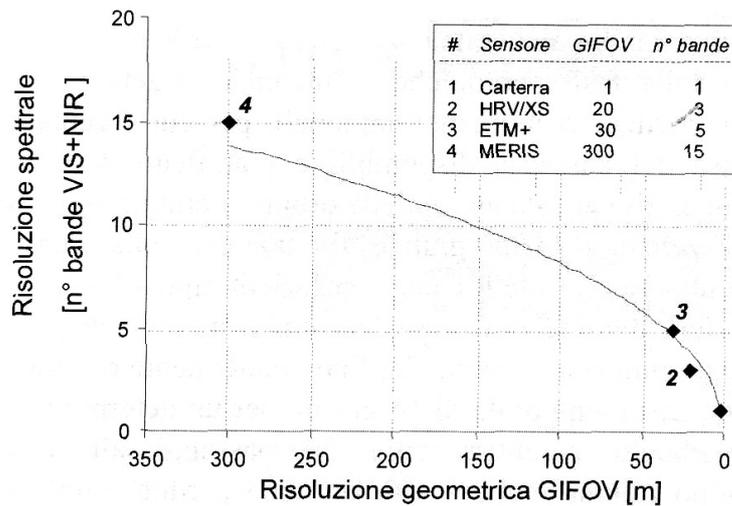


Fig. 2.15 – Dipendenza tra risoluzione geometrica e spettrale.

Alla luce di tutto ciò risulta più comprensibile la ridotta risoluzione spaziale del TIR, 90 m, rispetto a quelle degli altri due sottosistemi: considerando che il contenuto energetico diminuisce all'aumentare della lunghezza d'onda, i sensori termici possiedono, in genere, un IFOV (Instantaneous Field of View: cono angolare di visibilità dello strumento) elevato, per garantire che la radiazione captata dal detector sia sufficiente per l'affidabilità della misura.

2.5 Il futuro dell'analisi multispettrale

Come testimoniato da Michael Abrams, ricercatore scientifico e US ASTER Scienze Team Leader presso Jet Propulsion Laboratory (California Institute of Technology, Pasadena, California), nel suo intervento al convegno “I dati satellitari Aster: studi territoriali” (Bologna, 2007) e da Kenneth A. Duda, coordinatore scientifico NASA Land Processes DAAC presso USGS EROS (Sioux Falls, South Dakota), incontrato nel corso del 32nd ISRSE 2007, “Sustainable Development Through Global Earth Observations” (San Josè Costa Rica, 25-29 giugno 2007), per il satellite Terra, e in particolare per il sensore Aster, non è previsto un diretto successore come è accaduto per la missione Landsat. Nonostante sia già stato superato il numero di anni previsto per la sua attività, Aster è attualmente operativo ma in futuro, una volta dismesso, non sarà sostituito.

Anche se ora l'interesse statunitense è rivolto soprattutto alle missioni su Marte, è stata comunque programmata la ripresa della missione Landsat: Landsat Data Continuity Mission (LDCM), che subentrerà al Landsat 7, verrà lanciato nell'estate del 2011. Per le caratteristiche di questo sistema a confronto con quelle LANDSAT e ASTER si vedano le figure 2.16 e 2.17 rispettivamente.

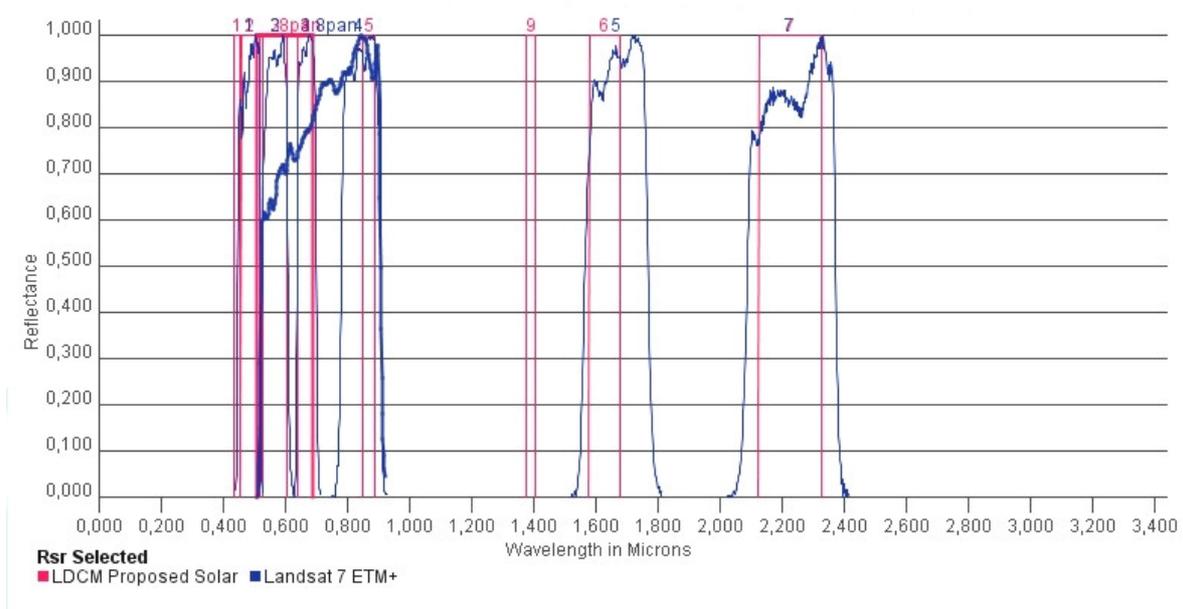


Fig. 2.16 – Confronto tra le bande LDCM e quelle Landsat 7

I satelliti Landsat hanno acquisito con continuità immagini multispettrali di quasi tutta la superficie terrestre dal 1972, generando l'archivio più ampio di riprese del nostro pianeta dallo spazio, unico per qualità, dettaglio e copertura. LDCM perseguirà obiettivi comuni ad Aster, primo fra tutti l'individuazione e la caratterizzazione dei cambiamenti terrestri naturali e di origine antropica, in termini di copertura e uso del suolo, dei quali le principali conseguenze sociali influenzano la produzione alimentare, le alterazioni climatiche, il consumo e la qualità dell'acqua e la salute umana.

La NASA ha scelto la Ball Aerospace and Technologies Corporation of Boulder (Colorado) per sviluppare quella parte di strumentazione destinata all'acquisizione delle immagini, che avverrà nel visibile e nel vicino infrarosso.

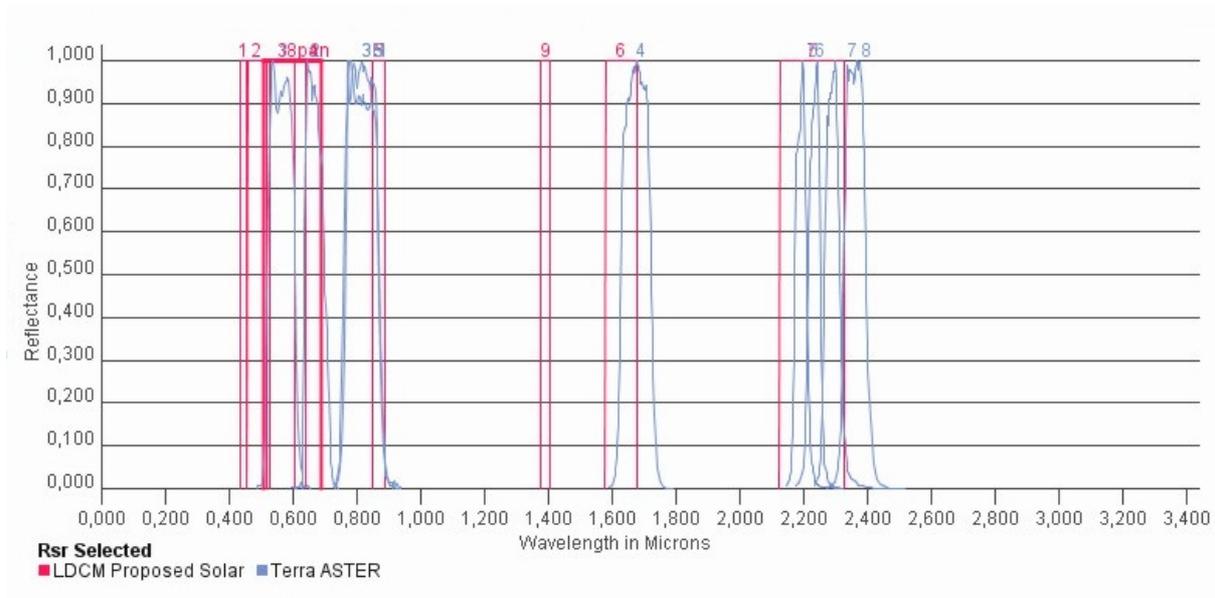


Fig. 2.17 – Confronto tra le bande LDCM e quelle Aster.

Le caratteristiche tecniche del nuovo sensore, ancora in fase progettazione, sono riportate nella seguente tabella 2.7.

NUMERO BANDA	BANDA	INTERVALLO DI	RISOLUZIONE SPAZIALE
		LUNHEZZA D'ONDA (µm)	AL NADIR (m)
1	Coastal/Aerosol	0.433 – 0.453	30
2	Blue	0.450 – 0.515	30
3	Green	0.525 – 0.600	30
4	Red	0.630 - 0.680	30
5	NIR	0.845 – 0.885	30
6	SWIR 1	1.560 – 1.660	30
7	SWIR 2	2.100 – 2.300	30
8	Panchromatic	0.500 – 0.680	15
9	Cirrus	1.360- 1.390	30

Tab. 2.7 – Principali caratteristiche del sensore LDCM

Le differenze sostanziali che emergono dal confronto tra i dati LDCM e Aster, e che influenzeranno anche le successive indagini scientifiche, sono la minore risoluzione per le bande dell'infrarosso (da 1 a 30 m) e la mancanza dell'infrarosso termico. Le immagini LDCM rimangono comunque le candidate più probabili per l'applicazione del metodo testato circa il riconoscimento delle colture agricole.

CAPITOLO 3: TECNICHE DI CLASSIFICAZIONE

3.1 Introduzione alle tecniche di classificazione

La classificazione di un'immagine telerilevata è il processo con il quale viene prodotta una cartografia tematica. La procedura operativa più comunemente impiegata, soprattutto in campo forestale e dell'uso del suolo, è la fotointerpretazione manuale (Lillesand et al., 2004). Come già descritto in precedenza, questa tecnica è anche prevista, attualmente, durante le fasi di verifica delle dichiarazioni PAC degli agricoltori (Politica Agricola Comune), eseguite dagli organi di controllo predisposti. Tuttavia, bisogna sottolineare che tale procedura è comunque condizionata da un certo grado di soggettività, sia per quanto riguarda la scelta dell'attribuzione dei poligoni o pixel alle diverse classi tematiche, sia per quanto riguarda la delimitazione del perimetro dei poligoni stessi. Per superare tali limiti sono stati sviluppati metodi di classificazione automatici (*unsupervised*) e semiautomatici (*supervised*) nel tentativo di ottenere prodotti più oggettivi, in tempi più brevi e a costi più contenuti di quelli derivati per fotointerpretazione.

Le tecniche di classificazione semiautomatica di dati telerilevati sono state ampiamente sperimentate per la realizzazione di cartografie di uso/copertura del suolo. I metodi *supervised* si basano sull'acquisizione del numero di pixel campione necessario per creare una firma spettrale da confrontare con quella dei pixel incogniti, mentre per quelli *unsupervised*, che non richiedono questa fase iniziale, le classi individuate vengono identificate direttamente dall'operatore al termine della classificazione, assimilabile ad una *cluster analysis*.

Purtroppo, però, le classificazioni *pixel oriented* hanno lo svantaggio di produrre cartografie in formato raster, che risultano poco fruibili da parte degli utenti finali, abituati a gestire mappe in formato vettoriale. Negli ultimi anni, sono state sviluppate tecniche di classificazione *object-oriented* (Batz e Schäpe, 1999), particolarmente utili per il trattamento di dati satellitari. I sistemi *object-oriented* permettono di ottenere un prodotto finale di tipo vettoriale, dove i poligoni, generati attraverso un processo automatico di *segmentazione* delle immagini, possono essere classificati con algoritmi di tipo *supervised* e *unsupervised*, oppure per fotointerpretazione manuale. A differenza dei classici algoritmi *pixel-oriented* che classificano separatamente ciascun pixel dell'immagine, un

vantaggio dell'approccio *object-oriented* risiede nel maggior contenuto informativo associabile ai poligoni vettoriali rispetto a quello derivabile per i singoli pixel dell'immagine.

Uno degli obiettivi della Tesi di Dottorato è confrontare i risultati ottenuti con l'applicazione di entrambe le tecniche sulle diverse immagini, per stabilire, eventualmente, se l'una sia preferibile all'altra, sia in termini di accuratezza che di costi e tempi, o per verificare l'indifferenza nella scelta di una specifica metodologia.

I software utilizzati per testare le due tipologie di classificazioni sono di tipo commerciale e da diverso tempo presenti sul mercato. Per la *classificazione per pixel* è stato impiegato IDRISI 32, realizzato da Clark University (Worcester, Massachusetts - USA), mentre per la *object-oriented* si è ricorsi al software per eccellenza in questo ambito, ovvero eCognition, creato da Definiens Imaging GmbH (München, Germany). In previsione di un utilizzo pubblico della metodologia sperimentata in questo lavoro per il monitoraggio delle colture agricole, lo scopo ultimo da raggiungere è la possibilità di sostituire tali software con delle valide versioni gratuite, ad esempio Multispec (School of Electrical and Computer Engineering, Purdue University, Indiana – USA) o Spring 4.0 (Istituto Nacional de Pesquisas Espaciais, San Josè dos Campos, Brazil).

3.2 Classificazione per pixel

La classificazione utilizza le B bande dello spazio multispettrale come assi ortogonali di uno spazio fittizio B -dimensionale lineare, che viene convenzionalmente chiamato SPAZIO MULTISPETTRALE (*multispectral space*) dell'immagine (Fig. 3.1).

Ogni pixel, cui corrispondono i valori x^1, x^2, \dots, x^B registrati nelle bande, può essere visto come un punto appartenente a tale spazio: x^1, x^2, \dots, x^B rappresentano le sue coordinate cartesiane e il vettore posizione è definito da

$$\mathbf{x} = \begin{bmatrix} x^1 \\ x^2 \\ \cdot \\ \cdot \\ x^B \end{bmatrix}$$

Il punto essenziale di ogni procedura di classificazione, è quello di assegnare i pixel alla medesima classe, quando essi sono reciprocamente vicini e a classi differenti, se sono distanti. La scelta, che caratterizza ogni approccio, avviene con la definizione di *distanza* e col modo con cui le regioni dello spazio multispettrale sono associate alle diverse classi di copertura del territorio.

In particolare, nella classificazione supervisionata, l'area circostante ciascun pixel di classe nota, e quindi i pixel incogniti contenuti in essa, vengono associati alla medesima categoria; il problema in questo caso è nel metodo seguito per definire i confini fra le differenti classi.

La separazione dello spazio multispettrale in classi o aree, è anche il punto fondamentale della classificazione non supervisionata. In tal caso, però, vi è una differenza fondamentale: le aree che definiscono i gruppi di pixel (CLUSTER) vengono individuate senza l'ausilio di altre informazioni. Una volta formati i cluster, la loro classe viene identificata mediante la conoscenza di quella classe di alcuni pixel appartenenti a ciascun cluster. Tale approccio è per sua natura legato a metodi di tipo iterativo, ove i cluster vengono uniti o divisi in base a uno schema predefinito di classificazione.

Con riguardo ai criteri di selezione della distanza nello spazio multispettrale, o equivalentemente di definizione dei confini fra le regioni corrispondenti a classi particolari, possiamo distinguere, in funzione dell'utilizzo o meno degli strumenti della statistica, fra *metodi deterministici* e *metodi statistici* di classificazione.

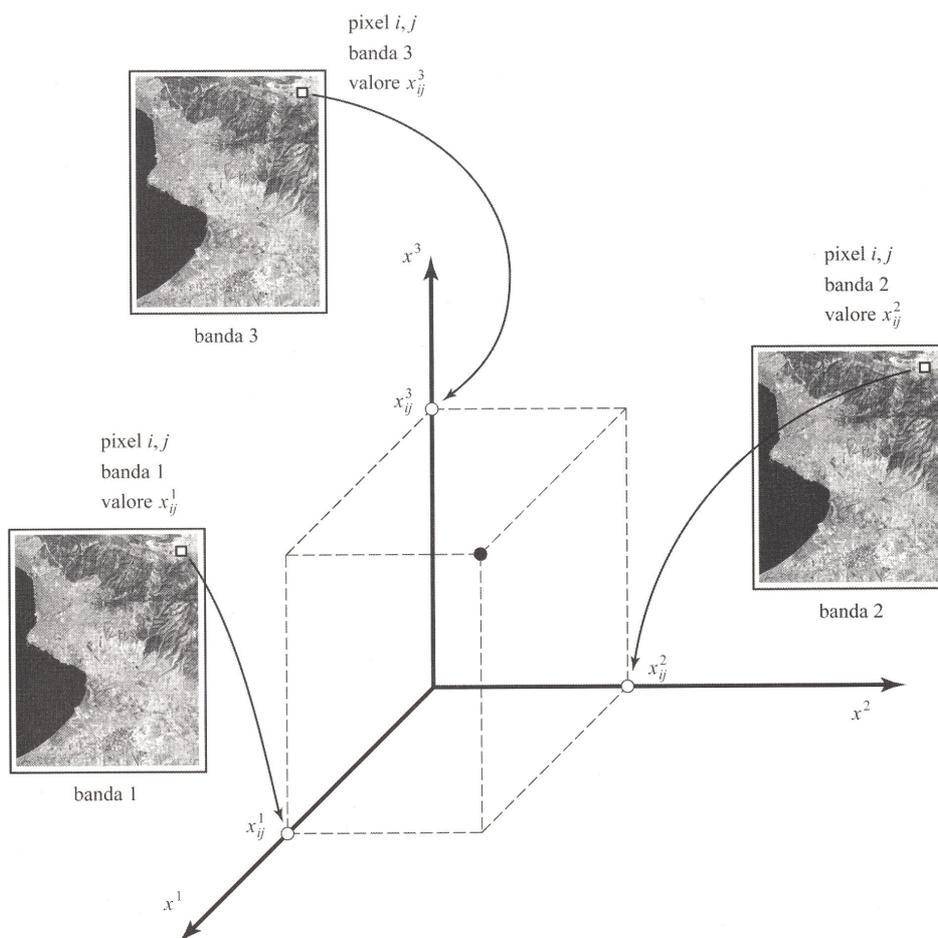


Fig. 3.1 – Concetto di spazio multispettrale in relazione ai valori assunti dai pixel dell'immagine nelle sue diverse bande (caso a tre bande)

I metodi statistici considerano, appunto, il comportamento statistico dei pixel presenti in ciascuna classe. Idealmente, la funzione multivariata di densità di probabilità $p_i(x)$ di ciascuna classe i dovrebbe essere nota; in genere, si utilizzano unicamente il vettore media m_i (contenente i valori medi di tutte le B bande) e la matrice di covarianza C_i per ciascuna classe: la scelta nasce dall'ipotesi che, per ciascuna classe, i pixel seguano una statistica multivariata normale e che, quindi, media e covarianza siano sufficienti a descrivere la funzione $p_i(x)$. Considerazioni statistiche possono anche portare a metodi di classificazione che minimizzino, in senso lato, il rischio (probabilità) di un'assegnazione sbagliata.

Infine, oltre alle k classi specifiche di copertura del suolo $\omega_1, \omega_2, \dots, \omega_k$ nelle quali si vuole separare lo spazio multispettrale, si deve considerare una classe aggiuntiva ω_0 per i pixel non classificati; ciò è necessario per evitare assegnazioni caratterizzate da alta probabilità di errore, perché cadono in mezzo a classi differenti.

3.2.1 Classificazione supervisionata

Nel corso di studi precedenti, è stata approfondita la conoscenza delle tecniche non supervisionate per il riconoscimento delle colture agricole (Verrucci, 2005): risultati soddisfacenti sono stati raggiunti solo con l'applicazione iterativa degli algoritmi e con il confronto con i risultati ottenuti contemporaneamente da classificazioni supervisionate sulle medesime immagini. Nell'ottica di creare una metodologia semplice e veloce a disposizione di utenti generici, i metodi *unsupervised* non rispondono, quindi, ai requisiti richiesti in termini di dispendio di tempo e di costi. Di conseguenza, l'attenzione è stata rivolta esclusivamente agli algoritmi *supervised*.

La *classificazione supervisionata* (supervised classification) si basa sull'ipotesi di conoscere a priori le classi di appartenenza per un numero sufficiente di pixel, chiamati *pixel campione* (sample). I pixel restanti, comunque in numero molto maggiore, vengono quindi classificati in base alla loro somiglianza con i pixel campione. Il concetto di somiglianza viene esteso a tutte le bande, ovvero viene valutata la vicinanza nello spazio multispettrale. Diverse misure di distanza generano specifici algoritmi di classificazione, che differiscono tra loro per semplicità computazionale ed efficacia teorica. Gli algoritmi più validi (da un punto di vista teorico) richiedono maggiore tempo di calcolo; vi sono, però, casi in cui gli algoritmi più semplici forniscono i risultati migliori. In sostanza una grande esperienza, maturata nell'analisi di tipi differenti di dati e di applicazioni, è condizione necessaria per una scelta ottimale del metodo da utilizzarsi in ogni caso specifico.

Un elemento essenziale nella classificazione supervisionata è la definizione delle classi $\omega_1, \omega_2, \dots, \omega_k$ cui dovranno essere assegnati i pixel dell'immagine. Per lo scopo generale della produzione di

cartografia tematica, si possono definire tipologie standard di classi; rimane, comunque, da decidere quali categorie includere nell'elaborazione, in base alla disponibilità di pixel campione. Vi sono, inoltre, applicazioni con obiettivi molto specifici, ove la selezione di classi dovrebbe essere lasciata al giudizio dell'utente. Per esempio, l'identificazione delle parti di foresta distrutte da incendi, l'analisi delle terre coperte da esondazioni, il monitoraggio di culture specifiche, sono tutte applicazioni finalizzate all'individuazione di un particolare tipo di classe e non alla distinzione delle altre, che possono essere genericamente accorpate.

Altro fattore critico per il successo della classificazione, è l'identificazione nell'immagine dei pixel campione. Questi, devono essere in numero sufficiente per ogni classe e la loro assegnazione non deve essere ambigua. Ciò richiede una valutazione delle aree omogenee, per evitare la presenza di pixel misti e per localizzare correttamente i pixel sul terreno. Si ricorda, infatti, che, nella maggior parte dei casi, è necessario un rilievo sul campo per l'identificazione delle classi; inoltre, per le classi corrispondenti a coperture del suolo variabili nel tempo, come le attività agricole, il rilievo sul campo deve essere condotto approssimativamente nel medesimo periodo dell'osservazione delle immagini.

I pixel noti, per ciascuna delle classi predefinite $\omega_1, \omega_2, \dots, \omega_k$, formano rispettivamente insiemi di campioni S_1, S_2, \dots, S_k (*training sites*), che contengono un numero di pixel pari a n_1, n_2, \dots, n_k . 'Training site' significa *luogo di addestramento* in quanto serve ad addestrare il classificatore su cosa cercare.

3.2.1.1 Logica generale

DETERMINAZIONE DEI TRAINING SITES

Il primo passo, richiesto da questa classificazione, consiste nel definire le aree da usare come *training sites* per ciascuna classe di copertura. Per rendere agevole l'identificazione, tra le immagini a disposizione, viene scelta la banda che presenta il maggiore contrasto (in genere il vicino-infrarosso) o un'immagine a *color composite*. I migliori training sites sono quelli che contengono i pixel più puri: ad esempio, per definire il site della foresta decidua, è importante scegliere un'area che non contenga anche conifere, suolo visibile e altra vegetazione. Inoltre, esiste il rischio di includere pixel appartenenti alle adiacenti coperture. Inconvenienti di questo tipo possono essere evitati, almeno in parte, utilizzando delle funzioni di zoom che permettono una visione più chiara dell'area considerata. Materialmente, vengono tracciati dei poligoni in corrispondenza delle aree omogenee e viene assegnato un *identificatore* unico per ciascun tipo di copertura. Il numero di pixel per ogni training site deve essere adeguato per la caratterizzazione statistica: una regola generale

prevede che tale numero non sia inferiore a dieci volte il numero di bande utilizzate. Qualora risulti difficile raggiungere questo limite con un unico training site, si possono prevedere più poligoni con lo stesso identificatore (Fig. 3.2).

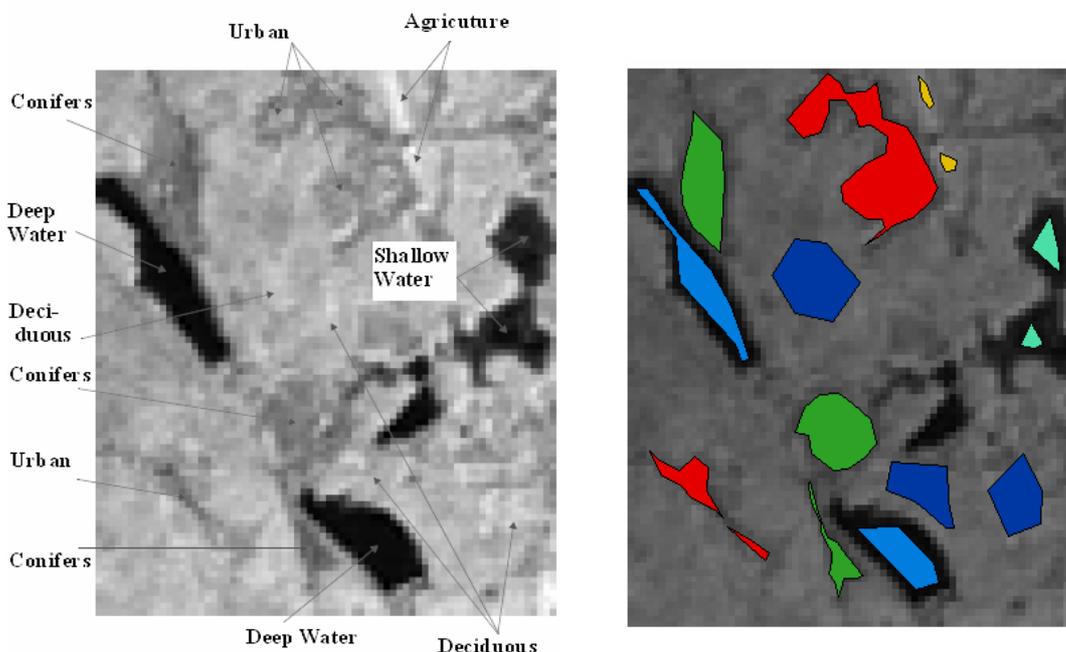


Fig. 3.2 – Esempio di estrazione dei training sites (a sinistra) dall’immagine originale (a destra).

ESTRAZIONE DELLA ‘FIRMA’

Il passo successivo richiede le caratterizzazioni statistiche (media, covarianza) dei valori di riflettanza (x) dei pixel contenuti nei training sites, per ciascuna classe.

Devono essere forniti al programma gli insiemi S_1, S_2, \dots, S_k , le bande usate per lo sviluppo della firma e l’elenco dei nomi delle classi, ottenendo come risultato le *firme (signatures)*.

Si precisa che vengono calcolate le stime $\hat{m}_1, \hat{m}_2, \dots, \hat{m}_i$ dei corrispondenti valori medi m_1, m_2, \dots, m_k mediante

$$\hat{m}_i = \frac{1}{n_i} \sum_{x \in S_i} x \quad i=1,2,\dots,k$$

e, come richiede la maggior parte degli algoritmi, la stima delle matrici di covarianza

$$\hat{C}_i = \frac{1}{n_i} \sum_{x \in S_i} (x - \hat{m}_i)(x - \hat{m}_i)^T \quad i = 1,2,\dots,k$$

con n_i pari al numero di pixel per ciascuna classe i .

La distinzione fra le stime \hat{m}_i , \hat{C}_i e i corrispondenti valori veri m_i , C_i è essenziale, poiché i primi sono calcolati a partire da un numero sostanzialmente piccolo di campioni in ciascuna classe, i secondi, secondo un approccio probabilistico rigoroso, dovrebbero corrispondere all'insieme infinito di tutti i pixel della classe. Comunque, dato che gli effetti atmosferici, che variano da immagine a immagine, influiscono anche sulla singola immagine, si considerano \hat{m}_i e \hat{C}_i come rappresentativi di tutti i pixel di una particolare immagine. Per semplicità, gli algoritmi vengono presentati utilizzando i valori teorici di media e covarianza, ma rimane sottinteso che, nel processo di stima, essi vengono sostituiti con le stime campionarie disponibili.

A seconda che siano presenti o meno impurità all'interno dei pixel, tale operazione è svolta da funzioni diverse, dando luogo a classificazioni differenti.

CLASSIFICAZIONE DELL'IMMAGINE

La classificazione dell'intera immagine avviene attraverso il confronto della firma di ciascun pixel, considerato singolarmente, con ognuna delle firme conosciute. Sono chiamati *hard classifications* i risultati ottenuti dall'assegnazione di ogni pixel al tipo di copertura che ha la firma più simile. Sono, invece, dette *soft classifications* quelle che valutano il grado di appartenenza del pixel a tutte le classi considerate. In entrambi i casi, la decisione circa la somiglianza della firma viene presa tramite un'analisi statistica prevista da molte tecniche differenti. Tuttavia senza una conoscenza dettagliata della zona di studio e di dati rilevati sul campo non è possibile definire quale classificazione dia risultati migliori. In genere si utilizzano più classificatori per valutare le caratteristiche delle firme e la natura del paesaggio.

Nel corso della stesura della Tesi di Laurea sono state utilizzate sia tecniche *hard* che *soft*, con l'obiettivo di discriminare non solo tra loro le diverse coperture presenti, in particolare le colture agricole, ma anche di individuare, all'interno di queste, delle variazioni di comportamento spettrale, causate, ad esempio, dalla presenza di piante malate o stressate, ecc (De Giglio, 2004). I classificatori *soft*, che prevedono comunque più passaggi nell'elaborazione dell'immagine, sono risultati i più idonei al raggiungimento di tale livello di dettaglio: purtroppo però, come già detto in precedenza per le tecniche *unsupervised*. Essi non si collocano perfettamente all'interno di una metodologia che deve essere usufruibile da ogni genere di utente, che richiede quindi una procedura quanto più possibile semplice e lineare, oltre che veloce.

L'algoritmo che, anche nella bibliografia esistente da alcuni anni in materia, ha presentato i migliori risultati dal punto di vista dell'accuratezza è quello della Massima Verosimiglianza che viene brevemente illustrato di seguito.

Se si accetta l'ipotesi che la distribuzione dei pixel di ciascuna classe segua una distribuzione multivariata normale (Gaussiana), è possibile determinare la funzione densità di probabilità, o funzione di massima verosimiglianza, $l_i(\mathbf{x})$

$$l_i(\mathbf{x}) = \frac{1}{(2\pi)^{B/2} |\mathbf{C}_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m}_i)^T \mathbf{C}_i^{-1}(\mathbf{x}-\mathbf{m}_i)\right]$$

Nel metodo di classificazione di MASSIMA VEROSIMIGLIANZA (maximum likelihood) ogni pixel \mathbf{x} viene assegnato alla classe ω_i cui corrisponde il valore massimo della funzione $l_i(\mathbf{x})$:

$$l_i(\mathbf{x}) > l_k(\mathbf{x}) \quad , \quad \forall k \neq i \quad \Leftrightarrow \quad \mathbf{x} \in \omega_i$$

In linea generale si può anche utilizzare, al posto di $l_i(\mathbf{x})$, una qualche funzione alternativa $d_i(\mathbf{x})$: è ammesso che, per ogni coppia $i \neq k$, valga la

$$l_i(\mathbf{x}) > l_k(\mathbf{x}) \quad \Leftrightarrow \quad d_i(\mathbf{x}) > d_k(\mathbf{x})$$

Facendo riferimento all'esempio riportato nei classificatori precedenti, il risultato finale può essere rappresentato con una zona ellittica di caratterizzazione della firma, (Fig 3.3).

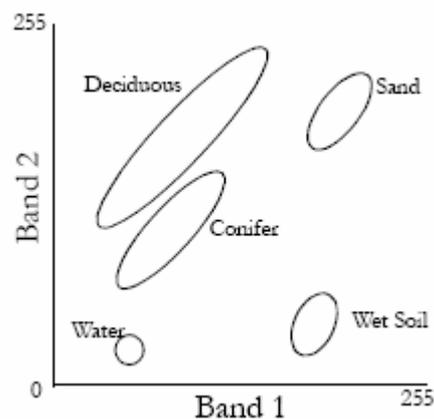


Fig. 3.3 – Metodo di massima verosimiglianza.

Questa procedura cerca di compensare le mancanze dei metodi della minima distanza e del parallelepipedo: in particolare la distanza euclidea e quella di Mahalanobis ne rappresentano dei casi semplificati. Quando è certa la qualità dei training sites, è il migliore dei classificatori hard, ma anche il più lento.

Teoria di probabilità bayesiana

Si considera ora un approccio statistico più generale: la *teoria di probabilità bayesiana*.

Dati i seguenti parametri:

N : numero totale di pixel nell'immagine (per ogni banda);

B : numero di bande;

$\omega_1, \omega_2, \dots, \omega_k$: le K classi presenti nell'immagine;

n_i : numero di pixel appartenenti alla classe ω_i ($i = 1, 2, \dots, K$);

n_x : numero di pixel con valore x (vettore dei valori in tutte le bande);

n_{xi} : numero dei pixel con valore x e appartenenti alla classe ω_i ($i = 1, \dots, k$);

si introducono le seguenti relazioni

$$\sum_x n_x = N, \quad \sum_i n_i = N, \quad \sum_x n_{xi} = n_i, \quad \sum_i n_{xi} = n_x$$

Si possono identificare le probabilità con le frequenze: si pone, cioè, che la frequenza con la quale si presenta una certa proprietà dell'immagine sia uguale alla probabilità che un pixel selezionato in modo casuale soddisfi tale proprietà. Seguono i diversi concetti di probabilità:

$$p(\omega_i) = \frac{n_i}{N} \quad \text{probabilità che un pixel appartenga alla classe } \omega_i \text{ (probabilità anteriore);}$$

$$p(\mathbf{x}) = \frac{n_x}{N} \quad \text{probabilità che un pixel abbia il valore } \mathbf{x} = [x^1 x^2 \dots x^B]^T ;$$

$$p(\mathbf{x}|\omega_i) = \frac{n_{xi}}{n_i} \quad \text{probabilità che un pixel appartenente alla classe } \omega_i \text{ assuma valore } x$$

(probabilità condizionata o posteriore);

$$p(\omega_i|\mathbf{x}) = \frac{n_{xi}}{n_x} \quad \text{probabilità che un pixel con valore } x \text{ appartenga alla classe } \omega_i \text{ (probabilità}$$

condizionata o posteriore);

$$p(\mathbf{x}, \omega_i) = \frac{n_{xi}}{N} \quad \text{probabilità che un pixel abbia valore } x \text{ e appartenga alla classe } \omega_i$$

(probabilità congiunta).

Come già visto, la funzione di probabilità $p = (\bar{x} / \omega_i)$ viene stimata dall'analisi statistica dei valori dei pixel campione dei training sites. La pratica comune è quella di calcolare la media e la matrice di covarianza per ogni classe e quindi di derivare la funzione di densità di probabilità assumendo che la distribuzione dei pixel sia normale.

Si può quindi calcolare la probabilità $p(\mathbf{x} | \omega_i)$, dopo aver introdotto l'ovvia identità

$$\frac{n_{xi}}{n_x} = \frac{\frac{n_{xi}}{N}}{\frac{n_x}{N}} = \frac{\frac{n_{xi}}{N_i} \frac{N_i}{N}}{\frac{n_x}{N}}$$

dalla quale segue la *formula di Bayes*

$$p(\omega_i | \mathbf{x}) = \frac{p(\mathbf{x} | \omega_i) p(\omega_i)}{p(\mathbf{x})}$$

posta in forma utile per gli scopi della classificazione di dati telerilevati, rappresenta la possibilità che il pixel studiato, con valore di riflettanza x , rientri nella categoria ω_i .

Siano A e B due eventi stocasticamente dipendenti: la forma originale della formula di Bayes postula che la probabilità che si realizzi A , qualora sia noto che si è realizzato B , è data dalla

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

A causa della simmetria vale anche

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}$$

quindi

$$\Pr(A|B)\Pr(B) = \Pr(A \cap B) = \Pr(B|A)\Pr(A)$$

e infine

$$\Pr(B|A) = \frac{\Pr(A|B)\Pr(B)}{\Pr(A)}$$

L'equivalenza tra i due casi si ottiene identificando l'evento A con x e l'evento B con la classe ω_i .

Si utilizza la probabilità congiunta $p(\mathbf{x}, \omega_i)$ per derivare forme alternative della formula di Bayes.

La relazione

$$p(\mathbf{x}, \omega_i) = \frac{n_{xi}}{N} = \frac{n_{xi}}{N_i} \frac{N_i}{N} = p(\mathbf{x} | \omega_i) p(\omega_i)$$

fornisce la

$$p(\mathbf{x}|\omega_i) = \frac{p(\mathbf{x}, \omega_i)}{p(\omega_i)}$$

mentre la relazione

$$p(\mathbf{x}, \omega_i) = \frac{n_{xi}}{N} = \frac{n_{xi}}{n_x} \frac{n_x}{N} = p(\omega_i|\mathbf{x}) p(\mathbf{x})$$

porta alla forma

$$p(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}, \omega_i)}{p(\mathbf{x})}$$

La conoscenza della probabilità $p(\mathbf{x})$ non è necessaria poiché rappresenta un denominatore comune per tutte le $p(\omega_i|\mathbf{x})$, $i = 1, 2, \dots, K$. Quindi in genere si scrive la formula di Bayes nella forma

$$p(\omega_i|\mathbf{x}) \propto p(\mathbf{x}|\omega_i) p(\omega_i)$$

ove \propto indica la proporzionalità tra i due membri. In ogni caso $p(\mathbf{x})$ può essere calcolato mediante la

$$p(\mathbf{x}) = \frac{n_x}{N} = \frac{1}{N} \sum_i n_{xi} = \sum_i \frac{n_{xi}}{N_i} \frac{N_i}{N} = \sum_i p(\mathbf{x}|\omega_i) p(\omega_i)$$

Dall'ultima relazione si ottiene la formula di Bayes in forma completa

$$p(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i) p(\omega_i)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\omega_i) p(\omega_i)}{\sum_i p(\mathbf{x}|\omega_i) p(\omega_i)}$$

che, a partire dalle probabilità $p(\mathbf{x}|\omega_i)$ stimate nel processo di addestramento, può essere utilizzata per calcolare le probabilità $p(\omega_i|\mathbf{x})$ da usarsi nella classificazione.

Quando le informazioni complete sono già disponibili o presupposte, la teoria di probabilità bayesiana valuta il rapporto fra la *prova indiretta* e l'*insieme di decisioni*. Quest'estensione della teoria di probabilità classica permette di combinare la nuova prova \mathbf{x} (associata al pixel sconosciuto), relativa a un'ipotesi ω_i , con la conoscenza a priori $p(\omega_i)$, per arrivare a stimare la probabilità che quell'ipotesi sia vera. Uno dei casi più semplici è dato dalla scelta fra due sole possibilità: ω_1 e ω_2 , con $p(\omega_1)$ e $p(\omega_2)$ le rispettive probabilità. Ad esempio, si vuole assegnare

un'area a una delle due classi date: inizialmente si verifica la conoscenza anteriore che può condurre alla probabilità che una o l'altra ipotesi sia vera. In caso di sua assenza tutte le ipotesi sono considerate ugualmente probabili. Il termine $p(\mathbf{x}|\omega_i)$ esprime la probabilità di trovare la prova se è accertata l'ipotesi. Si chiama anche probabilità condizionata ed è calcolata sulla base di aree, in cui è nota la veridicità dell'ipotesi, raccogliendo i dati utili per valutare la probabilità che la nuova prova sia coerente con l'ipotesi. Invece, il termine $p(\omega_i|\mathbf{x})$ è una probabilità posteriore creata dopo il confronto tra la combinazione della conoscenza a priori e della prova e il set di decisioni. Incrementando le notizie utili circa le ipotesi, la probabilità di ciascuna ipotesi viene modificata per incorporare le nuove informazioni. Il presupposto del teorema di Bayes, è che le informazioni complete siano realizzabili e che, un loro mancato raggiungimento di valutazione di probabilità, sia dovuto alla sola mancanza di prova: aggiungendo più prova alla conoscenza, teoricamente, si può raggiungere una reale valutazione per tutte le ipotesi.

In conclusione, nella classificazione bayesiana un pixel con valore x viene assegnato alla classe ω_i quando, per ogni $k \neq i$

$$p(\omega_i|\mathbf{x}) > p(\omega_k|\mathbf{x})$$

3.3 Classificazione Object-Oriented

3.3.1 Segmentazione

Il processo di classificazione di immagini digitali in un processo basato sugli oggetti, implica una pre-classificazione dell'immagine basata sulle caratteristiche di omogeneità della stessa. Questo processo è denominato SEGMENTAZIONE ed implica l'aggregazione dei singoli pixel in poligoni a diversi valori di scala ("oggetti simili"), sulla base di informazioni-spettrali e di forma (Herold et al., 2002).

Gli obiettivi dell'analisi di un'immagine possono essere diversi: molto frequenti sono gli studi sull'uso del suolo e delle coperture tramite tecniche di classificazione oppure l'estrazione automatica di oggetti. In ogni caso, gli oggetti rappresentati in un'immagine sono sempre molto eterogenei da un punto di vista geometrico e della risposta spettrale fornita.

Nell'approccio *object-oriented* implementato in eCognition, gli oggetti identificati dalla procedura di segmentazione possono essere considerati prima di tutto come delle primitive geometriche su cui poi si andranno ad effettuare procedure di classificazione tramite le quali verrà loro assegnata una classe di appartenenza.

Per avere buoni risultati devono essere tenuti in considerazione i seguenti elementi nelle prime fasi del lavoro di segmentazione:

- ☞ Il processo di segmentazione dovrebbe produrre segmenti altamente omogenei per l'individuazione di regioni dell'immagine con caratteristiche simili;
- ☞ La scala media degli oggetti estratti dall'immagine deve essere coerente con gli scopi dell'indagine, dal momento che nei dati sono presenti molte informazioni geometriche associate a diverse dimensioni reali degli oggetti;
- ☞ Quasi tutti gli attributi degli oggetti contenuti in un'immagine – toni, texture, forme, relazioni con oggetti adiacenti – sono più o meno dipendenti dalla scala. Solo strutture di scala simile sono di qualità confrontabile ed hanno attributi anch'essi confrontabili. Per questa ragione, gli oggetti risultanti da un'analisi devono essere all'incirca della stessa grandezza;
- ☞ La procedura di segmentazione dovrebbe essere generalizzabile, ovvero applicabile in modo efficace indipendentemente dalla tipologia di dati su cui si deve lavorare;
- ☞ L'osservazione remota terrestre spesso produce grandi data sets. La procedura di segmentazione dovrebbe essere, quindi, il più veloce possibile.

3.3.1.1 Procedura

La segmentazione è una tecnica di tipo *bottom up* che porta alla separazione di regioni sull'immagine a partire dall'oggetto minimo, caratterizzato da un singolo pixel (*region-merging*). Tali regioni si formano per accrescimento (*merging*) nelle successive fasi di analisi dei dati. In queste procedure di clustering, il criterio di base è quello della minimizzazione dell'eterogeneità media degli oggetti, pesata in funzione della loro dimensione (Fig. 3.4): eterogeneità che può essere definita come nh dove n è la dimensione dell'oggetto ed h l'eterogeneità definita in qualche modo arbitrario. L'unione delle frazioni segmentate procede in modo da minimizzare il valore di eterogeneità, che non deve superare una soglia prestabilita. In tal caso il processo di accrescimento termina.

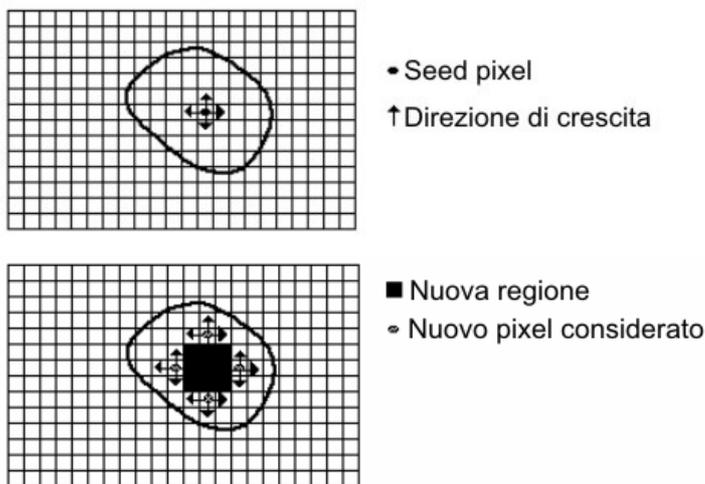


Fig. 3.4 - Schema delle fasi della tecnica region-merging

Di seguito, si riportano le fasi previste dalla procedura *region-merging*:

1. Un set iniziale di piccole aree viene iterativamente accorpato nel rispetto di alcuni vincoli di similarità;
2. Il processo comincia con la scelta arbitraria del cosiddetto “seme” (*seed pixel*) che viene confrontato con i pixel circostanti;
3. La regione cresce in dimensione a partire da questi semi attraverso l’unione dei pixel confinanti che rispettano le condizioni di somiglianza;
4. Quando la crescita di una regione si ferma verrà scelto un altro seme che non appartiene ancora ad alcuna regione ed il processo ricomincia;
5. Il processo di region-merging termina quando tutti i pixel sono assegnati ad una regione.

Si distinguono due tipologie diverse di vicinanza al pixel: piana e diagonale (Fig. 3.5). La vicinanza piana prevede come vicini solo quelli che hanno un lato in comune, mentre quella diagonale prevede come vicini quelli che hanno un lato o uno spigolo in comune.

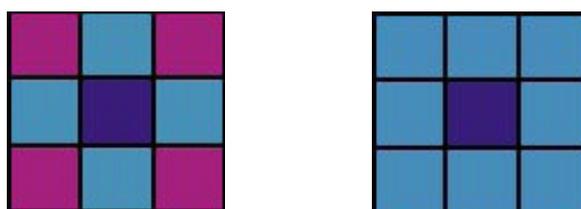


Fig. 3.5 - Esempio di vicinanza piana (a sinistra), nel quale i pixel comuni sono 4, e di vicinanza diagonale (a destra), nel quale i pixel comuni sono 8

La scelta della tipologia di “vicini” condiziona il processo di segmentazione, in quanto essa condiziona tutte le fasi di elaborazione dei dati. In molte applicazioni, la vicinanza piana rappresenta la scelta migliore, essendo quella che porta ad una maggiore velocità di calcolo, rispetto alla condizione di vicinanza diagonale. Invece, in immagini a media-bassa risoluzione, con oggetti di interesse che possono essere di dimensione simile alla scala del pixel (ad esempio dati LANDSAT o ASTER), è utile utilizzare la vicinanza diagonale. Si pensi ad un oggetto sull’immagine (ad esempio una strada) connesso ad un altro solo attraverso uno spigolo. Come esemplificato nella figura 3.6, la condizione di vicinanza deciderà se accorpare i due oggetti o meno.

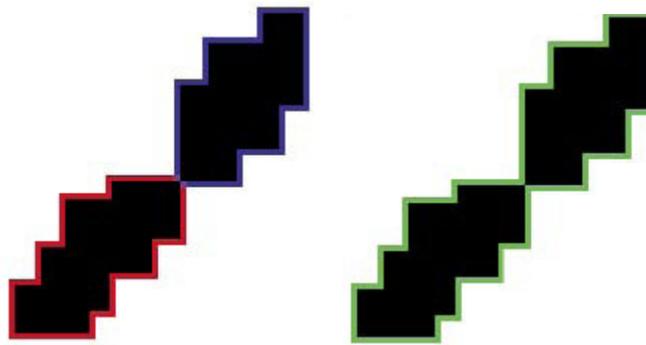


Fig. 3.6 – L’oggetto segmentato è composto di due parti nel caso di vicinanza piana (immagine a sinistra) e da una unica nel caso di vicinanza diagonale ((immagine a destra)

Nel caso di vicinanza diagonale il processo di analisi richiede più tempo essendo superiori le condizioni da analizzare.

CRITERI DI ETEROGENEITÀ

Il processo di segmentazione coinvolge diversi parametri fra loro interconnessi (Batz et al, 2004).

Il primo parametro da stabilire è lo *scale*, responsabile della dimensione media degli oggetti che saranno creati. In altre parole, quando, ad esempio, il valore di *scale* è 20, ciò significa che gli oggetti prodotti conterranno almeno 20 pixel, con differenti valori, che saranno raggruppati a formare un oggetto, in termini vettoriali, una *feature*. In questo modo, in quella parte di una immagine caratterizzata da elevata omogeneità, come in un appezzamento, si genereranno oggetti relativamente grandi, mentre dove è presente una elevata eterogeneità, come in una foresta, si produrranno oggetti relativamente piccoli (Mognol, 2007). Il parametro di scala definisce anche un criterio di rottura. Infatti è anche una misura della massima variazione dei valori di eterogeneità che si verificano quando due oggetti vengono fusi in uno unico. Tale valore, generalmente presente al quadrato, viene utilizzato proprio come soglia per la conclusione del processo di segmentazione.

Quando l'algoritmo valuta la possibilità di procedere con il merging di due oggetti, viene calcolato il parametro di scala quadratico e fatto il confronto con l'oggetto nato dalla fusione dei due.

Dal punto di vista matematico il parametro *scale* si può esprimere come segue:

$$Scale = w_{color} \cdot h_{spectral} + (1 - w_{color}) \cdot h_{shape}$$

Il valore *scale* è quindi una media pesata (w_{color} esprime il peso) tra un fattore di eterogeneità spettrale ($h_{spectral}$) ed un valore di eterogeneità di forma (h_{shape}). Se si pone w_{color} pari ad 1 significa dare peso solamente all'eterogeneità spettrale, mentre se si pone uguale a 0 si privilegia solo la forma. Il fattore spettrale (o di colore) $h_{spectral}$ si definisce come la sommatoria della deviazione standard (σ_c) dei *pixel* costituenti un oggetto, moltiplicata per un fattore di peso relativo alle diverse bande (w_c):

$$h_{spectral} = \sum w_c \cdot \sigma_c$$

Grazie a questo, è possibile utilizzare dati provenienti da fonti diverse, caratterizzate da differenti risoluzioni spaziali e radiometriche, come, ad esempio, un *DTM* insieme ad una immagine multispettrale.

In molti casi la sola minimizzazione dell'eterogeneità spettrale porta a fenomeni di ramificazione delle parti segmentate, oppure ad oggetti con contorni molto irregolari (Fig. 3.7). Questo fenomeno è evidente soprattutto in dati altamente texturizzati quali i dati radar.

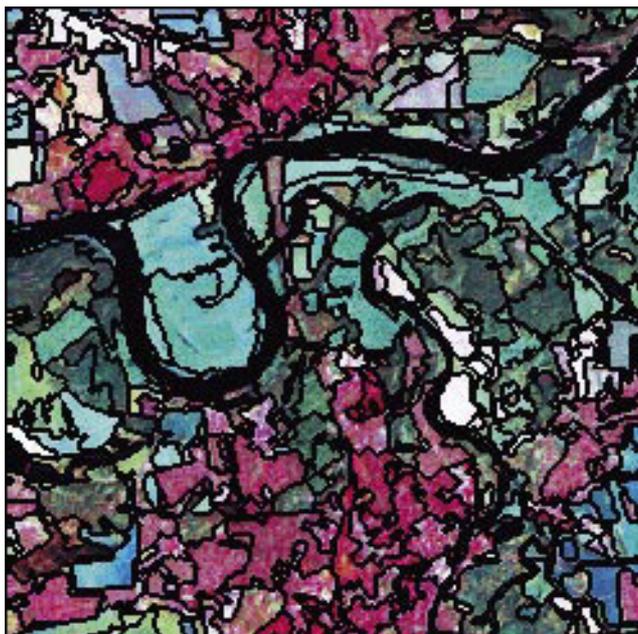


Fig. 3.7 – Esempio di segmentazione di dati LANDSAT tramite il solo criterio dell’ottimizzazione dell’omogeneità spettrale

Per questa ragione, spesso, è utile utilizzare criteri misti che coinvolgono l’eterogeneità spettrale e l’eterogeneità spaziale, in modo da ridurre i casi di allontanamento da forme di perimetri altamente irregolari (Fig. 3.8). L’eterogeneità, come deviazione da una forma compatta, può essere descritta dal rapporto tra il perimetro dell’oggetto (l) e la radice quadrata del numero di pixel che compongono l’oggetto stesso (n):

$$h_{compact} = \frac{l}{\sqrt{n}}$$

Una modalità per descrivere l’eterogeneità geometrica è rappresentata dal rapporto tra il perimetro dell’oggetto (l) ed il perimetro minore (b) dato calcolato sulla base del rettangolo che racchiude la geometria con lati paralleli a quelli dei pixel dell’immagine raster:

$$h_{smooth} = \frac{l}{b}$$

L’applicazione di questa definizione di eterogeneità ottimizza il livello di smoothness (addolcimento) della forma dell’oggetto risultante. La figura successiva mostra il risultato della segmentazione effettuata nella medesima area, ma tramite combinazione di criteri su base geometrica e radiometrica. La forma risultante è più compatta, con margini più regolari rispetto al caso precedente. Sebbene gli oggetti individuati sembrano non soddisfare pienamente il rispetto dei colori, appare “a prima vista” un risultato migliore.

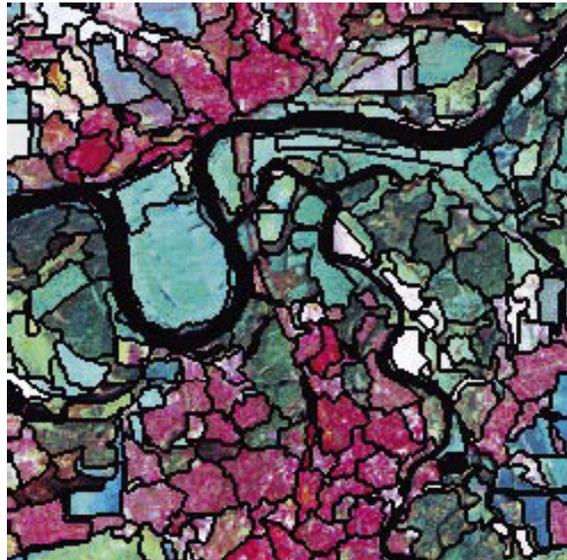


Fig. 3.8 - Esempio di segmentazione di dati LANDSAT tramite i criteri dell'ottimizzazione dell'omogeneità spettrale e spaziale

Questi criteri possono essere applicati in varie forme, combinate anche con criteri aggiuntivi stabiliti sulla base delle geometrie degli oggetti segmentati. In dati fortemente texturizzati, l'aggiunta di un criterio legato alla forma evita il formarsi di segmenti dai perimetri molto irregolari.

Riassumendo, il criterio di eterogeneità consiste di due parti: la prima relativa alla componente radiometrica e la seconda a quella geometrica. Il criterio spettrale si basa sulla variazione dell'eterogeneità che avviene quando due oggetti vengono fusi e descritta dalla variazione della deviazione standard pesata dei valori spettrali. Il criterio geometrico è basato su valori che descrivono il miglioramento della forma con riferimento a due differenti modelli relativi a forme ideali, e consiste in due sub-criteri relativi alle componenti smoothness (addolcimento) e compactness (compattezza). Il calcolo di questi criteri è anch'esso basato sui pesi imposti dall'utente:

$$h_{shape} = w_{compactness} \cdot h_{compact} + (1 - w_{compactness}) \cdot h_{smooth}$$

Anche qui, la variazione dell'eterogeneità geometrica, causata dall'operazione di merging di più oggetti, viene calcolata attraverso la differenza tra i valori precedenti e quelli successivi l'operazione stessa.

Attraverso quindi la procedura di segmentazione, l'intera immagine viene segmentata e diversi oggetti vengono generati sulla base dei criteri impostati di omogeneità ed eterogeneità spettrale e geometrica.

In Figura 3.9 è riportato un esempio di immagine segmentata.



Fig. 3.9 – Risultato di una segmentazione su dati ad alta risoluzione

Le regioni del dato con differenti *textures* vengono separate anche quando queste possiedono simili caratteristiche spettrali (per esempio gli alberi vengono separati dai prati): sono estratte regioni che presentano un contrasto locale. L’algoritmo utilizza una delle descrizioni di eterogeneità pesata in funzione della taglia degli oggetti sull’immagine. Aree di dimensione ridotta, che mostrano una deviazione nei valori di tonalità rispetto ai pixel circostanti, possono essere fuse se vengono rispettati altri criteri imposti dall’utente. Se, invece, queste piccole aree sono di interesse e si vuole conservarle, occorrerà impostare un parametro relativo alla scala relativamente piccolo. Questo consente di estrarre oggetti di dimensioni medie inferiori.

Il risultato di una procedura di segmentazione potrebbe essere il seguente: estrazione di oggetti omogenei di grandi dimensioni (quali ad esempio i prati), estrazione di oggetti piccoli ed eterogenei (quali alberi) ed estrazione di oggetti piccoli ed omogenei caratterizzati da un alto contrasto. Inoltre, la segmentazione ricava oggetti a scala simile, ed è per questa ragione che nell’immagine di figura 3.9 le aree erbose vengono rappresentate da oggetti diversi ed omogenei. Tuttavia, questi oggetti possono, successivamente, essere assegnati al medesimo utilizzo dei suolo con la procedura di classificazione.

Ovviamente, la segmentazione di un’immagine prevede una moltitudine di diverse soluzioni, che dipendono dalle condizioni imposte dall’utente attraverso i parametri di scala ecc. La segmentazione procede in modo autonomo, cercando di minimizzare l’eterogeneità pesata degli

oggetti che vengono identificati. In molti casi, oggetti simili potranno essere assegnati alla stessa classe, soprattutto quando i margini tra aree leggermente diverse non sono così netti, ma gli stessi oggetti potranno essere distinti se la segmentazione procede per una via differente. I risultati, quindi, non possono considerati univoci. Saranno invece più riproducibili gli oggetti che presentano margini molto ben distinti e, se il contrasto non rientra nei parametri imposti dall'utente, questi potranno anche essere estratti scegliendo un parametri di scala inferiore.

In generale, si può dire che per dati omogenei si avrà, nel criterio di omogeneità, un maggiore peso dell'omogeneità spettrale rispetto a quella geometrica, ed un più alto livello di riproducibilità dell'analisi. Va notato che neppure la digitalizzazione manuale fornisce risultati univoci e riproducibili.

3.3.2 Classificazione

Dopo la segmentazione, ovvero l'associazione di elementi dell'immagine sulla base delle loro caratteristiche geometriche e spettrali, sintetizzate dai parametri appena discussi, occorre associare ai poligoni le classi relative agli oggetti contenuti, che in questo caso sono riferite all'uso del suolo.

E' stata effettuata una classificazione supervisionata mediante algoritmo Nearest Neighbor, attraverso la definizione di training sites (sample) sull'immagine segmentata.

Contrariamente a quanto avviene nelle applicazioni più classiche del telerilevamento qui la somiglianza tra training sites ed aree segmentate viene valutata sia in base a criteri geometrici sia radiometrici. Infatti, in riferimento a un determinato livello di segmentazione, le informazioni che possono essere associate ai singoli poligoni sono di tipo:

- ✓ spettrale: dati statistici calcolati per ogni poligono a partire dai valori di *Digital Number* dei pixel in essi inclusi, per ciascuna delle bande spettrali disponibili (media, varianza, ecc.);
- ✓ geometrico: derivate dalla segmentazione, le caratteristiche geometriche più comuni sono l'area, il perimetro, la dimensione frattale (rapporto area/perimetro) o altri coefficienti di forma;
- ✓ tessiturale: sono informazioni che derivano dall'analisi della struttura spaziale interna ai poligoni derivanti dalla segmentazione;
- ✓ gerarchico: Sono statistiche legate alle relazioni esistenti tra oggetti nello stesso livello e fra livelli a differenti valori di scala;
- ✓ tematico: sono informazioni derivanti da strati informativi accessori (es. DEM), ancillari a quelle ottenibili dall'immagine telerilevata.

3.3.2.1 Logica generale

Operativamente la procedura consiste nelle seguenti fasi:

1. definizione del sistema di nomenclatura da applicare alla classificazione;
2. definizione del livello della segmentazione da classificare;
3. selezione delle informazioni da associare ai poligoni da classificare;
4. acquisizione manuale di un certo numero di poligoni di *training*;
5. applicazione di un algoritmo di classificazione.

Tutto il processo può essere reiterato, se l'accuratezza della classificazione non soddisfa, acquisendo nuovi poligoni di *training* o cambiando la base informativa (o il peso relativo delle diverse informazioni). Il criterio di classificazione utilizzato è quello del *Nearest Neighbor* (NN).

NEAREST NEIGHBOR

Il criterio NN quindi è utilizzato per valutare la somiglianza tra diversi oggetti sulla base di una serie di *features* associate agli oggetti stessi. Le Figure 3.10 e 3.11 consentono di illustrare la procedura di classificazione NN in modo schematico.

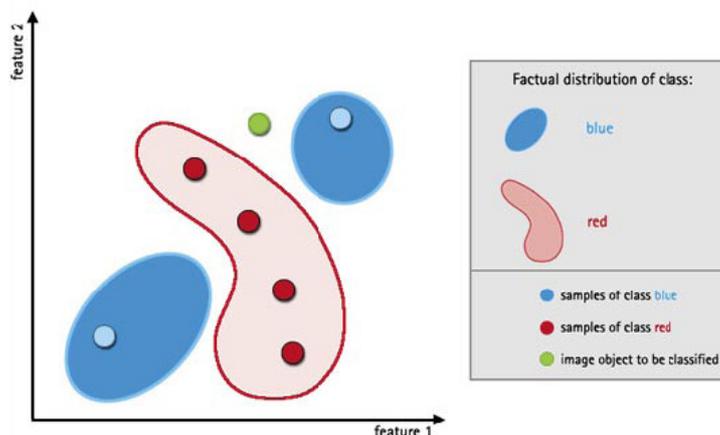


Fig. 3.10 – Fase iniziale della classificazione NN. L'oggetto identificato dal cerchio verde deve essere associato ad una delle classi blu o rossa in base alle sue caratteristiche

Limitando a due sole *features* i parametri utilizzati per la classificazione, i loro valori permettono di individuare uno spazio cartesiano, ove ogni punto del piano corrisponde ad una combinazione X, Y; ogni poligono di training (rosso e blu) è caratterizzato da un determinato valore X (1) e da un determinato valore Y (2) e si identifica come punto (tutti i poligoni di una stessa classe permettono

di individuare un intervallo di valori x_1 e x_2 tipico della classe): l'algoritmo di classificazione associa, quindi, ogni poligono da classificare alla classe più vicina nello spazio individuato dai parametri.

In un primo passo vengono dichiarati alcuni oggetti campioni (*samples*) per ogni classe (Fig.3.10). In seguito, l'algoritmo cerca, come detto, il campione più vicino ad un generico oggetto effettuando la ricerca nel dominio delle features. Se un oggetto non classificato è molto prossimo ad un campione, identificato dall'utente nella fase di addestramento, questo viene assegnato proprio a quello stesso campione. Gli oggetti che non sono stati assegnati a nessuna classe, e quelli che sono stati erroneamente classificati, possono essere scelti come campioni aggiuntivi per migliorare il risultato della classificazione. Anche gli oggetti che si trovano al limite di una classe, o nel prolungamento di un'altra, possono essere integrati nella classificazione.

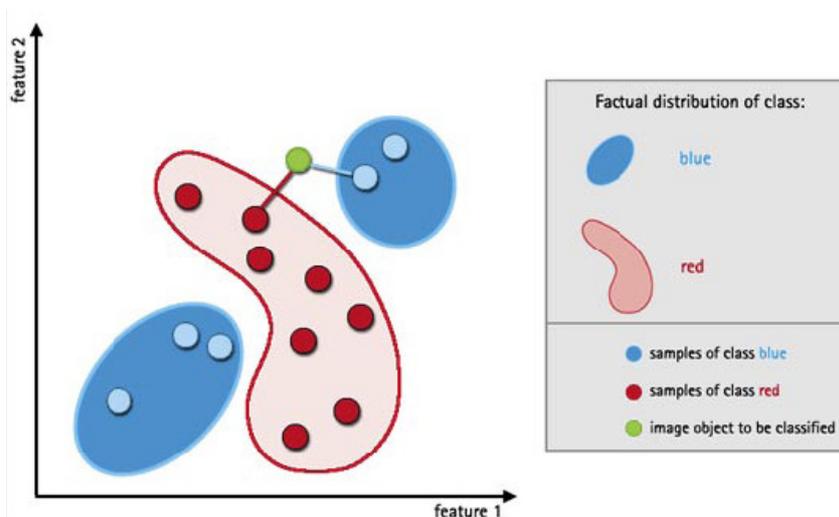


Fig. 3.11 – Procedura di affinazione di una classificazione tramite NN. L'oggetto verde può essere ri-assegnato ad una delle classi.

PARTE II

Introduzione alla parte II

La seconda parte della Tesi, compresa in un unico grande capitolo, è dedicata alla descrizione di tutte le informazioni necessarie per concretizzare le fasi che precedono la classificazione vera e propria. L'obiettivo è quello di evidenziare la complessità che comporta la raccolta e la gestione di tutti i dati, la difficile reperibilità dei quali a volte può compromettere i risultati finali delle elaborazioni. Infatti, per quanto pratico e poco teorico, questo stadio dell'intera procedura risulta essere quello più delicato perché richiede un notevole impegno, sia in termini di tempo che di risorse.

Dopo una breve illustrazione delle caratteristiche principali dell'area monitorata, ovvero la fascia costiera emiliano-romagnola estesa dal Delta del Po alla provincia di Ravenna, vengono introdotte le diverse fonti, di origine sia cartacea che digitale, dalle quali è stato possibile estrarre le informazioni fondamentali per ricostruire l'uso del suolo, in particolare quello agricolo.

Particolare attenzione è stata poi dedicata ad alcune problematiche relative alla disponibilità di dati a terra affidabili ed aggiornati, alla loro raccolta ed alla necessità di un'integrazione finalizzata alle fasi di istruzione degli algoritmi e di verifica dei risultati, con l'obiettivo di mettere a punto metodologie valide, esportabili ed il più possibile automatizzate per il monitoraggio del territorio. Le principali condizioni da soddisfare riguardano l'omogeneizzazione e la georeferenziazione dei dati in un unico sistema di riferimento, la contemporaneità delle informazioni al suolo con il dato registrato dal sensore, e la preventiva analisi del fenomeno reale oggetto dello studio, in funzione della quale deve essere scelto il momento ottimale di acquisizione dell'immagine.

CAPITOLO 4: VALUTAZIONE DELLE INFORMAZIONI NECESSARIE

4.1 Area di studio

Estesa dal Delta del Po al confine meridionale della provincia di Ravenna, l'area costiera osservata (Fig. 4.1) propone una notevole complessità eco-sistemica per la presenza di elementi naturali altamente sensibili, inseriti in un contesto antropico con il quale interagiscono continuamente.

Attualmente il quadro che si presenta è quello di un territorio molto frammentato ed eterogeneo, che ha perso gradualmente i suoi caratteri di naturalità, a causa di un processo di trasformazione molto intenso, operato dall'uomo, che ha contribuito, anche per effetto di innovazioni nelle tecniche di produzione agricola, a cancellare o a ridurre drasticamente l'identità di diversi luoghi e paesaggi naturali ben distinguibili nelle carte di fine ottocento - inizi novecento. Le principali fasi evolutive del Paesaggio, come per il resto della Romagna, hanno visto la realizzazione di importanti opere di bonifica, a partire dal periodo romano, per proseguire poi nel medioevo, con interventi costanti, e spesso contrastanti, volti alla costruzione di canali di scolo artificiali, alla diversione dei corsi d'acqua principali, alla realizzazione di casse di colmata, fino all'epoca moderna (inizio novecento). Un radicale cambiamento è stato segnato, in quel momento, dai nuovi criteri di bonifica che hanno introdotto lo scolo meccanico (canalizzazioni ed impianti idrovori), offrendo immediate possibilità di redimere zone malsane con l'acquisizione di nuove terre per l'agricoltura e per l'insediamento. Tale metamorfosi ha subito una notevole accelerazione nel periodo successivo alla seconda guerra mondiale, quando il territorio viene pervaso da una rapida trasformazione fondiaria, con la meccanizzazione dei mezzi di lavoro, con le nuove concimazioni e con la valorizzazione e il più razionale sfruttamento delle favorevoli condizioni agronomiche del terreno, con impianti estesi di frutteti e vigneti. Le carte dell'IGM dei primi anni cinquanta rivelano già un'area nella quale le opere di bonifica sono ormai completate con ampi territori coltivati a seminativo e in cui centri urbani principali assumono dimensioni consistenti. I fenomeni più evidenti del profondo cambiamento, avvenuto contemporaneamente ad un progressivo impoverimento dei livelli di naturalità, sono:

- ☞ l'accorpamento dei poderi con la trasformazione progressiva dei seminativi arborati (la piantata padana) in frutteti e vigneti o in seminativi semplici;
- ☞ la creazione di vaste aree di frange urbane in cui la transizione tra città e campagna assume diversi caratteri, difficilmente risolvibili e in contrasto tra loro sia per quanto riguarda il paesaggio che per gli equilibri ecologici.

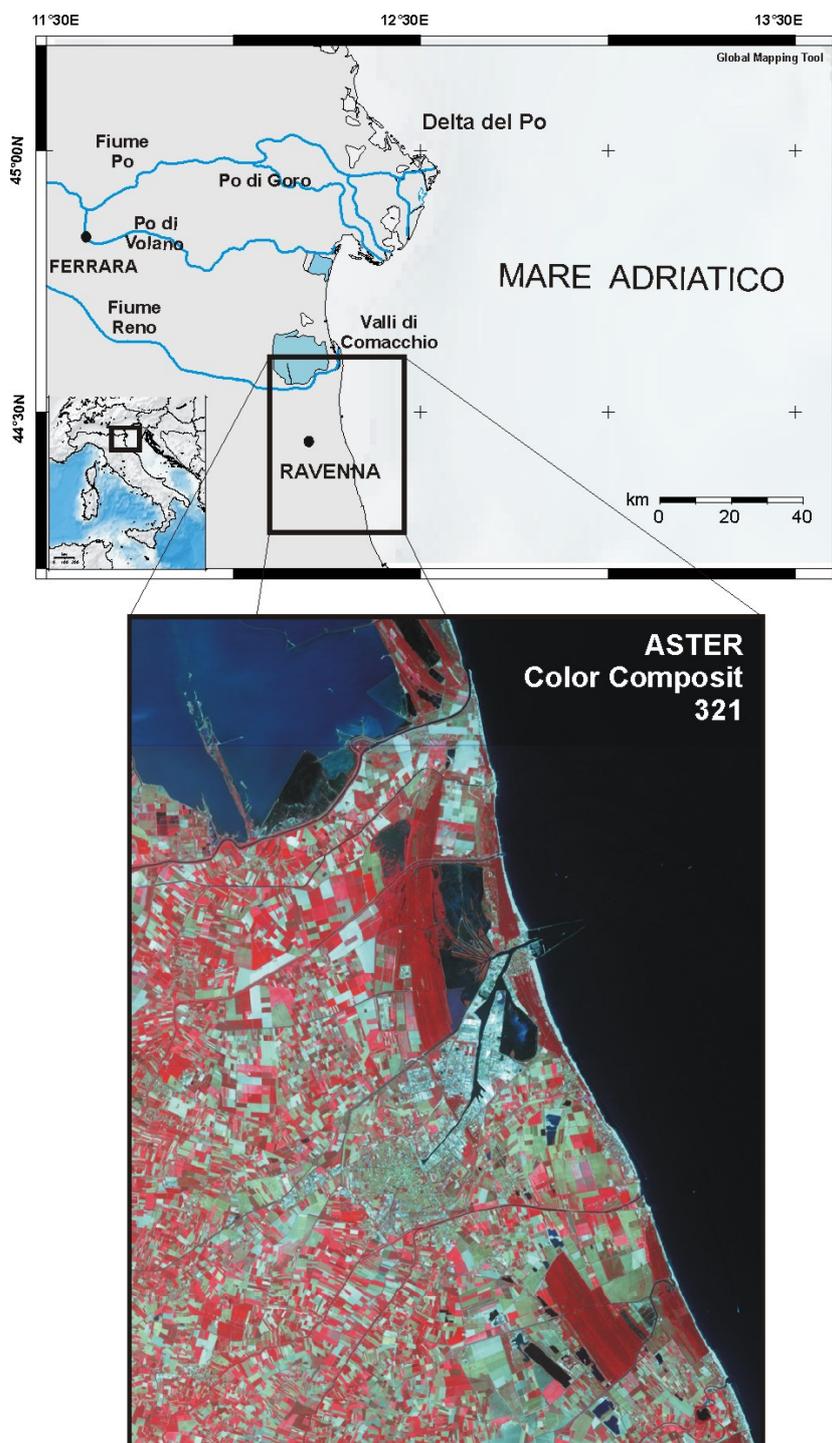


Fig. 4.1 – Inquadramento dell'area di studio.

Per quanto fortemente antropizzata e modificata dall'uomo nel corso dei secoli, l'area fortunatamente conserva ancora preziosi ambienti vallivi, pinete, boschi planiziali e piccoli tratti di litorale intatto. La presenza di molte aree umide rende inoltre diversi luoghi di notevole attrazione per l'avifauna locale: sono presenti infatti 146 specie di uccelli nidificanti ed altre 151 svernanti, per le quali la zona rappresenta un'ideale sosta durante le migrazioni. Per tali peculiarità, l'intero territorio monitorato è sottoposto ai vincoli del Parco Regionale del Delta del Po (attualmente inserito nella Lista del Patrimonio dell'Umanità dell'Unesco). Quest'ultimo è costituito da corpi idrici costieri (lagune, saline, piallassi, ecc.) in cui si mescolano le acque dolci continentali con quelle salate marine e, in funzione delle maree e delle portate dei fiumi, la salinità dell'acqua assume valori diversi. Tali ambienti consentono la sopravvivenza di sole piante capaci di tollerare alte concentrazioni di cloruro sodico (es. salicornie), mentre i canneti popolano gli specchi d'acqua dolce. Attorno agli stagni, nei quali vegetano le ninfee, si sviluppa un rigoglioso bosco di pioppi, salici, frassini ed altre specie arboree tipiche dei siti umidi riparati. Le pinete, costituite in prevalenza da pino domestico e da pino marittimo, occupano invece buona parte del paesaggio del litorale a ridosso delle dune e delle spiagge, dove le condizioni di vita sono imposte da vento, aridità e salsedine.

Nelle aree a maggiore sfruttamento agricolo infine, il territorio risulta oggi suddiviso in un numero elevato di terreni, di dimensioni variabili e con colture differenti, sia arboree che erbacee, caratterizzati da un'alta variabilità stagionale. In particolare, cereali (come mais, grano e simili), legumi e ortaggi hanno brevi cicli di maturazione e sono sottoposti a rotazione durante l'anno. Grandi appezzamenti sono invece occupati da alberi da frutto con cicli produttivi annuali.

All'interno di questo delicato equilibrio si inserisce l'industria, prevalentemente petrolchimica, concentrata a nord-est di Ravenna.

4.2 Tipi di dati utilizzati

Relativamente all'area oggetto dello studio sono disponibili diversi documenti, dai quali è stato possibile estrarre le informazioni necessarie alla realizzazione dei training site e alla fase di verifica dei risultati:

- ☞ la Carta dell'Uso del Suolo
- ☞ la Carta della Vegetazione
- ☞ le dichiarazioni PAC (Politica Agricola Comune)
- ☞ le mappe catastali

Gli algoritmi di classificazione sono invece stati applicati alle immagini multispettrali a media risoluzione ASTER acquisite nel periodo compreso tra il 2000-2006. Nei paragrafi successivi vengono descritti gli aspetti principali dei dati sopraelencati. Si ricorda inoltre il contributo ottenuto, data la maggiore risoluzione, dalla consultazione dell'archivio QuickBird disponibili su Google Earth.

4.2.1 Dati ASTER

L'archivio disponibile dei dati ASTER comprende le 10 immagini riportate in tabella 4.1.

	CODICE ID IMMAGINE	DATA DI ACQUISIZIONE
1)	AST_L1A.003:2010671217	17/09/2000
2)	AST_L1A.003:2006739690	23/04/2002
3)	AST_L1A.003:2007679574	19/06/2002
4)	AST_L1A.003:2024902175	01/07/2004
5)	AST_L1A.003:2029583275	11/06/2005
6)	AST_L1A.003:2029290435	26/05/2005
7)	AST_L1A.003:2030270406	29/07/2005
8)	AST_L1A.003:2030727388	30/08/2005
9)	AST_L1A.003:2034225193	13/05/2006
10)	AST_L1A.003:2034625019	14/06/2006

Tab. 4.1 – Archivio dei dati Aster

Un esempio di immagine Aster è visibile in figura 4.2 nella quale è riportata la combinazione delle tre bande del sottosistema VNIR (color composite: verde, rosso, infrarosso vicino).



Fig. 4.2 – Esempio di dato Aster: color composite 3,2,1 (acquisizione del 14/06/06)

Dataset Attribute	Attribute Value
Level 1A Scene ID	SC:AST_L1A.003:2034625019
Acquisition Date	2006/06/14
WRS-2 Path	192
WRS-2 Row	029
Upper Left Corner	44°39'29"N, 11°47'48"E
Upper Right Corner	44°33'05"N, 12°34'24"E
Lower Left Corner	44°06'23"N, 11°36'53"E
Lower Right Corner	44°00'02"N, 12°23'04"E
Scene Center	44°19'45"N, 12°05'32"E
Scene Cloud Cover	0%
SWIR Mode	ON
TIR Mode	ON
VNIR1 Mode	ON
VNIR2 Mode	ON
Day or Night	Day
Orbital Direction	Descending
Sun Elevation	65.276904
Sun Azimuth	143.589000
Acquisition Time	10:09:40.6980000
VNIR Pointing Angle	-5.7270000
TIR Pointing Angle	-5.7000000
SWIR Pointing Angle	-5.6689999

Tab. 4.2 – Esempio di Metadata del dato del 14/06/06

Il prodotto rilasciato dalla NASA, come singola immagine, è costituito dal dato ASTER in formato HDF-EOS con le bande scritte in band sequential (BSQ) format: viene cioè creato un file per ogni banda con aggiunto un file ausiliario contenente HDF e ECS metadata (Tab. 4.2). Purtroppo non è stato possibile applicare l'intera procedura sperimentata a tutte le immagini dell'archivio: infatti per il dato del 17/09/00 non sono risultate reperibili le informazioni relative all'uso agricolo dal suolo dell'anno corrispondente. Invece per le immagini del 23/04/02 e del 19/06/02, il numero esiguo di campioni estraibili dalle varie fonti non ha consentito di effettuare la delicata fase della verifica delle classificazioni ottenute. In particolare, nella prima delle due immagini, la presenza di una certa copertura nuvolosa in prossimità delle aree interessate, ha accentuato le difficoltà.

4.2.2 Carta dell'uso del suolo

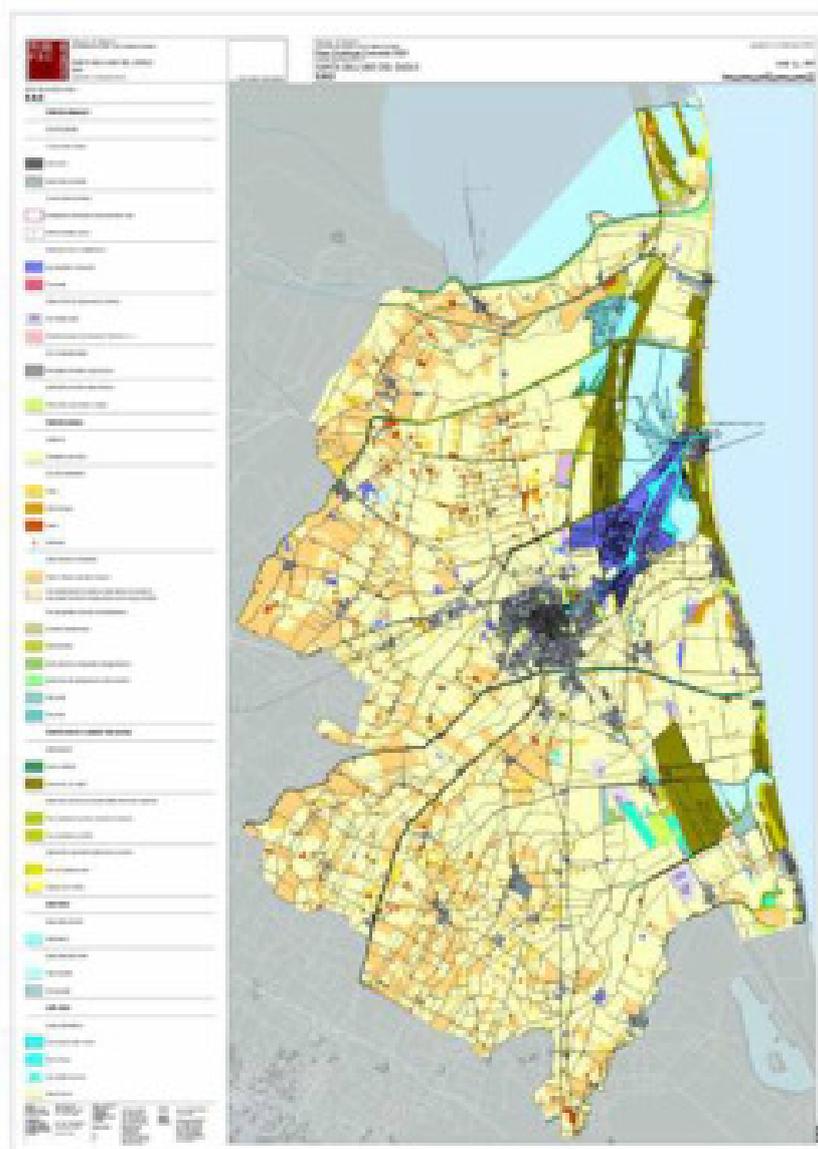
La “Carta dell’Uso del Suolo” del Comune di Ravenna (mostrata nella figura 4.3) è stata realizzata dall’Ufficio SIT (Sistema Informativo Territoriale) del comune stesso, per adempiere a quanto previsto dalla legge regionale del 4 Marzo 2000, in materia di aggiornamenti degli strumenti urbanistici. L’elaborato è il frutto di una fotointerpretazione accurata di ortofoto, integrata con le conoscenze territoriali e le dotazioni di vario genere in possesso dell’Ufficio: va ricordato che questo comune è uno dei più vasti in Italia (circa 66000 ha) ed ha una struttura insediativa-territoriale-ambientale complessa, comprendendo zone molto diverse fra loro (il litorale con 35 Km di costa, le zone ambientali di valore naturalistico, le Pinete e le Valli, il sistema delle zone agricole per circa 50000 ha, il comparto del Porto e le zone urbanizzate, costituite da oltre 60 centri, fra i quali il capoluogo).

Per la redazione della consistente legenda, è stato preso come riferimento il *Programma Corine* (COoRdination de l' INformation sur l' Environnement): varato dal Consiglio delle Comunità Europee nel 1985, che ha lo scopo primario di verificare dinamicamente lo stato dell'ambiente nell'area comunitaria, al fine di orientare le politiche comuni, controllarne gli effetti, proporre eventuali correttivi. Obiettivi secondari, ma non per questo meno validi, sono la formazione e la diffusione di standard e metodologie comuni e la promozione di contatti e scambi internazionali, per facilitare la realizzazione di iniziative intercomunitarie: all'interno del programma CORINE, il progetto CORINE-Land Cover è specificamente destinato al rilevamento e al monitoraggio, ad una scala compatibile con le necessità comunitarie, delle caratteristiche del territorio, con particolare attenzione alle esigenze di tutela.

I quarantaquattro tematismi, contenuti nel progetto citato, sono stati applicati ad un’area campione per selezionare quelli utili alla descrizione del territorio comunale: da tre gradi base di dettaglio, con l’introduzione dove necessario di approfondimenti superiori, la legenda si articola su quattro livelli. L’uso del suolo del Comune di Ravenna viene così definito da trentasei tematismi suddivisi in cinque classi dettagliate, ognuna delle quali permette di avere una lettura differenziata a seconda del dato cercato: ad ogni livello corrisponde una visione particolareggiata e una “disgregazione” dei sistemi territoriali, che illustrano via via ciò che è presente sulla zona studiata in maniera puntuale.

Per l’assegnazione dei colori sono state adattate delle scale cromatiche, che meglio facilitano la lettura dei gruppi principali (rosso, blu, marrone, verde), strutturate in palette di colori primari, secondari e complementari. Si legge, quindi, un territorio in cui l’uso agricolo (gamma dei gialli-arancio-marroni) è prevalentemente occupato da colture seminate, comprendente un insieme

particellare complesso in aree storicamente consolidate; i sistemi ambientali (verde-blu) si concentrano fra la linea di costa e il capoluogo con andamento longitudinale e lungo le principali aste idrografiche; i territori urbanizzati (grigio-viola) si insediano prevalentemente nella città di Ravenna e nelle zone industriali-portuali ad essa connesse.



*Fig. 4.3 – Carta dell’uso del suolo di Ravenna,
(<http://www.esriitalia.it/conferenza2003cd/content/documenti/9aprile/malkoswski9>)*

L’importanza di questo elaborato consiste nelle potenzialità che potrà avere in un futuro immediato: attuando una politica di scambio di informazioni, la carta dell’uso del suolo fornirà un aggiornamento, per esempio, sull’utilizzo agricolo, mettendo in evidenza come esso si evolve (per tipo di coltura, per dimensioni degli appezzamenti, ecc.) in funzione dello sviluppo economico, agro-alimentare o industriale delle coltivazioni; allo stesso modo potrà essere un database per il

monitoraggio dell'evoluzione delle biodiversità e dei sistemi ambientali in relazione allo sviluppo urbanistico e alle scelte di pianificazione.

Parallelamente, l'Amministrazione sta procedendo alla digitalizzazione dei dati del Catasto Terreni, i cui lavori sono in corso. La disponibilità di informazioni sull'intero complesso dei fondi agricoli costituirà un notevole potenziamento e sinergismo del sistema, in quanto le variazioni periodiche delle colture potranno essere direttamente aggiornate per via gabellare, tramite l'incrocio dei dati SIT con quelli derivanti dagli organismi che gestiscono i dB alfanumerici delle attività agricole.

4.2.3 Carta della vegetazione

La Carta della Vegetazione del Parco del Delta del Po è stata realizzata nei primi anni '90 in scala 1:25.000, grazie ad una convenzione fra la Regione Emilia-Romagna e l'Università di Ferrara, nell'ambito del più vasto programma di redazione della cartografia della vegetazione dei Parchi Regionali, avviato nel 1988 ed oggi completato. In particolare sono state pubblicate 5 carte della vegetazione che comprendono le 6 stazioni in cui il Parco è diviso: Volano – Mesola - Goro, Campotto di Argenta, Centro storico di Comacchio, Valli di Comacchio, Pineta S.Vitale e Piallasse di Ravenna, Pineta di Classe e Saline di Cervia. Gli autori (Piccoli et al., 1996) hanno assegnato un indicatore di pregio naturalistico ad ogni tipologia della vegetazione.

Per la compilazione della carta sono state usate sia fotografie aeree per le scale di dettaglio, sia immagini da satellite per le carte a piccola scala. Il materiale aerofotogrammetrico è stato utilizzato attingendo a tutta la gamma di strumenti disponibili anziché limitare la scelta ad un solo prodotto: l'uso di immagini effettuate con differenti tecniche e/o in periodi diversi ha permesso spesso di arrivare all'interpretazione delle situazioni più difficili. Nella prima fase del programma sono state perciò impiegate le fotografie aeree tradizionali disponibili in varie riprese: a colori in scala 1:13.000 circa, in bianco/nero in scala 1:30.000 circa e all'infrarosso/falso colore in scala 1:30.000 circa. Successivamente, a partire dal 1995, si sono usate anche le fotografie ad alta definizione, effettuate da alta quota nel 1994, e le immagini da satellite. Per la restituzione cartografica sono state utilizzate le basi topografiche prodotte dalla Regione Emilia-Romagna in scala 1:10.000 e 1:25.000 facilitando anche tutte le operazioni di digitalizzazione e di elaborazione relative alle carte della vegetazione.

L'uso di ingrandimenti fotografici del Volo Italia del 1994, che data l'alta definizione delle immagini ha permesso di ottenere prodotti validi per lo studio vegetazionale sia alla scala 1:25.000 che 1:10.000, ha consentito la messa a punto di un sistema molto preciso di riporto diretto del tematismo, specialmente per i territori non molto acclivi. La sovrapposizione della base topografica

su pellicola trasparente all'ingrandimento fotografico permetteva, specialmente lavorando su tavolo luminoso, il riporto diretto a vista delle principali campiture. Tale riporto è stato fatto a china su una pellicola trasparente punzonata alla base topografica e georeferenziata, grazie all'apposizione di un certo numero di croci di riferimento in punti prestabiliti. Oltre che per la precisione ottenuta, tale procedura presentava il vantaggio di ottenere tutti i limiti tematici su un supporto trasparente indeformabile che è particolarmente adatto per le procedure di digitalizzazione.

Infine, l'interpretazione delle varie tipologie è stata comunque eseguita col metodo tradizionale, sia mediante osservazione comparata dei vari tipi di fotografie con lo stereoscopio, sia mediante controlli di campagna.

4.2.4 Dichiarazioni degli agricoltori

Un ulteriore grado di conoscenza sull'uso del suolo si è raggiunto con la consultazione di alcuni dati, relativi alle dichiarazioni necessarie per ottenere gli aiuti economici previsti dalla PAC (Politica Agricola Comune), forniti dall'Associazione Agricoltori della Provincia di Ravenna (chiamata anche Unione Agricoltori Provinciale Ravenna). Si tratta di un'organizzazione sindacale apartitica, facente capo alla Confagricoltura, che ha come scopo fondamentale la tutela e la rappresentanza dei propri associati: è composta da imprenditori agricoli o persone giuridiche, coltivatori diretti e non, che globalmente rappresentano circa il 43% della superficie provinciale. L'Unione analizza i problemi riguardanti l'agricoltura, sviluppando progetti e proposte con l'intenzione di migliorare costantemente il settore agricolo locale; cura gli interessi degli associati e della categoria agricola nelle relazioni con gli enti pubblici, attraverso l'intervento dei propri rappresentanti all'interno degli Enti, Commissioni e Organismi; promuove ed attua servizi di assistenza economica tecnica, contabile, tributaria, legale, contrattuale e previdenziale. In particolare, segue i diversi soci in tutte le pratiche per accedere ai finanziamenti stabiliti dai regolamenti comunitari, tra i quali quelli della PAC.

Prima politica integrata a livello europeo, la PAC è stata creata progressivamente, a partire dal gennaio 1960, con la creazione delle Organizzazioni Comuni dei Mercati (OCM) e del Fondo Europeo Agricolo di Orientamento e Garanzia (FEAOG). I meccanismi della PAC si fondano essenzialmente su prezzi comuni garantiti e sulle OCM, che differiscono tra loro in funzione delle condizioni di produzione e di vendita dei prodotti. Schematizzando, si possono distinguere quattro tipi di meccanismi:

- ☞ il primo prevede aiuti diretti al produttore, legati a fattori come la terra o il bestiame, calcolati su basi storiche, e all'esistenza di meccanismi per il

controllo della produzione. Riguarda in particolare i cereali, i semi oleosi, le colture proteiche, nonché la carne bovina e ovina;

☞ il secondo tipo, relativo ad aiuti proporzionali alla produzione ottenuta, interessa l'olio di oliva, il tabacco, il cotone, alcuni prodotti ortofrutticoli trasformati, derivati da agrumi, pomodori o prugne e, in una certa misura, il vino da tavola;

☞ il terzo tipo prevede un sostegno principalmente a carico del consumatore ed interessa i prodotti lattiero-caseari e lo zucchero;

☞ il quarto tipo tende essenzialmente ad un adeguamento del mercato in funzione della domanda e dell'offerta, con interventi pubblici estremamente contenuti; coinvolge i prodotti ortofrutticoli, i vini di qualità, la carne suina, il pollame, le uova e il miele.

Nel bilancio comunitario, le spese agricole sono garantite dal FEAOG che, come indicato dal nome, si compone di due sezioni: il settore "Garanzia" finanzia la maggior parte delle uscite, quali aiuti diretti al reddito, sovvenzioni per agevolare lo smercio dei prodotti, regolarizzazione del mercato interno tramite gli acquisti degli organismi di intervento, spese di magazzinaggio, restituzioni all'esportazione e dal 1992, misure a favore dell'imboschimento, aiuti al prepensionamento e misure agroambientali. Il reparto "Orientamento" amministra i fondi comunitari destinati alla politica delle strutture agricole: azioni di ammodernamento, compensazioni per gli svantaggi naturali, aiuto all'insediamento dei giovani agricoltori, sostegno alla trasformazione e alla commercializzazione dei prodotti agricoli, partecipazioni a progetti di interesse collettivo (come aiuto alle regioni in ritardo di sviluppo e alle zone rurali in difficoltà, ecc.). Oltre agli agricoltori, anche consumatori, imprese e società in generale, traggono vantaggio dalla PAC.

I dati analizzati sono riferiti alle tre zone in cui è suddivisa il Comune di Ravenna: S.Alberto (a nord), Ravenna (al centro) e Savio (a sud). Segue l'esempio di uno schema utilizzato dall'Unione Agricoltori per organizzare le dichiarazioni annuali raccolte (Fig. 4.4), proveniente dalla versione cartacea utilizzata fino al 2004.

Valutazione delle informazioni necessarie

fg part	mais	grano t	soia	p.proteico	set_aside	bietola	p.industria	tare	laghi etc	totale	controllo
94 6								1400	2200	3600	
118								100	100	200	
120	20000			12500	18900			1100		52500	
179		75400	159400				61500	10500	34200	341000	
tot fg.	20000	75400	159400	12500	18900	0	61500	13100	36500	397300	397300
96 2								500	1300	1800	
3						141300	77700	5700	9500	234200	
11							31200		500	31700	
12		37900							300	38200	
21	88400	102000				108700		4100	600	303800	
24						1000				1000	
tot fg.	88400	139900	0	0	0	251000	108900	10300	12200	610700	610700
97 1	159900	79900		15700	42300	7800		6900	3200	315700	
2	200			1600	4300			700	1700	8500	
3	73000			60400	53600			1800		188800	
4	6300						14500	2000	24700	47500	
5								200	1000	1200	
6								100	900	1000	
8	2400						5200			7600	
10	93000						84300	2300		179600	
11									600	600	
12								100	600	700	
20	11900								7200	19100	
21								800		800	
tot fg.	346700	79900	0	77700	100200	7800	104000	14900	39900	771100	771100
tot sr	455100	295200	159400	90200	119100	258800	274400	38300	88600	1779100	

Fig. 4.4 – Esempio di dichiarazione PAC cartacea

Nelle prime due colonne sono indicati i numeri del *foglio* e della *particella catastale* in cui sono collocati gli ettari delle colture riportate nelle restanti colonne: sono comprese anche le tare, ovvero le superfici non considerate per la compensazione del reddito, spettante invece alle aree effettivamente investite da una delle coltivazioni previste dal regolamento comunitario.

Dal 2004, invece, le dichiarazioni vengono consegnate all'Unione Agricoltori anche in formato digitale dando così all'associazione l'opportunità di creare un archivio per una migliore gestione dei dati: il conseguente vantaggio ai fini della ricerca è stata una reperibilità delle informazioni più veloce e semplice rispetto al passato (Fig. 4.5).

ANNO DI UTILIZZAZIONE DELLE SUPERFICI AZIENDALI											
COMUNE		DATI CATASTALI				tipo di coltivazione	SUPERFICIE IN CONDUZIONE ettari/are	COLTIVAZIONE		SUPERFICIE UTILIZZATA ettari/are	COLTURA / UTILIZZO
FR	descrizione	sezione	foglio	particella	sub			codici			
								COLTURA	VARIETA' UTILIZZO		
RA	RAVENNA	B	115	4	000	001	0,35	157	000	0,35	USO NON AGRICOLO - FABBRICAT
RA	RAVENNA	B	115	26	000	001	0,57	157	000	0,57	USO NON AGRICOLO - FABBRICAT
RA	RAVENNA	B	115	30	000	001	4	158	000	0,2	USO NON AGRICOLO - TARE ED I
RA	RAVENNA	B	115	30	000	001	4	240	000	1,89	BARBABIETOLA DA ZUCCHERO
RA	RAVENNA	B	115	30	000	001	4	587	000	1,91	GRANO (FRUMENTO) TENERO
RA	RAVENNA	B	115	31	000	001	39,07	110	000	2,31	BOSCO
RA	RAVENNA	B	115	31	000	001	39,07	158	000	1,21	USO NON AGRICOLO - TARE ED I
RA	RAVENNA	B	115	31	000	001	39,07	3	003	2,46	SUPERFICI RITIRATE DALLA PRO INTENZIONE DI SEMINA DOPO IL
RA	RAVENNA	B	115	31	000	001	39,07	4	000	9,25	SOIA - FAVE
RA	RAVENNA	B	115	31	000	001	39,07	20	000	8,22	PISELLI ALLO STATO SECCO
RA	RAVENNA	B	115	31	000	001	39,07	240	000	8,96	BARBABIETOLA DA ZUCCHERO
RA	RAVENNA	B	115	31	000	001	39,07	587	000	6,42	GRANO (FRUMENTO) TENERO
RA	RAVENNA	B	115	31	000	001	39,07	673	000	0,24	SUSINO
RA	RAVENNA	B	115	34	000	001	19,16	110	000	0,44	BOSCO
RA	RAVENNA	B	115	34	000	001	19,16	158	000	1,77	USO NON AGRICOLO - TARE ED I

Fig. 4.5 – Esempio di dichiarazione PAC in formato digitale

Dal 2007 la PAC è stata sostituita dalla Domanda di Premio Unica: si tratta dell'attuale modello dichiarativo che il titolare di un'azienda agricola deve utilizzare per chiedere di essere ammesso ai benefici comunitari previsti dal Regolamento (CE) n. 1782/03 e pertanto la presentazione della stessa è l'unica modalità che consente di ottenere il pagamento degli aiuti di cui si ha diritto. La Domanda Unica favorisce la semplificazione amministrativa e riduce gli adempimenti a carico degli agricoltori poiché permette di richiedere contestualmente più pagamenti attraverso la compilazione e la presentazione, appunto, di un'unica istanza.

L'azienda agricola che intende beneficiare dei sostegni comunitari deve costituire presso il CAA (Centro di Assistenza Agricola, in questo caso l'Unione Agricoltori), a cui ha conferito mandato, il proprio fascicolo aziendale, che è il registro in cui vengono conservate le informazioni e la documentazione relativa all'azienda e alla sua consistenza. Solo dopo aver costituito il proprio fascicolo, l'azienda può presentare la Domanda Unica. Il fascicolo aziendale può essere aggiornato in qualsiasi momento dell'anno rivolgendosi al CAA oppure direttamente ad Agea (Agenzia per l'Erogazione in Agricoltura). Gli agricoltori che hanno conferito apposito mandato devono presentare la Domanda Unica al proprio CAA che provvederà ad inserire l'istanza direttamente nel sistema informatico di Agea.

4.2.5 Mappe catastali

Per rendere i dati provenienti dagli agricoltori utili al proseguimento della tesi, si è rivelato indispensabile il ricorso al patrimonio cartografico catastale (sia informatizzato che cartaceo), per mezzo del quale è stato possibile materializzare la correlazione tra la serie di numeri riportati nelle dichiarazioni PAC e l'immagine di partenza. In tale contesto infatti, un terreno rappresenta un'unità immobiliare, individuabile tramite il comune, il foglio di appartenenza e la particella (Fig. 4.6).

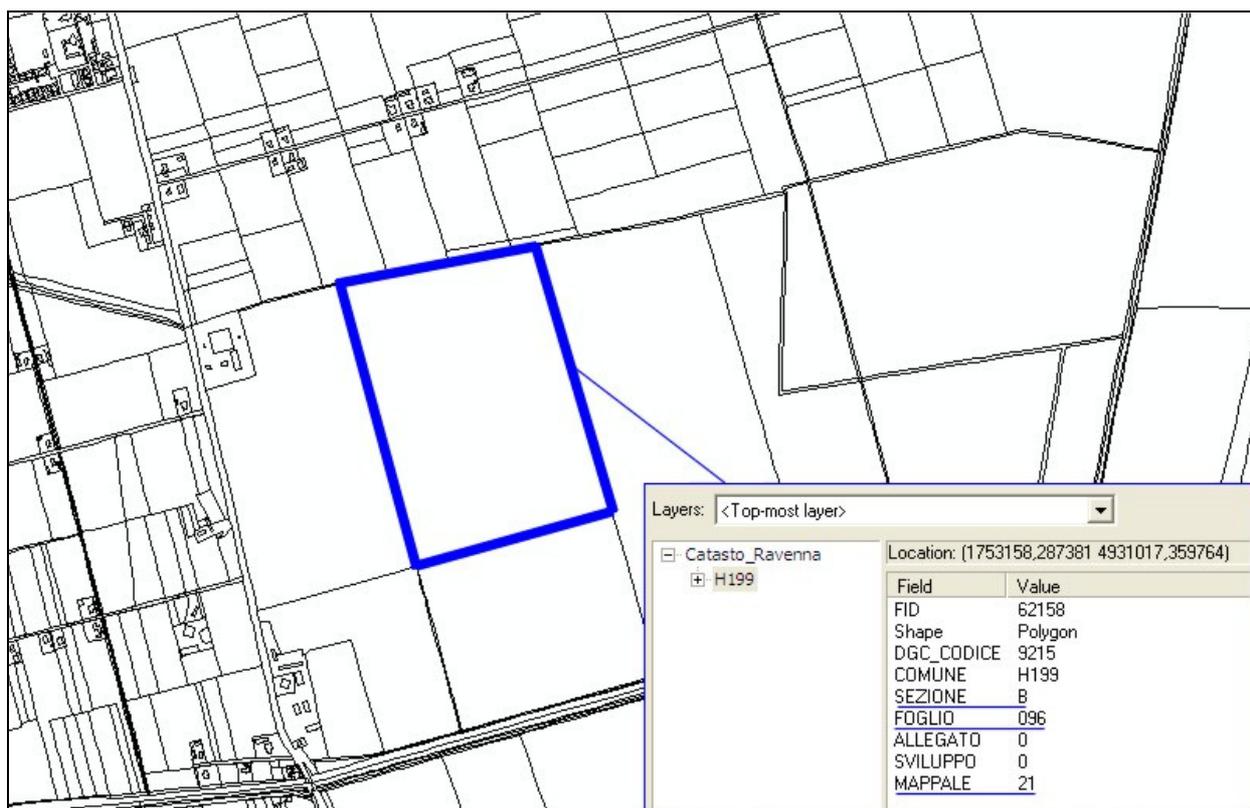


Fig. 4.6 – Estratto di mappa catastale in formato vettoriale con riportati gli attributi relativi al mappale (o particella) evidenziato

Nella versione cartacea, ciascun foglio, suddiviso appunto in particelle (corrispondenti a terreni o immobili), restituisce in scala 1:2000 un'area di circa 2.0x1.4 Km, nel sistema di riferimento Cassini-Solder. Invece, nel formato vettoriale, il sistema di riferimento utilizzato è l'UTM. Entrambe le versioni sono state fornite dall'*Agenzia del Territorio, Ufficio Provinciale di Ravenna, Direzione e Servizi catastali*.

Esiste già comunque un parziale uso sinergico delle dichiarazioni PAC e delle mappe catastali: è il GIS dell'AGREA (Agenzia per le Erogazioni in Agricoltura della Regione Emilia-Romagna) che, solo per le aziende iscritte, consente di visualizzare la sovrapposizione della mappa catastale ai fotogrammi georeferenziati nel sistema cartografico Gauss-Boaga (Fig. 4.7). L'intero progetto

Valutazione delle informazioni necessarie

consiste nella gestione dell'anagrafe dei beni immobili, con relativo inserimento e aggiornamento in mappa (catasto), dei titolari del possesso (conservatoria) e della conduzione (registro contratti), nonché di tutte le funzionalità utili per registrare con semplicità e sicurezza da internet le variazioni catastali, le compravendite e gli affitti, ma anche i dati ricavati da fotointerpretazione o da rilievi sul campo relativi alla consistenza di terreni e fabbricati.



Fig. 4.7 – Viewer del GIS dell'Agrea (<http://gisagrea.regione.emilia-romagna.it/siticatasto/>)

L'accesso a tale servizio però non è libero: è rivolto principalmente alle Amministrazioni Comunali, ai Consorzi di Bonifica, alle Province, alle Regioni, agli uffici tecnici, alle società di servizio, alle associazioni di categoria. L'Agrea comunque, dopo la presa visione del progetto di ricerca in corso, ha fornito una password per entrare nel sistema e visualizzarlo.

4.3 Principali problematiche relative ai dati

4.3.1 Armonizzazione dei dati

I dati fino ad ora descritti provengono da enti differenti, pubblici e privati, che gestiscono le informazioni a loro disposizione in modo diverso, archiviandoli in formati non sempre compatibili tra loro. Inoltre spesso questi documenti non sono di libero accesso al pubblico. Di seguito si riassumono tutti i sistemi di riferimento presenti nelle varie mappe:

TIPO DI DATO	DATO	SISTEMA DI RIFERIMENTO
RASTER	IMMAGINI ASTER	UTM - WGS84
	CARTA DELL'USO DEL SUOLO	UTM*
VETTORIALE	CARTA DELLA VEGETAZIONE	UTM*
	MAPPE CATASTALI	UTM e CASSINI - SOLDNER
ALTRO	DICHIARAZIONI PAC	RIFERIMENTO CATASTALE

* UTM*: è una variazione del sistema UTM, realizzata dalla Regione Emilia-Romagna, che consiste nell'omissione della prima cifra, in questo caso il 4, dalla coordinata di latitudine.

L'utilizzo simultaneo di "verità a terra" e di dati satellitari, in questo caso immagini Aster proiettate nel sistema UTM-WGS84, deve necessariamente prevedere una fase di omogeneizzazione delle informazioni che, per la loro stessa natura, sono caratterizzate da risoluzioni spaziali, accuratezza ed affidabilità differenti. I prodotti raster e vettoriali coinvolti nelle varie fasi dell'elaborazione dei dati devono essere georeferenziati in un comune sistema cartografico (Fig. 4.8). Il materiale di origine cartografica in questa fase non presenta particolari difficoltà in quanto si può georeferenziare con un'accuratezza finale dipendente dalla scala di rappresentazione del materiale originale oppure dall'accuratezza dei rilievi effettuati. Due problemi, invece, nascono con l'uso di dati satellitari a media risoluzione: il primo è relativo all'accuratezza generale della georeferenziazione delle immagini satellitari e, quindi, alla possibilità di un loro utilizzo congiunto a quella delle informazioni di diversa natura. Ad esempio, nella selezione dei training site dovrà esserci corrispondenza geografica tra l'area di addestramento selezionata sull'immagine satellitare e l'uso del suolo dedotto derivato da mappe di altra origine.

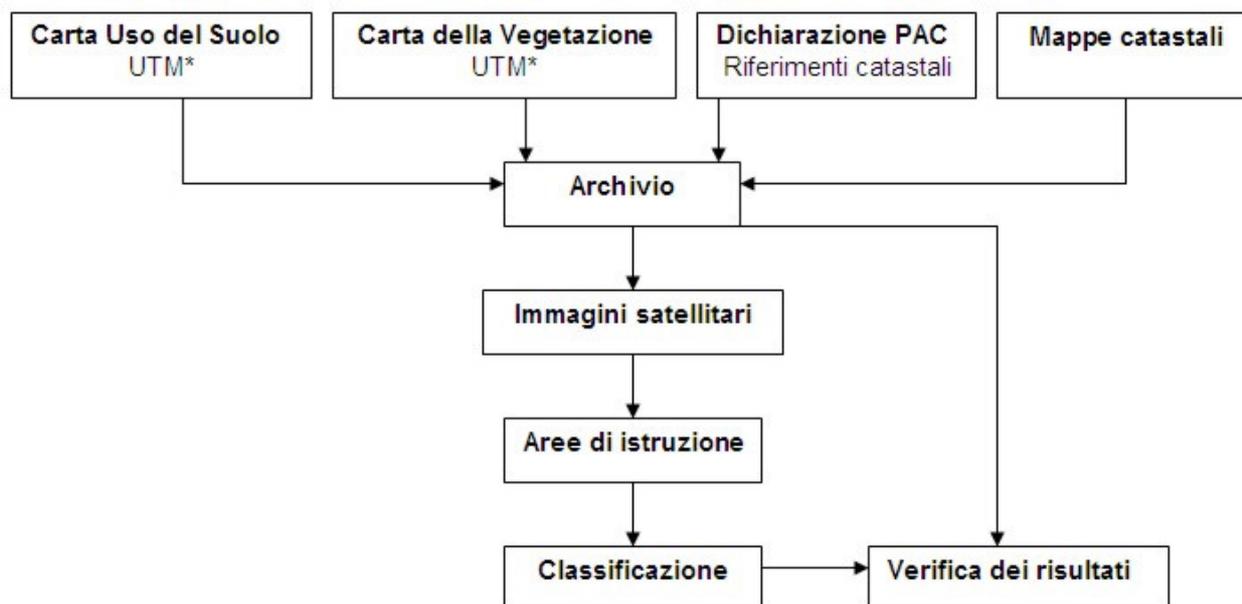


Fig. 4.8 – Schema relativo all'integrazione dei dati finalizzata alla classificazione

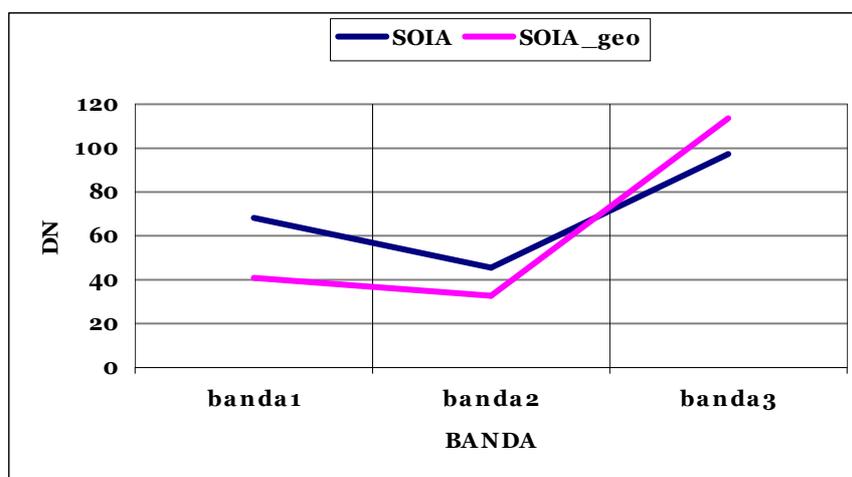
Il secondo problema nasce dalla considerazione che, in questo lavoro, i dati satellitari vengono utilizzati per la classificazione. Per questo motivo le immagini non vengono preventivamente georeferenziate, operazione che richiederebbe necessariamente una fase di ricampionamento geometrico e radiometrico (anche se conservativi) che potrebbe compromettere l'efficacia degli algoritmi di classificazione. Infatti, ogni volta che su un'immagine è stata applicata una trasformazione geometrica, sia per effetto di correzioni sistematiche, sia per cambiamenti di scala, sia per raddrizzamento o registrazione, il risultato è che la griglia dei centri dei pixel si trova in una nuova posizione, più accurata dal punto di vista geometrico, ma il cui valore radiometrico non rappresenta più i valori originali misurati sulla scena (Brivio et al.,2006).

A testimonianza di quanto sopra riportato si veda la Fig. 4.9, nella quale vengono confrontate le firme relative alla stessa coltura (soia), estratte dall'immagine Aster non georeferenzata e dalla corrispondente immagine georeferita nel sistema UTM*.

Considerando l'oggetto dell'indagine, ovvero le coltivazioni, queste variazioni potrebbero essere una fonte di errore, insieme ad altri fattori che verranno successivamente descritte, nella discriminazione delle colture: in generale infatti le firme spettrali sono molto simili, specialmente per quelle che presentano il ciclo vegetativo contemporaneo (Fig. 4.10).

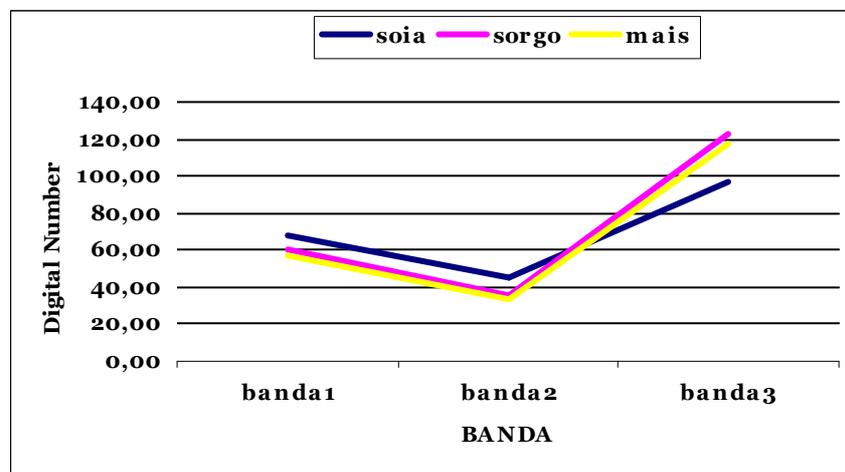
Il confronto è stato eseguito solo sulle prime tre bande del sistema VNIR (verde, rosso e infrarosso) perché, oltre ad essere quelle maggiormente rappresentative per la vegetazione, sono le uniche sulle quali non viene effettuato nessun ricampionamento prima della classificazione: si ricorda, infatti,

che le bande SWIR hanno un a risoluzione doppia rispetto alle precedenti, e di conseguenza le dimensioni del pixel devono essere dimezzate.



Soia	Mean	Stdev	Soia	Mean	Stdev
Band 1	<u>67.8</u>	1.9	Band 1	<u>41.2</u>	1.2
Band 2	<u>45.2</u>	2.4	Band 2	<u>33.0</u>	1.7
Band 3	<u>97.1</u>	3.4	Band 3	<u>113.5</u>	4.0

Fig. 4.9 – Confronto tra la firma della soia estratta dal dato Aster non georeferenziato e quella ottenuta dalla corrispondente immagine georeferita nel sistema UTM*.



Band	Coltura	Mean	Stdev	Coltura	Mean	Stdev	Coltura	Mean	Stdev
Band 1	SOIA	67.8	1.9	SORGO	60.0	1.1	MAIS	56.7	1.3
Band 2	SOIA	45.2	2.4	SORGO	35.3	1.1	MAIS	33.1	1.1
Band 3	SOIA	97.1	3.4	SORGO	122.9	5.7	MAIS	117.6	2.1

Fig. 4.10- Confronto tra le firme della soia, del sorgo e del mais estratte dal dato Aster non georeferenziato

4.3.2 Contemporaneità dei dati

Il miglior approccio ad un'indagine finalizzata al riconoscimento dell'uso del suolo, prevede un'attenta analisi dei dati relativi alla realtà osservata a terra, in seguito alla quale deve essere scelto o programmato il momento di acquisizione dell'immagine satellitare. Tale metodologia si scontra generalmente con una serie di problemi pratici legati al reperimento dei dati telerilevati presso gli enti spaziali che li producono: la scelta a priori dell'istante di osservazione, inteso come giorno, mese, anno e ora, comporta infatti un ordine preventivo con il quale si chiede all'Agenzia coinvolta di pianificare, per quella data, il passaggio del satellite in una determinata area. E' necessario pertanto prendere in considerazione due importanti aspetti. Il primo riguarda la priorità che viene assegnata all'ordine all'interno di un certo numero di richieste provenienti da tutto il mondo, con il conseguente rischio di una lunga attesa, anche di mesi, prima della sua esecuzione. Si rivela utile, quindi, progettare l'acquisto con largo anticipo rispetto al periodo di interesse, compatibilmente con la dinamicità del fenomeno oggetto dello studio. Una seconda valutazione deve essere effettuata tenendo conto degli imprevisti di diversa natura che si possono verificare, primo fra tutti le condizioni atmosferiche non favorevoli nella giornata per la quale è stata pianificata l'acquisizione.

In funzione della collaborazione con l'azienda agricola Ca'Bosco di Ravenna che, con la mediazione dell'Unione Agricoltori di Ravenna, ha messo a disposizione sia personale tecnico e che documentazione circa la gestione dei propri terreni, oltre a concedere la possibilità di effettuare sopralluoghi, è stato ideato un progetto per un'acquisizione programmata, accettata dall'Ente spaziale contattato, ma purtroppo non perfettamente eseguita. L'obiettivo era quello di studiare la risposta spettrale di alcune delle colture presenti sul terreno, nel periodo suggerito dall'azienda in base allo stadio avanzato del ciclo fenologico. La disponibilità di informazioni simultanee all'acquisizione sul fenomeno in corso può, infatti, consentire delle concrete valutazioni sull'esito finale della classificazione.

4.3.2.1 Programmazione dell'acquisizione di un dato Aster

Nel caso in cui i dati Aster disponibili, consultabili negli archivi EDG (Earth Observing System Data Gateway - USA) e GDS (Ground Data System - Japan), non rispondano ai requisiti geografici o temporali o, in generale, di acquisizione voluti dall'utente, è possibile programmare una specifica ripresa seguendo la procedura prevista dal sito JPL (<http://asterweb.jpl.nasa.gov/NewReq.asp>). La richiesta prevede essenzialmente due passaggi: il primo consiste nella sottomissione di una breve proposta on-line, che, se accettata dopo una revisione, permette di ottenere un account di accesso al servizio predisposto *Data Acquisition Request (DAR)*. Il passo successivo è l'utilizzo del *Data*

Acquisition Request Tool per confermare e dettagliare la domanda iniziale. DAR Tool è il riferimento per l'acquisizione di un numero di scene inferiore a 10, altrimenti per numeri superiori è necessario procedere con una *STAR (Science Team Acquisition Request)* entrando in contatto direttamente con lo Science Team responsabile. L'agenzia si raccomanda di verificare con accuratezza che l'area d'interesse non sia già stata ripresa in precedenza, dato che tutte le acquisizioni, comprese quelle pianificate dagli utenti, sono rese disponibili entro pochissimi giorni dalla loro generazione. Si riportano di seguito tutte le fasi previste:

☞ **Compilazione Short Proposal:** in tale proposta deve essere indicato il tipo di utilizzo previsto per l'immagine richiesta. In particolare, se gli scopi non sono scientifici o didattici, ma solo commerciali, la domanda deve essere inoltrata direttamente al Japan GDS, ovvero al partner orientale dell'intero progetto, andando a costituire la *General Users Observation Request (GPR)*, le cui modalità sono riportate nel corrispondente sito giapponese (<https://imsweb.aster.ersdac.orc.or.jp>). Negli altri casi, invece, deve essere descritta brevemente la ricerca o l'ambito scientifico all'interno del quale si colloca l'uso dell'immagine Aster, precisando a quali elaborazioni si intende sottoporla e indicando in modo approssimativo l'area geografica coinvolta, oltre al periodo di ripresa desiderato (Fig. 11)

☞ **Valutazione Short Proposal:** la richiesta viene valutata da *NASA's ASTER User Authorization Committee* che, in caso di esito positivo, lo comunica all'utente mediante mail elettronica, nella quale preannuncia l'invio entro breve tempo delle istruzioni necessarie per effettuare la registrazione e per inoltrare la richiesta definitiva. Riporta, inoltre, l'estensione approssimativa della zona che dovrà essere ripresa, precisando che potrebbe essere contenuta in più immagini separate. Nel caso di rifiuto, all'utente vengono comunque trasmesse le motivazioni alla base della decisione. Viene sottolineato che, per varie ragioni, l'accettazione della proposta non è una garanzia della effettiva acquisizione del dato e della sua qualità. Le due cause principali della mancata ripresa del dato richiesto o della non corrispondenza con i requisiti sottomessi sono:

1. le risorse del satellite destinate all'osservazione sono limitate e le richieste degli utenti possono anche essere in contrasto tra loro in termini di utilizzo dello stesso sensore allo stesso tempo;

2. l'effetto di disturbo legati agli agenti atmosferici (ad esempio una densa copertura nuvolosa) o a problemi tecnici non prevedibili

Jet Propulsion Laboratory
California Institute of Technology

[+ View the NASA Portal](#)

JPL HOME EARTH SOLAR SYSTEM STARS & GALAXIES TECHNOLOGY

ASTER

Advanced Spaceborne Thermal Emission and Reflection Radiometer

Proposal to NASA to be authorized to request ASTER data acquisitions

Please completely fill out the form below.

First Name: Last Name:

Institution:

Address:

City: State (US Only):

Zip Code: Country:

Email address:

Telephone: Fax:

Type of Proposal: Research
 Educational
 Operational/Commercial Demonstration

Primary Area of Study:

Affiliation:

Text of proposal, including:

- proposed research (or other use).
- possible geographic areas and times of year for ASTER observations, and
- reasons why data from ASTER (rather than some other instrument) are required.

[Maximum of 2 pages (approximately 4000 ASCII characters.)]

My research deals with the use of ASTER satellite image for the monitoring of several environmental aspects, with particular focus on the land use and agriculture. Supervised and unsupervised classification produces are being used and already tested on ASTER data over an area where ground truths were available. ASTER bands can be used to study the way several vegetal components can affect the spectral signatures of vegetation. The main purpose of the sperimentation is the mapping of agricultural productivity and vegetation health investigation sported by the analysis of some ground properties.

The area of interest includes the coastal zone of Emilia-Romagna (Italy), its inland and the easternmost part of the Po Valley. The period for the daytime ASTER

Amount of ASTER data required :

Total number of scenes being requested is composed of multitudes of 60km by 60km area. Area less then 3600 sq-km is still considered to be one full scene. Example. a box area of 7000 sq-km would be considered 2 scenes.

Total Number of Scenes (60km x 60km)

Fig. 11 – Modulo on-line per compilare la Short Proposal (<http://asterweb.jpl.nasa.gov/proposal.asp>).

☞ **Registrazione DAR:** l'autorizzazione concessa consente la creazione di un LP DAAC User Account per accedere al DAR Tool, introducendo le informazioni richieste nell'apposita pagina web dell'*EOS Data Gateway* (EDG). Per attivare il nuovo account, è necessario contattare la segreteria automatica predisposta, via telefono o via mail. Dopodiché tutti i dati relativi al nuovo utente vengono trasferite al GDS in Giappone, il quale provvederà a comunicare l'avvenuta registrazione: ogni eventuale modifica o aggiornamento apportato successivamente dall'utente dovrà seguire lo stesso percorso. L'attivazione avviene in genere entro tre giorni dall'invio dei dati.

☞ **Compilazione DAR:** le principali indicazioni che devono essere riportate, e in base alle quali vengono organizzate le richieste, sono relative all'area di interesse (AOI, *Area of Interest*) e alla finestra temporale utile all'acquisizione (*Acquisition Window*). Poi, una varietà di altri parametri, come ad esempio la percentuale massima di copertura nuvolosa accettata, deve essere specificata per formare lo *scheduling* finale. Durante la creazione di una DAR, può capitare che siano inseriti dei parametri impossibili da soddisfare. Per evitare tale inconveniente, sul sito vengono spiegati brevemente gli elementi ai quali gli utenti devono prestare attenzione, assieme agli accorgimenti utili per ottimizzare le possibilità di ottenere il prodotto richiesto. A meno di particolari necessità da prevedere durante l'acquisizione, lo *Science Scheduling Support Group* (SSSG) consiglia di scegliere parametri standard di acquisizione da fornire al DAR Tool. Le uniche eccezioni sono relative a: *Maximum Acceptable Cloud Cover* (copertura nuvolosa massima consentita) e *VNIR and SWIR Gain Settings* (settaggi dei valori del gain) per acquisizioni di oggetti che si prevedono molto o molto poco riflettenti. Di seguito si riportano le specifiche relative ai principali parametri:

- **Livello di urgenza:** richiedendo il massimo livello di urgenza, ovvero quello "Urgent", si intende indicare che il fenomeno osservato a terra si presenta solo in un periodo di tempo limitato ed in una limitata estensione geografica. Questa scelta assegna all'acquisizione una alta priorità e quindi aumenta la probabilità che l'acquisizione venga inserita nello scheduling della missione con particolare riferimento all'area di interesse indicata (AOI, *Area Of Interest*). Automaticamente vengono definiti due aspetti: la durata massima dell'ordine in 18 giorni e il costo in termini di risorse tecniche ed economiche.

- Area di interesse: esiste una doppia opzione, *Normal* o *Sampled*. Nel primo caso viene acquisita un'immagine delle dimensioni a terra di tipo standard mentre nella seconda modalità si richiede di effettuare dei campionamenti di estensione limitata su una AOI di notevoli estensioni. Tali aree di campionamento possono essere definite tramite sottoaree o singoli ordini riferiti a porzioni interne dell'area principale. Gli utenti ASTER vengono informati che la frammentazione dell'acquisizione dovrebbe sempre essere consentita, salvo rigide esigenze scientifiche che necessitano di acquisire tutta l'area di interesse simultaneamente. Selezionando "No" in questa opzione diminuisce la probabilità di successo dell'ordine, perchè la richiesta potrebbe entrare in contrasto con esigenze legate al rispetto dell'orbita stabilita per il satellite e per problemi relativi allo scheduling delle successive e precedenti acquisizioni. Ovviamente per AOI molto estese la frammentazione è necessaria (Fig. 4.12).
- Finestra temporale utile: deve essere selezionato l'intervallo temporale (la durata) durante il quale il sensore dovrebbe acquisire i dati con i parametri di acquisizione richiesti. Il periodo utile deve essere almeno di 16 giorni, ovvero il tempo di rivisitazione del sensore lungo l'orbita (parametri orbitali EOS-AM1). Per periodi inferiori il sensore potrebbe non trovarsi mai al nadir del luogo osservato, soprattutto alle basse latitudini. In ogni caso è consultabile il calendario delle acquisizioni per verificare se nel momento desiderato vi sono, e quante sono, altre riprese programmate in precedenza. In alcuni casi è possibile richiedere dati con una certa ciclicità: in questa casistica rientrano le acquisizioni finalizzate al monitoraggio di fenomeni che esigono una osservazione regolare su un certo intervallo temporale. Risulta fondamentale, quindi, evidenziare il termine ultimo del periodo di acquisizione. L'utente può anche indicare il numero massimo di riprese da effettuare.
- Copertura nuvolosa massima: sebbene sia preferibile l'assenza di coperture nuvolose, gli utenti possono anche accettare coperture che superano normali livelli soprattutto in zone caratterizzate da stagioni ad alta nuvolosità. La nuvolosità è espressa in percentuale sul totale dell'area rappresentata dall'immagine e potrebbe, comunque, coprire la particolare zona di interesse. Va notato che l'algoritmo che regola le acquisizioni (*ASTER Scheduler*) assegna una bassa priorità all'ordine di quelle scene per le quali viene previsto

un alto valore di copertura nuvolosa (in base alle previsioni meteo globali della NOAA). Fig. 4.13.

- **Geometria di acquisizione**: in questa fase entra in gioco la competenza dell'utente di dati ASTER. Infatti, la scelta dei parametri di acquisizione può essere in disaccordo con altri relativi alla tempistica dell'acquisizione: in particolare deve esserci accordo tra la geometria di presa e gli aspetti relativi alle finestre temporali ed alla frammentazione o meno dell'immagine. Per le osservazioni diurne gli angoli consigliati per l'elevazione del sole sull'orizzonte sono 0 e 90° rispettivamente per il valore minimo e massimo. Inoltre, è necessario rispettare un limitato range di valori per l'angolo di vista, ricordando che valori bassi garantiscono una geometria migliore da un punto vista prospettico. Tuttavia, anche l'imposizione di un angolo di vista particolarmente basso potrebbe generare una condizione impossibile per l'algoritmo che compone lo *scheduling* dell'acquisizione dei dati ASTER. Le probabilità di ottenere dei dati diminuisce imponendo vincoli di questo tipo.
- **Gain**: dato che i sottoinsiemi SWIR e VNIR del sensore ASTER utilizzano bande con una limitata risoluzione spettrale occorre specificare anche i settaggi relativi al gain da utilizzare. Questi valori dipendono principalmente dall'uso che deve essere fatto dei dati, ed alcuni settaggi particolari e "tipici" possono essere contenuti nella DAR stessa. La Tab. 4.3 fornisce i valori di default relativi ai parametri del gain per alcuni dei target classici presenti al suolo (sottoinsiemi VNIR e SWIR).

Banda	Possibili Valori	Deserto	Ghiaccio	Altri
VNIR #1	high, normal, low – 1	normal	low-1	high
VNIR #2	high, normal, low – 1	normal	low-1	high
VNIR #3	high, normal, low – 1	normal	normal	normal
SWIR #4	high, normal, low – 1, low - 2	normal	high	normal
SWIR #5	high, normal, low – 1, low - 2	normal	high	normal
SWIR #6	high, normal, low – 1, low - 2	normal	high	normal
SWIR #7	high, normal, low – 1, low - 2	normal	high	normal
SWIR #8	high, normal, low – 1, low - 2	normal	high	normal
SWIR #9	high, normal, low – 1, low - 2	normal	high	normal

Tab. 4.3 - Parametri di default del gain (<http://dartool.ecs.nasa.gov/>)

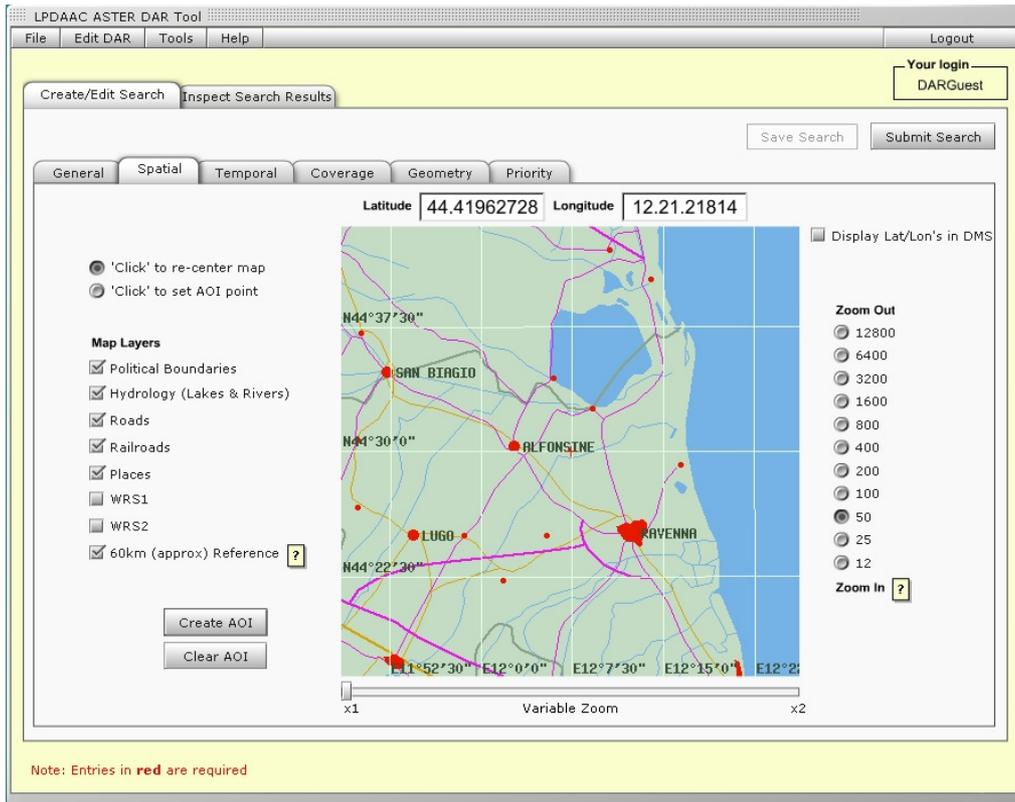


Fig. 4.12 – ASTER DAR Tool: area di interesse (<http://dartool.ecs.nasa.gov/>)

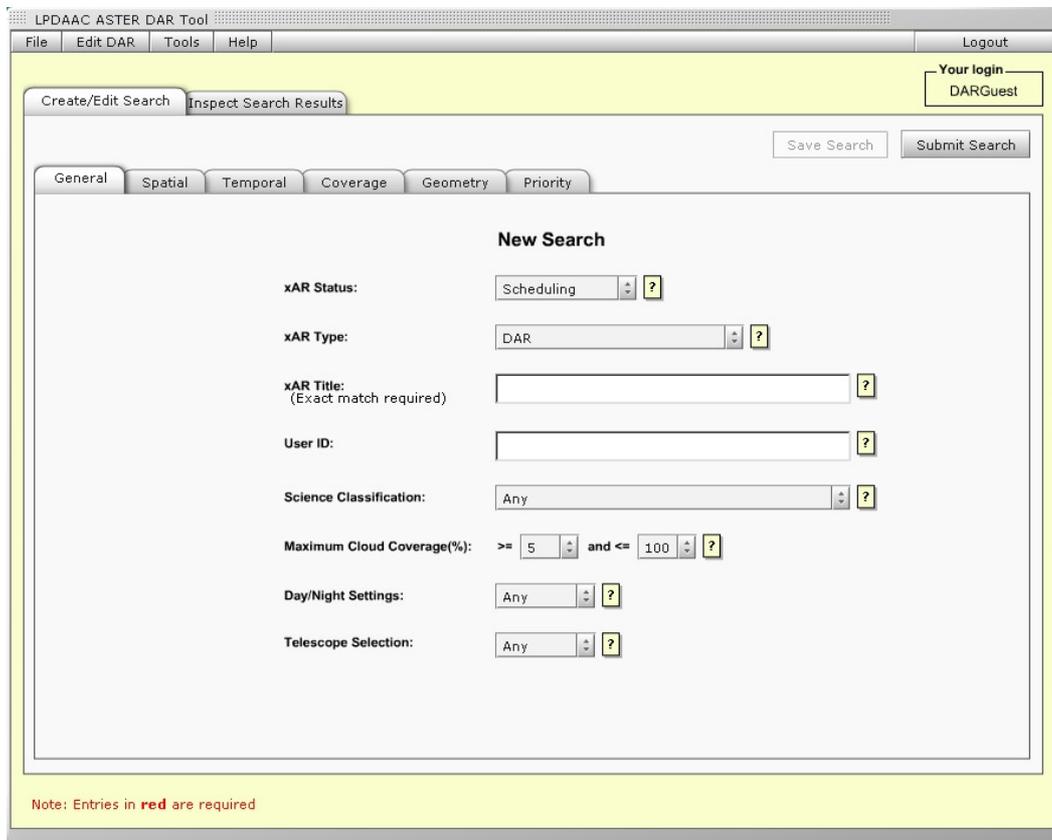


Fig. 4.13 – ASTER DAR Tool (<http://dartool.ecs.nasa.gov/>)

☞ Valutazione DAR: il DAR Tool consente alla NASA di sottomettere la richiesta al *Ground Data System* (GDS) in Giappone, che, ogni giorno, analizza il database delle DAR effettuate dagli utenti e pianifica le acquisizioni (lo scheduling appunto) per le successive 27 ore di osservazione. Le riprese ASTER vengono gestite dal GDS tramite l'algoritmo *ASTER Scheduler*, al quale viene demandata la scelta delle scene da acquisire in modo da ottimizzare il ritorno scientifico della missione spaziale su base giornaliera. Lo *Scheduler* analizza tutte le richieste DARs e STARs in modo da massimizzare il numero di acquisizioni anche in funzione delle modalità di funzionamento richieste per i tre sottoinsiemi. Infatti a livello di GDS vengono anche decisi quali sottosistemi devono essere utilizzati nelle acquisizioni (VNIR, SWIR, TIR), in relazione alle richieste pervenute che, per esempio, potrebbero prevedere delle riprese notturne solo tramite sensore TIR. Queste combinazioni vengono definite *Observing mode*. Le modalità di osservazione possono essere le seguenti:

- Modalità TIR, SWIR e VNIR: la maggior parte delle acquisizioni diurne vengono effettuate in questa modalità; il sensore VNIR acquisisce anche in modalità stereo grazie alla registrazione di due immagini nella banda 3 (810 nm) con orientamento del telescopio nelle posizioni “nadir” e “back-looking”, l'ultima delle quali acquisita circa 60 secondi dopo quella nadirale.
- Modalità TIR e SWIR
- Modalità TIR: in questa modalità vengono acquisite le immagini notturne.
- Modalità VNIR (inclusiva dell'acquisizione stereo nella banda 3 (810-nm)).
- Modalità VNIR (solo stereo).

Vengono quindi assegnate delle priorità e viene verificata la compatibilità con i parametri richiesti dagli utenti ASTER.

☞ Approvazione DAR: dopo una prima mail, generata in automatico, nella quale vengono riassunti tutti i parametri costituenti la DAR sottomessa, LPDAAC ne invia una seconda per confermare la registrazione della richiesta. Non essendo garantito il successo dell'operazione, per la programmazione dell'acquisizione non è previsto nessun costo. Solo dopo aver visionato l'immagine ripresa, l'utente

decide se acquistarla. In ogni caso il dato (indipendentemente dalla qualità) viene inserito nell'archivio generale e risulta disponibile per tutti gli utenti;

☞ Conferma dell'acquisizione: avviene mediante mail inviata da ASTER-GDS nella quale sono indicate le coordinate dell'area di interesse e i riferimenti per le fasi successive di ordine dell'immagine e di pagamento.

Il tempo necessario per eseguire tutta la procedura sopradescritta, esclusa l'avvenuta acquisizione, è di circa 45-50 giorni, in funzione del numero di richieste che ogni giorno viene fatto da tutte le parti del globo. Per quanto riguarda il costo, non è specificato sul sito: viene però indicato nella mail di conferma dell'acquisizione che è prevista una spesa aggiuntiva minima rispetto il prezzo di base, che si aggira intorno agli 80\$.

4.3.2.2 Risultato della programmazione di acquisizione del dato Aster

Come già anticipato nel paragrafo 4.3.2, la programmazione effettuata per l'acquisizione dell'immagine in corrispondenza dell'area di studio oggetto della tesi, nella finestra temporale compresa tra la fine giugno 2006 e i primi di luglio 2006, non ha avuto esito positivo. Infatti, confrontando le coordinate dei quattro vertici delle scene riprese in quel periodo (ASTL1A 06063010094806070305666 e ASTL1A 060630100939060703025), di cui LPDACC ha comunicato il successo dell'acquisizione

1) ASTL1A 0606301009480607030566

```
C.Latitude 44.392702, C. Longitude 11.628927
Start Data& Time 2006/06/30 10 :09 :48z
UPPERLEFT = (44.720983, 11.321528)
UPPERRIGHT= (44.615691, 12.111446)
LOWERLEFT= (44.168774, 11.143667)
LOWERRIGHT= (44.064384, 11.926296)
```

2) ASTL1A 0606301009390607030265

```
C.Latitude 44.920946, C. Longitude 11.803972
Start Data& Time 2006/06/30 10 :09 :39z
UPPERLEFT= (45.249573, 11.494187)
UPPERRIGHT= (45.143390, 12.291274)
LOWERLEFT= (44.697579, 11.313938)
LOWERRIGHT= (44.592325, 12.103542)
```

con le coordinate fornite nella DAR iniziale, appare evidente la ridotta sovrapposizione:

DAR Title: barbarella_DAR

DAR ID: 000045634
Right Upper Latitude: 44.74 deg
Right Upper Longitude: 12.38 deg
Left Lower Latitude: 44.39 deg
Left Lower Longitude: 12.01 deg.

La stessa agenzia ha riconosciuto l'errore, dopo aver verificato la correttezza dei dati di input inviati con la DAR, individuandolo direttamente nell'*ASTER DAR Tool Scenes Viewer* (Fig. 4.14).

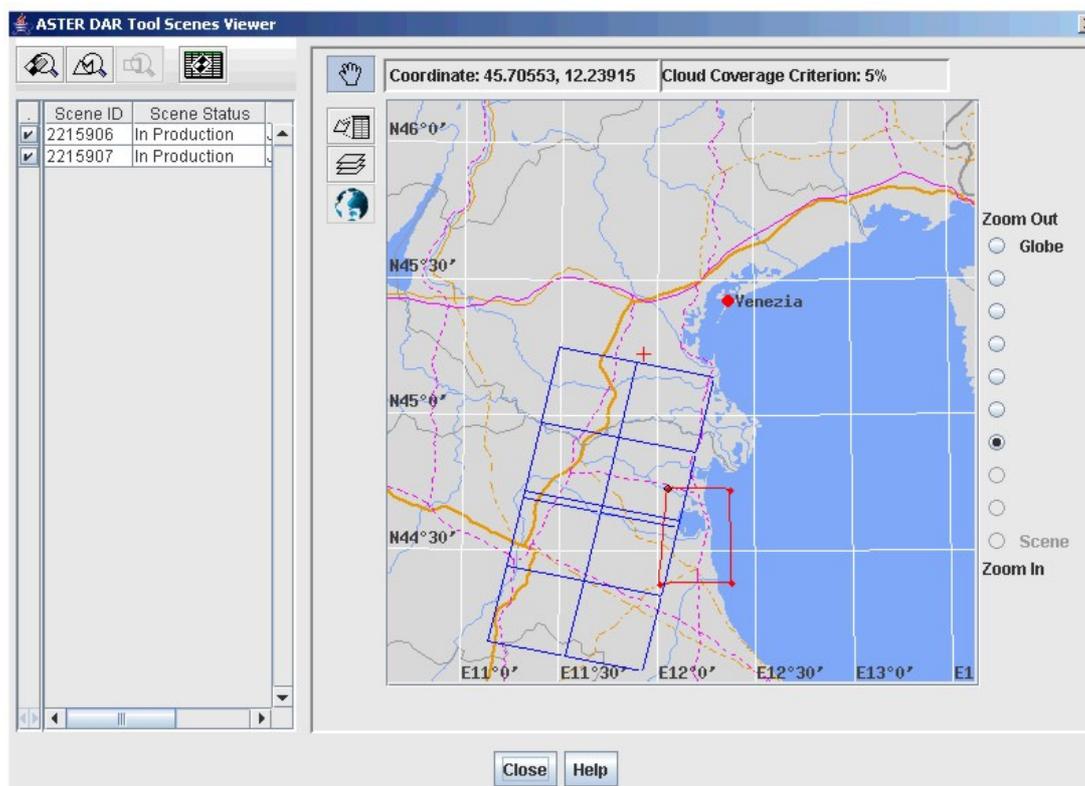


Fig. 4.14 – Area effettivamente acquisita

Il successivo acquisto dell'immagine non è quindi avvenuto. Fortunatamente però la zona di interesse era presente nelle scene acquisite durante passaggio precedente del satellite, verificatosi il 14/06/06, all'indomani del primo sopralluogo fatto nell'azienda. Di conseguenza, ai fini dell'indagine svolta, è stato possibile recuperare parte del progetto ideato, prevedendo un cambio nel tipo di coltura, da analizzare nel dettaglio, legato alla fase di maturazione delle piante: invece che sulla soia l'attenzione è stata concentrata soprattutto sulla barbabietola da zucchero. E' importante sottolineare infatti che un periodo di quindici giorni rispetto al ciclo vitale di una coltivazione è un intervallo di tempo consistente, nel quale diverse caratteristiche, come la

copertura fogliare, l'attività fotosintetica, ecc., mutano in modo significativo, incidendo pesantemente sulla risposta spettrale.

4.3.2.3 Aster Data Available at No Cost

Aster Data Available at No Cost è un progetto attraverso il quale il LPDAAC mette a disposizione gratuitamente i dati Aster per fini di ricerca, di didattica e simili. E' il NASA Office of Earth Science a realizzare e aggiornare continuamente la lista degli utenti autorizzati. Per ottenere l'accesso al servizio è necessario compilare un modulo on-line (<http://edcsgs16.cr.usgs.gov/landdaac/aster/application.html>) nel quale bisogna precisare, oltre ad illustrare brevemente il contesto scientifico nel quale si intende utilizzare le immagini Aster, l'appartenenza ad una delle tre seguenti categorie:

- ✓ ricercatori che ricevono fondi di ricerca NASA tramite terzi
- ✓ ricercatori che collaborano con partner internazionali della NASA
- ✓ docenti, PhD student e simili (sponsorizzati o meno dalla NASA)

La NASA, come membro del Working Group on Education and Training of the Committee on Earth Observation Satellites (CEOS), valuta e approva le richieste e le corrispondenti ricerche, soprattutto quelle accessibili tramite web. E' importante ricordare che ricercatori e educatori possono diffondere le immagini ottenute gratuitamente entro determinati limiti, in particolare è vietata la distribuzione al pubblico, pena la revoca del privilegio acquisito.

Il periodo di tempo (utile) tra la domanda e la successiva risposta è dell'ordine di 5-6 mesi: ad esempio la nostra richiesta è stata fatta nel novembre 2005 e la mail di approvazione è datata 12/06/06. Lo scarico dei dati può essere effettuato tramite FTP entro 48 ore dall'invio dell'ordine.

Grazie all'adesione a tale progetto è stato realizzato un archivio con le 10 immagini Aster precedentemente elencate (paragrafo 4.2.1), acquisite tra il 2000 e il 2006, sulle quali sono stati applicati gli algoritmi di classificazione.

4.3.3 Analisi del fenomeno studiato

Parallelamente alla scelta del dato satellitare, deve essere condotto uno studio sul fenomeno monitorato, finalizzato all'individuazione del periodo di acquisizione migliore per estrarre il maggior numero di informazioni possibile dall'immagine. L'informazione telerilevata selezionata dipende soprattutto dall'obiettivo perseguito nell'indagine, oltre che dalla disponibilità di archivio. Nel lavoro presentato, lo scopo della classificazione è, come già detto, il riconoscimento dell'uso del suolo di un'area prevalentemente agricola: la conoscenza iniziale richiesta non si limita ai tipi di colture presenti e alla loro localizzazione sul terreno, fase comunque molto delicata, ma comprende anche i singoli cicli di maturazione, le modalità di coltivazione, ecc. L'insieme di tutte queste indicazioni risulta fondamentale per la discriminazione delle firme spettrali che contraddistinguono i singoli oggetti. In generale, la risposta spettrale della vegetazione dipende dal tipo di pigmenti, soprattutto clorofilla, nel visibile, dalla struttura della foglia nell'infrarosso vicino e dal contenuto d'acqua nell'infrarosso a media lunghezza d'onda. In particolare, ulteriori tre fattori influenzano il comportamento delle coltivazioni: lo stato fenologico della coltura, perché il livello di energia emesso o riflesso durante la fase di emergenza è differente da quello registrato in avanzato stadio di maturità, lo stato di salute delle piante, perché la presenza di malattie o di stress modifica la firma diversificandola dalla corrispondente in condizioni normali, e infine la densità di copertura fogliare, in quanto, nel caso di una coltura rada, la sua risposta spettrale ha un andamento intermedio tra quello della pianta e quello del suolo nudo sottostante. È importante sottolineare inoltre che lo sfasamento dei cicli di vita delle diverse specie rende complessa l'identificazione del momento ottimale per la ripresa dell'immagine: infatti non è univoco l'istante in cui il fenomeno è più evidente (Barbarella et al., 2006).

4.3.3.1 Cicli di maturazione delle colture erbacee

MAIS

Il mais è una pianta di origine tropicale, tipicamente macroterma e brevidiurna. Grazie alla forte variabilità esistente all'interno della specie, il mais ha allargato moltissimo la sua area di distribuzione, per auto-adattamento o per selezione antropica. Secondo le condizioni locali di fotoperiodo, di temperatura, di umidità e di durata della stagione vegetativa si sono andate formando popolazioni diversissime tra loro, caratterizzate, di conseguenza, da firme spettrali simili ma mai coincidenti.

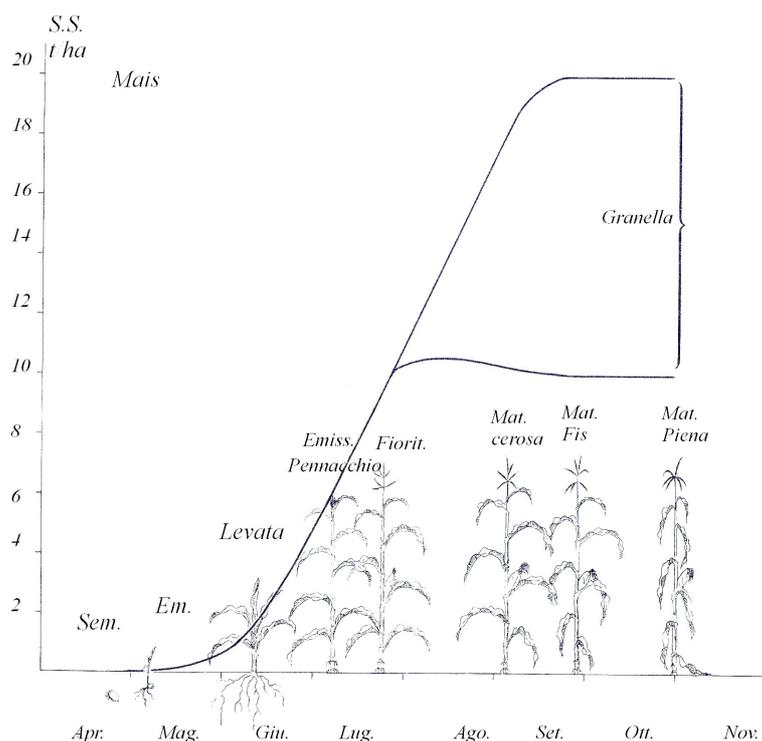


Fig. 4.15 - Schema del ciclo biologico del mais (da Bonciarelli, 2001)

Data la varietà, non esiste un clima ideale che assicuri la produzione, che può avvenire fino a 50° di latitudine Nord. Tuttavia il mais, in ragione della sua macrotermia, esige temperature elevate per tutto il suo ciclo vitale, durante il quale manifesta esigenze via via crescenti. Il mais non germina e non si sviluppa quando le temperature sono inferiori a 10°C; per assicurare una nascita tempestiva, è necessario aspettare, dunque, che la temperatura del terreno abbia raggiunto i 12°C. La temperatura ottimale per l'accrescimento è di 22-24°C, per la fioritura è di circa 26°C. Alle nostre latitudini, è necessario aspettare il mese di Aprile, per procedere alla semina, a seguito della quale l'emergenza si verifica tempestivamente, intorno alla seconda decade di Maggio. Con l'inizio dell'attività fotosintetica, la pianta accumula le sostanze nutritive necessarie all'accrescimento fino alla completa maturazione, che avviene solitamente agli inizi di Novembre (Fig. 4.15).

SORGO

È coltura a ciclo primaverile-estivo che, data la resistenza all'aridità, è praticabile sia in asciutto che in irriguo ed è pertanto molto adatta alle condizioni climatiche dei nostri ambienti. Il sorgo è specie termofila, quindi piuttosto esigente in fatto di temperatura fin dalla germinazione, fase in cui possono arrecarle danno temperature inferiori a 10 °C. Le temperature ottimali per l'intero ciclo vegetativo si aggirano tra i 25 ed i 30 °C, mentre valori superiori a 38-40 °C possono mettere in crisi la pianta per eccesso termico, soprattutto in ambiente arido. Estese sono le possibilità di destinazione produttiva di questa pianta, idonea sia per la produzione di granella, che per quelle di

foraggio (erbaio o stadio ceroso), fibra e zucchero. Coltivato per la produzione di granella prende nell'avvicendamento il posto della coltura principale, mentre per le altre destinazioni è preferibile l'inserimento come intercalare. Nel primo caso, la semina avviene generalmente in maggio (quando più difficile è l'eventualità di temperature inferiori allo zero), su terreno ben lavorato con aratura profonda, frangizollatura e/o fresatura, disponendo il seme a 2-3 cm di profondità (profondità superiori potrebbero rendere difficoltosa l'emergenza) con seminatrice di precisione, ad un investimento di 30 piante per m².

BARBABIETOLA DA ZUCCHERO

La barbabietola è una pianta a ciclo biennale: nell'anno di semina rimane in fase vegetativa formando un ricco apparato fogliare, nel secondo anno, dopo aver subito la vernalizzazione, fruttifica. Questa fase interessa solo le colture da seme, mentre nelle colture destinate alla produzione di zucchero, la pianta è lasciata compiere solo la fase vegetativa e viene raccolta quando è massimo l'accumulo di saccarosio nella radice.

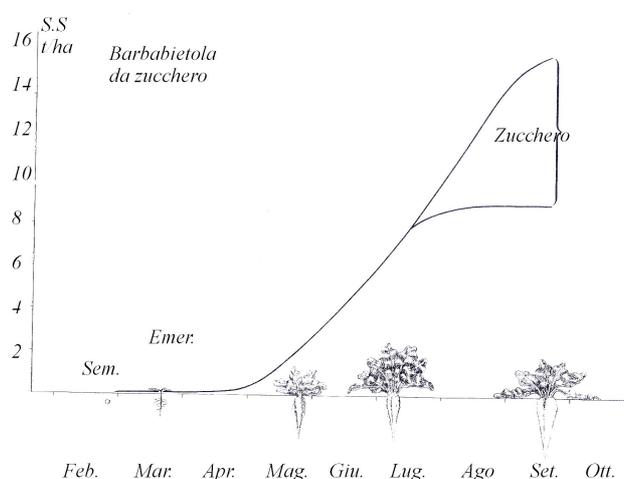


Fig. 4.16 - Schema del ciclo biologico della barbabietola da zucchero (da Bonciarelli, 2001).

Anche se la germinazione del seme di barbabietola può avvenire a temperature piuttosto basse (3 - 4°C), per evitare nascite troppo lente e stentate è conveniente seminare solo dopo che la temperatura del terreno è giunta a 6°C: fenomeno che, alle nostre latitudini, si verifica normalmente a fine Febbraio-Marzo. In condizioni normali di temperatura, tra la semina e l'emergenza trascorrono 15-20 giorni circa. Segue un periodo di 50-55 giorni di crescita relativamente lenta, durante il quale la radice cresce poco, ma viene a differenziarsi un ricco apparato fogliare. La

crescita delle foglie è facilitata dalla temperata mite del periodo che si attesta intorno ai 20°C (Fig. 4.16).

GRANO

Si diversifica in più varietà tra cui il grano duro, originario dell'Africa mediterranea, della Grecia e dell'Etiopia e il grano tenero, originario dell'Asia sud-occidentale. Entrambi sono destinati all'alimentazione umana. Il grano duro produce semole dai granuli grossi con spigoli netti, mentre dal grano tenero si ottengono farine dai granuli tondeggianti. La pianta del grano è considerata “sfruttante” e viene avvicinata con colture da rinnovo (mais, barbabietola) e da leguminose - foraggere (erba medica, trifoglio, ecc.).

Il grano duro appartiene ai cereali autunno-vernini, categoria che raggruppa le specie competitive, dotate di elevata capacità di adattamento a differenti situazioni pedo-climatiche, anche in presenza di forti fattori limitanti. Dal punto di vista foto-periodico è una specie longigiurna, mentre da quello termico è una specie microterma, cioè non ha bisogno di alte temperature per crescere, svilupparsi e produrre. Le sue esigenze termiche minime sono relativamente basse: 2-3°C per la germinazione, 18-20°C per la maturazione. Le alte temperature, infatti, non sono favorevoli a questa varietà di frumento, poiché non soddisfano le esigenze di vernalizzazione. Per questi motivi il grano duro viene coltivato, in tutte le regioni con un inverno sufficientemente mite, tale da assicurare la sopravvivenza delle piante, provvedendo alla semina autunnale e alla raccolta in primavera-estate. Durante l'emergenza, infatti, la resistenza al freddo delle piantine è abbastanza limitata, 6-8°C sotto lo zero (Fig. 4.17).

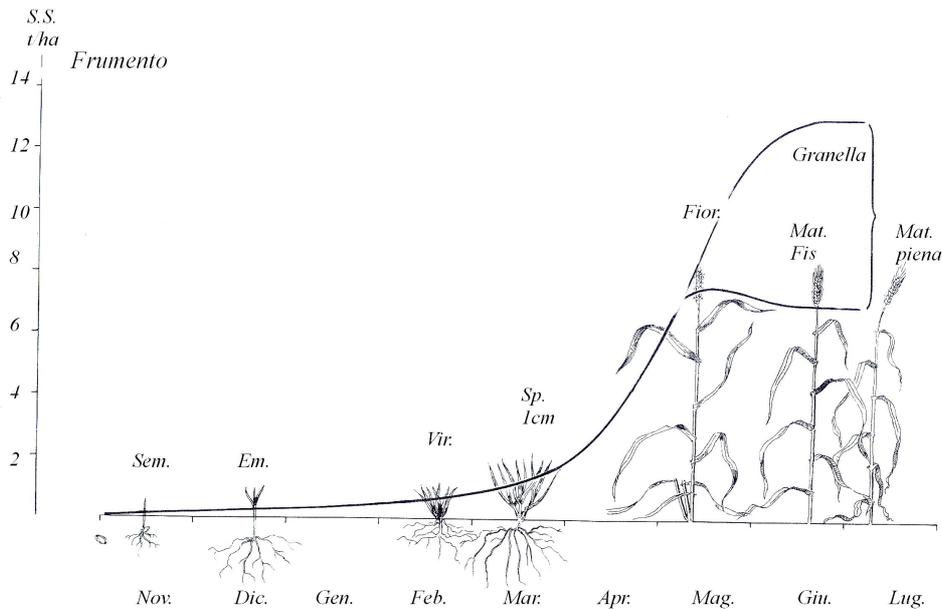


Fig. 4.17 - Schema del ciclo biologico del grano (da Bonciarelli, 2001)

ERBA MEDICA

Si tratta di una varietà delle colture foraggere poliennali falciabili. La coltura dell'erba medica è localizzata nei terreni alluvionali, profondi e freschi dell'Italia centro-settentrionale, è coltivata su circa 1.3 milioni di ettari, concentrati prevalentemente in Emilia-Romagna, Lombardia, Veneto e Lazio. Il seme dell'erba medica germina in condizioni ottimali a 18°C, la ripresa vegetativa delle piante già seminate inizia, invece, quando la temperatura media raggiunge 8-10°C; nell'Italia settentrionale questa temperatura può verificarsi intorno al mese di Marzo. La fioritura inizia indicativamente verso il 15 Maggio, dopo di che si procede al primo taglio, l'ultimo si ha verso Novembre (Fig. 4.18).

Per quanto riguarda le esigenze ambientali, questa pianta può essere coltivata dalla latitudine del Circolo polare artico alle regioni equatoriali. Tuttavia, i maggiori centri di coltivazione sono le zone a clima temperato caldo. Verificandosi più tagli durante l'anno, è possibile trovare nello stesso momento, su terreni differenti e quindi sottoposte a gestioni agricole diverse, coltivazioni in fasi fenologiche non coincidenti: per tale motivazione in alcune leggende l'erba medica può comparire come tipo A e tipo B.

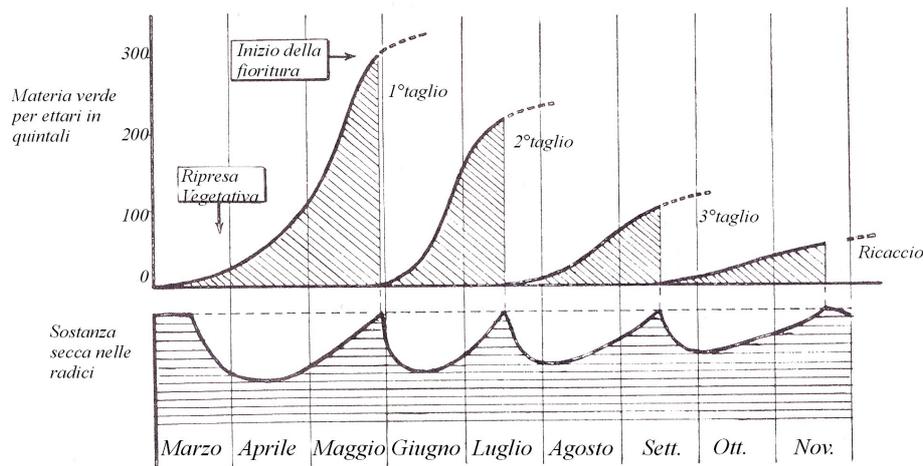


Fig. 4.18 – Andamento della vegetazione di un prato di erba medica
(da Bonciarelli, 2001)

PISELLO

Il pisello è una leguminosa annuale, con uno stelo cilindrico sottile e debole, di lunghezza variabile da 30 cm a 3 metri a seconda della varietà: nani, semi-nani e rampicanti. E' una pianta microterma, con limitate esigenze di temperatura per crescere e svilupparsi, e rifugge dai forti calori e dalle siccità. Per questo la coltura del pisello può essere fatta con successo negli ambienti e nelle stagioni fresche. In Italia, la semina è autunnale nelle regioni a inverno mite, mentre in quelle settentrionali deve essere fatta dopo passati i rigori dell'inverno.

Il pisello germina con accettabile prontezza anche con temperature del terreno intorno a 4°C, mentre la temperatura ottimale per il compimento del ciclo vitale è compresa tra 15°C e 18°C.

SOIA

La soia è una pianta erbacea annuale, estiva, alta da 70 a 130 cm, a portamento eretto più o meno cespuglioso. Il seme di soia, per nascere con ragionevole rapidità, ha bisogno di trovare nel terreno temperature stabilmente attestate oltre i 12-13°C. Con temperature superiori, fino a 32°C, come si possono riscontrare nel terreno di seconda coltura, la germinazione e la nascita avvengono anche in 3-5 giorni. Trascorsa la fase dell'emergenza, la soia sopporta, senza grave danno, abbassamenti di temperatura di poco sotto lo zero (-2°C), ma lo sviluppo vegetativo si arresta quando la temperatura scende sotto 10°C. I limiti minimo e massimo per l'accrescimento sono 18 e 30°C, con un intervallo ottimale tra 22 e 25°C.

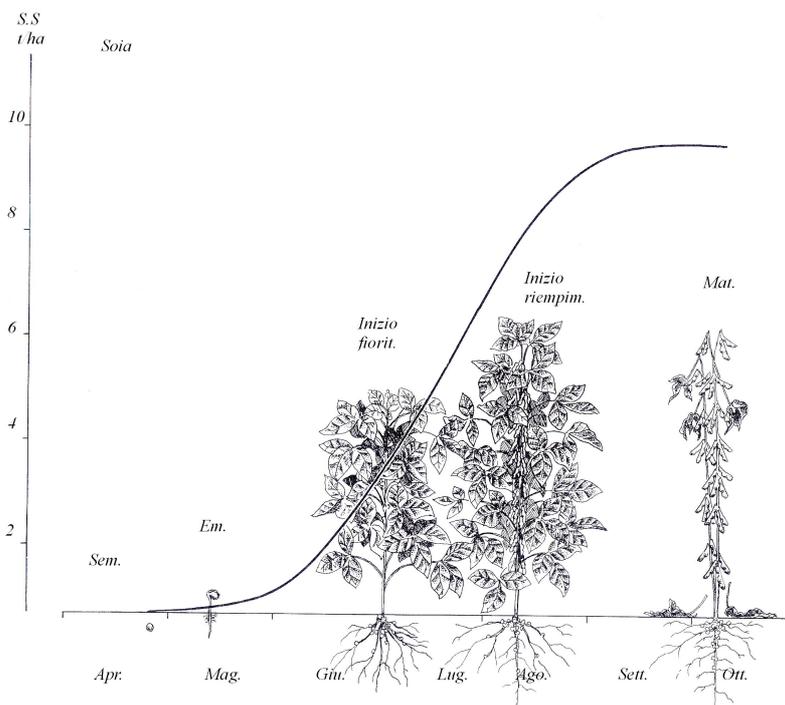


Fig. 4.19 - Schema del ciclo biologico della soia (da Bonciarelli, 2001)

La soia è una pianta brevidiurna, ciò significa che per passare alla fase di differenziazione dei fiori necessita di un ritmo giornaliero di illuminazione e di oscurità particolare, esigendo un periodo giornaliero di buio superiore ad una certa soglia critica, caratteristica di ogni varietà. In certe varietà, il periodo di buio minimo indispensabile per fiorire è molto lungo, fino alle undici ore, in altre, invece è relativamente breve, circa sette ore.

Le piante della varietà del primo tipo, se seminate in primavera a una latitudine abbastanza elevata, non fioriscono durante i mesi estivi, a causa del lungo periodo giornaliero di luce, di oltre 15 ore. Si trovano, dunque, nelle condizioni per fiorire solo a stagione molto inoltrata, verso la seconda decade di Settembre (Fig. 4.19).

Queste varietà, proprio in ragione della “tardività” che le caratterizza, non si prestano ad essere coltivate ad alta latitudine. In tali zone, infatti, si preferisce coltivare piante della seconda varietà: trovano soddisfatto il loro bisogno di buio anche durante le notti estive, caratterizzate da una durata minima di 8 ore di buio ed arrivano a fioritura in tempi abbastanza rapidi, i quali si aggirano intorno alle sei-sette settimane dopo la nascita.

Per quanto riguarda le esigenze climatiche, quelle della soia sono molto simili a quelle del mais, con un maggior bisogno però di calore nelle fasi della germinazione e della maturazione. Successivamente le temperature devono attestarsi a valori medi durante la fase di fioritura, per non compromettere i processi fecondativi.

POMODORO

Il pomodoro è una pianta a fusto sarmentoso, che si ramifica abbondantemente nelle parti più basse. Originaria dell'America Latina fu, per lungo tempo, coltivata in Europa a solo scopo ornamentale, poiché i frutti non erano ritenuti commestibili. Solo con l'inizio del secolo, nei paesi mediterranei, iniziò la coltivazione intensiva e la trasformazione e conservazione industriale del prodotto.

Da allora, il pomodoro è diffuso come coltura ortiva in tutta Italia, ma in pieno campo è coltivato soprattutto in Puglia, Campania, Emilia-Romagna, Calabria e Sicilia.

Si tratta di una pianta da clima temperato-caldo, che esige una temperatura minima di germinazione di 12-13°C e di 22-25°C per svilupparsi e produrre. Ci sono numerose varietà che si distinguono in funzione della forma e della grandezza dei frutti.

Alle nostre latitudini, si procede solitamente alla semina, che avviene in un luogo a riparo dalle temperature più rigide, verso la seconda decade di Febbraio o agli inizi di Marzo, quando la stagione invernale è stata particolarmente rigida. Si attende, dunque, che la fase dell'emergenza sia ultimata e che la pianta raggiunga una resistenza tale da consentire il trapianto, che avviene, solitamente, a Maggio. In condizioni di temperatura non avverse, i frutti sono prodotti in poche settimane e maturano, abbastanza velocemente, nel corso dell'estate. Il pomodoro da mensa viene raccolto a maturazione incompleta, quello da industria quando il frutto è maturo; la raccolta si esegue a partire da Luglio per le varietà più precoci e si protrae a tutto Agosto - Settembre.

GIRASOLE

Il girasole è una pianta tipicamente a ciclo primaverile-estivo: i fattori ecologici prevalenti che influenzano la tecnica di coltivazione e di sviluppo sono, perciò, la temperatura minima necessaria per la germinazione, responsabile dell'epoca della semina, e le disponibilità idriche durante il ciclo vegetativo. Il girasole è una pianta che sopporta bene la siccità e, essendo originaria dei climi temperati, ha esigenze termiche piuttosto modeste. Anche se la temperatura standard per la germinazione è di 20-30°C, il seme di girasole può germinare a temperature piuttosto basse, per evitare nascite troppo lente e scalari. Si approfitta di questa proprietà per anticiparne la crescita nel periodo in cui le temperature del suolo sono di circa 10°C: questo significa che può essere seminato piuttosto presto, a fine inverno o inizio primavera (Fig.4. 20). Ciò è consentito anche dalla particolare resistenza della pianta alle temperature rigide: allo stadio dell'emergenza, il girasole resiste, senza danno, a temperature di parecchi gradi sotto lo zero. Sotto l'azione di queste, subisce un momentaneo ingiallimento, cui segue, però, una completa e rapida ripresa appena le condizioni termiche tornano normali.

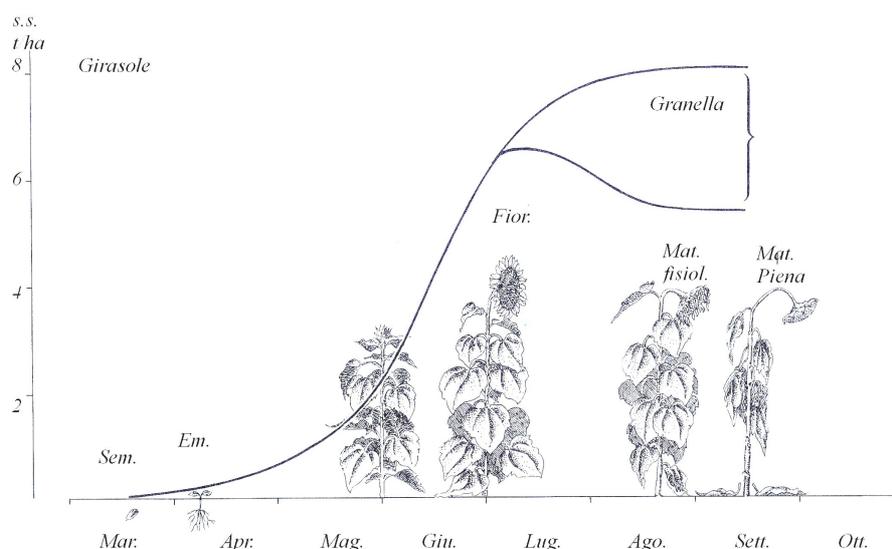


Fig. 4.20 - Schema del ciclo biologico del girasole.(da Bonciarelli, 2001)

RAVANELLO

Quest'ortaggio, originario delle zone della Cina e del Giappone, viene coltivato principalmente per le radici, la parte commestibile, che possono essere di diverso colore (rosso, bianco, verde), forma e grandezza. Si tratta di una annuale, con ciclo vegetativo molto breve. La parte aerea si presenta con delle foglie lobate abbastanza piccole coperte da una peluria sulla pagina superiore. Alla base della pianta, invece, le foglie sono più piccole e picciolate. I fiori, insignificanti, di colore bianco violaceo raggruppati in gruppi, sbocciano 3 mesi dopo la semina e da questi si formano dei frutti a siliqua di colore rosso scuro. La pianta va seminata nel periodo che va da febbraio fino a luglio. Nei primi due mesi la semina deve essere effettuata in un luogo protetto, mentre nei mesi successivi può essere effettuata all'aperto, sia a spaglio che a file. Nel primo caso, la semente dovrà essere distribuita il più uniformemente possibile sul terreno, evitando di spargere la semente in modo troppo fitto, perché ciò impedirebbe la crescita di molte piante. Nella semina a file, invece, è opportuno lasciare 7-8 cm di distanza tra una fila e l'altra, per favorire la crescita delle radici. In entrambi i casi si consiglia di effettuare un diradamento delle piante lasciando solo le più vigorose. Questa essenza, visto la velocità di crescita, può essere seminata anche tra altri ortaggi a ciclo vegetativo più lungo. La pianta teme le temperature molto elevate e le siccità: per questo motivo è opportuno irrigare frequentemente. La raccolta avviene un mese dopo la semina, quando le radici sono ben sviluppate. Se la semina viene effettuata durante il periodo estivo il tempo di raccolta viene posticipato di 40 giorni.

LEGUMINOSE DA SEME

Esistono numerose specie di leguminose da granella molto diverse tra loro sia come botanica che come origine: in Italia vengono coltivati principalmente Fagioli, Ceci, Fave, Lenticchie e Piselli. Nelle dichiarazioni PAC è stato però possibile distinguere solo le colture di Pisello, tutte le altre rientrano sotto la voce Leguminose da seme.

4.3.3.2 Indice di vegetazione: NDVI

L'Indice di Vegetazione, nella sua formulazione come NDVI (*Normalized Difference Vegetation Index*), sfrutta la diversa risposta della copertura vegetale alle bande spettrali del visibile (rosso) e del vicino infrarosso, e fornisce un valore numerico adimensionale, teoricamente compreso tra -1 e +1. Tale valore è stato dimostrato essere in stretta relazione con lo stato di salute della vegetazione, intesa come biomassa e area fogliare (*Leaf Area Index*), e con i processi biochimici ad essa correlati (attività fotosintetica). La caratteristica del comportamento dei pigmenti delle foglie non sottoposte a stress è infatti quella di riflettere soltanto circa il 10% della radiazione ricevuta nella regione spettrale della luce visibile (rosso), ed allo stesso tempo di riflettere oltre il 40% di quella ricevuta nell'infrarosso vicino. Nella mappe elaborate, valori bassi di NDVI si verificano in aree a bassa o assente copertura vegetale, o dove la vegetazione presente è senescente o sofferente, mentre gli alti valori dell'indice rispecchiano una situazione di forte attività fotosintetica e quindi elevata presenza di biomassa. Nella realtà osservata, i valori di NDVI variano prevalentemente in un range compreso tra -0.2 e +0.6 sulle terre emerse (con i corpi d'acqua, la neve e le nubi che mostrano valori vicini allo zero o leggermente negativi). La formula utilizzata per il calcolo dell'indice NDVI è la seguente:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

nella quale NIR indica la banda dell'infrarosso vicino (banda 3N - VNIR) e RED indica la banda del rosso (banda 2 - VNIR).

Con il monitoraggio delle mappe di NDVI vengono tenuti sotto costante osservazione fenomeni su scala continentale, quale quello dell'avanzamento della desertificazione in Africa, da parte della Nazioni Unite e della F.A.O.; è in grande crescita inoltre, in ogni parte del mondo, l'utilizzo di questo strumento da parte degli enti per la programmazione e l'assistenza tecnica in campo agricolo su scala nazionale e regionale, con particolare riferimento alla valutazione dell'impatto sulle produzioni agricole a seguito di particolari eventi ambientali (siccità, inondazioni, infestazioni etc.).

E' importante precisare che l'utilizzo dell'Indice di Vegetazione per il monitoraggio delle condizioni della copertura vegetale ha una valenza assolutamente "qualitativa", e non consente, allo stato attuale delle conoscenze (ma molte ricerche vengono oggi dirette in questo senso), una valutazione "quantitativa" di tale stato. In particolare il confronto a fini agrometeorologici tra le diverse mappe dell'indice nei diversi periodi dell'anno, richiede un'analisi approfondita delle differenze nella tipologia di vegetazione e di suolo presenti, al fine di differenziare localmente i diversi stadi fenologici dall'influenza dei differenti tipi di suolo.

A questo riguardo è opportuno evidenziare che la scala dei colori scelta per rappresentare i diversi valori dell'indice è puramente arbitraria e non ha nessun riferimento all'apparenza reale della vegetazione.

Di seguito si riportano gli indici NDVI calcolati per le diverse immagini di archivio confrontando tra loro quelle relative allo stesso anno. Di default il software utilizzato (Idrisi) consente di associare una legenda suddivisa in 255 livelli: però nell'immagine-risultato distribuisce le gradazioni di colore solo all'interno del range di livelli realmente presenti. Per rendere confrontabili i diversi indici è stato necessario ridistribuire i colori sull'intera scala.

ANNO 2002

Considerando che tra la metà e la fine di aprile gran parte delle coltivazioni seminate devono ancora entrare nella fase di emergenza del terreno, risulta evidente, dalla Fig. 4.21, l'assenza di vegetazione su numerosi appezzamenti presenti nel dettaglio dell'immagine prodotta come NDVI. Mentre una maggiore attività fotosintetica appare nell'immagine di giugno, anche se alcune colture sono già state raccolte.

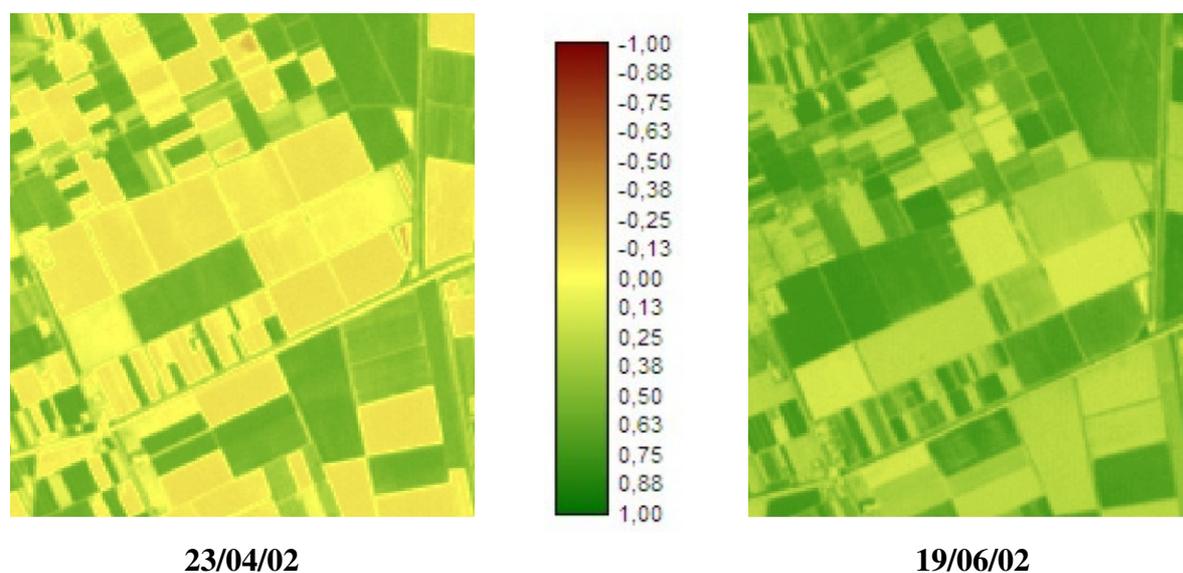


Fig. 4.21 – Confronto tra gli NDVI calcolati per l'immagine del 23 aprile 2002 e per l'immagine del 19 giugno 2002

ANNO 2005

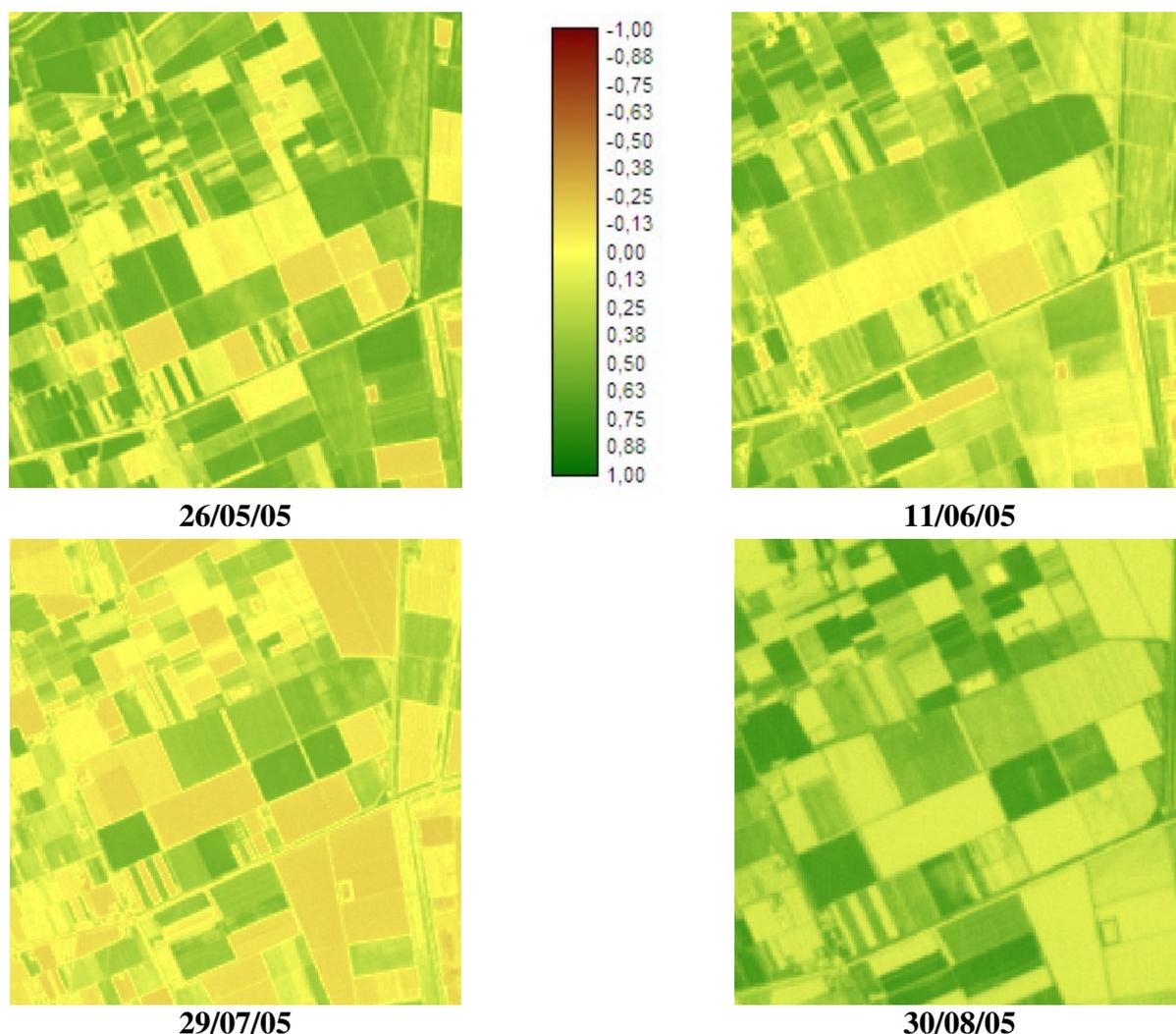


Fig. 4.22 – Confronto tra gli NDVI calcolati per le immagini del 26 maggio 2005, 11 giugno 2005, 29 luglio 2005 e 30 agosto 2005

Per il 2005 è stato possibile confrontare la situazione presente sul terreno in quattro momenti diversi: fine maggio, metà giugno, fine luglio e fine agosto (Fig. 4.22). Il periodo corrisponde all'intero ciclo di maturazione di alcune specie, come la soia e il mais, e a quello parziale di altre come il grano e l'erba medica. Come già detto in precedenza è difficile, infatti, stabilire quando il fenomeno studiato, data la sua grande variabilità, è maggiormente rappresentativo dal punto di vista spettrale. Nei dettaglio del 30/08/05 la gradazione del colore sembra essere concentrata solo intorno al verde: questa differenza rispetto alle altre immagini è legata, oltre al fattore temporale, ad una differente calibrazione nella registrazione dei valori radiometrici, visibile anche nel corrispondente color composite.

Per comprendere, poi, come la risposta spettrale di una coltura cambia rapidamente nel corso di quattro mesi, si osservi la variazione subita dai valori NDVI per la coltivazione della soia (Fig. 4.23).

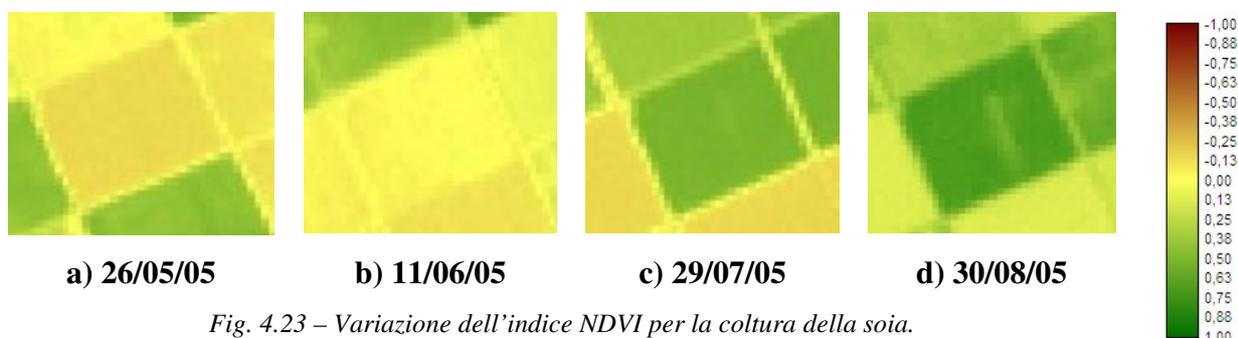


Fig. 4.23 – Variazione dell'indice NDVI per la coltura della soia.

Dopo la semina, che avviene ad aprile, la fase di emergenza si verifica rapidamente verso la fine di giugno: questo è dimostrato anche dal dettaglio del 11/06/05 (b), nel quale i valori da negativi (a) cominciano a superare lo 0, indicando l'inizio di un processo vegetativo. A fine luglio (c) la pianta ha raggiunto quasi l'80% della sua altezza e la maturazione procede per tutto il mese di agosto (d) fino alla maturazione, prevista per la fine di settembre.

ANNO 2006

L'intervallo di tempo che intercorre tra le due immagini riportate in Fig. 4.24 è di un mese: il cambiamento che maggiormente colpisce è quello relativo alle condizioni del terreno, seminato a mais o soia a seconda degli appezzamenti, che presenta valori negativi a metà giugno e valori superiori allo 0 a metà luglio, quando le piante sono prossime ad emergere dal suolo.

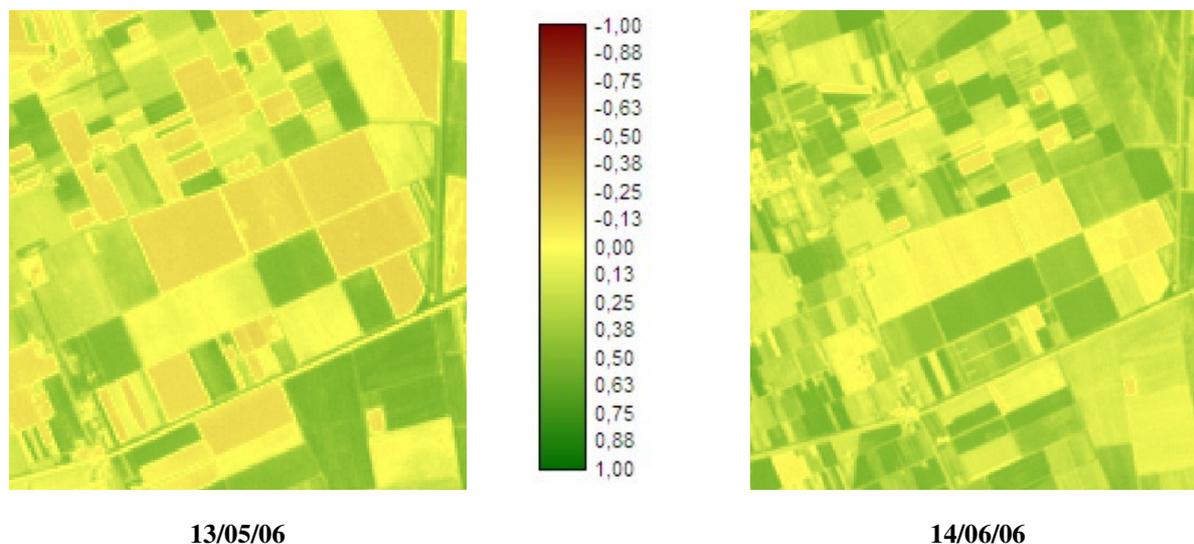


Fig. 4.24 – Confronto tra gli NDVI calcolati per le immagini del 23 aprile 2002 e 19 giugno 2002

PARTE III

Introduzione alla parte III

Per verificare l'eventuale raggiungimento dell'obiettivo principale della Tesi, ovvero il riconoscimento delle colture agricole mediante immagini satellitari Aster, nella terza ed ultima parte sono riportati i risultati delle elaborazioni, ottenute sia da classificazioni per pixel sia da classificazioni object-oriented. Alla luce delle problematiche descritte nel Capitolo 4 dell'elaborato, circa la raccolta e la gestione dei dati dai quali estrarre le informazioni sull'uso del suolo, vengono inizialmente mostrati alcuni esempi di tracciamento delle aree di addestramento, con un particolare riferimento ai casi di dubbia interpretazione.

Avendo poi a disposizione una serie di dati telerilevati relativi ad un intervallo temporale di cinque anni (2002-2006), è stato possibile effettuare due tipi di controlli qualitativi: uno tra classificazioni di immagini riprese nello stesso anno ma in mesi diversi, utile per ricostruire l'intero ciclo di vita di alcune colture mediante l'osservazione dei singoli stati fenologici, e l'altro tra classificazioni relative allo stesso periodo di maturazione ma in anni diversi, necessario per studiare tutti quei fattori che influenzano la risposta spettrale di una stessa coltura. Il confronto visivo tra le metodologie è comunque possibile farlo mediante gli esempi, sia di successo che di insuccesso di classificazione, distribuiti nei vari paragrafi.

Da un punto di vista quantitativo, l'accuratezza delle singole elaborazioni è stata valutata mediante il calcolo di una serie di parametri, tutti derivati dalla matrice di confusione, realizzata per ogni classificazione. L'attenzione è stata rivolta soprattutto alla determinazione del coefficiente K, che tende a depurare la stima dell'accuratezza dall'effetto legato all'accordo prodotto dalle coincidenze casuali fra la classificazione e la mappa di riferimento, basata sulle verità a terra. Tale coefficiente permette di valutare l'accuratezza sia a livello globale che di singole classi. Il confronto tra le due metodologie di classificazione utilizzate invece è stato realizzato con un test statistico, Z test, in grado di stabilire se vi sono grosse differenze tra i risultati relativi ad una stessa immagine. Lo scopo del test è stabilire se una metodologia è migliore dell'altra ai fini della discriminazione delle firme spettrali delle colture agricole, o se invece è consentito un loro indifferente utilizzo.

Infine, nell'ottica di progettare una procedura generale semplice e il più possibile veloce per il riconoscimento delle coltivazioni, viene suggerito un flusso di lavoro per eseguire le diverse operazioni, che si traduce in una schematizzazione dei capitoli operativi dell'intero elaborato di Tesi.

CAPITOLO 5: CLASSIFICAZIONI

5.1 La fase di addestramento

5.1.1 Classi individuate

Dai documenti a disposizione, (la Carta dell'Uso del Suolo, la Carta della Vegetazione, le dichiarazioni PAC e le mappe catastali), dalla consulenza con membri dell'Unione Agricoltori di Ravenna e dai sopralluoghi effettuati in alcune aziende agricole, è stato possibile individuare un certo numero di classi, variabile da anno ad anno. Di seguito, si riportano le tabelle (5.1, 5.2, 5.3 e 5.4) con le categorie presenti nei singoli dati, distinti per anno di acquisizione.

La determinazione dei training sites, usati con entrambe le metodologie previste, ha comportato una serie di problemi, descritti in seguito, legati soprattutto alla difficile reperibilità dei dati a terra. Per la realizzazione materiale dei training sites, è stato creato un vector layer apposito, uno per ciascuna immagine, da aggiungere all'immagine ASTER presa come riferimento, ovvero il color composite (ottenuto dalla combinazione delle bande verde, rosso e infrarosso vicino). La scelta è ricaduta su quest'ultima perché risulta la migliore per il riconoscimento visivo della suddivisione del territorio.

L'analisi che segue ha il duplice scopo di descrivere nel dettaglio le procedure svolte e di evidenziare le difficoltà incontrate, il cui superamento a volte parziale ha influenzato la precisione delle elaborazioni.

DATI ACQUISITI NEL 2002	
ID	CLASSE
1	MAIS
2	BIETOLA
3	GRANO
4	ERBA MEDICA
5	PISELLO
6	SOIA
7	FRUTTETO
8	VIGNETO
9	COLTURE DA LEGNO
10	BOSCHI MISTI O DI CONIFERE
11	SPIAGGIA
12	BACINI D'ACQUA SALATA
13	PALUDE SALMASTRA
14	ACQUA DI BASSO FONDALE
15	ACQUA DI MEDIO FONDALE
16	ACQUA DI ALTO FONDALE
17	FIUMI
18	ZONA INDUSTRIALE
19	ZONA URBANA
20	POMODORO
21	VEGETAZIONE SOMMERSA ACQUA DOLCE
22	VEGETAZIONE SOMMERSA ACQUA SALMASTRA
23	GIRASOLE
24	SET-ASIDE
25	PESCHETO
26	PERETO
27	ACQUA DI FOCE

Tab. 5.1 – Classi individuate dalle dichiarazioni PAC relative al 2002 e dagli altri documenti disponibili

DATI ACQUISITI NEL 2004	
ID	CLASSE
1	MAIS
2	SORGO
3	BIETOLA
4	GRANO
5	VIGNETO
6	ERBA MEDICA_A
7	SOIA
8	POMODORO
9	RAVANELLO
10	PISELLO
11	GRANO TENERO
12	ERBA MEDICA_B
13	PESCHETO
14	PERETO
15	FIUMI
16	ACQUA DI BASSO FONDALE
17	ACQUA DI MEDIO FONDALE
18	ACQUA DI ALTO FONDALE
19	PALUDE CHIUSA
20	PALUDE SALMASTRA
21	BACINI D'ACQUA SALATA
22	SPIAGGIA
23	BOSCHI MISTI O DI CONIFERE
24	SALINE
25	VEGETAZIONE SOMMERSA ACQUA DOLCE
26	VEGETAZIONE SOMMERSA ACQUA SALMASTRA
27	ZONA URBANA
28	ZONA INDUSTRIALE
29	VEGETAZIONE ALOFILA
30	VEGETAZIONE PALUSTRE ACQUA DOLCE

Tab. 5.2 – Classi individuate dalle dichiarazioni PAC relative al 2004 e dagli altri documenti disponibili

DATI ACQUISITI NEL 2005	
ID	CLASSE
1	MAIS
2	SOIA
3	ERBA MEDICA
4	GRANO
5	PISELLO
6	BIETOLA
7	LEGUMINOSE DA SEME
8	PERETO
9	PESCHETO
10	VIGNETO
11	POMODORO
12	FIUMI
13	ACQUA DI BASSO FONDALE
14	ACQUA DI MEDIO FONDALE
15	ACQUA DI ALTO FONDALE
16	PALUDE CHIUSA/ALLEVAMENTI
17	PALUDE SALMASTRA
18	BACINI D'ACQUA SALATA
19	SPIAGGIA
20	BOSCHI MISTI O DI CONIFERE
21	VEGETAZIONE SOMMERSA ACQUA DOLCE
22	VEGETAZIONE SOMMERSA ACQUA SALMASTRA
23	VEGETAZIONE PALUSTRE ACQUA DOLCE
24	VEGETAZIONE ALOFILA
25	ZONA URBANA
26	ZONA INDUSTRIALE

Tab. 5.3 – Classi individuate dalle dichiarazioni PAC relative al 2005 e dagli altri documenti disponibili

DATI ACQUISITI NEL 2006	
ID	CLASSE
1	MAIS_A
2	SOIA
3	ERBA MEDICA_A
4	GRANO TENERO
5	PISELLO_A
6	BIETOLA
7	GIRASOLE
8	PERETO
9	PESCHETO
10	VIGNETO
11	POMODORO
12	FIUMI
13	ACQUA DI BASSO FONDALE
14	ACQUA DI MEDIO FONDALE
15	ACQUA DI ALTO FONDALE
16	PALUDE CHIUSA/ALLEVAMENTI
17	PALUDE SALMASTRA
18	BACINI D'ACQUA SALATA
19	SPIAGGIA
20	BOSCHI MISTI O DI CONIFERE
21	VEGETAZIONE SOMMERSA ACQUA DOLCE
22	VEGETAZIONE SOMMERSA ACQUA SALMASTRA
23	VEGETAZIONE PALUSTRE ACQUA DOLCE
24	VEGETAZIONE ALOFILA
25	ZONA URBANA
26	ZONA INDUSTRIALE
27	ERBA MEDICA_A
28	GRANO DURO
29	PISELLO_B
30	MAIS_B

Tab. 5.4 – Classi individuate dalle dichiarazioni PAC relative al 2006 e dagli altri documenti disponibili

5.1.2 Determinazioni dei campioni

Per riconoscere e discriminare, in una immagine, le classi definite nella legenda, il classificatore deve essere a tal fine addestrato: in base alla disponibilità delle fonti informative relative all'area studiata, l'operatore individua dei training samples per ogni classe, sui quali vengono calcolate le firme. Dovendo essere rappresentativi, risulta importante che i campioni siano scelti (Brivio et al., 2006):

- ✓ in aree omogenee al loro interno, per meglio rappresentare il prototipo o modello della classe; sono da evitare inoltre le aree di confine fra elementi diversi sul terreno per limitare il problema dei pixel misti;
- ✓ in zone diverse della stessa immagine, in modo che nella costruzione del prototipo della classe si tenga conto della variabilità propria della classe;
- ✓ in numero sufficiente per garantire dei risultati ragionevoli usando tecniche che ricorrono a parametri statistici, come media e covarianza.

Considerando l'oggetto di studio della ricerca, il principale obiettivo è stato quello di estrarre dalle varie fonti il maggior numero di informazioni possibile circa le colture agricole, presenti sul terreno nei diversi anni: procedendo però con classificazioni supervisionate, è risultato necessario prevedere i campioni di tutte le altre coperture costituenti l'area di interesse. Come già descritto nel capitolo precedente, la zona presenta una grande varietà di paesaggi, sia naturali che antropici.

5.1.2.1 Utilizzo della Carta dell'Uso del Suolo di Ravenna

Nonostante il suo carattere relativamente generico, rispetto ai fini ultimi della tesi, la Carta dell'Uso del Suolo è risultata molto utile per la determinazione diretta di tre classi relative alle colture arboree (FRUTTETI, VIGNETI, COLTURE DA LEGNO), e come fonte di consultazione ausiliaria per le rimanenti coperture.

L'estensione mediamente grande e la varietà limitata di questa tipologia di colture (vigneti, pescheti, pereti, pruneti sono le principali), hanno giocato un ruolo importante nella redazione della carta, prima, e nella loro identificazione sul dato ASTER dopo (Figg. 5.1, 5.2 e 5.3).

Dei restanti tematismi, alcuni prevedono più usi del suolo sotto un'unica voce (esempio: "sistemi colturali e particellari complessi") e, quindi, non rispettando il dettaglio richiesto per le colture erbacee, sono stati scartati. Altri, come "verde urbano, aree sportive e ricettive" o "reti stradali e ferroviarie e spazi accessori", a causa della risoluzione minima di 15 m, non sono facilmente individuabili nel color composite. Sono invece servite, nella fase di verifica, le classi "boschi misti o di conifere", "Spiagge, dune e sabbia", "Aree portuali", "Ambiti urbani consolidati". Inoltre, la

suddivisione dei corpi idrici adottata ha suggerito la scelta dei campioni relativi utilizzati nelle classificazioni.



Fig. 5.1 – Particolare della Carta dell'Uso del Suolo: secondo la legenda redatta dal SIT di Ravenna, l'arancione corrisponde alle colture da legno.



Fig. 5.2- Dettaglio dell'immagine QuickBird (Google Earth) in cui si riconosce la colture di legno.

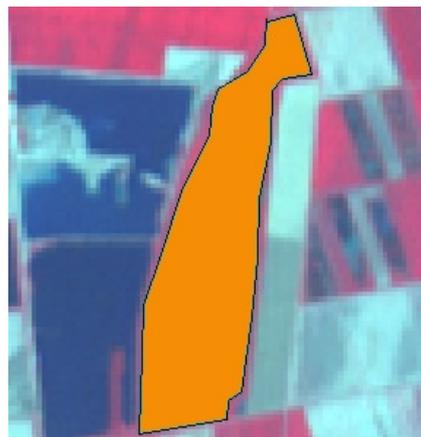


Fig. 5.3 – Dettaglio dell'immagine ASTER (color composite) con riportato il training site della classe COLTURE DI LEGNO

5.1.2.2 Utilizzo della Carta della Vegetazione di Ravenna

Dalla Carta della Vegetazione è stato possibile estrarre i campioni relativi ai diversi ambienti naturali presenti nella zona, come PALUDE SALMASTRA, BACINI D'ACQUA SALATA, SPIAGGIA, BOSCHI MISTI O DI CONIFERE, VEGETAZIONE SOMMERSA ACQUA DOLCE, VEGETAZIONE SOMMERSA ACQUA SALMASTRA, VEGETAZIONE ALOFILA (Fig. 5.4). Si ricorda che alcuni corpi idrici, come la Piallassa Baiona e la Piallassa Piombone, sono caratterizzate da un'alta variabilità del livello dell'acqua che conferisce una certa dinamicità all'estensione delle zone occupate dalla vegetazione: di conseguenza la risposta spettrale di tali aree cambia a seconda del periodo di acquisizione dell'immagine.

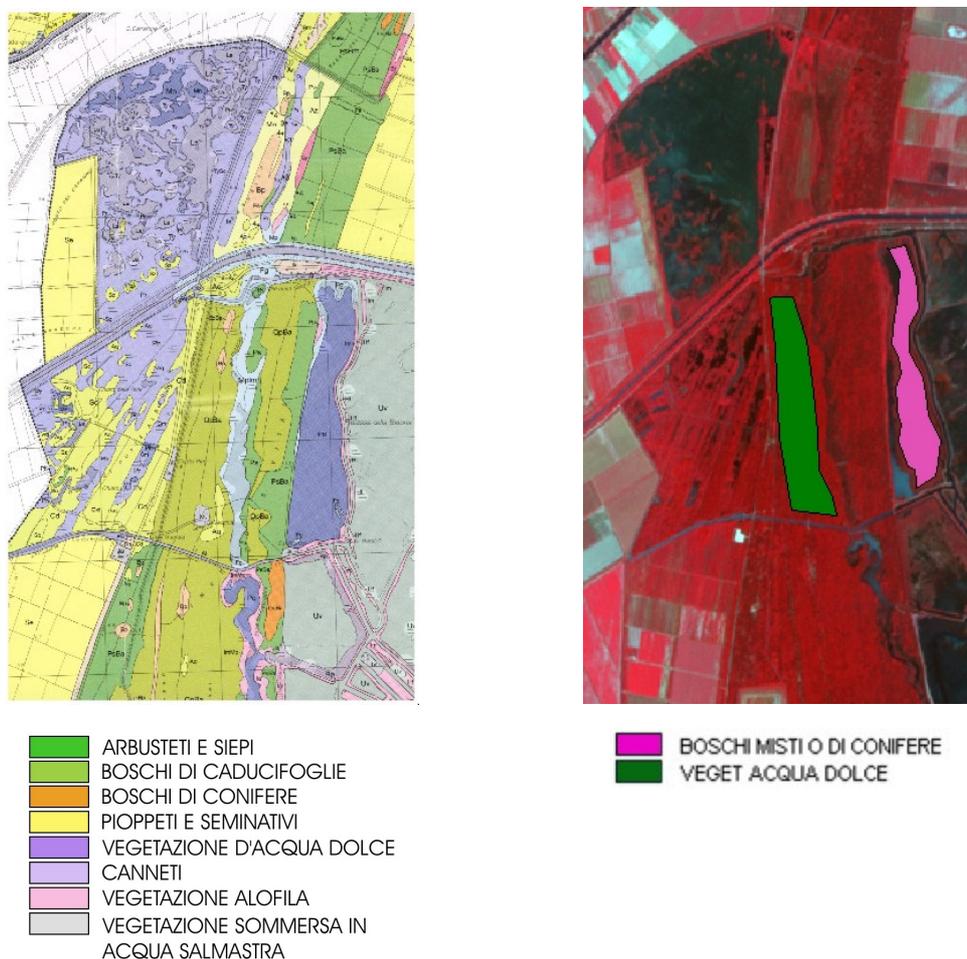


Fig. 5.4 – Esempio di estrazione di training site dalla Carta della Vegetazione

Inoltre, questa fonte è risultata utile anche per individuare i campioni dell'area urbana, comunque già visibile, almeno per i centri più grossi, nel color composite (Fig. 5.5).

I training site prelevati per l'area urbana e l'area industriale, racchiudono al loro interno situazioni altamente disomogenee: essendo però centri grandi e totalmente differenti dal contesto in cui sono inseriti, una loro mancata considerazione tra le aree di addestramento causerebbe dei problemi di confusione agli algoritmi propri della sperimentazione.



Fig. 5.5 – Esempio di estrazione del training site dell'area urbana dalla Carta della Vegetazione

5.1.2.3 Utilizzo delle dichiarazioni PAC e delle mappe catastali

Le dichiarazioni PAC e le mappe catastali rappresentano il contributo più importante per la determinazione dei campioni relativi alle colture erbacee. Di seguito, si riporta la procedura generale seguita.

La figura 5.6 rappresenta l'estratto della dichiarazione PAC del 2006 di un'azienda di Ravenna, nella quale per il rispetto della privacy sono stati cancellati tutti i dati fiscali, situata nella sezione B del comune di Ravenna (S.Alberto): quelle messe in risalto, sono le colture coltivate nella particella 34 del foglio 115, cioè PERA, PISELLO, BIETOLA, GRANO, VITE e PESCA, assieme ai relativi ettari di superficie occupata. Si ricorda che non sono stati presi in considerazione terreni con una superficie inferiore ai 3 ettari, pari a 30000 m²: in tal modo, data la risoluzione massima di 15 m di un pixel, sono stati ammessi, salvo eccezioni strettamente necessarie, solo training site costituiti da un numero di pixel superiore a 100, sufficienti per la determinazione dei parametri statistici nella generazione della firma spettrale. I manuali dei software consigliano un numero di pixel pari a 10 volte il numero di bande usate (Eastman J.R., 1999).

PIANO DI UTILIZZAZIONE DELLE SUPERFICI AZIENDALI																
FR	COMUNE	descrizione	DATI CATASTALI				tipo di coltivazione	SUPERFICIE IN COLTIVAZIONE		COLTIVAZIONE		SUPERFICIE UTILIZZATA		COLTURA UTILIZZO	ONE FRAC.	Stamenti
			sezione	foglio	particella	sub.		ettariare	metri	CODICE	VARIETA' UTILIZZO	ettariare	metri			
1	RA	RAVENNA	B	115	34	000	001	19,16	802	000	5,33	PERA				
2	RA	RAVENNA	B	115	34	000	001	19,16	20	000	2,53	PISELLO				
3	RA	RAVENNA	B	115	34	000	001	19,16	240	000	3,14	BIETOLA				
4	RA	RAVENNA	B	115	34	000	001	19,16	587	000	2,00	GRANO				
5	RA	RAVENNA	B	115	34	000	001	19,16	673	000	2,20	VITE				
6	RA	RAVENNA	B	115	34	000	001	19,16	801	000	3,07	PESCA				
7	RA	RAVENNA	B	115	35	000	001	3,22	801	000	2,83	PESCA				
8	RA	RAVENNA	B	115	35	000	001	3,22	158	000	0,39	USO NON AGRICOLO - TARE ED I				
9	RA	RAVENNA	B	115	36	000	001	0,17	673	000	0,17	SUSINO				
10	RA	RAVENNA	B	115	37	000	001	5,65	3	003	0,3	SUPERFICI RITIRATE DALLA PRO INTENZIONE DI SEMINA DOPO IL				
11	RA	RAVENNA	B	115	37	000	001	5,65	110	000	1,05	BOSCO				
12	RA	RAVENNA	B	115	37	000	001	5,65	158	000	0,36	USO NON AGRICOLO - TARE ED I				
13	RA	RAVENNA	B	115	37	000	001	5,65	240	000	2,05	BARBABIETOLA DA ZUCCHERO				
14	RA	RAVENNA	B	115	37	000	001	5,65	587	000	1,89	GRANO (FRUMENTO) TENERO				
15	RA	RAVENNA	B	115	38	000	001	2,95	802	000	2,89	PERO				

Fig. 5.6 – Estratto di dichiarazione PAC, relativo alle particella del foglio 115 (Sezione B – S.Alberto)

Dalla figura 5.7 risulta evidente che le informazioni relative alla particella 34 non soddisfano le reali esigenze di dettaglio, relativamente alla sua composizione, necessarie in questa fase: spesso per determinare la collocazione delle diverse colture all’interno delle particella (Fig 5.8), si sono rivelati fondamentali la collaborazione con i tecnici dell’azienda e i sopralluoghi nelle aziende agricole.



Fig. 5.7 – Dettaglio della particella 34 (Foglio 115, Sezione B) nella mappa catastale (a sinistra), con la relativa tabella degli attributi, e nel color composite Aster (a destra)

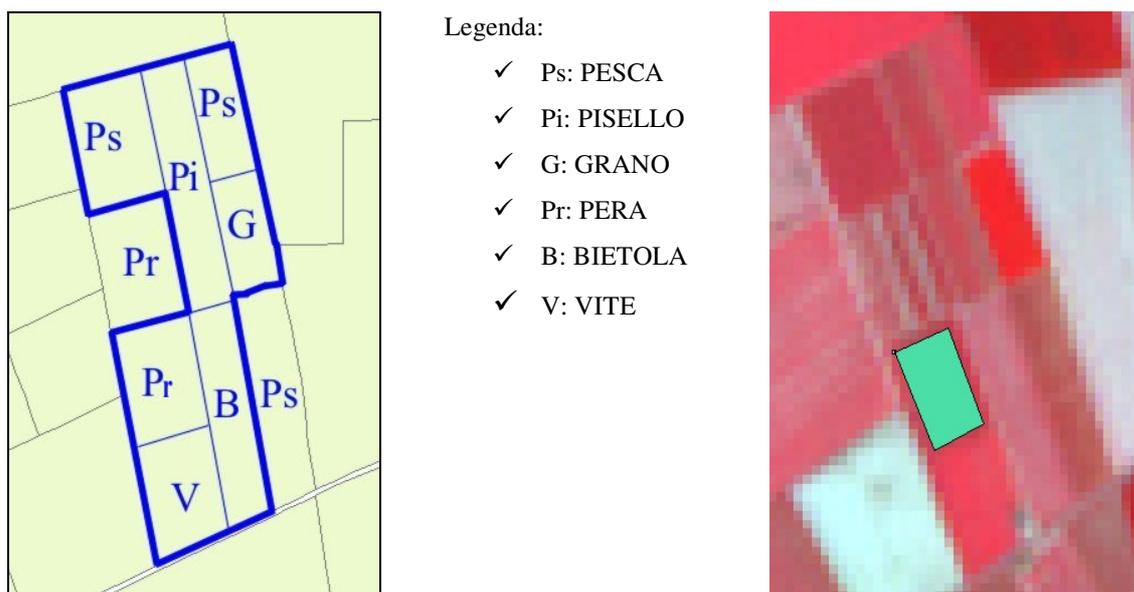


Fig. 5.8 – Suddivisione in terreni della particella 34 con le rispettive colture (a sinistra) e individuazione del training site PERA sul color composite Aster (a destra)

Si sono comunque verificati alcuni casi nei quali i dati presenti nelle dichiarazioni PAC si sono rivelati sufficienti:

- ✓ coincidenza del singolo appezzamento (di superficie estesa) con il mappale (o particella) catastale (caso raro);
- ✓ presenza sul terreno di colture con ciclo completamente sfasato e quindi non tutte contemporaneamente presenti nel momento di acquisizione;
- ✓ confronto con altre dichiarazioni PAC.

5.1.3 Difficoltà della fase di addestramento

Alla luce delle considerazioni evidenziate all’inizio del paragrafo 5.1.2 e delle problematiche descritte nel Capitolo 4, si sottolinea che le difficoltà pratiche incontrate in questa importante fase sono state diverse e di natura differente:

- ✓ Mancata corrispondenza tra la mappa catastale e la suddivisione reale degli appezzamenti. Come già accennato nel paragrafo precedente, le particelle catastali non coincidono con i singoli terreni coltivati; si possono presentare le seguenti due situazioni: più terreni costituiscono un’unica particella oppure uno stesso appezzamento appartiene a più di una particella (Fig. 5.9). Di conseguenza, il contributo delle mappe catastali, relative alle dichiarazioni PAC, spesso è stato limitato alla sola localizzazione della zona all’interno della quale l’azienda possiede e coltiva i propri terreni.



Fig. 5.9 – Esempio di mancata corrispondenza tra le particelle catastali (nero) e gli appezzamenti (verde)

- ✓ Descrizione delle coltivazioni sul sito di AGREA non sufficiente. L'accesso on-line al GIS dell'Agrea non è risultato particolarmente utile ai fini della determinazione dei training site delle colture erbacee: infatti, come visibile in Fig. 5.10, per diversi appezzamenti coltivati viene indicata, oltre alla superficie in m², solo la definizione di SEMINATIVO DA FOTOINTERPRETAZIONE, operazione che forse verrà effettuata in un secondo momento.

Soggetto: RAVENNA-SANT'ALBERTO (RA) [039014B - N177] - Foglio: 96
 X: 1752470 Y: 4931538
 Clicca per selezionare un oggetto

Pz.	N	El.	Ut.	Dettaglio	Area (mq)	Tara	U.A.	Cons.
4	1	17		666-SEMINATIVO DA FOTOINTERPRETAZIONE	29.142	0	NO	
3	1	17		666-SEMINATIVO DA FOTOINTERPRETAZIONE	20.976	0	NO	
7	1	17		666-SEMINATIVO DA FOTOINTERPRETAZIONE	17.506	0	NO	
2	1	17		666-SEMINATIVO DA FOTOINTERPRETAZIONE	12.652	0	NO	

Data Situazione ATTUALE 0,5 km

Fig. 5.10 – GIS Agrea

- ✓ Incompatibilità tra la realtà a terra e l'immagine acquisita. E' capitato che, una volta accertata la fase matura di una coltura nel periodo della ripresa dell'immagine, non ci sia stata corrispondenza tra la denuncia degli agricoltori e gli oggetti rilevati dal sensore: le motivazioni potrebbero essere diverse, come una, calibrazione difettosa dei sottosistemi, o una dichiarazione non totalmente attinente alla realtà, o un problema legato all'azione di agenti atmosferici o alla gestione tecnica dell'azienda, che hanno compromesso il normale ciclo delle piante, ecc. La verifica della natura dell'incongruenza è stata possibile solo in pochi casi e, soprattutto in quelli riferiti a immagini più recenti, quando l'azienda ha permesso di effettuare sopralluoghi nei propri terreni o di consultare uno dei suoi tecnici.



Fig. 5.11 - Esempio di piante di SOIA non rilevate dal sensore

- Un esempio di quanto detto sopra, è riportato in Fig. 5.11: nel color composite la vegetazione compare rossa perché la banda dell'infrarosso, in corrispondenza della quale vi è una maggiore riflessione da parte delle piante, viene inserita nel canale del rosso. E' evidente che sull'appezzamento cerchiato non è presente attività fotosintetica: sul terreno, però, al momento dell'acquisizione (fine giugno) era presente la coltura della soia, già in fase avanzata di maturazione.
- ✓ Colture con più raccolti durante l'anno
Un altro tipo di dubbio circa la modalità con cui estrarre il campione, si è presentato con colture, come l'erba medica, che possono avere più raccolti nello stesso anno ed effettuati in momenti diversi nelle varie aziende: si rischia così di avere una firma spettrale che è una media di diversi stadi fenologici (Fig. 5.12).

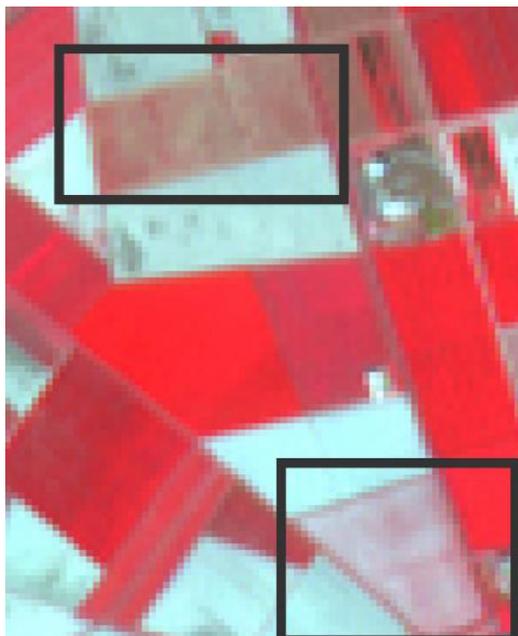
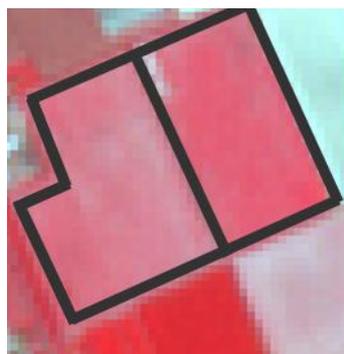


Fig. 5.12 – Esempio di due appezzamenti di ERBA MEDICA coltivati in modo differente

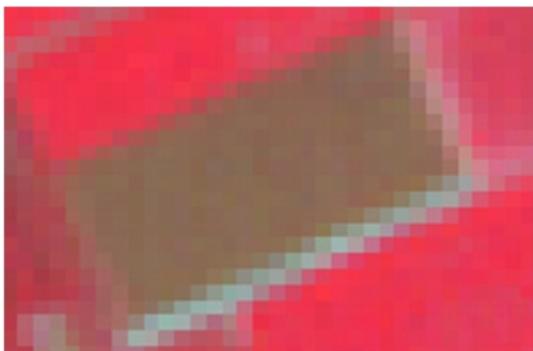
✓ Diverso comportamento di una stessa coltura



*Fig. 5.13 – Comportamento
differente di due terreni coltivati a
MAIS*

Si è verificato che uno stesso tipo di coltura, in alcuni casi appartenente alla medesima azienda e in altri no, abbia avuto comportamenti spettrali differenti, forse influenzati dalla diversa struttura o composizione del terreno: gli appezzamenti riportati nella figura 5.13 sono coltivati entrambi a mais, con le stesse tecniche e la medesima tempistica, come confermato dai tecnici dell'azienda coinvolta, ma già nel color composite appaiono differenti.

✓ Comportamento simile di coperture diverse



*Fig. 5.14 – Esempio di un comportamento spettrale
omogeneo dato da due coperture diverse*

Alcune colture, come il grano, una volta subito il taglio, vengono lasciate a seccare sul terreno, disposte secondo delle strisce, dette andane. A livello di risposta spettrale, questo tipo di copertura può essere facilmente confuso con la vegetazione naturale o artificiale che cresce nei terreni messi a riposo (set-aside). Dalla Fig. 5.14 questo fenomeno risulta già evidente nel color composite: l'appezzamento è ricoperto per metà di andane di grano e per metà è lasciato a set-aside.

5.2 Classificazioni

Sia la classificazione per *pixel* che quella *object-oriented*, hanno prodotto risultati variabili, ciascuno condizionato da diversi fattori, fra i quali il più importante è rappresentato dalla data di acquisizione dell'immagine. Infatti, dal momento della ripresa dipendono sia il numero di colture presenti effettivamente sul terreno, e quindi l'errore legato alla classificazione di appezzamenti lasciati al set-aside (ricoperti di vegetazione spontanea), sia lo stato di maturazione della pianta, e quindi l'errore dovuto a firme simili, come quelle che caratterizzano le colture nei primi giorni di crescita o nella fase terminale di essiccamento al sole della paglia.

Per entrambi i metodi sono state utilizzate le seguenti 5 bande:

	BANDA	INTERVALLO SPETTRALE	SOTTOSISTEMA	RISOLUZIONE
1	1	Verde	VNIR	15 m
2	2	Rosso	VNIR	15 m
3	3N	Infrarosso vicino	VNIR	15 m
4	4	Infrarosso a breve lunghezza d'onda	SWIR	30 m
5	5	Infrarosso a breve lunghezza d'onda	SWIR	30 m

La motivazione di tale scelta è legata all'esistenza di forti livelli di correlazione tra le bande di una stessa immagine multispettrale: ad esempio, alti valori di riflettanza per una data copertura in una certa banda, quasi sicuramente comporteranno valori simili nelle bande più prossime. Nei casi estremi, due bande perfettamente correlate potrebbero descrivere la stessa informazione, che ai fini della classificazione diventa ridondante.

La disponibilità di 14 bande nel dato Aster, è alla base di diverse correlazioni tra le bande appartenenti allo stesso sottosistema e il contemporaneo utilizzo di tutte è stato spesso in grado di compromettere il buon funzionamento del software, problema presentatosi soprattutto con Idrisi (che non ammette comunque più di sette bande). La preferenza di alcune bande rispetto ad altre è avvenuta mediante il metodo *Principal Components Analysis*: la procedura consiste nella

trasformazione dell'insieme delle bande originali in un nuovo insieme (components) di bande non correlate tra loro, ordinate in termini di quantità di variazione di informazione che contengono. Nella matrice di correlazione che viene fornita in output si riconoscono quelle con i valori più alti e di conseguenza è possibile valutare quali bande scartare (De Giglio, 2004).

Si ricorda, inoltre, che i software usati richiedono la co-registrazione delle bande prima dell'applicazione degli algoritmi: quindi le bande 4 e 5 sono state ricampionate da 30 m a 15 m e poi adattate alle immagini VNIR corrispondenti. Come riportato nel paragrafo 4.3.1 (Cap. 4), tale elaborazione influenza negativamente la stabilità del contenuto radiometrico del pixel, aspetto di cui si deve tener conto nella fase di classificazione vera e propria. Alla luce di quanto detto e di prove effettuate in precedenti studi, non sono state considerate le bande termiche, la cui risoluzione di 90 m risulta troppo piccola rispetto a quella di riferimento: le classificazioni comprendenti anche le bande termiche presentano, infatti, disturbi da un punto di vista qualitativo e errori notevoli da quello quantitativo.

5.2.1 Confronto multitemporale su base mensile

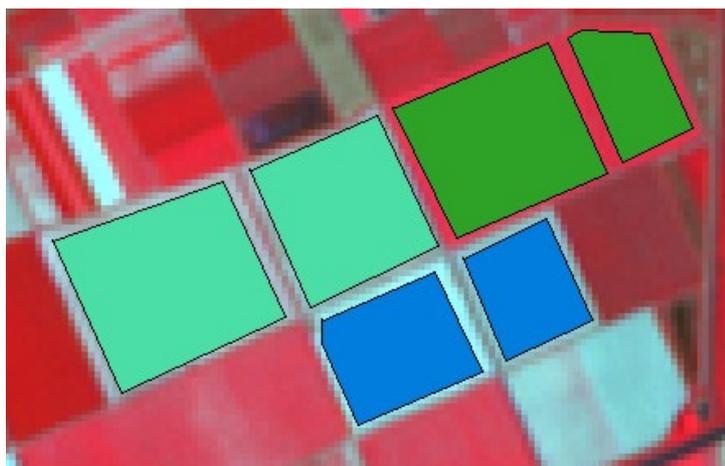
La disponibilità di più di un'immagine per anno, escluso il 2004, consente di verificare se e come avviene il riconoscimento delle colture nelle diverse fasi del ciclo fenologico. A tale scopo, molto utile risulta l'andamento della firma spettrale che, come appare nelle figure successive, viene influenzata diversamente a seconda del momento della crescita della pianta considerato.

Come anticipato nel Capitolo 3, la classificazione per pixel è stata realizzata con l'algoritmo della Massima Verosimiglianza, previsto anche dal software Idrisi. La classificazione ad oggetti invece ha previsto l'uso del Neighbourhood del software eCognition: la scelta dei parametri ($h_{\text{shape}}=0.7$; $h_{\text{compact}}=0.1$; $h_{\text{smooth}}=0.9$) è stata condizionata da numerose prove e da riferimenti bibliografici (in particolare Gusella et al., 2005).

Di seguito si riportano alcuni esempi dei risultati ottenuti con i due metodi, suddivisi per anno.

ANNO 2005

Anche se, cronologicamente, il 2005 si trova a metà del periodo studiato, si è scelto di analizzarlo per primo, perché è l'anno nel quale sono state eseguite più acquisizioni non influenzate negativamente dai vari fattori esterni. In particolare, sono stati analizzati i dati del 26/05/05, del 11/06/05, del 29/07/05 e del 30/08/05. Di seguito si riportano alcuni risultati ottenuti dalla classificazione per pixel.

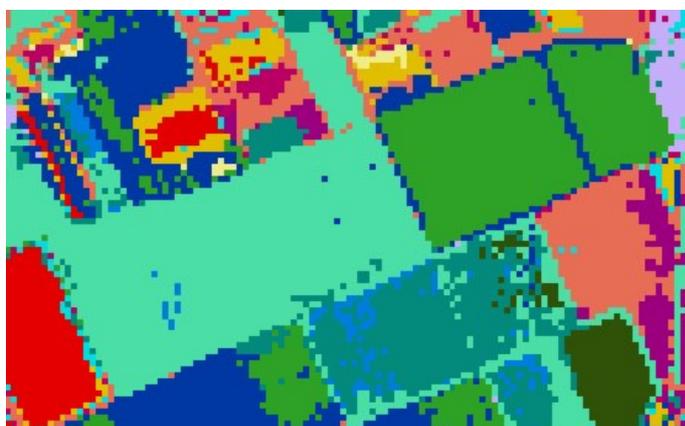


- | | |
|----------------|---------------------|
| MAIS | ACQUA MEDIO F |
| SOIA | ACQUA ALTO F |
| ERBA MEDICA | PALUDE CHIUSA |
| GRANO | PALUDE SALMASTRA |
| PISELLO | BACINO ACQUA SALATA |
| BIETOLA | SPIAGGIA |
| LEGUM. DA SEME | BOSCHI MISTI |
| PERETO | VEG. SOMM. A. DOLCE |
| PESCHETO | VEG. SOMM. A SALM. |
| VIGNETO | VEG. PAL. A. DOLCE |
| POMODORO | VEG. ALOFILA |
| FIUMI | ZONA URBANA |
| ACQUA BASSO F | ZONA INDUSTRIALE |

Fig. 5.15 – Campioni di verità a terra delle colture di mais, soia e bietola

Tra le colture in gioco, l’attenzione è stata concentrata su tre, MAIS, SOIA E BIETOLA, i cui cicli di maturazione hanno una durata sufficientemente lunga da coprire l’intero periodo di acquisizione considerato. In Fig. 5.15 sono riportati i campioni delle verità a terra delle tre colture.

MAIS: dopo la semina che avviene generalmente ad aprile, le piante di mais subiscono tempestivamente la fase di emergenza del terreno e, a fine maggio (Fig. 5.16), si trovano già nel pieno dell’attività fotosintetica, durante la quale accumulano le sostanze nutritive necessarie all’accrescimento. La firma spettrale si avvicina già all’andamento tipico noto in bibliografia, che è



- | | |
|----------------|---------------------|
| MAIS | ACQUA MEDIO F |
| SOIA | ACQUA ALTO F |
| ERBA MEDICA | PALUDE CHIUSA |
| GRANO | PALUDE SALMASTRA |
| PISELLO | BACINO ACQUA SALATA |
| BIETOLA | SPIAGGIA |
| LEGUM. DA SEME | BOSCHI MISTI |
| PERETO | VEG. SOMM. A. DOLCE |
| PESCHETO | VEG. SOMM. A SALM. |
| VIGNETO | VEG. PAL. A. DOLCE |
| POMODORO | VEG. ALOFILA |
| FIUMI | ZONA URBANA |
| ACQUA BASSO F | ZONA INDUSTRIALE |

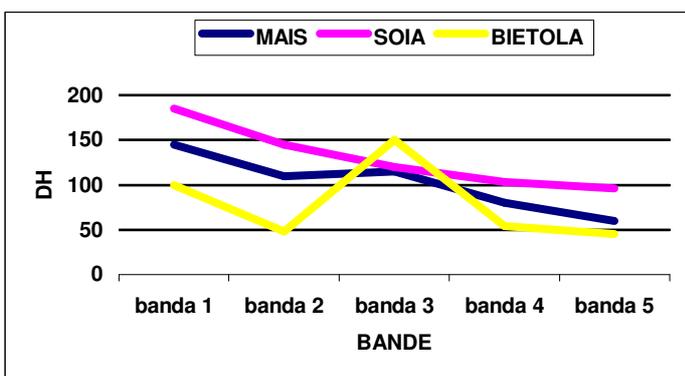


Fig. 5.16 – Dato acquisito il 26/05/05. In alto: risultato della classificazione per pixel (Massima Verosimiglianza). In basso: firme spettrali delle colture MAIS, SOIA E BIETOLA



- | | |
|----------------|---------------------|
| MAIS | ACQUA MEDIO F |
| SOIA | ACQUA ALTO F |
| ERBA MEDICA | PALUDE CHIUSA |
| GRANO | PALUDE SALMASTRA |
| PISELLO | BACINO ACQUA SALATA |
| BIETOLA | SPIAGGIA |
| LEGUM. DA SEME | BOSCHI MISTI |
| PERETO | VEG. SOMM. A. DOLCE |
| PESCHETO | VEG. SOMM. A. SALM. |
| VIGNETO | VEG. PAL. A. DOLCE |
| POMODORO | VEG. ALOFILA |
| FIUMI | ZONA URBANA |
| ACQUA BASSO F | ZONA INDUSTRIALE |

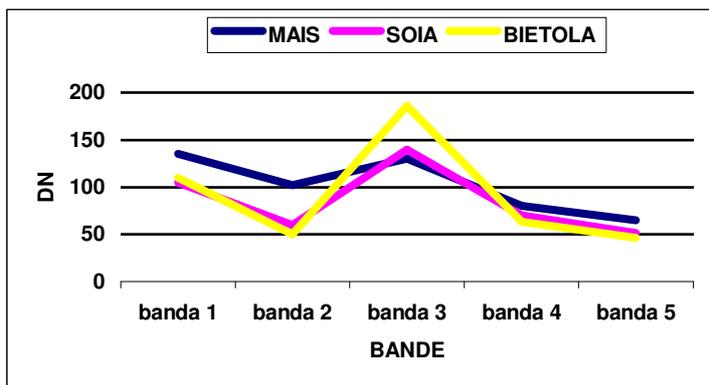


Fig. 5.17 – Dato acquisito il 11/06/05. In alto: risultato della classificazione per pixel (Massima Verosimiglianza). In basso: firme spettrali delle colture MAIS, SOIA E BIETOLA



- | | |
|----------------|---------------------|
| MAIS | ACQUA MEDIO F |
| SOIA | ACQUA ALTO F |
| ERBA MEDICA | PALUDE CHIUSA |
| GRANO | PALUDE SALMASTRA |
| PISELLO | BACINO ACQUA SALATA |
| BIETOLA | SPIAGGIA |
| LEGUM. DA SEME | BOSCHI MISTI |
| PERETO | VEG. SOMM. A. DOLCE |
| PESCHETO | VEG. SOMM. A. SALM. |
| VIGNETO | VEG. PAL. A. DOLCE |
| POMODORO | VEG. ALOFILA |
| FIUMI | ZONA URBANA |
| ACQUA BASSO F | ZONA INDUSTRIALE |

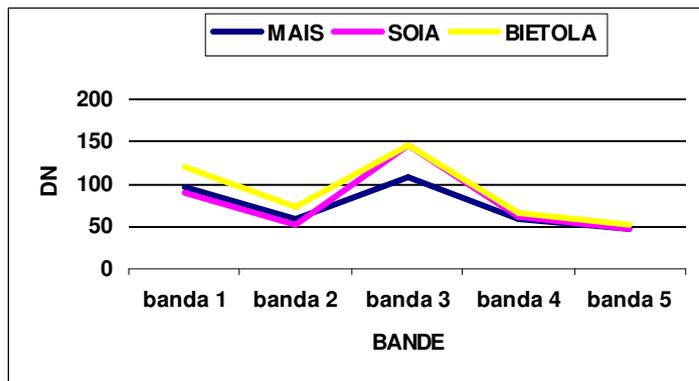


Fig. 5.18 – Dato acquisito il 29/07/05. In alto: risultato della classificazione per pixel (Massima Verosimiglianza). In basso: firme spettrali delle colture MAIS, SOIA E BIETOLA



- | | |
|----------------|---------------------|
| MAIS | ACQUA MEDIO F |
| SOIA | ACQUA ALTO F |
| ERBA MEDICA | PALUDE CHIUSA |
| GRANO | PALUDE SALMASTRA |
| PISELLO | BACINO ACQUA SALATA |
| BIETOLA | SPIAGGIA |
| LEGUM. DA SEME | BOSCHI MISTI |
| PERETO | VEG. SOMM. A. DOLCE |
| PESCHETO | VEG. SOMM. A SALM. |
| VIGNETO | VEG. PAL. A. DOLCE |
| POMODORO | VEG. ALOFILA |
| FIUMI | ZONA URBANA |
| ACQUA BASSO F | ZONA INDUSTRIALE |

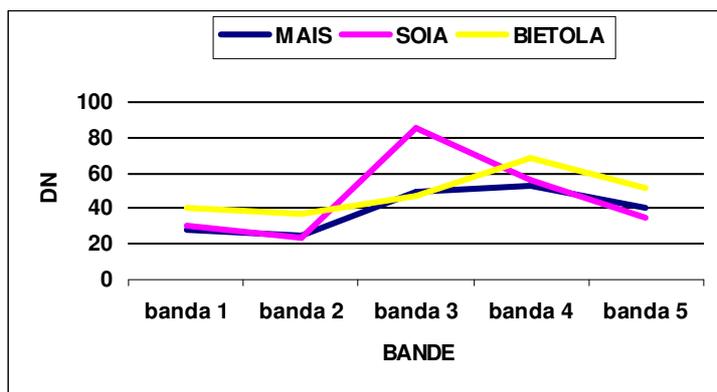


Fig. 5.19 – Dato acquisito il 30/08/05. In alto: risultato della classificazione per pixel (Massima Verosimiglianza). In basso: firme spettrali delle colture MAIS, SOIA E BIETOLA

meglio rappresentato in Fig. 5.17: nel verde (banda 1) la riflettività è legata alla clorofilla e ai pigmenti in genere, che raggiungono il valore massimo a fine luglio (Fig. 5.18), mentre la struttura fogliare è responsabile del comportamento spettrale della piante nell’infrarosso vicino (banda 3). A fine agosto – inizio settembre (Fig. 5.19) livelli più alti si hanno in corrispondenza delle bande 4 e 5, nelle quali l’influenza maggiore è data dal contenuto d’acqua.

SOIA: la soia invece non è presente sul terreno fino alla seconda decade di giugno. Infatti nelle classificazioni di Fig. 5.16 e Fig. 5.17 non è riconosciuta e le firme spettrali individuate non appartengono a questa coltura. Una volta entrata nella fase di emergenza, la pianta cresce rapidamente per raggiungere la maturazione piena a fine agosto (Fig. 5.19), e anche l’andamento della sua firma cambia.

BIETOLA: la bietola, invece, viene riconosciuta nelle prime tre immagini (Fig. 5.16, Fig. 5.17 e Fig. 5.18), con una firma che varia di poco tra maggio e luglio, mentre nel dato di fine agosto viene confusa con la classe ZONA INDUSTRIALE. Molto probabilmente era già provveduto alla raccolta, essendo stato già raggiunto il massimo accumulo di saccarosio nella radice.

ANNO 2002

Le immagini disponibili per il 2002 sono state acquisite il 23/04/02 e il 19/06/02. Gli esempi riportati sotto provengono da classificazioni ad oggetti.



- | | |
|----------------------------|---------------------|
| MAIS | ACQUA ALTO F. |
| BARBABIETOLA | FIUMI |
| GRANO TENERO | ZONA INDUSTRIALE |
| ERBA MEDICA | ZONA URBANA |
| PISELLO I | POMODORO |
| SOIA | VEG. SOMM. A. DOLCE |
| FRUTTETI | VEG. SOMM. A. SALM. |
| VITE | GIRASOLE |
| COLTURE DA LEGNO | SET ASIDE |
| BOSCHI MISTI O DI CONIFERE | PESCA |
| SPIAGGIA | PERA |
| BACINO ACQUA SALATA | ACQUA FOCE |
| PALUDE SALMASTRA | |
| ACQUA BASSO F. | |
| ACQUA MEDIO F. | |

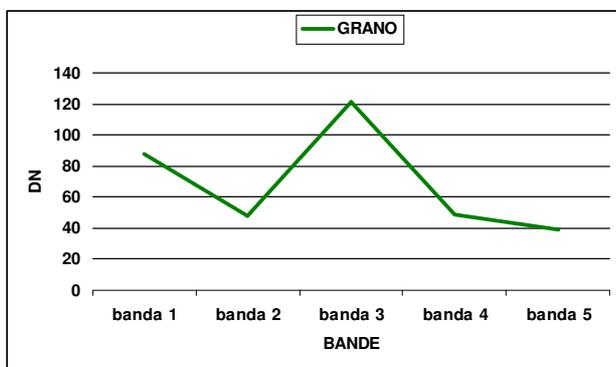
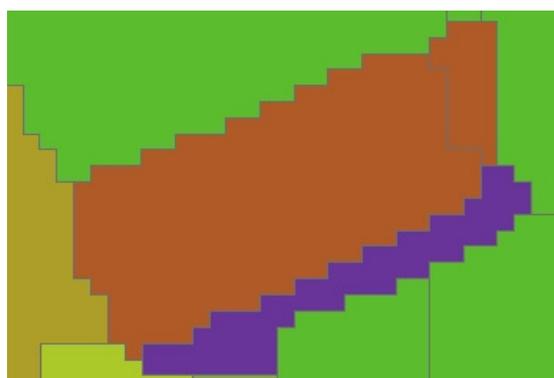


Fig. 5.20 – Dato acquisito il 23/04/02. In alto: risultato della classificazione object-oriented (Nearest Neighbor). In basso: firma spettrale della coltura del GRANO



- | | |
|----------------------------|---------------------|
| MAIS | ACQUA ALTO F. |
| BARBABIETOLA | FIUMI |
| GRANO TENERO | ZONA INDUSTRIALE |
| ERBA MEDICA | ZONA URBANA |
| PISELLO I | POMODORO |
| SOIA | VEG. SOMM. A. DOLCE |
| FRUTTETI | VEG. SOMM. A. SALM. |
| VITE | GIRASOLE |
| COLTURE DA LEGNO | SET ASIDE |
| BOSCHI MISTI O DI CONIFERE | PESCA |
| SPIAGGIA | PERA |
| BACINO ACQUA SALATA | ACQUA FOCE |
| PALUDE SALMASTRA | |
| ACQUA BASSO F. | |
| ACQUA MEDIO F. | |

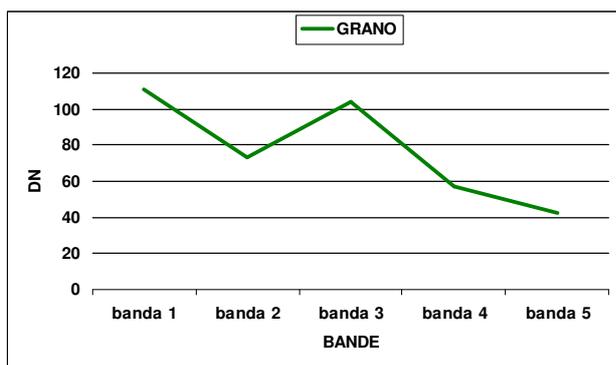


Fig. 5.21 – Dato acquisito il 19/06/02. In alto: risultato della classificazione object-oriented (Nearest Neighbor). In basso: firma spettrale della coltura del GRANO

Appartenendo ai cereali autunno-vernini, le piante di GRANO, nella seconda metà di aprile, hanno già subito la prima fioritura e si trovano a metà della loro altezza totale (70-80 cm): infatti, questa è una delle poche colture di cui è possibile discriminare la firma in questo periodo (Fig. 5.20). A metà giugno invece, una volta prodotti i loro semi, le piante cominciano a terminare il loro ciclo vitale diventando marroni e seccandosi: di conseguenza si nota un abbassamento della riflettività nell'infrarosso e un leggero aumento nella banda del rosso (Fig. 5.21). Appena i semi sono sufficientemente secchi, la raccolta può avere inizio.

ANNO 2006

Le immagini disponibili per il 2006 sono state acquisite il 13/05/06 e il 14/06/06.

Anche per il GIRASOLE è stato possibile studiare l'evoluzione della firma spettrale. Tra le due estratte (Fig. 5.22 e Fig. 5.23), quella significativa è quella relativa al 14/06/06, quando ormai la pianta è entrata nella fase di fioritura. La suddivisione dello stesso appezzamento in numerosi oggetti ottenuta dalla segmentazione è dovuta alla presenza all'interno della piantagione di un sito archeologico, parzialmente coperto da una tettoia, e di zone nelle quali la vegetazione è più rada, come visibile nel color composite di Figura 5.24.

Gli esempi mostrati provengono da classificazioni ad oggetti.



- | | |
|----------------|----------------------------|
| MAIS | PALUDE CHIUSA/ALLEVAMENTI |
| SOIA | PALUDE SALMASTRA |
| ERBA MEDICA | BACINO ACQUA SALATA |
| GRANO TENERO | SPIAGGIA |
| PISELLO I | BOSCHI MISTI O DI CONIFERE |
| BARBABIETOLA | VEG. SOMM. A. DOLCE |
| GIRASOLE | VEG. SOMM. A. SALM. |
| PERA | VEG. PAL. A. DOLCE |
| PESCA | VEG. ALOFILA |
| VITE | ZONA URBANA |
| POMODORO | ZONA INDUSTRIALE |
| FIUMI | ERBA MEDICA II IMPIANTO |
| ACQUA BASSO F. | GRANO DURO |
| ACQUA MEDIO F. | PISELLO II |
| ACQUA ALTO F. | |

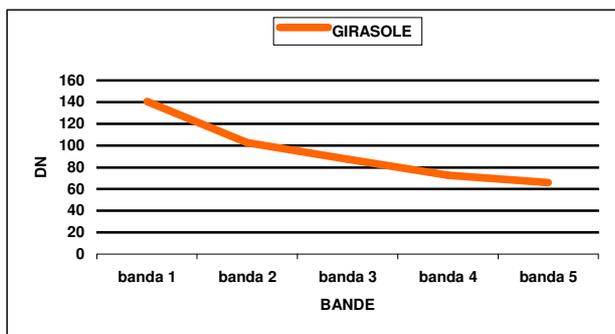


Fig. 5.22 – Dato acquisito il 13/05/06. In alto: risultato della classificazione object-oriented (NN). In basso: firma spettrale della coltura del GIRASOLE

L'area archeologica è stata classificata correttamente come AREA URBANA solo nell'immagine del 13/05/05.

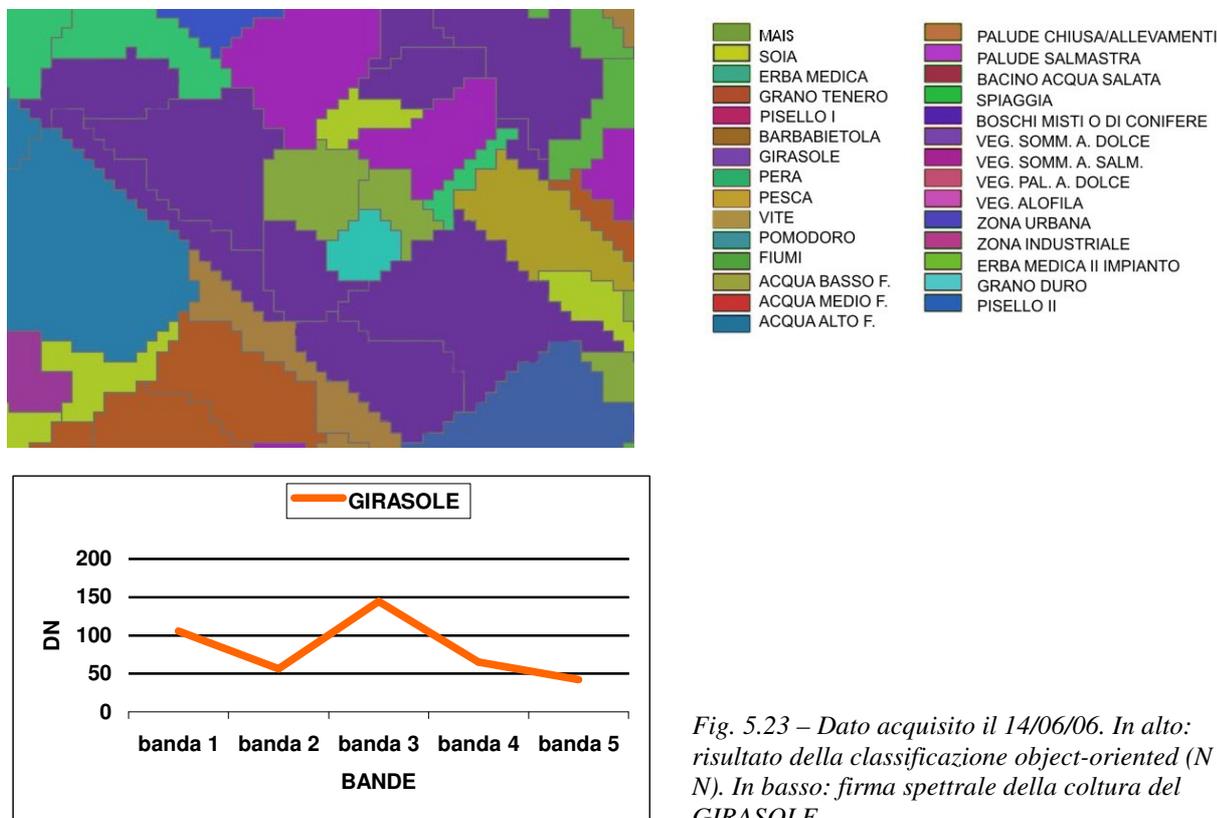


Fig. 5.23 – Dato acquisito il 14/06/06. In alto: risultato della classificazione object-oriented (N N). In basso: firma spettrale della coltura del GIRASOLE.



Fig. 5.24 – Sito archeologico all'interno dell'appezzamento coltivato a GIRASOLE

5.2.2 Confronto multitemporale su base annuale

La disponibilità di più immagini acquisite nello stesso periodo, ma in anni diversi, consente di confrontare la firma spettrale di una stessa coltura anno per anno, per verificare e studiarne eventuali variazioni subite. Di seguito, si riportano i grafici relativi agli andamenti delle firme spettrali della BIETOLA e dell'ERBA MEDICA, estratte dai dati del 19/06/02, del 11/06/05 e del 14/06/06.

Per quanto riguarda la BIETOLA (Fig. 5.25), si può osservare che la differenza più consistente si ha in corrispondenza della banda del verde (banda 1), nella quale la riflettività è legata alla presenza di

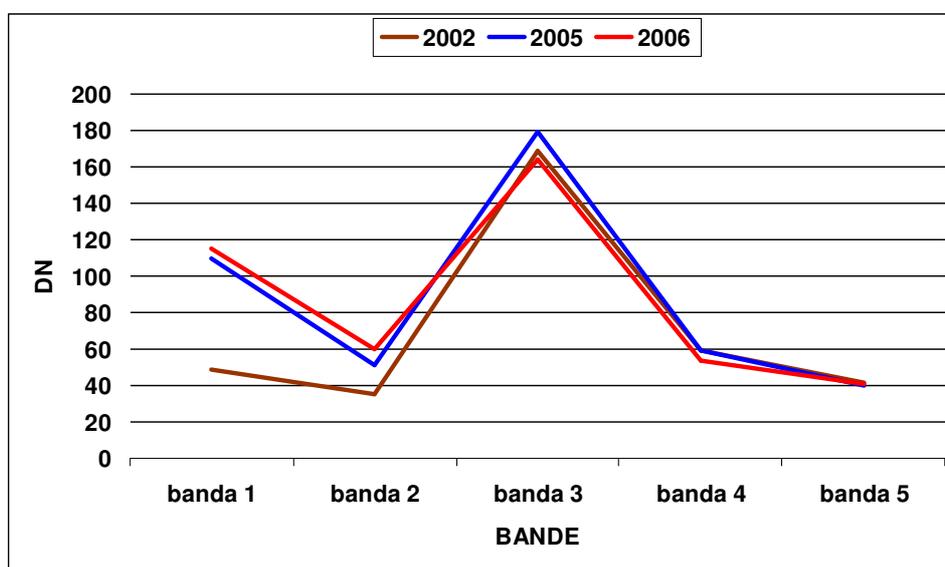


Fig. 5.25 – Firme spettrali della BIETOLA relative ai dati del 19/06/02, 11/06/05 e del 14/06/06

clorofilla, ed è la conseguenza di diverse condizioni di attività fotosintetica. In generale, a metà giugno le piante di bietola, in pieno rigore vegetativo, ricoprono tutto il terreno anche con foglie nuove, che crescono a discapito della riserva zuccherina della radice (Fig. 5.26).



Fig. 5.26 – Coltivazione di BIETOLA: metà giugno 2006 (Azienda Agricola Ca’ Bosco)

La somiglianza tra le firme è, invece, favorita dal fatto che derivano da tre campioni estratti all’interno della stessa azienda agricola, la Ca’ Bosco di Ravenna (Fig. 5.27): questo aspetto implica che, al di là del clima, le altre variabili in gioco sono state pressoché le stesse, in termini di tipo di terreno e di modalità di coltivazione.

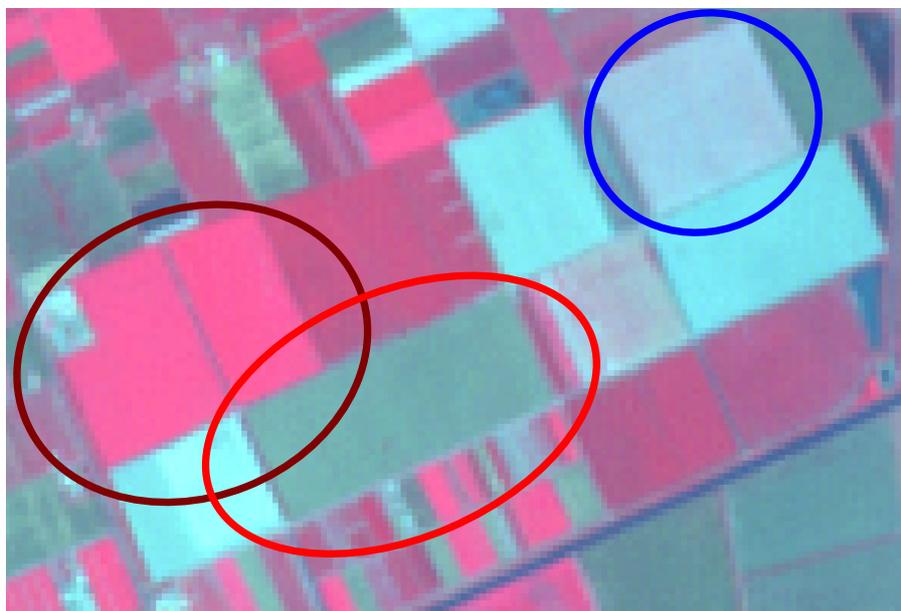


Fig. 5.27 – Appezamenti di terreno di una stessa azienda agricola, coltivati a BIETOLA in anni diversi. I colori utilizzati corrispondono a quelli usati per distinguere i diversi anni nel grafico di Fig. 5.25 (marrone = 2002, blu = 2005 e rosso = 2006).

Un diverso comportamento negli anni si osserva, invece, per la coltura dell’ERBA MEDICA. Le cause della variazione rappresentata nel grafico di Figura 5.28, sono principalmente due: la prima risiede nella differente localizzazione dei training nell’area di interesse, che coinvolge più aziende

agricole caratterizzate da diversi schemi di gestione delle colture. La seconda risente della frequenza con cui avvengono i tre tagli dell'erba medica durante l'anno: in generale sono effettuati nei mesi di maggio, luglio e settembre ma non sono previste delle date precise. La concomitanza di questi due fattori rende, quindi, non univoca la firma dell'erba medica, che risulta così più facilmente confondibile con quella di altre coperture.

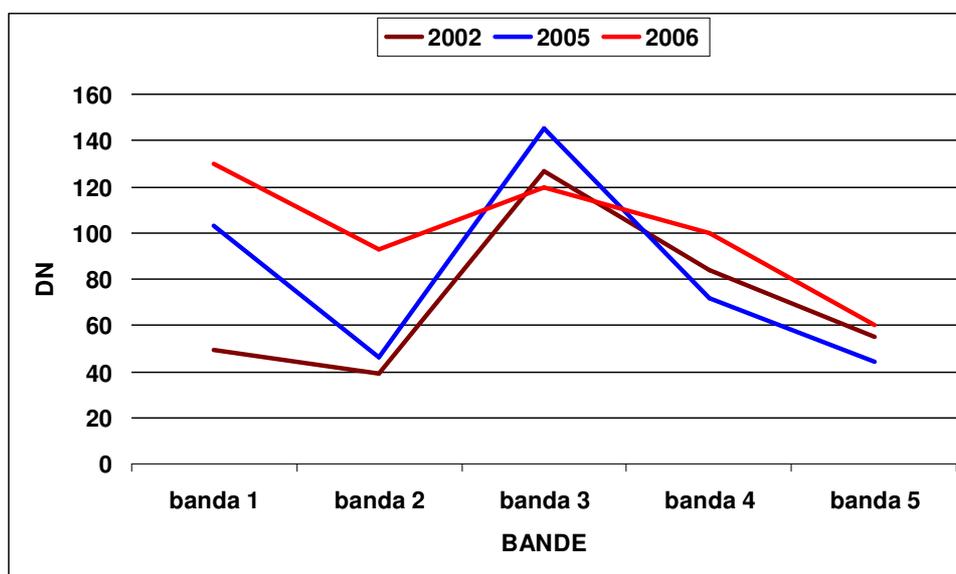


Fig. 5.28 – Firme spettrali dell'ERBA MADICA relative ai dati del 19/06/02, 11/06/05 e del 14/06/06

5.2.3 Confronto qualitativo tra i metodi di classificazione

Per confrontare da un punto di vista qualitativo le due tecniche utilizzate, ovvero la classificazione per pixel e la classificazione object-oriented, si è scelto di illustrare due casi particolari di dubbia interpretazione. Molto preziosa si è rivelata la consulenza fornita dal responsabile dell'Azienda Agricola Ca'Bosco di Ravenna, con il quale è stato possibile ricostruire la storia di alcuni appezzamenti.

Nel dettaglio mostrato in Fig. 5.29, appartenente all'immagine del 26/06/05 e relativo ad una coltivazione di PISELLO, viene messa in evidenza, con un cerchio nero, un'area a diverso comportamento spettrale: causa del fenomeno è una semina spazialmente disomogenea provocata da problemi di calibrazione della macchina seminatrice, utilizzata in momenti successivi.

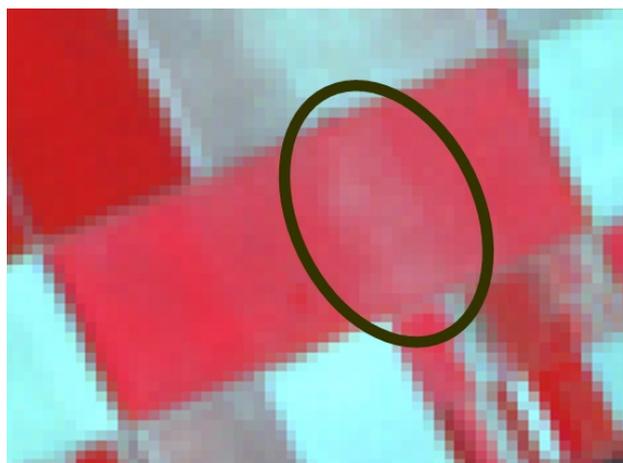


Fig. 5.29 – Coltivazione di PISELLO con problemi di semina spazialmente disomogenea

Con entrambe le metodologie sperimentate, si sono verificati errori di classificazione. Con la classificazione per pixel (Fig. 5.30a), la zona soggetta a diversa modalità di semina non è stata riconosciuta come PISELLO ma come BARBABIETOLA; con la tecnica ad oggetti (Fig. 5.30b), invece, sono stati individuati due errori: come prima, l'errata classificazione del PISELLO come BARBABIETOLA, ma anche il mancato riconoscimento come coltura dell'area a destra della zona studiata, confusa con la classe ZONA INDUSTRIALE. Si ricorda che il range nel quale è definita la firma spettrale della zona industriale è sufficientemente ampio da comprendere anche le altre firme considerate, data la variabilità di materiali che sono racchiusi nel suo campione di addestramento.

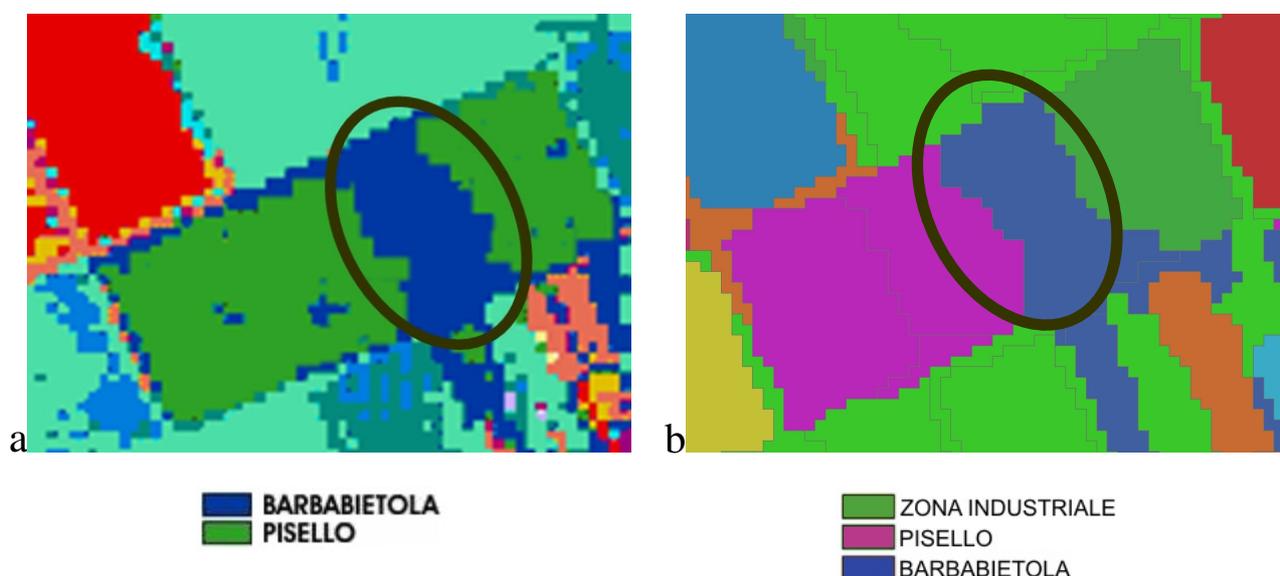


Fig. 5.30 – Confronto tra la classificazione per pixel (a) e la classificazione ad oggetti (b) di un'area dubbia

Nel secondo esempio (Fig. 5.31) invece viene riportato un particolare dell'immagine del 14/06/06, relativo ad una coltivazione di BARBABIETOLA (o BIETOLA), all'interno della quale si può osservare un comportamento anomalo delle piante.

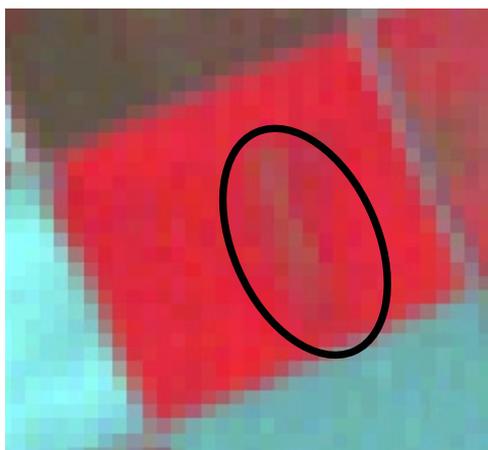


Fig. 5.31 – Anomalia all’interno di una coltivazione di BARBABIETOLA

L’anomalia è discriminata, all’interno dell’appezzamento, sia con la classificazione per pixel, nella quale viene riconosciuta come LEGUMINOSE DA SEME, sia con la classificazione object-oriented secondo cui rientra nella classe POMODORO (vedi Fig. 5.32).

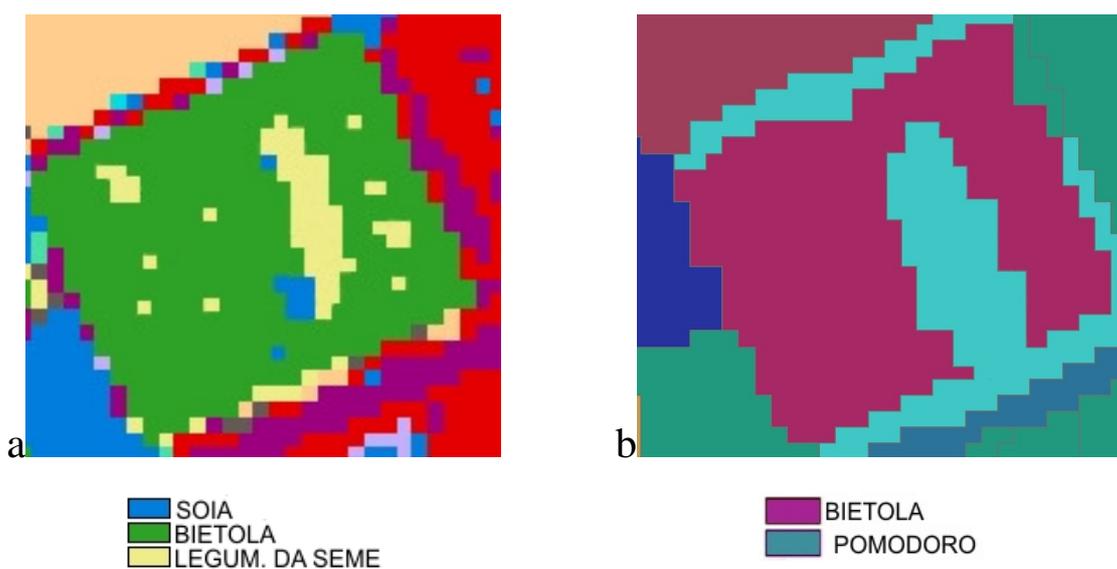


Fig. 5.32 – Confronto tra la classificazione per pixel (a) e la classificazione ad oggetti (b) di un’area dubbia

E’ importante sottolineare come entrambi i classificatori, pur attribuendola erroneamente a determinate coperture, sono in grado di distinguere la risposta spettrale della zona osservata da quella dell’appezzamento circostante. Tale variazione è causata dalla *Cercospora della barbabietola*, una malattia che si manifesta con areole puntiformi, in corrispondenza delle quali i tessuti dapprima scoloriscono e successivamente imbruniscono. Potendo comparire su tutte le parti verdi della pianta, le lesioni col tempo si espandono nelle caratteristiche macchioline, per lo più rotondeggianti di 2-3 mm di diametro, circondate da un alone bruno-rossastro. Confluendo tra loro, le lesioni si possono trasformare in ampie aree disseccate che interessano buona parte del lembo

fogliare. In condizioni climatiche favorevoli si arriva al completo disseccamento delle foglie, come visibile in Figura 5.33: la fotografia è stata scattata nel corso del primo sopralluogo nell'Azienda Agricola Ca' Bosco, avvenuto quasi contemporaneamente all'acquisizione del dato.



Fig. 5.33 – Pianta di BARBABIETOLA colpita da Cercospora

CAPITOLO 6: ACCURATEZZA DELLE CLASSIFICAZIONI

6.1 Introduzione

Da un punto di vista qualitativo è difficile giudicare se le due tecniche di classificazione utilizzate, per pixel e per oggetti, forniscono risultati molto differenti tra loro, in termini di riconoscimento delle colture: in alcuni casi coincidono, soprattutto quando sono favorevoli tutte le condizioni necessarie per la discriminazione delle firme, in altri invece classificano lo stesso appezzamento con due colture diverse o addirittura con tipologie di coperture che appartengono ad ambiti differenti. Inoltre, un confronto di questo tipo diventa soggettivo nel momento in cui si assegnano delle priorità ad alcuni tipi di colture, come ad esempio a quelle in avanzato stato di crescita durante l'acquisizione dell'immagine.

Risulta utile quindi ricorrere ad un confronto oggettivo e quantitativo: gli strumenti per realizzarlo sono dati dalla statistica tramite una serie di parametri, in parte già calcolati dai software, che verranno di seguito descritti.

Lo scopo è quello di comprendere se una delle due tecniche risulta più economica dell'altra, a livello di tempi e costi di realizzazione, o se entrambe possono inserirsi facilmente nella metodologia di lavoro messa a punto in questo elaborato.

6.2 Valutazione dell'accuratezza

La produzione di mappe tematiche provenienti da classificazione possiede una grande varietà di ambiti di applicazioni, sia scientifici che di tipo decisionali, come l'uso del suolo, la mappatura dei tipi di suoli, gli inventari forestali, ecc.: in ogni caso una mappa non risultata utile e affidabile se non è sottoposta ad una rigorosa valutazione di accuratezza. Tale aspetto risulta importante sia per gli utenti finali (*Users*) delle informazioni ricavate dal processo di classificazione, sia per i produttori delle mappe tematiche (*Producers*) che devono valutare la qualità del loro prodotto.

Si definisce *accuratezza* di una classificazione la misura dell'accordo tra la mappa tematica, ottenuta dalla classificazione, e un riferimento che si assume corretto, cioè la "verità a terra".

Quanto più l'immagine classificata si avvicina al riferimento tanto più è accurata (Brivio et al., 2006).

E' importante non confondere l'accuratezza con il concetto di precisione, associato al livello di dettaglio: è più facile infatti realizzare una classificazione in cui si riconoscono categorie più grossolane, come foresta e prato, rispetto ad una che richiede un maggior dettaglio, come la distinzione tra conifere giovani e adulte, o come nel caso in esame tra le diverse colture. Maggiore è la precisione richiesta, maggiori sono le possibilità di errore e quindi di un'accuratezza inferiore.

L'errore di classificazione è l'assegnazione, durante il processo di classificazione, di un pixel ad una categoria o classe diversa da quella cui appartiene nel riferimento (verità a terra): è presente in ogni processo di classificazione. La natura degli errori, la loro frequenza, la grandezza e la sorgente da cui derivano sono i principali tipi di informazione che ricerca di estrarre dall'analisi dell'accuratezza:

- ✓ la natura degli errori riguarda le classi che vengono confuse nella mappa, prendendo in considerazione le classi i cui elementi sono assegnati ad una categoria mentre in realtà appartengono ad un'altra;
- ✓ la frequenza indica quanto spesso si verificano queste situazioni di confusione fra la classificazione e il riferimento;
- ✓ la grandezza è un concetto più sottile, nel senso che alcuni tipi di disaccordo possono essere considerati più gravi di altri. Ad esempio può essere giudicato più grave come errore classificare una zona di acqua come bosco di conifere che confondere un bosco di latifoglie con un bosco di conifere;
- ✓ la sorgente da cui provengono gli errori interessa principalmente agli esperti che generano la classificazione. Nella classificazione automatica di immagini digitali, gli errori derivano da un complesso di fattori interagenti, quali la struttura spaziale del paesaggio (forma, dimensione e disposizione delle parcelle di un certo territorio), la risoluzione del sensore (spaziale, spettrale e radiometrica), gli algoritmi di pre-elaborazione e la procedura di classificazione. La natura discreta stessa dei dati numerici rappresentati dai pixel dà origine ai pixel misti, soprattutto ai bordi fra una classe e l'altra sul terreno, che sono facilmente confusi anche dalle procedure più robuste e affidabili.

La valutazione dell'accuratezza viene realizzata mediante il confronto tra il risultato del processo di classificazione e la mappa di riferimento, costruita sulla base di informazioni più dettagliate e più accurate rispetto a quelle che devono essere valutate, ottenute in genere da osservazioni fatte

direttamente sul terreno durante sopralluoghi o da interpretazioni di carte ufficiali. Per rendere possibile il confronto fra le due mappe, è necessario che esse siano compatibili fra di loro rispetto alla scala, cioè devono essere registrate fra loro, rispetto al livello di dettaglio e infine rispetto al numero e identità di categoria.

6.3 Matrice di confusione

6.3.1 Teoria

Il metodo standard di accumulazione delle informazioni, ottenute dal confronto pixel a pixel della classificazione da valutare rispetto alla mappa di riferimento, prevede la costruzione della *matrice di errore (error matrix)* o *matrice di confusione (confusion matrix)*, sulla quale si basano i test multinomiali. Per studiare gli errori associati a classi singole infatti è necessario utilizzare test di questo tipo che, oltre a rispondere alle esigenze di definizione della frequenza e della natura dell'errore, hanno lo scopo di chiarire quali classi non sono attribuite correttamente.

La matrice di errore (**X**) ha dimensioni $r \times r$, dove r è il numero delle classi in gioco: convenzionalmente le colonne rappresentano i dati del riferimento (osservati), mentre le righe rappresentano i risultati della classificazione (mappati). L'elemento i e j della matrice contiene il numero di pixel attribuiti alla classe i in realtà appartenenti alla classe j nel campione di riferimento; quindi la diagonale rappresenta l'accordo (x_{ii}) tra i risultati della classificazione e le osservazioni a terra che si possono considerare prive di errore (come in un test binomiale) mentre fuori dalla diagonale compaiono le classificazioni errate di osservazioni associate a classi differenti (Fig. 6.1).

La matrice può essere espressa sia in termini assoluti che in termini di frequenze relative, dividendo per n , il numero totale dei pixel, i valori assoluti.

		Reference Class				Total
		A	B	C	D	p_{i+}
Mapped Class	A	35 <i>0.2147</i>	14 <i>0.0859</i>	11 <i>0.0675</i>	1 <i>0.0061</i>	61 <i>0.3742</i>
	B	4 <i>0.0245</i>	11 <i>0.0675</i>	3 <i>0.0184</i>	0 <i>0.0000</i>	18 <i>0.1104</i>
	C	12 <i>0.0736</i>	9 <i>0.0552</i>	38 <i>0.2331</i>	4 <i>0.0245</i>	63 <i>0.3865</i>
	D	2 <i>0.0123</i>	5 <i>0.0307</i>	12 <i>0.0736</i>	2 <i>0.0123</i>	21 <i>0.1288</i>
Total p_{+j}		53 <i>0.3252</i>	39 <i>0.2393</i>	64 <i>0.3926</i>	7 <i>0.0429</i>	163 <i>1.0000</i>

Fig. 6.1 – Esempio di matrice di confusione di dimensione 4x4

In una situazione ideale di perfetta attribuzione del pixel, tutti gli elementi della matrice che si trovano al di fuori della diagonale principale dovrebbero essere nulli, indicando così una situazione di perfetto accordo e che nessun pixel è stato erroneamente classificato.

I totali di riga (marginali di riga) $\sum_{j=1}^r x_{ij} = x_{i+}$ rappresentano il totale dei pixel della classificazione che appartengono alla classe i , mentre i totali di colonna (marginali di colonna) $\sum_{i=1}^r x_{ij} = x_{+j}$ rappresentano i pixel che appartengono alla classe vera j .

Dato n il numero totale di osservazioni (pixel), si determinano le frequenze relative $p_{ij} = \frac{x_{ij}}{n}$, che è la percentuale delle osservazioni relative all'elemento della riga i e colonna j sul totale, $p_{i+} = \frac{x_{i+}}{n}$, ovvero la proporzione di area che nella mappa tematica appartiene alla classe i , e infine $p_{+j} = \frac{x_{+j}}{n}$, ovvero la vera proporzione di area che appartiene alla classe j .

Come strumento di valutazione dell'accuratezza, la matrice di errore si rivela molto efficace, poiché contiene non solo le informazioni sugli elementi correttamente classificati, ma anche sugli errori e come sono distribuiti fra le diverse classi. Nella Tabella 6.1 sono riassunte le principali notazioni della matrice di confusione:

Simbolo	Significato	Tipo di calcolo
X	Matrice di Confusione → righe [1...r] sono i dati classificati (mappati) → colonne [1...r] sono i dati di riferimento (veri)	
<i>r</i>	Numero di righe e colonne di X	
x_{ij}	Numero di osservazioni in corrispondenza della riga <i>i</i> , colonna <i>j</i> . Oggetto della classe riferimento <i>j</i> mappato come <i>i</i> .	osservazione
x_{i+} x_{+j}	Sommatoria degli elementi della riga <i>i</i> (mappati) Sommatoria degli elementi della colonna <i>j</i> (riferimento)	$\sum_{j=1}^r x_{ij}$ $\sum_{i=1}^r x_{ij}$
<i>n</i>	Numero totale di osservazioni oppure oppure	$\sum_{i=1}^r \sum_{j=1}^r x_{ij}$ $\sum_{i=1}^r x_{i+}$ $\sum_{j=1}^r x_{+j}$
P	Matrice delle frequenze relative	X , <i>n</i>
p_{ij}	percentuale delle osservazioni relative all'elemento della riga <i>i</i> e colonna <i>j</i> sul totale	x_{ij}/n
p_{i+}	percentuale delle osservazioni relative alla riga <i>i</i> -esima sul totale	x_{i+}/n
p_{+j}	percentuale delle osservazioni relative alla colonna <i>j</i> -esima sul totale	x_{+j}/n

Tab. 6.1 – Notazioni usate nella matrice di confusione

6.3.2 Misure di accuratezza

Dalla matrice di errore si possono ricavare una serie di *misure di accuratezza*:

- ✓ Accuratezza globale (A_0), definita dal rapporto fra il totale degli accordi (traccia della matrice) e il numero *n* degli elementi dei campioni di test (somma di tutti gli elementi della matrice):

$$\sum_{i=1}^r x_{ii} / n = \sum_{i=1}^r p_{ii}$$

Rappresenta la probabilità che un pixel estratto a caso sia classificato correttamente nell'immagine classificata. In modo simile è possibile determinare l'accuratezza delle singole classi.

- ✓ Accuratezza Utente (*User Accuracy*) per la classe i (C_i), definita dal rapporto fra il totale dei pixel corretti della classe i e il numero totale di pixel che sono stati assegnati a quella classe dalla classificazione (totale di riga):

$$C_i = \frac{x_{ii}}{x_{i+}} = \frac{p_{ii}}{p_{i+}}$$

Indica la probabilità che un pixel scelto a caso tra quelli appartenenti alla classe i nella classificazione, appartenga alla classe i nei dati di riferimento. All'utente, che non conosce la verità, interessa sapere in quale proporzione le aree indicate sulla mappa tematica come appartenenti ad una certa classe, corrispondano "veramente" quella classe.

- ✓ Accuratezza Produttore (*Producer Accuracy*) per la classe j (O_j), definita dal rapporto fra il totale dei pixel corretti della classe j e il numero totale di pixel di quella classe presenti nel riferimento (totale di colonna):

$$O_j = \frac{x_{jj}}{x_{+j}} = \frac{p_{jj}}{p_{+j}}$$

Indica la probabilità che un pixel scelto a caso tra quelli appartenenti alla classi j nel riferimento, appartenga alla classe j nella classificazione, ossia la probabilità che i pixel del riferimento siano classificati correttamente. L'esperto che realizza la classificazione è interessato a sapere in quale proporzione il suo classificatore ha riconosciuto correttamente una certa classe j .

- ✓ Errore di Commissione (*Commission error*) per la classe i (\bar{C}_i), associato all'Accuratezza Utente, indica la probabilità che scelto a caso un pixel assegnato alla classe i nella mappa temetica, appartenga alla classe j ($i \neq j$) nei dati di riferimento:

$$\bar{C}_i = 1 - C_i = 1 - \frac{p_{ii}}{p_{i+}}$$

Si tratta di un errore del 2° tipo, poiché si accetta come vero ciò che è falso.

- ✓ Errore di Omissione (*Omission error*) per la classe j (\bar{O}_j), associato all'Accuratezza Produttore, indica la probabilità che scelto a caso un pixel appartenente alla classe j nei dati di riferimento, sia assegnato alla classe i ($i \neq j$) nella mappa tematica:

$$\bar{O}_j = 1 - O_j = 1 - \frac{p_{jj}}{p_{+j}}$$

Si tratta di un errore del 1° tipo, poiché si rifiuta ciò che è vero.

La terminologia tradizionale riguardante la definizione di errori è a volte confusa ma ben consolidata. Essa viene definita dal punto di vista di chi esegue la classificazione. Osservando attraverso le righe (classi mappate), si ha un errore di commissione quando viene classificata una classe in una posizione dove essa non esiste. Questo porta ad un valore di Accuratezza Utente inferiore. Invece osservando la matrice attraverso una colonna (classi osservate sul campo) si ha un errore di omissione quando il calcolo non riesce a riconoscere la classe che è effettivamente presente. Questo porta ad una diminuzione dell’Affidabilità del Produttore.

In altre parole, gli elementi sulle righe, fuori dalla diagonale, sono gli errori di commissione o inserimento, sono cioè pixel che per verità appartengono alle classi B e C, ma il classificatore erroneamente inserisce come membri della classe A. Analogamente gli elementi sulle colonne, fuori della diagonale, sono gli errori di omissione o esclusione, sono cioè pixel che appartengono per verità alla classe A, ma le vengono impropriamente sottratti non essendo riconosciuti tali dal classificatore che li assegna alle classi B e C.

Simbolo	Significato	Calcolo
C_i	Accuratezza utente per la classe i (mappata)	x_{ii} / x_{i+}
\bar{C}_i	Errore di commissione per la classe i (mappata)	p_{ii} / p_{i+} $1 - C_i$
O_j	Affidabilità del produttore per la classe j (riferimento)	x_{jj} / x_{+j}
\bar{O}_j	Errore di omissione per la classe j (riferimento)	p_{jj} / p_{+j} $1 - O_j$
A_o	Accuratezza globale	$\sum_{i=1}^r x_{ii} / n$
\bar{A}_o	Errore complessivo	$\sum_{i=1}^r p_{ii}$ $1 - A_o$

Tab. 6.2 - Misure di accuratezza derivate dalla matrice di confusione

La Tabella 6.2 mostra le operazioni statistiche che derivano dalla matrice di confusione.

In Figura 6.2 sono invece riportati le accuratezze utente e produttore relative alla matrice di confusione di Figura 6.1. Si supponga di valutare la capacità del classificatore di riconoscere la classe B: l’accuratezza del produttore è pari al 61,11%, mentre quella utente è pari solo al 28,21%. Significa che, anche se il 61,11% delle aree B siano state riconosciute dal classificatore, solo il 28,21% delle aree indicate come tali sulla mappa tematica sono realmente B.

Class	C_i
A	0.5738
B	0.6111
C	0.6032
D	0.0952

Class	O_j
A	0.6604
B	0.2821
C	0.5938
D	0.2857

Overall accuracy: $A_o = 0.5276$

Fig. 6.2 – Accuratezza Utente e Accuratezza Produttore relativi alla matrice di confusione riportata in Fig. 6.1

L'accuratezza complessiva A_o può essere stimata anche considerando tutti gli elementi fuori dalla diagonale che rappresentano appunto degli errori nella operazione di classificazione. Queste matrici possono essere studiate con vari approcci statistici tra i quali:

1. naïve measures of agreement (parametri elementari di associazione);
2. kappa;
3. tau.

In questo lavoro è stato studiato e applicato solo il coefficiente kappa, perché ritenuto il più idoneo al tipo di indagine svolta.

6.4 Coefficiente K

6.4.1 Introduzione

Dalla matrice di errore si possono ricavare altre misure di accuratezza, oltre a quelle già descritte, tra le quali la principale è il *coefficiente di accordo KAPPA* (k) o *KIA*, basato su tecniche discrete multivariate. Ogni singola classe infatti può essere rappresentata da una distribuzione di tipo binomiale: ogni elemento cioè ha solo due possibilità, quella di essere classificato correttamente (probabilità di successo p) o quella di essere classificata erroneamente (probabilità di insuccesso $q=1-p$). Nel loro insieme hanno quindi un distribuzione multinominale.

Introdotta per altri scopi nel 1960, il coefficiente k tende a depurare la stima dell'accuratezza dall'effetto legato all'accordo prodotto dalle coincidenze casuali fra la classificazione e la mappa di riferimento; misura cioè lo scostamento da una classificazione casuale.

Kappa è stato sviluppato da Jacob Cohen e altri (Cohen et al, 1960) all'interno di un contesto relativo a diagnosi psicologiche e psichiatriche; il loro obiettivo era quello di verificare se due classificazioni riguardanti lo stesso gruppo di soggetti risultavano in accordo o meno: in altre parole volevano confrontare le diagnosi provenienti da due medici differenti per valutarne l'eventuale concordanza. In caso contrario, avrebbero potuto sostenere l'errore di una delle due diagnosi o che le categorie classificate non erano realmente distinguibili dalle competenze mediche professionali coinvolte.

K è stato successivamente introdotto nell'ambito del telerilevamento come un utile indice di accuratezza delle classificazioni: in analogia con lo scopo originale, viene utilizzato per confrontare i risultati di due elaborazioni, la prima automatica e la seconda prodotta da un utente esperto. Nel nostro caso sarà utilizzato, come appare nella parte sperimentale della tesi, per confrontare i risultati ottenuti tramite due classificatori automatici diversi o per valutare l'accuratezza della classificazione prodotta dal software con la rappresentazione della realtà a terra creata manualmente dall'operatore.

Il concetto alla base del coefficiente k è che una parte dell'accuratezza della classificazione sia dovuta alla casualità (Congalton, 1983): aspetto rilevante quando diverse classi hanno più probabilità di altre di essere incontrate durante il campionamento sul terreno. Ad esempio, si consideri una mappa nella quale la classe A ricopre il 90% dell'intera area e il restante 10% è occupato dalla classe B. Tale suddivisione può essere usata come stima a priori della copertura, e un generico elemento sul terreno avrebbe il 90% di chance di appartenere alla classe A. La probabilità associata del punto di essere assegnato casualmente (e correttamente) alla classe A è pari a 0.81. Allo stesso modo la probabilità congiunta associata alla classe B è pari a 0.01, così come l'accuratezza complessiva attesa è dell'82% semplicemente per la possibilità di avere assegnazioni favorevoli indipendentemente dal vero pattern mostrato sulla mappa. Questo rende difficile la discriminazione tra un buon risultato ed un risultato fortunato. Questo tipo di distinzione è necessaria se si vogliono confrontare metodi di classificazione diversi attraverso le relative accuratezze. Il coefficiente *Kappa* può essere adottato proprio a questo scopo.

6.4.2 La stima di k

L'utilizzo del coefficiente Kappa si basa sull'assunzione che le frequenze siano note al classificatore prima di assegnare le osservazioni ad una classe particolare. Questo principio deriva dall'utilizzo del coefficiente nelle scienze psichiatriche, dove le percentuali relative delle possibili condizioni che si possono presentare sono note a priori, sulla base di precedenti studi e popolazioni simili. Nelle scienze del rilevamento la presenza relativa delle singole classi può essere definita in base al campionamento a terra, che si può assumere corretto e rappresentativo della situazione rilevata dall'immagine satellitare. Quindi la percentuale delle osservazioni relative all'elemento della riga i (p_{i+}) possono essere ragionevolmente considerate come le probabilità a priori per proporzioni lungo la colonna (p_{+j}) che costituisce il riferimento. Prima della valutazione dell'accuratezza sono perciò definite le proporzioni che si dovrebbero trovare. Ma il coefficiente Kappa assume anche siano note le frequenze, il che non rappresenta certamente il caso classico di un classificatore automatico. La formula di k più comunemente usata prevede che tutti gli errori abbiano lo stesso peso. Sono di seguito riportati i 2 coefficienti intermedi necessari alla costruzione della formula di K:

$$\theta_1 = \frac{1}{n} \sum_{i=1}^r x_{ii} = \sum_{i=1}^r p_{ii}$$
$$\theta_2 = \frac{1}{n^2} \sum_{i=1}^r x_{i+} \cdot x_{+i} = \sum_{i=1}^r p_{i+} \cdot p_{+i}$$

Il primo coefficiente θ_1 rappresenta l'accuratezza globale A_0 . Il secondo coefficiente θ_2 rappresenta l'accuratezza complessiva nel caso in cui esiste la possibilità di avere accordo anche casuale tra le osservazioni di riferimento e quelle ottenute dal calcolo. Per capire meglio il significato dei prodotti $p_{i+} \cdot p_{+i}$ nella stima dell'accordo casuale atteso può essere utile considerare la situazione nel suo contesto spaziale: i totali di riga e i totali di colonna rappresentano rispettivamente i pixel assegnati dal classificatore alle diverse classi nella classificazione e i pixel occupati dalle diverse classi nel riferimento. La probabilità che un generico pixel appartenga ad una classe è data dalla proporzione fra l'area occupata da quella classe rispetto al tutto. Ad esempio siano A la classe nella mappa tematica e A' la classe nel riferimento: la probabilità di coincidenza fra due classi qualsiasi nelle due mappe è la probabilità congiunta, cioè il prodotto delle proporzioni delle aree occupate dalle due classi nelle mappe, $P(A)$ totale di riga e $P(A')$ totale di colonna. La probabilità congiunta di A e A', cioè $P(A \cap A') = P(A) \cdot P(A')$, è il valore atteso delle coincidenze casuali delle due mappe (accordo casuale atteso).

Da questi coefficienti nasce anche la definizione della stima di Kappa (Bishop et al., 1975):

$$\hat{k} = \frac{\theta_1 - \theta_2}{1 - \theta_2} = \frac{n \sum_{i=1}^r x_{ii} - \sum_{i=1}^r x_{i+} x_{+i}}{n^2 - \sum_{i=1}^r x_{i+} x_{+i}} = \frac{\sum_{i=1}^r p_{ii} - \sum_{i=1}^r p_{i+} p_{+i}}{1 - \sum_{i=1}^r p_{i+} p_{+i}}$$

Il Kappa può essere considerato come il rapporto che ha al numeratore la differenza tra l'accordo ottenuto con il classificatore e quello derivante dal puro caso, e al denominatore la probabilità di disaccordo dovuto puramente al caso. Si noti il denominatore corrisponde al massimo valore che il numeratore può assumere.

Se $k=1$ significa che la classificazione ottenuta coincide con la realtà al suolo, nei limiti dei campioni di verifica considerati; se $k=0$ significa che il mancato accordo che si verifica è il medesimo che ci si può aspettare da un'assegnazione casuale dei pixel alle diverse classi; infine, se k è minore di 0, significa che l'accordo complessivo raggiunto dal classificatore è inferiore a quello che si otterrebbe con una classificazione casuale dei pixel ed è indicativo di una classificazione veramente scadente. In ogni caso, il limite inferiore k dipende dalla distribuzione dei valori al di fuori della diagonale.

Per il calcolo della varianza di k , riportato nel paragrafo successivo, è necessario calcolare altri due coefficienti:

$$\theta_3 = \frac{1}{n^2} \sum_{i=1}^r x_{ii} \cdot (x_{i+} + x_{+i}) = \sum_{i=1}^r p_{ii} \cdot (p_{i+} + p_{+i})$$

$$\theta_4 = \frac{1}{n^3} \sum_{i=1}^r \sum_{j=1}^r x_{ij} \cdot (x_{+i} + x_{j+})^2 = \sum_{i=1}^r \sum_{j=1}^r p_{ij} \cdot (p_{+i} + p_{j+})^2$$

Si noti che nella formula per θ_4 la quantità x_{+i} per la classe i è sommata alla quantità x_{j+} per la classe di riferimento j : quindi tali quantità non rappresentano i totali marginali per la cella ma piuttosto quelli della matrice trasposta X^T . La statistica associata al Kappa può risultare di difficile comprensione. L'incremento dell'accuratezza della classificazione con l'aumentare di K non è una relazione diretta come quella dell'accuratezza complessiva. In particolare essa dipende dalle frequenze assolute. Anche se non è semplice dire quale sia un buon valore per il coefficiente Kappa si può affermare che questo deve essere significativamente diverso da zero, il che indicherebbe l'incompetenza del classificare. Questo caso si presenta raramente.

6.4.2.1 La distribuzione della variabile statistica Kappa

La distribuzione di Kappa è stata studiata da più autori, dai quali si è dedotto che conviene considerare, per l'applicazione in oggetto, il comportamento asintotico della variabile.

La distribuzione asintotica di Kappa può essere ritenuta normale $k \approx N(0, \sigma_k^2)$, con una varianza derivante dall'applicazione, per un campione grande, del criterio della Massima Verosimiglianza (Fleiss et al., 1969, Bishop et al., 1975, Hudson et al, 1987, Rossiter, 2004):

$$\sigma^2 \left[\hat{k} \right] = \frac{1}{n} \left[\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2^2)}{(1-\theta_1)^4} \right]$$

Si sottolinea che in letteratura spesso si trovano delle espressioni della varianza con riportati degli errori; la formula qui riportata è quella corretta.

Data quindi la distribuzione asintotica di k, si può utilizzare una variabile normale standardizzata $N(0,1)$, per studiare la significatività di k: $z_k = \frac{k}{s_k}$. Il test consiste nel formulare un'ipotesi (ipotesi nulla, H_0) e nel verificare se con i dati a disposizione è possibile rifiutarla o meno, dopo aver stabilito il livello di significatività α .

Nella sperimentazione, riportate nell'ultimo paragrafo di questo capitolo, è stata scelto α tale da avere una probabilità di rifiuto di 0.05% quando H_0 è vera:

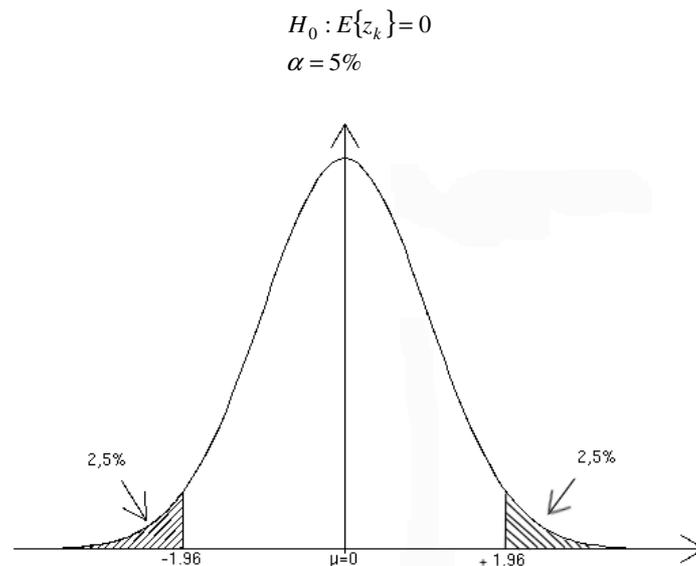


Fig. 6.3 – Test di significatività con $\alpha=5\%$

È quindi immediato calcolare il valore degli estremi dell'intervallo di confidenza, una volta scelto il livello di significatività (probabilità di errore di I tipo): se il valore calcolato per la statistica ricade nella zona di rigetto, l'ipotesi nulla viene rifiutata.

Un esempio numerico dei coefficienti sopra descritti è riportato nella tabella 6.3, nella quale i valori sono stati ricavati dalla matrice di confusione mostrata in Figura 6.1:

Parametro	Valore
θ_1	0.5276
θ_2	0.3054
θ_3	0.3575
θ_4	0.4037
k	0.3199
$s^2[k]$	0.00274
s[k]	0.05234

Tab. 6.3 – Parametri relativi alla matrice di confusione di Figura 6.1

In questo esempio $k \approx 0.32$ è sostanzialmente basso e l'accuratezza globale del 0.53 circa che indica la possibile presenza di accordo casuale.

6.4.2.2 Conditional kappa

Il Kappa può anche essere calcolato per ogni classe sia lungo una riga (Accuratezza Utente) sia lungo una colonna (Affidabilità del produttore). Questo ci consente di determinare quale classe è meglio classificata. Il calcolo segue la stessa logica del Kappa complessivo ma riferendo l'analisi ad una riga o ad una colonna.

$$\hat{k}_{i+} = \frac{(p_{ii} / p_{i+}) - p_{+i}}{1 - p_{+i}} = \frac{C_i - p_{+i}}{1 - p_{+i}} = \frac{p_{ii} - (p_{i+} \cdot p_{+i})}{p_{i+} - (p_{i+} \cdot p_{+i})}$$

$$\hat{k}_{+j} = \frac{(p_{jj} / p_{j+}) - p_{+j}}{1 - p_{j+}} = \frac{O_j - p_{j+}}{1 - p_{j+}} = \frac{p_{jj} - (p_{+j} \cdot p_{j+})}{p_{+j} - (p_{+j} \cdot p_{j+})}$$

Il Conditional Kappa ricade nell'intervallo $[0... p_{+i}]$ (utente) o $[0... p_{j+}]$ (produttore). In generale $\hat{k}_{i+} = C_i$ se $C_i = 1$, il che equivale ad avere un'accuratezza utente assoluta per la classe. Per un certo livello di accordo C_i , invece il valore di \hat{k}_{i+} è inferiore alla misura che la proporzione marginale p_{+i}

della classe di riferimento fornisce. Quindi quanto più probabile diventa la classe di riferimento, maggiormente affidabile dovrebbe essere l'accordo apparente rappresentato dal valore C_i .

La varianza del Conditional Kappa per una classe utente si calcola nel modo seguente

$$\sigma^2[\hat{k}_{i+}] = \frac{1}{n} \left\{ \frac{p_{i+} - p_{ii}}{p_{i+}^3 \cdot (1 - p_{+i})^3} \cdot [(p_{i+} - p_{ii}) \cdot (p_{i+} \cdot p_{+i} - p_{ii}) + p_{ii} \cdot (1 - p_{i+} - p_{+i} + p_{ii})] \right\}$$

La stessa equazione vale per la classe produttore ma invertendo le sommatorie calcolate per righe e colonne, ovvero p_{i+} con p_{+i} .

6.5 Z Test

Come già visto, la verifica della qualità di una classificazione può essere eseguita proprio sulla base della statistica associata alla procedura stessa. Per studiare l'effettivo risultato di una singola classificazione, si può testare l'ipotesi nulla che k valga 0, come visto in precedenza.

Anche il confronto tra due prodotti può essere effettuato confrontando le statistiche associate, consentendo all'utente di scegliere quello che mostra, in termini di qualità dei dati, il migliore risultato. Altra situazione che può presentarsi è quella che comporta la necessità di confrontare due mappe su base quantitativa. Queste possono essere entrambe accurate, se prese singolarmente, ma con errori diversi nei due casi.

Il confronto più semplice tra due prodotti della classificazione prevede la verifica dei valori di accuratezza ovvero la verifica di quale prodotto presenta la più bassa frequenza d'errore sull'intera mappa. L'ipotesi nulla è che i valori attesi delle due statistiche siano uguali:

$$H_0 : E\{k_1\} = E\{k_2\}$$

ovvero considerare la variabile differenza $d = k_1 - k_2$, che è ancora distribuita normalmente (essendo combinazione lineare di variabili casuali normali), a media 0 e varianza σ_d^2 . L'ipotesi nulla per la variabile differenza diviene:

$$H_0 : E\{k_1 - k_2\} = 0$$
$$d \sim N(0, \sigma_d^2)$$

Ipotizzando k_1 e k_2 indipendenti, la varianza della differenza è semplicemente: $\sigma_d^2 = \sigma_1^2 + \sigma_2^2$.

Per cui si può calcolare la seguente variabile normale normalizzata

$$z = \frac{k_1 - k_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad , \quad z \sim N(0,1)$$

ed eseguire il test nel modo consueto: fissato α si individua il limite di confidenza τ e si prende la decisione: se $|z| \leq \tau$ allora l'ipotesi nulla è accettata, altrimenti è rigettata con probabilità di errore di I tipo di α . Per $\alpha = 0.05$, $\tau = 1,96$.

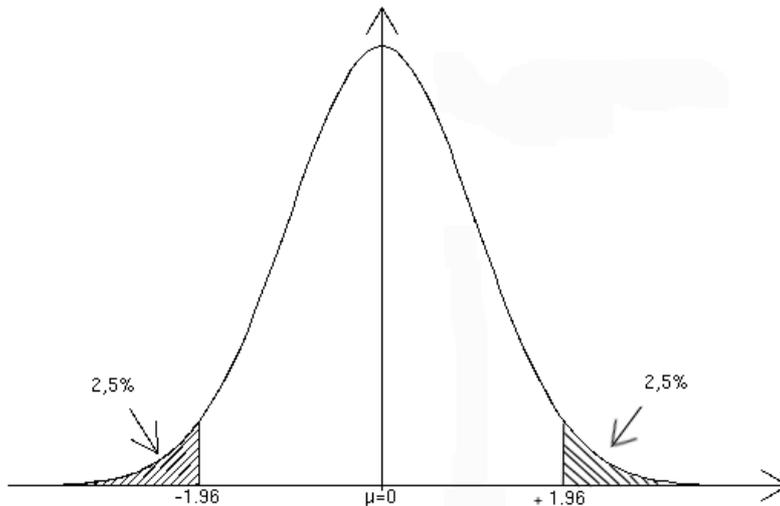


Fig. 6.4 – Test di significatività con $\alpha=5\%$

In realtà k_1 e k_2 sono coefficienti ottenuti da due processi di classificazione differenti che però dipendono dallo stesso set di training site: di conseguenza l'ipotesi di indipendenza dei due coefficienti non è soddisfatta. Il metodo di confronto dei k appena descritto non dovrebbe essere usato per coefficienti k relativi allo stesso set di aree di addestramento (Foody, 2004). Evidentemente, se si conoscesse la correlazione tra i kappa derivante dall'uso degli stessi dati, σ_{12} , potrebbe essere calcolato il valore corretto della varianza della differenza, e applicare il test alla variabile normale normalizzata:

$$z = \frac{k_1 - k_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}} \quad (\text{Donner et al., 2000})$$

La determinazione della covarianza tra i k non è però ampiamente discussa in letteratura, per cui in questo contesto la correzione per la correlazione non è considerata.

6.6 Applicazione ai casi reali

Dato che i software utilizzati, Idrisi e eCogniton, forniscono il calcolo della matrice di confusione, del coefficiente k sia globale che condizionato, ma non i valori di varianza associati, è stato implementato un programma in Fortran che ne permette il calcolo a partire dalla matrice di confusione. Inoltre viene effettuato direttamente il calcolo della variabile z .

Di seguito si riportano i valori delle misure di accuratezza sopradescritte, esclusi i dati relativi alle immagini del 23/04/02 e 19/06/02, per le quali le informazioni a disposizione si sono rivelate insufficienti per la determinazione di campioni di verifica.

6.6.1 Significatività dei singoli k globali

Data quindi la distribuzione asintotica normale di k , $k = N(0, \sigma_k^2)$, è stata calcolata, per ogni classificazione effettuata, la variabile normale standardizzata $z_k \approx N(0,1)$: $z_k = \frac{k}{s_k}$. Considerando che

su ciascuna immagine è stata realizzata sia la classificazione per pixel che quella object-oriented, e che i dati analizzati dal punto di vista statistico sono 7, il totale di variabili normali z trovate è di 14.

Il valore di significatività è pari a $\alpha = 5\%$ ($\tau = 1,96$), mentre l'ipotesi nulla è $H_0: E\{z_k\} = 0$. Se il valore atteso di z_k è nullo, allora anche $k=0$, cioè il risultato della classificazione è puramente casuale. Dai dati di seguito riportati (Tab. 6.4), si può vedere che per ogni classificazione l'ipotesi H_0 è rifiutata, cioè i classificatori forniscono risultati non casuali.

01/07/2004					
k1	0,57065018	σ_1	0,00235977	z1	241,8245
k2	0,53354974	σ_2	0,00237214	z2	224,9234
26/05/2005					
k1	0,81225279	σ_1	0,00285335	z1	284,6664
k2	0,73624002	σ_2	0,0032066	z2	229,6015
11/06/2005					
k1	0,75050352	σ_1	0,00328601	z1	228,3936
k2	0,72819726	σ_2	0,00338834	z2	214,9127
29/07/2005					
k1	0,85644863	σ_1	0,00288721	z1	296,6354
k2	0,77379332	σ_2	0,00343556	z2	225,2306
30/08/2005					
k1	0,75814569	σ_1	0,0032069	z1	236,4108
k2	0,75529302	σ_2	0,00328694	z2	229,7861
13/05/2006					
k1	0,81882398	σ_1	0,0030459	z1	268,8283
k2	0,84441388	σ_2	0,00288022	z2	293,1769
14/06/2006					
k1	0,90023622	σ_1	0,00239075	z1	376,5497
k2	0,83509775	σ_2	0,00293346	z2	284,6801

Tab. 6.4 – Z test applicato al singolo k

Con l'indice 1 si fa riferimento alle classificazioni eseguite con il software eCognition, mentre con l'indice 2 si indicano le elaborazioni effettuate con Idrisi.

Alla luce dei risultati sopra riportati si può affermare che quasi tutte le classificazioni effettuate, considerate singolarmente, risultano accurate, dato i valori dei k non scendono mai al di sotto di 0,73. L'unica eccezione l'immagine del luglio 2004 per la quale i k sono prossimi al valore 0.5.

6.6.2 Z Test tra le coppie di classificazioni

Dall'applicazione dello Z test è emerso che le classificazioni per pixel e quella ad oggetti, relative alla stessa immagine, sono diverse tra loro, infatti l'ipotesi nulla H_0 , $H_0: E\{k_1 - k_2\} = 0$, si può accettare solo in un caso, ovvero per il dato acquisito il 30/08/05 (Tab. 6.5).

01/07/2004					
k1	0,57065018	σ_1	0,00235977	zd	11,0880766
k2	0,53354974	σ_2	0,00237214		
26/05/2005					
k1	0,81225279	σ_1	0,00285335	zd	17,70907019
k2	0,73624002	σ_2	0,0032066		
11/06/2005					
k1	0,75050352	σ_1	0,00328601	zd	4,72587181
k2	0,72819726	σ_2	0,00338834		
29/07/2005					
k1	0,85644863	σ_1	0,00288721	zd	18,41837868
k2	0,77379332	σ_2	0,00343556		
30/08/2005					
k1	0,75814569	σ_1	0,0032069	zd	0,62120015
k2	0,75529302	σ_2	0,00328694		
13/05/2006					
k1	0,81882398	σ_1	0,0030459	zd	-6,10440088
k2	0,84441388	σ_2	0,00288022		
14/06/2006					
k1	0,90023622	σ_1	0,00239075	zd	17,21287793
k2	0,83509775	σ_2	0,00293346		

Tab. 6.5 – Z test applicato alle coppie di k relative alla stessa immagine

Considerando che le classificazioni hanno previsto la discriminazione di circa 27 classi in media, di cui solo una parte, in genere poco meno della metà, sono rappresentative di colture agricole, il k globale tiene conto della errata attribuzione di tutte le classi in gioco. Per meglio valutare come si comportano i due classificatori ai fini del riconoscimento delle sole coltivazioni, si è ritenuto opportuno analizzare le misure di accuratezza a livello di singole tipologie di copertura.

6.6.3 Confronto tra le singole classi

Per valutare quale dei due software, e quindi quale delle due tecniche di classificazione, meglio discrimina le colture presenti realmente sul terreno, si commentano di seguito i singoli risultati di accuratezza produttore (O_i) relativi a ciascuna immagine, che comunque non si discostano molto dai rispettivi k condizionati. Si fa riferimento a O_i perché in questa sede l'interesse è rivolto a sapere in quale proporzione il classificatore riconosce correttamente una certa classe j.

Si sottolinea però che spesso i valori dell'accuratezza, e quindi del coefficiente k condizionato, calcolati con entrambe le metodologie, sono apparsi diversi tra loro per la stessa coltura. Si è stabilito allora un criterio per ritenere riconosciute o meno le classi nella classificazione:

Si → $O_i > 0.70$

Si/No → $O_i > 0.55$

No → $O_i < 0.55$

L'insieme dei valori di tutti i parametri calcolati sono riportati nel paragrafo successivo.

ANNO 2004

Nel dato del 01/07/04 le classi agricole riconosciute da entrambi i classificatori sono il GRANO DURO e il PERETO. Mentre le colture non identificate sono ERBA MEDICA_A, PISELLO, GRANO TENERO. I risultati migliori sono stati ottenuti con la classificazione per pixel che ha individuato correttamente la metà delle colture coinvolte (Tab. 6.6).

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais	Si	Si/No
2	Sorgo	Si/No	Si/No
3	Bietola	Si	No
4	Grano Duro	Si	Si
5	Vigneto	No	Si
6	Erba Medica_A	No	No
7	Soia	Si	No
8	Pomodoro	Si	No
9	Ravanello	Si/No	No
10	Pisello	No	No
11	Grano tenero	No	No
12	Erba Medica_B	Si/No	No
13	Pescheto	Si	No
14	Pereto	Si	Si

*Tab. 6.6 – 07/01/04: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

ANNO 2005

Nell'immagine del 26/05/05 entrambi i classificatori riconoscono correttamente MAIS, BIETOLA, che nel periodo di acquisizione considerato presentano sul terreno delle piante già sviluppate, e la classe PERETO. In questo caso è il classificatore object-oriented a distinguere in modo corretto più colture (Tab. 6.7).

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais	Si	Si
2	Soia	No	Si/No
3	Erba Medica	No	No
4	Grano	Si	Si/No
5	Pisello	Si/No	No
6	Barbabietola	Si	Si
7	Leguminose da seme	Si/No	Si
8	Pereto	Si	Si
9	Pescheto	No	No
10	Vigneto	No	No
11	Pomodoro	Si/No	Si

*Tab. 6.7 – 26/05/05: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

Nell'immagine del 11/06/05 entrambi i classificatori riconoscono correttamente LEGMINOSE DA SEME, VIGNETO e POMODORO. La classificazione per pixel fornisce risultati nettamente migliori rispetto quella ad oggetti, identificando correttamente 7 classi. Le uniche due colture che confonde completamente sono SOIA e ERBA MEDICA che, a metà giugno, possono non essere presenti sul terreno (Tab. 6.8).

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais	Si	Si
2	Soia	No	No
3	Erba Medica	No	No
4	Grano	Si	No
5	Pisello	No	No
6	Barbabietola	Si/No	No
7	Leguminose da seme	Si	Si
8	Pereto	Si	No
9	Pescheto	Si	No
10	Vigneto	Si	Si
11	Pomodoro	Si	Si

*Tab. 6.8 – 11/06/05: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

Nell'immagine del 29/07/05 entrambi i classificatori riconoscono correttamente gran parte delle colture presenti; inoltre le classi riconosciute sono le stesse nelle due classificazioni effettuate (Tab. 6.9). ERBA MEDICA e GRANO, non identificate, potrebbero non esserci fisicamente sul terreno.

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais	Si	Si
2	Soia	Si	Si
3	Erba Medica	No	No
4	Grano	Si/No	No
5	Pisello	Si	Si
6	Barbabietola	Si	Si
7	Leguminose da seme	Si	Si
8	Pereto	No	No
9	Pescheto	No	No
10	Vigneto	Si	Si
11	Pomodoro	Si	Si

*Tab. 6.9 – 29/07/05: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

Nell'immagine del 30/08/05 gran parte delle colture non vengono riconosciute sia dal classificatore ad oggetti sia da quello per pixel: molte di loro potrebbero già essere state raccolte (Tab. 6.10). Questa immagine è l'unica che presenta due classificazioni molto simili, come testimoniato più avanti dal risultato dello Z test.

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais	Si	Si
2	Soia	Si	No
3	Erba Medica	No	No
4	Grano	No	No
5	Pisello	No	No
6	Barbabietola	No	No
7	Leguminose da seme	Si	Si
8	Pereto	No	No
9	Pescheto	Si/No	Si/No
10	Vigneto	No	No
11	Pomodoro	Si	Si

*Tab. 6.10 – 30/08/05: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

ANNO 2006

Nell'immagine del 13/05/06 entrambi i classificatori riconoscono correttamente gran parte delle colture presenti; inoltre le classi riconosciute sono le stesse nelle due classificazioni (Tab. 6.11).

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais	Si/No	Si/No
2	Soia	Si/No	Si/No
3	Erba Medica A	Si	Si
4	Grano Tenero	Si	Si
5	Pisello A	Si	Si
6	Barbabietola	Si	Si
7	Girasole	No	No
8	Pereto	Si	Si
9	Pescheto	No	No
10	Vigneto	No	Si
11	Pomodoro	Si	Si
27	Erba Medica B	Si	Si
28	Grano Duro	No	No
29	Pisello B	No	No

*Tab. 6.11 – 13/05/06: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

Anche per l'immagine del 14/06/06 entrambi i classificatori riconoscono correttamente gran parte delle colture; inoltre le classi riconosciute sono le stesse nelle due classificazioni (Tab. 6.12).

	Coltura	Pixel (Idrisi)	Object-Oriented (eCognition)
1	Mais A	Si	Si
2	Soia	Si	Si/No
3	Erba Medica A	Si/No	Si/No
4	Grano Tenero	Si	Si
5	Pisello	Si	Si
6	Barbabietola	Si	Si
7	Girasole	Si	Si
8	Pereto	No	No
9	Pescheto	No	No
10	Vigneto	Si	Si
11	Pomodoro	Si	Si
27	Erba Medica B	No	No
28	Grano Duro	Si	Si
29	Mais B	Si	Si

*Tab. 6.12 – 14/06/06: Accuratezza Produttore delle singole classi
(Si, $O_i > 0.70$; Si/No, $O_i > 0.55$; No, $O_i < 0.55$)*

6.6.3.1 Considerazioni

I risultati dello Z test applicato alle coppie di classificazione, riferite alla stessa immagine e realizzate con i due algoritmi descritti, indicano la presenza di molte differenze tra le elaborazioni, molto probabilmente legate alle teorie sulle quali i classificatori si basano.

Analizzando però la accuratezze produttore classe per classe di ogni dato, si è potuto constatare che non sempre le colture vengono identificate in modo diverso e spesso la confusione è dovuta principalmente alle caratteristiche della coltura (stato fenologico, contenuto d'acqua, ecc.) che al classificatore in sè. Dei sette confronti, indicati nel paragrafo precedente, due hanno attribuito una maggiore affidabilità alla tecnica per pixel, uno alla tecnica ad oggetti, mentre i rimanenti quattro appaiono indifferenti di fronte a tale scelta.

A prescindere dalle considerazioni fino ad ora fatte, da un punto di vista puramente funzionale e pratico, la classificazione per pixel risulta meno onerosa rispetto a quella ad oggetti perché non prevede la fase di segmentazione, che comunque comporta a volte un adattamento forzato dei training site non facendoli coincidere perfettamente con la realtà che devono rappresentare; questo problema si presenta nel caso in cui si utilizzano le medesime aree di addestramento per entrambi gli algoritmi, come avvenuto in questa sperimentazione.

Di seguito sono riportati gli output forniti dal software implementato per il calcolo delle varianze.

DATO DEL 01-07-04

test 010704 - 30 CLASSI
 NUMERO DI CLUSTER = 30
 NUMERO DI PIXEL = 47791
 FILE DATI : C:\MICHAELA\TESI_DOTTORATO_pc\PROGRAMMA_PROF_nuovo\010704_rifare_\accuracy_010704_eC.txt

SUM_ROW(I)
 958. 740. 224. 581. 371. 132. 296. 390. 0. 0. 12. 121. 405. 69. 4512. 3310. 2181. 5845. 4576. 1197. 911.
 2894. 961. 1241. 172. 4749. 2012. 5036. 802.
 2083.
 SUM_COL(J)
 1597. 490. 224. 315. 100. 191. 0. 0. 205. 252. 347. 360. 0. 69. 1877. 5075. 2300. 5845. 1578. 3599. 2441.
 3049. 1732. 1216. 587. 4673. 1534. 5514. 697.
 914.

USER'S ACCURACY E KAPPA PER RIGA			PRODUCER'S RELIABILITY E KAPPA PER COLONNA		
USER'S ACC	K_ROW	sqm	PRODUCER'S REL	K_COL	sqm
0.892484	0.888767	0.010331	0.535379	0.525875	0.012536
0.495946	0.490724	0.018425	0.748980	0.745032	0.019845
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.542169	0.539131	0.020691	1.000000	1.000000	0.000000
0.000000	-0.002097	0.000210	0.000000	-0.007824	0.000408
0.000000	-0.004013	0.000291	0.000000	-0.002770	0.000241
0.000000	0.000000	0.000000	-NaN	-NaN	-NaN
0.000000	0.000000	0.000000	-NaN	-NaN	-NaN
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.000000	-0.007314	0.000394	0.000000	-0.000251	0.000073
0.000000	-0.007590	0.000402	0.000000	-0.002538	0.000231
0.000000	0.000000	0.000000	-NaN	-NaN	-NaN
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.353945	0.327534	0.006951	0.850826	0.835274	0.009025
0.965861	0.961805	0.003522	0.629951	0.602414	0.006985
0.968822	0.967245	0.003904	0.918696	0.914808	0.005947
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.220717	0.194107	0.005870	0.640051	0.601936	0.013199
0.999165	0.999097	0.000903	0.332315	0.315162	0.007643
0.783754	0.772115	0.014313	0.292503	0.278755	0.009046
0.719765	0.700668	0.008762	0.683175	0.662753	0.008782
0.986472	0.985964	0.003865	0.547344	0.538055	0.012003
0.886382	0.883416	0.009216	0.904605	0.902062	0.008627
0.970930	0.970569	0.012968	0.284497	0.281913	0.018525
0.000000	-0.108377	0.001669	0.000000	-0.110334	0.001687
0.692843	0.682657	0.010486	0.908735	0.904724	0.007653
0.780183	0.751513	0.006435	0.712550	0.678692	0.006575
0.756858	0.753259	0.015309	0.870875	0.868671	0.012894
0.418627	0.407291	0.010738	0.954048	0.951954	0.007235

STAMPE DI CONTROLLO DEL CALCOLO

Accuratezza delle classificazioni

NUMERO DI CLUSTER = 30
 NUMERO DI PIXEL = 47791
 FILE DATI : C:\MICHAELA\TESI_DOTTORATO_pc\PROGRAMMA_PROF_nuovo\010704_rifare_\accuracy_010704_I.txt

```
SUM_ROW(I)
  890.  129.  1202.  1204.   14.    1.  345.  300.  136.   19.    3.  273.  276.   72.  7807.   715.  1714.  6091.  1346.   122.  4818.
3263. 1149. 1329.   537.  4318. 2059. 5024.  867.
 1768.
SUM_COL(J)
  368.  490.  1226.   790.   100.  191.  333.  289.  205.  252.  347.  360.   85.   69.  1877.  5075.  2300.  5845.  1578.  3599.  2441.
3049. 1732. 1216.   587.  4673. 1534. 5569.  697.
  914.
```

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.407865	0.403270	0.016417	0.986413	0.986155	0.006149
0.635659	0.631885	0.042768	0.167347	0.165093	0.016765
1.000000	1.000000	0.000000	0.980424	0.979919	0.004057
0.630399	0.624186	0.014012	0.960759	0.959745	0.007082
0.000000	-0.002097	0.000210	0.000000	-0.000293	0.000078
0.000000	-0.004013	0.000291	0.000000	-0.000021	0.000021
0.915942	0.915352	0.015034	0.948949	0.948578	0.012145
0.850000	0.849087	0.020721	0.882353	0.881610	0.019058
0.955882	0.955692	0.017683	0.634146	0.633102	0.033684
0.000000	-0.005301	0.000335	0.000000	-0.000398	0.000091
0.000000	-0.007314	0.000394	0.000000	-0.000063	0.000036
1.000000	1.000000	0.000000	0.758333	0.756945	0.022651
0.307971	0.306738	0.027726	1.000000	1.000000	0.000000
0.958333	0.958273	0.023582	1.000000	1.000000	0.000000
0.222493	0.190708	0.004268	0.925413	0.910850	0.007225
0.703497	0.668270	0.019022	0.099113	0.085430	0.003857
0.948658	0.946062	0.005590	0.706957	0.696056	0.009702
0.947956	0.940704	0.003220	0.987853	0.986079	0.001639
0.098811	0.068039	0.008225	0.084284	0.057746	0.007007
1.000000	1.000000	0.000000	0.033898	0.031426	0.002804
0.226858	0.185243	0.005877	0.447767	0.385853	0.010877
0.826234	0.814392	0.006998	0.884224	0.875740	0.006171
0.991297	0.990970	0.002843	0.657621	0.649187	0.011536
0.829195	0.824735	0.010542	0.906250	0.903568	0.008577
0.474860	0.468330	0.021688	0.434412	0.427985	0.020547
0.000000	-0.108377	0.001669	0.000000	-0.099326	0.001585
0.665372	0.654274	0.010586	0.893090	0.888277	0.008216
0.868033	0.850627	0.005325	0.783085	0.757603	0.006007
0.680507	0.675779	0.015976	0.846485	0.843648	0.013875
0.484729	0.474682	0.011884	0.937637	0.935241	0.008296

DATO DEL 26-05-05

TEST 260505 - 26 CLASSI

NUMERO DI CLUSTER = 26
NUMERO DI PIXEL = 20010
FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_260505_eC.txt

SUM_ROW(I)
1527. 406. 0. 465. 454. 1278. 83. 365. 357. 441. 413. 498. 36. 258. 745. 1305. 1132. 888. 305. 4363. 355.
591. 207. 168. 2413. 957.
SUM_COL(J)
1820. 614. 483. 828. 797. 807. 83. 32. 147. 42. 163. 501. 78. 258. 745. 1305. 1103. 1133. 310. 4362. 352.
407. 137. 328. 1738. 1437.

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.979044	0.976947	0.004026	0.821429	0.806676	0.009555
0.884236	0.880572	0.016342	0.584691	0.576089	0.020038
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
1.000000	1.000000	0.000000	0.561594	0.551164	0.017330
0.407489	0.382910	0.023696	0.232120	0.214294	0.014839
0.625978	0.610260	0.013760	0.991326	0.990734	0.003486
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.087671	0.086210	0.014580	1.000000	1.000000	0.000000
0.000000	-0.007401	0.000613	0.000000	-0.018165	0.000970
0.015873	0.013803	0.005842	0.166667	0.147887	0.058703
0.394673	0.389702	0.023945	1.000000	1.000000	0.000000
0.931727	0.929974	0.011572	0.926148	0.924263	0.011960
1.000000	1.000000	0.000000	0.461538	0.460568	0.056429
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.974382	0.972887	0.004962	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.783760	0.773718	0.012638
1.000000	1.000000	0.000000	0.983871	0.983621	0.007264
0.999771	0.999707	0.000293	1.000000	1.000000	0.000000
0.991549	0.991398	0.004945	1.000000	1.000000	0.000000
0.563452	0.554388	0.020554	0.818182	0.812648	0.019626
0.642512	0.640048	0.033416	0.970803	0.970498	0.014531
0.904762	0.903175	0.023006	0.463415	0.458871	0.027522
0.701202	0.672781	0.009820	0.973533	0.969903	0.004368
0.902821	0.895303	0.010266	0.601253	0.581224	0.013168

Accuratezza delle classificazioni

STAMPE DI CONTROLLO DEL CALCOLO

TETA= 0.828636 0.087260 0.163365 0.052158
 VAL_K= 0.812252794
 UMT1,UMT2,FA 0.171364318 0.912739642 0.205696343
 VA1,VA2,VA3,VAR_k 0.000008518 -0.000000422 0.000000046 0.000008142
 MATRICE DI CONFUSIONE

1495.	5.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	27.
0.	359.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	47.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	465.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
262.	0.	0.	0.	185.	7.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
58.	0.	0.	0.	420.	800.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	83.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1.	0.	191.	6.	0.	0.	0.	32.	87.	35.	0.	13.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	357.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	292.	0.	0.	0.	0.	0.	7.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	142.	0.	0.
0.	250.	0.	0.	0.	0.	0.	0.	0.	0.	163.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	464.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	34.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	36.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	258.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	745.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1305.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	29.	0.	0.	0.	1103.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	888.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	305.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	4362.	0.	0.	1.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	352.	0.	3.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	13.	0.	0.	0.	0.	245.	0.	0.	0.	333.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	74.	133.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	16.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	152.	0.
0.	0.	0.	0.	192.	0.	0.	0.	30.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1692.
4.	0.	0.	0.	0.	0.	0.	0.	30.	0.	0.	8.	0.	0.	0.	0.	0.	0.	5.	0.	0.	0.	0.	0.	46.

VAL_K,VAR_K,SQMK1= 0.8123 0.0000 0.0029

NUMERO DI CLUSTER = 26
 NUMERO DI PIXEL = 20010
 FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_260505_I.txt

SUM_ROW(I)
 1746. 62. 39. 825. 489. 1205. 64. 669. 156. 61. 200. 673. 74. 473. 530. 1101. 0. 2067. 308. 4182. 326.
 127. 611. 154. 2026. 1842.
 SUM_COL(J)
 1820. 614. 483. 828. 797. 807. 83. 32. 147. 42. 163. 501. 78. 258. 745. 1305. 1103. 1133. 310. 4362. 352.
 407. 137. 328. 1738. 1437.

Accuratezza delle classificazioni

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.945590	0.940146	0.005942	0.907143	0.898266	0.007388
0.838710	0.833604	0.048165	0.084691	0.081846	0.010955
0.102564	0.080366	0.049704	0.008282	0.006345	0.004043
0.953939	0.951951	0.007598	0.950483	0.948354	0.007847
0.856851	0.850913	0.016436	0.525721	0.513841	0.017784
0.667220	0.653235	0.013855	0.996283	0.996044	0.002279
0.828125	0.827409	0.047329	0.638554	0.637394	0.052823
0.044843	0.043313	0.007758	0.937500	0.935338	0.044266
0.294872	0.289653	0.036577	0.312925	0.307527	0.038350
0.327869	0.326455	0.060108	0.476190	0.474589	0.077215
0.815000	0.813481	0.027631	1.000000	1.000000	0.000000
0.739970	0.733293	0.017189	0.994012	0.993804	0.003566
1.000000	1.000000	0.000000	0.948718	0.948528	0.025063
0.545455	0.539517	0.022943	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.711409	0.703558	0.016867
0.999092	0.999028	0.000971	0.842912	0.833765	0.010547
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.534591	0.506657	0.011054	0.975287	0.972440	0.005136
1.000000	1.000000	0.000000	0.993548	0.993448	0.004618
0.999522	0.999388	0.000432	0.958276	0.947252	0.003788
0.990798	0.990633	0.005382	0.917614	0.916249	0.014876
0.000000	-0.020762	0.001040	0.000000	-0.006387	0.000569
0.216039	0.210635	0.016361	0.963504	0.962354	0.016521
0.740260	0.735931	0.035851	0.347561	0.342501	0.026213
0.766535	0.744328	0.010038	0.893556	0.881564	0.008151
0.535831	0.499918	0.011969	0.686848	0.655098	0.013160

STAMPE DI CONTROLLO DEL CALCOLO

TETA= 0.759170 0.086937 0.156063 0.050251
VAL_K= 0.736240023
UMT1,UMT2,FA 0.240829585 0.913063413 0.288873668
VA1,VA2,VA3,VAR_k 0.000010960 -0.000000761 0.000000083 0.000010282

Accuratezza delle classificazioni

1651.	61.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	12.	8.	14.	
10.	52.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	4.	32.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	3.	0.	0.	0.	0.	0.	
0.	0.	0.	787.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	38.	0.	0.	0.	0.	0.	
27.	8.	2.	0.	419.	3.	30.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
34.	0.	0.	0.	366.	804.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	
0.	0.	11.	0.	0.	0.	53.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
85.	0.	355.	0.	0.	0.	0.	30.	86.	22.	0.	0.	0.	0.	0.	0.	0.	0.	0.	64.	0.	0.	0.	27.	0.	
0.	0.	90.	6.	0.	0.	0.	0.	46.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	14.	0.	0.	
0.	0.	21.	0.	0.	0.	0.	2.	15.	20.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	2.	1.	0.	
0.	37.	0.	0.	0.	0.	0.	0.	0.	0.	163.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	498.	0.	0.	0.	0.	143.	28.	0.	0.	0.	4.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	74.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	258.	215.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	530.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1100.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	960.	1105.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	308.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	4180.	2.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	323.	0.	3.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	401.	132.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	2.	114.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	38.	1553.	435.
13.	456.	0.	3.	12.	0.	0.	0.	0.	0.	3.	0.	0.	0.	0.	1.	0.	0.	2.	66.	0.	2.	0.	121.	176.	987.
VAL_K, VAR_K, SQMK2=							0.7362		0.0000		0.0032														

CONFRONTO VALORI KAPPA

KAPPA 1= 0.81225279 +/- 0.00285335

KAPPA 2= 0.73624002 +/- 0.00320660

IMPOSTI I VALORI: COVAR = 0.00000000 CORRE = 0.00000000

VARIABILE NORMALIZZATA Z = 17.70907019

DATO DEL 11-06-05

TEST 110605 - 26 CLASSI

NUMERO DI CLUSTER = 26

NUMERO DI PIXEL = 18201

FILE DATI : C:\MICHAELA\TESI_DOTTORATO_pc\PROGRAMMA_PROF_nuovo\110605_rif\accuracy_110605_eC2.txt

```
SUM_ROW(I)
1846.  9.  0.  174.  777.  720.  425.  47.  53.  109.  540.  991.  111.  711.  1484.  751.  779.  1322.  587.  2007.  438.
1.  54.  846.  1517.  1902.
SUM_COL(J)
1652.  548.  502.  798.  169.  1509.  114.  34.  132.  30.  174.  574.  177.  673.  1457.  520.  779.  1609.  471.  2003.  438.
239.  109.  288.  1331.  1871.
```

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.894908	0.884417	0.007762	1.000000	1.000000	0.000000
0.000000	-0.031043	0.001347	0.000000	-0.000495	0.000165
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.804598	0.795638	0.031378	0.175439	0.167480	0.013102
0.000000	-0.009372	0.000724	0.000000	-0.044594	0.001635
1.000000	1.000000	0.000000	0.477137	0.455602	0.012795
0.265882	0.261255	0.021195	0.991228	0.991018	0.008942
0.000000	-0.001872	0.000321	0.000000	-0.002589	0.000378
1.000000	1.000000	0.000000	0.401515	0.399767	0.042605
0.275229	0.274033	0.042664	1.000000	1.000000	0.000000
0.322222	0.315680	0.019894	1.000000	1.000000	0.000000
0.551968	0.537378	0.015907	0.952962	0.950253	0.009332
1.000000	1.000000	0.000000	0.627119	0.624831	0.036437
0.908579	0.905069	0.011183	0.959881	0.958250	0.007860
0.981806	0.980223	0.003766	1.000000	1.000000	0.000000
0.692410	0.683364	0.017112	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.821628	0.807658	0.010123
0.718910	0.711443	0.018872	0.895966	0.892499	0.014497
0.998007	0.997761	0.001118	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.000000	-0.013306	0.000866	0.000000	-0.000055	0.000055
0.851852	0.850959	0.048613	0.422018	0.420298	0.047282
0.224586	0.212119	0.014054	0.659722	0.643135	0.029120
0.789717	0.773126	0.011083	0.900075	0.890989	0.008899
0.730810	0.699967	0.011000	0.742918	0.712918	0.010969

Accuratezza delle classificazioni

STAMPE DI CONTROLLO DEL CALCOLO

TETA= 0.766496 0.064100 0.112525 0.020922
 VAL_K= 0.750503516
 UMT1,UMT2,FA 0.233503654 0.935899575 0.266584675
 VA1,VA2,VA3,VAR_k 0.000011227 -0.000000446 0.000000018 0.000010798
 MATRICE DI CONFUSIONE

1652.	0.	190.	2.	0.	0.	1.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	9.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	140.	0.	0.	0.	0.	34.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	777.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	720.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	312.	0.	0.	0.	113.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	47.	0.
0.	0.	0.	0.	0.	0.	0.	0.	53.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	79.	30.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	197.	0.	0.	169.	0.	0.	0.	0.	0.	174.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	547.	0.	0.	0.	0.	0.	287.	0.	0.	0.	0.	59.	98.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	111.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	65.	646.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	27.	1457.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	520.	0.	0.	0.	0.	0.	0.	231.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	779.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1322.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	422.	0.	0.	0.	0.	0.	3.	162.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	2003.	0.	0.	4.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	8.	46.	0.	0.	0.
0.	0.	0.	648.	0.	0.	0.	0.	0.	0.	0.	8.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	190.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1198.	319.
0.	351.	0.	8.	0.	12.	0.	0.	0.	0.	9.	0.	0.	0.	0.	0.	0.	0.	49.	0.	0.	0.	0.	0.	83.	1390.

VAL_K,VAR_K,SQMK1= 0.7505 0.0000 0.0033

NUMERO DI CLUSTER = 26
 NUMERO DI PIXEL = 18201
 FILE DATI : C:\MICHAELA\TESI_DOTTORATO_pc\PROGRAMMA_PROF_nuovo\110605_rif\accuracy_110605_cor_I.txt

SUM_ROW(I)
 1584. 287. 0. 859. 808. 664. 411. 46. 141. 32. 218. 1942. 176. 659. 1465. 329. 0. 1254. 475. 1999. 428.
 0. 299. 352. 1589. 2184.
 SUM_COL(J)
 1652. 548. 502. 798. 169. 1509. 114. 34. 132. 30. 174. 574. 177. 673. 1457. 520. 779. 1609. 471. 2003. 438.
 239. 109. 288. 1331. 1871.

Accuratezza delle classificazioni

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.898359	0.888212	0.008273	0.861380	0.848166	0.009189
0.006969	-0.023858	0.005157	0.003650	-0.012313	0.002708
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.904540	0.900163	0.010437	0.973684	0.972381	0.005940
0.084158	0.075575	0.009468	0.402367	0.374604	0.039256
1.000000	1.000000	0.000000	0.440027	0.418824	0.012635
0.274939	0.270369	0.021798	0.991228	0.991025	0.008935
0.521739	0.520844	0.073700	0.705882	0.705137	0.078297
0.709220	0.707096	0.038436	0.757576	0.755683	0.037525
0.843750	0.843492	0.064275	0.900000	0.899824	0.054860
0.798165	0.796217	0.027380	1.000000	1.000000	0.000000
0.294027	0.271038	0.009858	0.994774	0.994149	0.003368
1.000000	1.000000	0.000000	0.994350	0.994295	0.005688
1.000000	1.000000	0.000000	0.979198	0.978416	0.005704
0.994539	0.994064	0.002092	1.000000	1.000000	0.000000
0.996960	0.996871	0.003124	0.630769	0.623972	0.021323
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
1.000000	1.000000	0.000000	0.779366	0.763040	0.010878
0.869474	0.866006	0.015813	0.876858	0.873558	0.015497
0.998999	0.998876	0.000794	0.997004	0.996635	0.001371
1.000000	1.000000	0.000000	0.977169	0.976619	0.007305
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.180602	0.175665	0.022081	0.495413	0.486985	0.048542
0.767045	0.763300	0.022789	0.937500	0.936267	0.014530
0.745752	0.725693	0.011514	0.890308	0.879816	0.009306
0.645147	0.604489	0.010902	0.753073	0.719403	0.011024

DATO DEL 29-07-05

TEST 290705 - 26 CLASSI

NUMERO DI CLUSTER = 26
NUMERO DI PIXEL = 15544
FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_290705_eC.txt

SUM_ROW(I)
1847. 599. 518. 0. 777. 635. 502. 80. 26. 60. 166. 634. 348. 149. 417. 827. 904. 1365. 0. 2085. 585.
143. 259. 421. 747. 1450.
SUM_COL(J)
1789. 610. 387. 518. 764. 637. 129. 19. 160. 27. 162. 529. 90. 149. 417. 827. 904. 1387. 366. 2085. 615.
491. 241. 406. 755. 1080.

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.956687	0.951053	0.005323	0.987703	0.986044	0.002953
1.000000	1.000000	0.000000	0.981967	0.981244	0.005600
0.000000	-0.025533	0.001314	0.000000	-0.034474	0.001541
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.983269	0.982404	0.004835	1.000000	1.000000	0.000000
0.996850	0.996716	0.002318	0.993721	0.993453	0.003262
0.256972	0.250754	0.019192	1.000000	1.000000	0.000000
0.000000	-0.001224	0.000281	0.000000	-0.005173	0.000580
0.000000	-0.010400	0.000826	0.000000	-0.001675	0.000329
0.450000	0.449043	0.064201	1.000000	1.000000	0.000000
0.975904	0.975650	0.012024	1.000000	1.000000	0.000000
0.829653	0.823651	0.015347	0.994329	0.994088	0.003403
0.000000	-0.005824	0.000616	0.000000	-0.022901	0.001242
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.934066	0.927606	0.007330	0.919250	0.911476	0.007959
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.951220	0.949312	0.009008
1.000000	1.000000	0.000000	0.291242	0.284661	0.020229
0.930502	0.929407	0.016032	1.000000	1.000000	0.000000
0.964371	0.963415	0.009267	1.000000	1.000000	0.000000
0.993307	0.992965	0.003135	0.982781	0.981912	0.004969
0.744828	0.725774	0.012000	1.000000	1.000000	0.000000

Accuratezza delle classificazioni

STAMPE DI CONTROLLO DEL CALCOLO

TETA= 0.865993 0.066490 0.126538 0.024207
 VAL_K= 0.856448634
 UMT1,UMT2,FA 0.134006691 0.933510383 0.153775864
 VA1,VA2,VA3,VAR_k 0.000008567 -0.000000241 0.000000010 0.000008336
 MATRICE DI CONFUSIONE

1767.	0.	0.	0.	0.	0.	0.	0.	80.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	599.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	518.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	764.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	13.	0.	0.
0.	2.	0.	0.	0.	633.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	373.	0.	0.	0.	129.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	80.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
17.	9.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	14.	0.	0.	0.	0.	19.	0.	27.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	4.	0.	0.	0.	0.	162.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	526.	0.	0.	0.	0.	0.	108.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	348.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	149.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	417.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	827.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	904.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	90.	0.	0.	0.	1275.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	2085.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	585.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	143.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	18.	0.	241.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	3.	0.	0.	0.	0.	0.	0.	0.	0.	12.	0.	0.	406.	0.	0.
5.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	742.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	4.	366.	0.	0.	0.	0.	0.	0.	1080.
VAL_K,VAR_K,SQMK1=					0.8564		0.0000		0.0029																

NUMERO DI CLUSTER = 26
 NUMERO DI PIXEL = 15544
 FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_290705_I.txt

SUM_ROW(I)	1759.	614.	229.	298.	767.	630.	451.	99.	84.	45.	173.	1955.	72.	149.	417.	721.	603.	306.	293.	2079.	509.				
518.	362.	451.	827.	1133.																					
SUM_COL(J)	1789.	610.	387.	518.	764.	637.	129.	19.	160.	27.	162.	529.	90.	149.	417.	827.	904.	1387.	366.	2085.	615.				
491.	241.	406.	755.	1080.																					

Accuratezza delle classificazioni

USER"S ACCURACY E KAPPA PER RIGA			PRODUCER"S RELIABILITY E KAPPA PER COLONNA		
USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.972143	0.968520	0.004419	0.955841	0.950206	0.005448
0.983713	0.983048	0.005313	0.990164	0.989759	0.004158
0.000000	-0.025533	0.001314	0.000000	-0.014953	0.000995
0.966443	0.965286	0.010785	0.555985	0.547306	0.021925
0.989570	0.989031	0.003856	0.993455	0.993116	0.003068
1.000000	1.000000	0.000000	0.989011	0.988547	0.004303
0.272727	0.266641	0.020699	0.953488	0.952099	0.019088
0.030303	0.029116	0.017145	0.157895	0.152497	0.084105
0.321429	0.314371	0.051299	0.168750	0.164234	0.029517
0.600000	0.599304	0.073072	1.000000	1.000000	0.000000
0.936416	0.935747	0.018734	1.000000	1.000000	0.000000
0.264962	0.239065	0.009362	0.979206	0.976214	0.007091
1.000000	1.000000	0.000000	0.800000	0.799069	0.042311
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.998613	0.998535	0.001464	0.870617	0.864323	0.012152
1.000000	1.000000	0.000000	0.667035	0.653597	0.015985
0.836601	0.820593	0.023129	0.184571	0.168196	0.009851
0.996587	0.996505	0.003489	0.797814	0.793930	0.021294
1.000000	1.000000	0.000000	0.997122	0.996678	0.001354
1.000000	1.000000	0.000000	0.827642	0.821807	0.015636
0.000000	-0.032618	0.001496	0.000000	-0.034474	0.001541
0.665746	0.660482	0.024986	1.000000	1.000000	0.000000
0.829268	0.824689	0.018102	0.921182	0.918827	0.013744
0.868198	0.861470	0.012275	0.950993	0.948240	0.008278
0.849073	0.837804	0.011299	0.890741	0.882151	0.010158

DATO DEL 30-08-05

TEST 300805 - 26 CLASSI

NUMERO DI CLUSTER = 26
NUMERO DI PIXEL = 18447
FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_300805_eC.txt

SUM_ROW(I)
1695. 235. 0. 0. 0. 167. 503. 574. 88. 0. 1019. 691. 680. 165. 293. 944. 753. 1087. 1295. 3942. 648.
34. 214. 485. 1210. 1620.
SUM_COL(J)
1697. 613. 494. 199. 733. 857. 126. 28. 168. 35. 164. 542. 96. 165. 293. 987. 753. 1162. 440. 3940. 663.
575. 311. 465. 1103. 1733.

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.933333	0.926579	0.006630	0.932233	0.925377	0.006675
1.000000	1.000000	0.000000	0.383361	0.375404	0.019481
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.000000	-0.048721	0.001704	0.000000	-0.009136	0.000710
0.250497	0.245342	0.019054	1.000000	1.000000	0.000000
0.000000	-0.001520	0.000288	0.000000	-0.032115	0.001362
1.000000	1.000000	0.000000	0.523810	0.521527	0.038548
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.158979	0.151435	0.011018	0.987805	0.987092	0.009071
0.764110	0.756969	0.016490	0.974170	0.973165	0.007073
0.141176	0.136684	0.012997	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.956434	0.954084	0.006832
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.935456	0.931415	0.007628
0.335907	0.319680	0.012811	0.988636	0.987778	0.005433
0.999493	0.999355	0.000456	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.977376	0.976552	0.005980
0.000000	-0.032173	0.001363	0.000000	-0.001847	0.000317
0.929907	0.928705	0.017737	0.639871	0.635645	0.027372
0.958763	0.957697	0.009252	1.000000	1.000000	0.000000
0.911570	0.905947	0.008629	1.000000	1.000000	0.000000
0.925926	0.918246	0.007132	0.865551	0.852607	0.008865

Accuratezza delle classificazioni

STAMPE DI CONTROLLO DEL CALCOLO

TETA= 0.777850 0.081472 0.155744 0.048550
VAL_K= 0.758145692
UMT1,UMT2,FA 0.222149943 0.918527957 0.263306420
VA1,VA2,VA3,VAR_k 0.000011103 -0.000000901 0.000000083 0.000010284
MATRICE DI CONFUSIONE

1582.	0.	0.	0.	0.	0.	0.	28.	0.	35.	1.	14.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	2.	0.	0.	33.
0.	235.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	93.	7.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	66.
0.	377.	0.	0.	0.	0.	126.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	494.	0.	0.	0.	0.	80.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	88.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	857.	0.	0.	0.	0.	162.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	528.	0.	0.	0.	0.	0.	75.	0.	0.	0.	0.	88.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	96.	0.	0.	9.	0.	0.	0.	0.	0.	0.	575.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	165.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	293.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	944.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	753.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1087.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	726.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	435.	0.	0.	0.	0.	0.	0.	134.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	3940.	0.	0.	2.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	648.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	34.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	15.	0.	199.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	20.	465.	0.	0.
0.	1.	0.	106.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1103.	0.
115.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	5.	0.	0.	0.	0.	0.	0.	1500.

VAL_K,VAR_K,SQMK1= 0.7581 0.0000 0.0032

NUMERO DI CLUSTER = 26
NUMERO DI PIXEL = 18447
FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_300805_I.txt

SUM_ROW(I)	1818.	354.	92.	71.	483.	118.	158.	77.	105.	17.	330.	787.	120.	180.	277.	1032.	736.	1321.	709.	3825.	649.			
5.	413.	602.	1141.	3027.																				
SUM_COL(J)	1697.	613.	494.	199.	733.	857.	126.	28.	168.	35.	164.	542.	96.	165.	293.	987.	753.	1162.	440.	3940.	663.			
575.	311.	465.	1103.	1838.																				

Accuratezza delle classificazioni

USER"S ACCURACY E KAPPA PER RIGA			PRODUCER"S RELIABILITY E KAPPA PER COLONNA		
USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.882838	0.870968	0.008206	0.945787	0.939860	0.006066
0.988701	0.988312	0.005810	0.570962	0.562568	0.020088
0.000000	-0.027516	0.001255	0.000000	-0.005012	0.000524
0.140845	0.131476	0.041599	0.050251	0.046582	0.015392
0.755694	0.745584	0.020227	0.497954	0.484455	0.018582
0.000000	-0.048721	0.001704	0.000000	-0.006438	0.000595
0.772152	0.770585	0.033533	0.968254	0.967980	0.015751
0.000000	-0.001520	0.000288	0.000000	-0.004192	0.000479
0.990476	0.990389	0.009565	0.619048	0.616867	0.037550
0.176471	0.174905	0.092565	0.085714	0.084871	0.047281
0.472727	0.467998	0.027467	0.951220	0.950331	0.017119
0.688691	0.679268	0.016776	1.000000	1.000000	0.000000
0.666667	0.664923	0.043164	0.833333	0.832242	0.038252
0.911111	0.910309	0.021384	0.993939	0.993880	0.006101
1.000000	1.000000	0.000000	0.945392	0.944560	0.013464
0.845930	0.837221	0.011768	0.884498	0.877654	0.010708
1.000000	1.000000	0.000000	0.977424	0.976485	0.005633
0.852385	0.842461	0.010302	0.969019	0.966629	0.005464
0.488011	0.475501	0.018849	0.786364	0.777824	0.020216
1.000000	1.000000	0.000000	0.970812	0.963177	0.003359
1.000000	1.000000	0.000000	0.978884	0.978114	0.005783
0.200000	0.174261	0.184603	0.001739	0.001468	0.001687
0.239709	0.226672	0.021007	0.318328	0.302717	0.026711
0.765781	0.759724	0.017570	0.991398	0.991108	0.004426
0.867660	0.859244	0.010580	0.897552	0.890798	0.009671
0.480674	0.423204	0.009191	0.791621	0.750716	0.011078

STAMPE DI CONTROLLO DEL CALCOLO

Accuratezza delle classificazioni

TETA= 0.777308 0.089964 0.158651 0.051587
 VAL_K= 0.755293023
 UMT1,UMT2,FA 0.222692037 0.910035501 0.268898276
 VAL1,VA2,VA3,VAR_k 0.000011331 -0.000000602 0.000000075 0.000010804

1605.	119.	20.	0.	0.	0.	0.	23.	7.	32.	0.	0.	0.	0.	0.	0.	0.	0.	2.	0.	0.	0.	1.	0.	9.	
0.	350.	0.	0.	0.	0.	4.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	92.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	10.	1.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	58.	
0.	0.	0.	68.	365.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	24.	0.	0.	0.	0.	0.	26.	
0.	0.	0.	13.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	17.	88.	
0.	36.	0.	0.	0.	0.	122.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
1.	1.	16.	0.	0.	0.	0.	0.	56.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	3.	0.	
0.	1.	0.	0.	0.	0.	0.	0.	104.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
12.	0.	0.	0.	0.	0.	0.	0.	1.	3.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	
0.	0.	0.	92.	0.	49.	0.	0.	0.	0.	156.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	33.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	542.	11.	0.	0.	0.	17.	8.	0.	0.	11.	198.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	80.	0.	0.	0.	22.	0.	0.	0.	18.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	164.	16.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	277.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	873.	0.	0.	0.	0.	158.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	736.	0.	0.	0.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	106.	0.	1126.	0.	0.	88.	0.	0.	0.	0.	
0.	0.	0.	0.	325.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	346.	0.	0.	0.	0.	0.	38.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	3825.	0.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	649.	0.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	3.	0.	0.	0.	1.	0.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	4.	0.	0.	7.	0.	1.	0.	0.	299.	99.	3.	0.	0.	
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	113.	14.	0.	14.	461.	0.	0.
3.	7.	0.	5.	0.	0.	0.	5.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	990.	131.	
76.	7.	458.	11.	42.	808.	0.	0.	0.	0.	7.	0.	0.	0.	0.	0.	0.	2.	69.	0.	0.	0.	0.	0.	92.	1455.

VAL_K,VAR_K,SQMK2= 0.7553 0.0000 0.0033

CONFRONTO VALORI KAPPA

KAPPA 1= 0.75814569 +/- 0.00320690

KAPPA 2= 0.75529302 +/- 0.00328694

IMPOSTI I VALORI: COVAR = 0.00000000 CORRE = 0.00000000

VARIABILE NORMALIZZATA Z = 0.62120015

DATO DEL 13-05-06

TEST 130506 - 29 CLASSI

NUMERO DI CLUSTER = 29

NUMERO DI PIXEL = 17278

FILE DATI : D:\E\KAPPA\EM_130506_-eCogniton-_tab.txt

```

SUM_ROW(I)
 623.  922.  222.  533.  316.  252.  117.   1.   9.  210.  399.  363.  138.  476.  649. 1364.  407. 1205.  236. 4244.  607.
79.  174.  196. 2544.  532.  313.  147.   0.
SUM_COL(J)
 337. 1190.  222.  240.  287.  299.  127.  25.  96.  38.  64.  363.  165.  476.  649.  942.  407. 1205.  415. 4169.  607.
474.  175.  332. 2043. 1286.  186.  382.  77.

```

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.333868	0.320617	0.018806	0.617211	0.602892	0.027261
0.830803	0.818287	0.013137	0.643697	0.623612	0.014296
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.450281	0.442538	0.021480	1.000000	1.000000	0.000000
0.908228	0.906678	0.016487	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.842809	0.840483	0.021303
0.000000	-0.007405	0.000659	0.000000	-0.006818	0.000632
0.000000	-0.001449	0.000290	0.000000	-0.000058	0.000058
0.111111	0.106145	0.105294	0.010417	0.009901	0.010312
0.142857	0.140968	0.023948	0.789474	0.786883	0.066917
0.155388	0.152248	0.017848	0.968750	0.968011	0.022261
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.836364	0.835046	0.028987
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.690616	0.672775	0.012905	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.864407	0.861070	0.022791	0.491566	0.484525	0.024575
0.982328	0.976708	0.002652	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.949367	0.947939	0.025357	0.158228	0.154361	0.016450
1.000000	1.000000	0.000000	0.994286	0.994228	0.005756
0.994898	0.994798	0.005188	0.587349	0.582615	0.027111
0.796777	0.769525	0.008759	0.992168	0.990816	0.002285
0.919173	0.912673	0.012735	0.380249	0.360560	0.013314
0.594249	0.589834	0.027850	1.000000	1.000000	0.000000
0.578231	0.568696	0.041502	0.222513	0.215842	0.021095
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000

STAMPE DI CONTROLLO DEL CALCOLO

Accuratezza delle classificazioni

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.755459	0.750594	0.028873	0.513353	0.506817	0.027329
0.838974	0.827064	0.012523	0.687395	0.668700	0.013926
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.436399	0.428460	0.021875	0.929167	0.927008	0.017048
0.666667	0.661036	0.023116	0.982578	0.982141	0.007915
0.964413	0.963786	0.011240	0.906355	0.904806	0.017099
0.000000	-0.007405	0.000659	0.000000	-0.001101	0.000253
0.215686	0.214550	0.040595	0.880000	0.879287	0.065367
0.250000	0.245810	0.217678	0.010417	0.010188	0.010308
0.000000	-0.002204	0.000358	0.000000	-0.002088	0.000348
0.927536	0.927267	0.031318	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.884758	0.881494	0.014105	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.904468	0.901109	0.011901
1.000000	1.000000	0.000000	0.998938	0.998877	0.001122
1.000000	1.000000	0.000000	0.987715	0.987422	0.005589
0.997429	0.997237	0.001593	0.965975	0.963511	0.005587
0.973615	0.972965	0.008431	0.889157	0.886671	0.015714
1.000000	1.000000	0.000000	0.993044	0.990852	0.001690
1.000000	1.000000	0.000000	0.988468	0.988053	0.004488
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.519288	0.514369	0.027236	1.000000	1.000000	0.000000
0.523885	0.514557	0.019965	0.990964	0.990623	0.005388
0.865215	0.847141	0.008185	0.914342	0.902110	0.006999
0.582734	0.549179	0.012502	0.755832	0.729742	0.013007
0.622881	0.618777	0.031727	0.790323	0.787419	0.030193
0.845361	0.841865	0.037508	0.214660	0.210226	0.020759
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000

STAMPE DI CONTROLLO DEL CALCOLO

DATO DEL 14-06-06

TEST 140606 - 30 CLASSI

NUMERO DI CLUSTER = 29
NUMERO DI PIXEL = 17265
FILE DATI : D:\PROGRAMMA_PROF_nuovo\accuracy_140606_eC.txt

```

SUM_ROW(I)
 267.  552.  419.  241.  268.  441.  251.   1.  28.  78.  155.  737.  526.  286.  371.  784.  754.  867.  406.  4702.  649.
482.  128.  404. 1529. 1238.   64.  383.  254.
SUM_COL(J)
 267.  585.  381.  240.  270.  442.   60.  48.  98.  26.  155.  410.  618.  286.  371.  784.  754. 1104.  411.  4702.  649.
500.  110.  402. 1621. 1227.  192.  382.  170.

```

USER'S ACCURACY E KAPPA PER RIGA			PRODUCER'S RELIABILITY E KAPPA PER COLONNA		
USER'S ACC	K_ROW	sqm	PRODUCER'S REL	K_COL	sqm
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.586957	0.572470	0.021402	0.553846	0.539111	0.020907
0.377088	0.363032	0.023845	0.414698	0.400140	0.025533
0.995851	0.995792	0.004199	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.992593	0.992476	0.005300
1.000000	1.000000	0.000000	0.997738	0.997678	0.002319
0.235060	0.232392	0.026559	0.983333	0.983087	0.016770
1.000000	1.000000	0.000000	0.020833	0.020777	0.020560
0.000000	-0.005709	0.000578	0.000000	-0.001624	0.000307
0.282051	0.280968	0.050864	0.846154	0.845456	0.071064
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
0.553596	0.542737	0.018395	0.995122	0.994904	0.003594
0.952471	0.950707	0.009607	0.810680	0.804730	0.016142
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.785326	0.773976	0.012828
0.766010	0.760304	0.021404	0.756691	0.750832	0.021548
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.964000	0.962966	0.008561
0.859375	0.858473	0.030892	1.000000	1.000000	0.000000
0.995050	0.994931	0.003575	1.000000	1.000000	0.000000
0.983649	0.981955	0.003574	0.927822	0.920809	0.007002
0.894184	0.886089	0.009337	0.902200	0.894646	0.009069
0.000000	-0.011246	0.000816	0.000000	-0.003721	0.000466
0.997389	0.997330	0.002666	1.000000	1.000000	0.000000
0.661417	0.658050	0.029838	0.988235	0.988060	0.008392

STAMPE DI CONTROLLO DEL CALCOLO

Accuratezza delle classificazioni

USER"S ACCURACY E KAPPA PER RIGA PRODUCER"S RELIABILITY E KAPPA PER COLONNA

USER"S ACC	K_ROW	sqm	PRODUCER"S REL	K_COL	sqm
0.941667	0.940750	0.015354	0.846442	0.844277	0.022321
0.641447	0.628872	0.016121	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.414698	0.409292	0.025143
0.990148	0.990009	0.007029	0.837500	0.835567	0.024042
1.000000	1.000000	0.000000	0.788889	0.786252	0.025063
1.000000	1.000000	0.000000	0.977376	0.976795	0.007250
0.266667	0.264109	0.029301	1.000000	1.000000	0.000000
0.000000	-0.002788	0.000403	0.000000	-0.001508	0.000296
0.166667	0.161909	0.088267	0.030612	0.029601	0.017325
0.833333	0.833082	0.076170	0.769231	0.768910	0.082715
1.000000	1.000000	0.000000	0.993548	0.993490	0.006488
0.249847	0.231599	0.010186	0.995122	0.994612	0.003800
1.000000	1.000000	0.000000	0.988673	0.988258	0.004411
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
1.000000	1.000000	0.000000	0.998724	0.998664	0.001335
1.000000	1.000000	0.000000	0.003979	0.003806	0.002193
0.978643	0.977184	0.005469	0.705616	0.691387	0.014103
0.935829	0.934264	0.012962	0.851582	0.848295	0.017860
1.000000	1.000000	0.000000	0.981923	0.975324	0.002637
1.000000	1.000000	0.000000	1.000000	1.000000	0.000000
-NaN	-NaN	-NaN	0.000000	0.000000	0.000000
0.824000	0.822871	0.034236	0.936364	0.935900	0.023435
0.490172	0.478018	0.017503	0.992537	0.992168	0.004504
0.978603	0.976386	0.003998	0.959284	0.955157	0.005385
0.883019	0.874069	0.009416	0.953545	0.949684	0.006486
0.231405	0.222761	0.038564	0.145833	0.139805	0.025410
0.974227	0.973644	0.008222	0.989529	0.989288	0.005327
0.971429	0.971144	0.012715	1.000000	1.000000	0.000000

STAMPE DI CONTROLLO DEL CALCOLO

Accuratezza delle classificazioni

TETA= 0.852129 0.103278 0.199382 0.085579
VAL_K= 0.835097754
UMT1,UMT2,FA 0.147871416 0.896721663 0.183894571
VAL1,VA2,VA3,VAR_k 0.000009076 -0.00000555 0.00000084 0.000008605

226.	0.	0.	0.	0.	0.	0.	10.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	2.	0.
7.	585.	197.	3.	52.	0.	0.	24.	31.	5.	0.	0.	0.	0.	0.	0.	0.	0.	0.	6.	0.	0.	0.	0.	0.	0.	0.	1.	0.	1.	0.
0.	0.	158.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	201.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	2.	0.	0.	0.	0.	0.
0.	0.	0.	0.	213.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	432.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	1.	60.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	164.	0.	0.
0.	0.	26.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1.	0.	0.	0.	0.	0.	0.	14.	3.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	4.	20.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	154.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	408.	2.	0.	0.	0.	751.	325.	0.	0.	0.	147.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	611.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	286.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	371.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	783.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	3.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	2.	5.	0.	0.	1.	0.	779.	0.	0.	0.	9.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	350.	0.	0.	0.	0.	0.	0.	24.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	4617.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	649.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	22.	103.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	30.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	79.	0.	299.	7.	399.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	6.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1555.	28.	0.	0.	0.	0.	0.
0.	0.	0.	0.	5.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	61.	0.	23.	0.	0.	65.	1170.	0.	0.	0.	0.	0.
32.	0.	0.	0.	0.	0.	0.	0.	60.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	28.	1.	0.	0.	0.
1.	0.	0.	0.	0.	9.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	378.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	170.	0.	0.

VAL_K,VAR_K,SQMK2= 0.8351 0.0000 0.0029

CONFRONTO VALORI KAPPA

KAPPA 1= 0.90023622 +/- 0.00239075

KAPPA 2= 0.83509775 +/- 0.00293346

IMPOSTI I VALORI: COVAR = 0.00000000 CORRE = 0.00000000

VARIABILE NORMALIZZATA Z = 17.21287793

CAPITOLO 7: SCHEMATIZZAZIONE DELLA PROCEDURA GENERALE DI LAVORO

Dai capitoli precedenti è emerso che il riconoscimento delle colture agricole tramite immagini satellitari telerilevate non costituisce un obiettivo semplice da raggiungere. Infatti, se da un punto di vista tecnico il lavoro effettuato dimostra come sia possibile ottenere l'identificazione di una coltura con un buon livello di affidabilità, è altrettanto vero che sono numerose le variabili in gioco. Tutte queste devono essere considerate per raggiungere un risultato veramente utile per l'utente finale. A questi fattori si aggiungono aspetti relativi alle diverse professionalità che vengono coinvolte: la doppia veste dell'"agronomo-ingegnere", in grado di manipolare dati satellitari ed interpretare correttamente le informazioni fornite dall'analisi, in relazione all'uso effettivo del suolo, non trova ancora oggi una risposta adeguata nei vari profili professionali delle offerte formative esistenti. E' altresì vero che la suddivisione dei compiti tra diversi esperti nei vari settori è difficilmente sostenibile dalle organizzazioni e/o associazioni che possono essere interessate, inizialmente in via solo sperimentale, ai risultati forniti dalle analisi discusse in questa tesi.

Visto che lo scopo di questo lavoro è anche quello di pensare ad un'applicazione che possa avere una ricaduta ed una utilità per le Associazioni e le Amministrazioni coinvolte nei temi della produzione agricola e della gestione degli incentivi economici, nazionali e comunitari, che, secondo vari canali, giungono ai produttori, si è cercato di valutare, oltre all'efficacia dei metodi investigati, anche la possibilità di giungere all'ingegnerizzazione del metodo.

In effetti, i motivi che hanno portato alla scelta di questo argomento per la dissertazione di tesi nascono anche da un precedente interessamento verso i metodi di analisi dell'uso del suolo da immagini satellitari, dell'Unione Agricoltori di Ravenna. L'Associazione, fortemente impegnata nella gestione delle domande redatte dagli agricoltori per accedere agli incentivi previsti dalla PAC (oggi Domanda di Premio Unica), si è mostrata fin dall'inizio molto interessata a questa applicazione. In particolare le sue aspettative sono rivolte alla realizzazione di una metodologia veloce ed affidabile per la classificazione delle colture agricole, al quale ricorrere in caso di contestazioni durante il processo di controllo delle dichiarazioni effettuate dai produttori.

D'altra parte i dati satellitari potrebbero rappresentare una valida alternativa agli strumenti attualmente utilizzati nelle verifiche, i sopralluoghi e le immagini aeree, la cui disponibilità è legata all'esecuzione di voli dedicati, da effettuare in periodi utili ai fini delle verifiche volute. Di fatto le indagini aeree non sono eseguite con regolarità e tanto meno in periodi efficaci per un'analisi che è fortemente legata alla stagionalità delle produzioni agricole coinvolte.

Per questo motivo l'Ente è alla ricerca di procedure di indagine sulle coltivazioni in corso che siano veloci, geograficamente estese e che presentino con un certo grado di affidabilità, comunque valutabile secondo procedure statistiche.

7.1 Flusso di lavoro

La procedura proposta è stata sviluppata nei Capitoli centrali della Tesi, anche se con un taglio più scientifico che pratico dal punto di vista operativo. Si è voluto allora riepilogare in uno schema a blocchi (Fig. 7.1) la metodologia sperimentata, risultando così più intuitiva. Di seguito sono riassunti gli aspetti fondamentali delle varie fasi, con particolare attenzione per le problematiche emerse e per le eventuali soluzioni proposte.

Il primo passo prevede l'individuazione dell'oggetto a terra da monitorare mediante il dato multispettrale satellitare: come si è già detto in precedenza, sono molti i fenomeni, sia naturali sia di origine antropica, che si prestano a questo tipo di approccio. Nel caso in esame l'attenzione è stata rivolta al riconoscimento delle colture, erbacee e arboree, presenti nel terreno, in una zona storicamente sfruttata dal punto di vista agricolo e che attualmente richiede un miglior controllo nella gestione delle risorse. Trattandosi di realtà soggette ad una grande variabilità stagionale, la raccolta delle informazioni relative alle coltivazioni risulta molto articolata, sia per la tipologia che per il numero di dati coinvolti: nel secondo livello il flusso di lavoro prosegue dividendosi in due strade parallele, ugualmente importanti data la criticità della fase.

Da un lato è necessario capire quali sono le fonti utilizzabili per ricostruire la situazione a terra dal quale estrarre i campioni e verificarne l'immediata disponibilità: le cartografie ufficiali e le dichiarazioni delle aziende agricole. Se il lavoro è svolto all'interno della stessa associazione degli agricoltori, il problema principale rimane la consultazione delle mappe (Carta dell'Uso del Suolo, Carta della Vegetazione, Mappe Catastali) che sono comunque reperibili presso gli enti che le producono, delle quali deve essere sempre verificato lo stato di aggiornamento. Dall'esterno invece, per gli studi a scopo scientifico, l'ostacolo maggiore risulta invece la visione dei documenti privati degli agricoltori, per la quale deve essere previsto un accordo con l'ente responsabile.

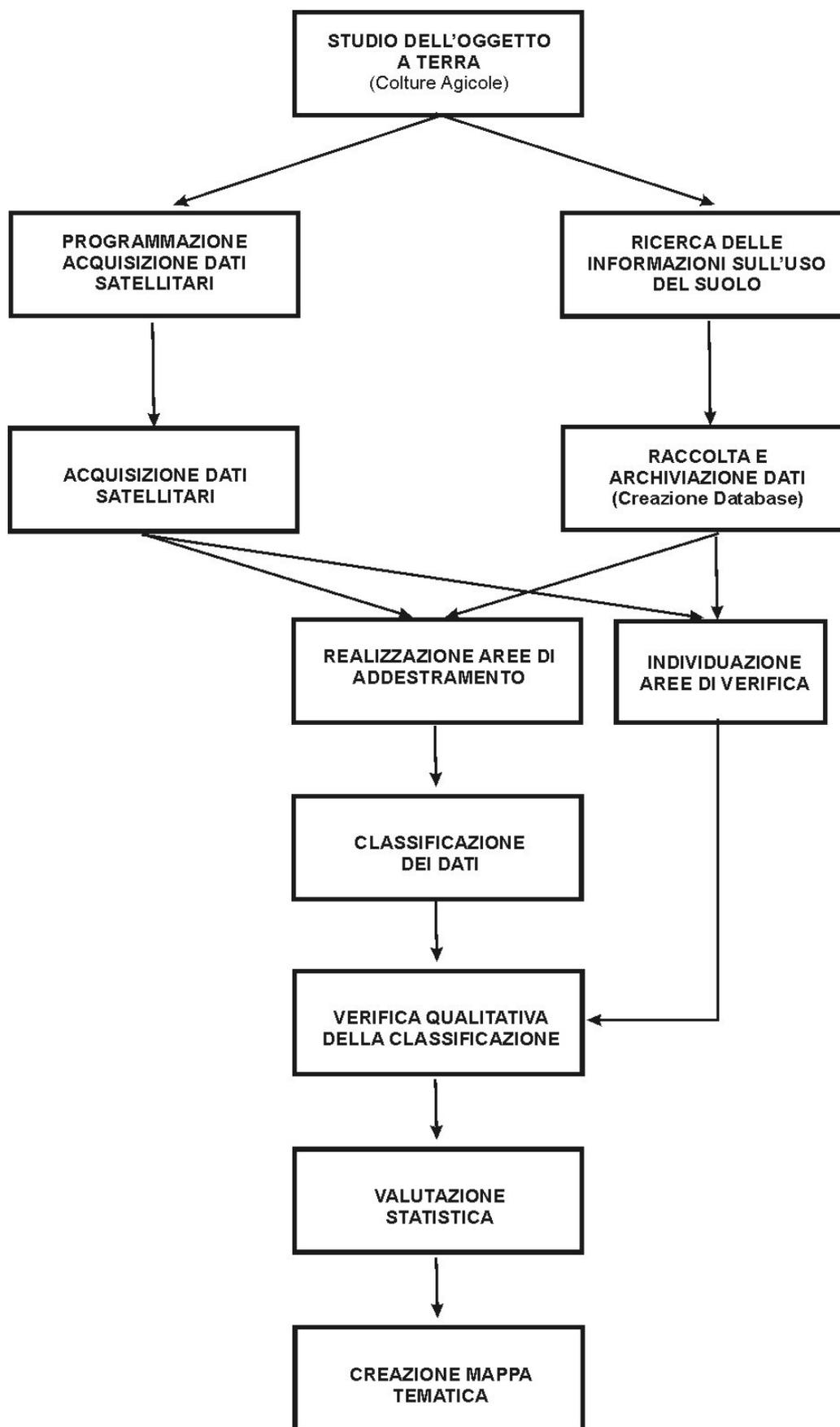


Fig. 7.1 – Flow-chart della procedura di lavoro

In genere i vari documenti vengono realizzati da soggetti differenti, anche in formati a volte tra loro non compatibili. Dal 2005 sono tutti comunque in versione digitale: la fase successiva prevista è la georeferenziazione dei raster e dei vettoriali raccolti in un unico sistema di riferimento, in modo da rendere agevole la sovrapposizione dei diversi strati. Il loro reale utilizzo integrato però necessita di un'organizzazione in un database geografico, nel quale archiviare anche tutte le informazioni associate ai dati, in questo caso agli appezzamenti di terreno (posizione, destinazione d'uso annuale o stagionale, dimensione, ecc.), continuamente e facilmente aggiornabile. Devono quindi essere previste una serie di tabelle degli attributi con l'obiettivo di rendere agevole, ma anche il più automatico possibile, la fase di individuazione delle aree di addestramento e di verifica. Dato il ruolo essenziale che questo database può svolgere in tutte le fasi del lavoro, serve inoltre il contributo di una figura esperta di trattamento di dati geografici e di GIS in generale.

Contemporaneamente alla ricerca dei dati sull'uso del suolo, l'oggetto dell'indagine deve essere approfondito in ogni suo aspetto per individuare il momento o, come più spesso accade, i momenti ottimali per l'acquisizione delle immagini satellitari. Infatti, analizzando le colture tipicamente coltivate nell'area in esame, la fascia costiera emiliano-romagnola, la risposta spettrale più rappresentativa per ciascuna tipologia di pianta si ottiene in periodi diversi, a seconda dello ciclo di maturazione, dello stato fenologico, della copertura fogliare e di tutti quei fattori che influenzano il comportamento della vegetazione. Di conseguenza bisogna valutare l'ipotesi di sottoporre più dati agli algoritmi di classificazione, diversificando per ognuno le colture che si possono riconoscere perché effettivamente sul terreno da quelle non distinguibili dal suolo sottostante.

La programmazione dell'acquisizione risulta il metodo più indicato per avere la contemporaneità tra l'immagine e la realtà a terra, in quanto verrebbe suggerita da tecnici che conoscono in modo dettagliato ciò che avviene sui terreni e la cui esperienza faciliterebbe la fase di interpretazione della classificazione. La procedura da seguire per ordinare un'immagine, descritta nel capitolo 4, impone comunque, almeno inizialmente, l'affiancamento di un esperto di dati satellitari in grado di scegliere in modo opportuno i diversi parametri richiesti. Nell'eventualità l'acquisizione non rispetti le condizioni iniziali è possibile ricorrere ad immagini di archivio, scaricabili direttamente dai siti interessati.

Si è scelto, in questo ultimo contesto, di non fare riferimento ad uno specifico dato satellitare, come ad esempio Aster, perché l'obiettivo inseguito è quello di creare una metodologia esportabile ad altri tipi di dati, da utilizzare indipendentemente da quello su cui è stata effettuata la sperimentazione: infatti si vuole in futuro poter effettuare una scelta in funzione, oltre che della disponibilità del sensore, dei costi e dei tempi che questo comporta.

Una volta immagazzinati tutti i dati in un database, acquisizioni satellitari comprese, si può procedere con la prima vera fase operativa della classificazione, ovvero la determinazione dei training site, non realizzabile comunque senza gli stadi precedenti. L'insieme delle informazioni disponibili è alla base anche dell'individuazione delle aree di verifica: infatti nella valutazione generale delle classi presenti sui terreni della zona esaminata, bisogna garantire lo stesso numero di coperture, attraverso campioni differenti, sia per l'addestramento del software sia per il controllo finale dei risultati. Durante l'esecuzione di questa delicata operazione, la consulenza di un agronomo diventa molto utile ai fini di un tracciamento corretto dei campioni, perché basato non solo sulle dichiarazioni degli agricoltori ma soprattutto sulla conoscenza delle numerose caratteristiche delle singole coltivazioni. L'efficacia di ogni sample verrebbe così direttamente valutata in funzione dei fattori in grado di influenzare la risposta spettrale dell'oggetto che contiene. Segue poi la classificazione delle immagini satellitari mediante gli algoritmi predisposti: la scelta tra la tecnica per pixel e quella object-oriented deve essere effettuata prendendo in considerazione le risorse a disposizione. Dato che, nella maggior parte delle elaborazioni effettuate, il confronto tra le due tecnologie non ha rivelato molte differenze nel riconoscimento delle singole colture, la decisione può essere presa valutando i costi relativi al software e all'addestramento del personale. Si ricorda che da un punto di vista operativo la classificazione per pixel richiede una procedura più veloce, non prevedendo la fase di segmentazione.

Vista l'esigenza di creare una procedura che possa essere utilizzata anche da utenti non altamente professionali occorre che questa sia il più possibile semplificata, ovviamente nel rispetto dei limiti relativi al successo dell'analisi. Questo si traduce anche nel ricorso ad un software di facile accesso. Comunque questa decisione deve essere presa a posteriori, quando un ben definito protocollo di analisi dei dati è stato scelto e validato dai risultati. Quindi, anche se in questa sede il lavoro è stato eseguito con software di una certa complessità, l'intenzione è quella di trasferire, successivamente, la metodologia su pacchetti dall'approccio più semplice e possibilmente scaricabili gratuitamente dalla rete.

I risultati della classificazione devono poi essere sottoposti ad una valutazione statistica che permetta di valutarne l'affidabilità. All'interno della procedura generale, questo passaggio si traduce con il calcolo della matrice di confusione, l'accuratezza globale, l'accuratezza utente, l'accuratezza produttore, errori di commissione, errori di omissione e infine l'indice K, per tenere conto dell'assegnazione casuale dei pixel nella classificazione. Considerando che sono concetti che richiedono determinate conoscenze nel trattamento di variabili statistiche deve essere prevista

un'adeguata preparazione in materia, che gli permetta di giudicare la validità di una classificazione all'interno delle contese nelle quali saranno chiamati in causa.

Il prodotto finale di tutto il protocollo impostato si può concretizzare con la stampa di una mappa aggiornata delle coperture agricole sul suolo: devono essere comunque applicate delle tecniche di post-elaborazione (ad esempio filtri per eliminare pixel isolati, filtri per uniformare i confini degli appezzamenti, ecc.) per rendere il risultato maggiormente leggibile dal punto di vista qualitativo.

In più livelli della procedura è stato fatto un richiamo circa la formazione che occorrerebbe per il personale che, all'interno della Società o Ente interessato, dovrebbe svolgere questo lavoro. Anche se tale aspetto non rappresenta, tuttavia, uno degli obiettivi della tesi. Si può comunque prevedere un necessario periodo di affiancamento con operatori più esperti, principalmente per la parte relativa al trattamento dei dati e alla conoscenza dei software. Del resto, il destinatario di questo periodo di training, potrebbe essere già in possesso, probabilmente, di un bagaglio culturale di tipo agronomo-forestale.

7.2 Studio di un'applicazione reale

In collaborazione con l'Ing. Michela Zanni, si è voluto studiare un caso di applicazione dei risultati della classificazione.

Lo scopo finale del lavoro riassunto in questo paragrafo è la progettazione di un Sistema Informativo Geografico per la gestione e il monitoraggio delle colture basato sull'elaborazione e sull'integrazione di differenti fonti informative, ovvero immagini satellitari, mappe tematiche, mappe catastali, dati sulle coltivazioni, ecc.

Lo strumento GIS permette infatti di archiviare e coordinare tutti i dati geografici disponibili per una specifica area, soprattutto quando queste informazioni provengono da enti diversi, sia pubblici che privati, sono rappresentati in differenti formati e sistemi di riferimento, e non sono aggiornati o facilmente consultabili, come accade per la zona interessata, la fascia costiera emiliano-romagnola, ovvero la medesima considerata nell'elaborato di Tesi.

Il processo di integrazione di tutte le informazioni richiede le seguenti fasi: riorganizzazione dei dati e loro conversione in formato digitale, georeferenziazione degli stessi in un unico sistema di riferimento cartografico, progettazione del database, digitalizzazione delle immagini QuickBird per la delimitazione dei singoli appezzamenti, classificazione dell'uso agricolo del suolo mediante immagini multispettrali ASTER, creazione della procedura automatica per l'aggiornamento delle singole colture.

Poiché la ricerca ha origine dalla necessità di studiare e monitorare costantemente il territorio e la sua vulnerabilità, essa è indirizzata a due tipologie principali di utenti: da un lato la comunità scientifica, che richiede la possibilità di accedere liberamente ad un sistema completo e aggiornato di dati per approfondire gli studi sull'uso del suolo e sulle sue variazioni, dall'altro le associazioni degli agricoltori che necessitano di un sistema per la gestione delle dichiarazioni dei loro associati che sia integrato con dati spaziali.

7.2.1 Metodologia

Vengono di seguito descritte le fasi ritenute più importanti per lo sviluppo del lavoro.

A) Riorganizzazione dei dati e progettazione del Database

Il primo passo nella creazione del GIS comporta la progettazione del Database che deve contenere tutte le informazioni relative alle colture e alle fonti utilizzate per estrarre i dati (Fig. 7.2). Tale database si potrebbe considerare come un ampliamento di quello previsto all'interno della procedura precedentemente descritta.

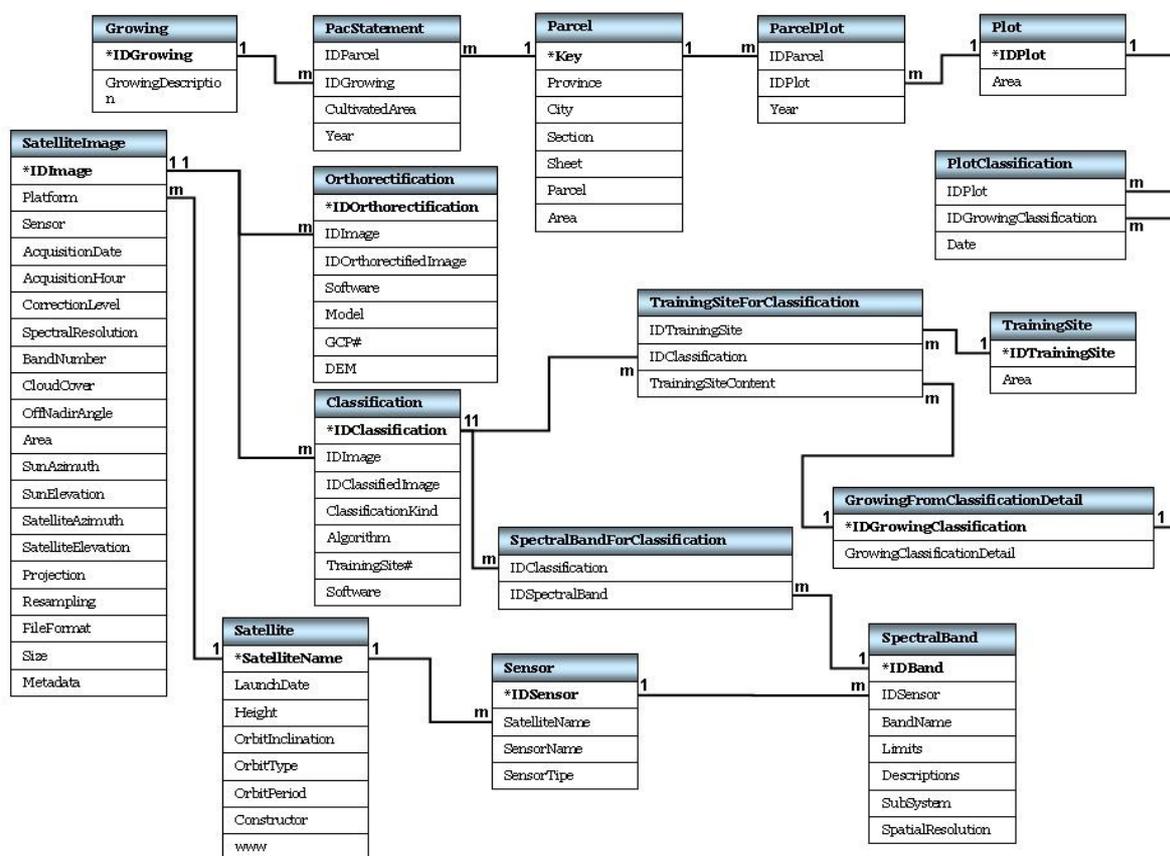


Fig. 7.2 – Struttura del Database

In particolare saranno gestite, con MS Access2003, sia le informazioni derivanti dalle dichiarazioni PAC, relative alle percentuali di particella catastale coltivata con una particolare coltura ogni anno, sia le indicazioni derivanti dalla classificazione delle immagini Aster. Si è ritenuto opportuno integrare nel Database anche le fonti informative (immagini satellitari e le tecniche di estrazione dei dati).

B) Classificazione delle immagini satellitari Aster

Sulla base di precedenti esperienze si è scelto di utilizzare la classificazione supervisionata per pixel di immagini satellitari come metodo di riconoscimento dell'uso del suolo: a partire da informazioni disponibili circa la realtà a terra, il classificatore associa ciascun pixel del dato alla classe che mostra qualità radiometriche simili. L'algoritmo applicato è quello della Massima Verosimiglianza (Idrisi 3.2).

C) Georeferenziazione dei dati

La maggior parte dei dati disponibili, relativi all'area di interesse e usati nell'applicazione, coincidono con quelli descritti nel Capitolo 4: la Carta dell'Uso del Suolo e la Carta della Vegetazione a scala 1:25000 rappresentate in falso UTM (indicato con UTM*, nel quale viene omessa la prima cifra della coordinata Nord), le mappe catastali, in origine in Cassini-Soldner, poi in UTM*, le immagini multispettrali a media risoluzione ASTER proiettate nel sistema UTM-WGS84 e acquisite nel periodo compreso tra il 2002-2006, le dichiarazioni PAC depositate dalle aziende agricole presso l'Unione Agricoltori di Ravenna, compilate con riferimento alle particelle catastali. Inoltre sono state utilizzate le ortoimmagini QuickBird e la Carta Tecnica Regionale a scala 1:5000 riferite al sistema di coordinate Gauss-Boaga, gentilmente concesse dalla Regione Emilia-Romagna.

Prima di archiviare tutti i dati nel Database è stato necessario riferirli ad un comune sistema di coordinate, e la scelta è ricaduta sul sistema cartografico nazionale Gauss-Boaga. La trasformazione dagli altri sistemi, UTM* e UTM-WGS84, è stata realizzata utilizzando gli stessi punti di controllo.

Le immagini Aster L1B sono state ortorettificate con PCI Geomatica OrthoEngine usando il modello rigoroso di Toutin, con DTM a 10 m e GCPs presi dalla Carta Tecnica Regionale. In ambiente ArcGis 9.1 invece, le immagini classificate sono state georeferenziate con una trasformazione polinomiale del primo ordine, mentre i vettoriali con la trasformazione affine prevista dal tool Spatial Adjustment.

D) Vettorializzazione delle immagini QuickBird

Le fasi successive sono la delimitazione degli appezzamenti esistenti e l'assegnazione ad essi del contenuto informativo derivante dalla classificazione delle immagini Aster. A tal fine, dato che la maggiore risoluzione del sensore Aster, generata dal sottosistema VNIR e pari a 15 m, non è sufficiente per identificare correttamente i confini dei singoli terreni, si è ritenuto opportuno utilizzare le ortofotografie pancromatiche QuickBird per generare lo shape dei perimetri. L'obiettivo è stato quello di ridurre il lavoro di fotointerpretazione e di editing manuale attraverso una procedura semi-automatica di estrazione dei limiti degli appezzamenti, ricorrendo ad una preliminare segmentazione delle immagini realizzata con il software *object-oriented eCognition*.

In questo contesto quindi il processo di segmentazione, che implica la costituzione di poligoni ("oggetti simili") sulla base dei valori assunti da parametri spettrali e di forma, non è stato utilizzato come fase precedente della classificazione *object-oriented*.

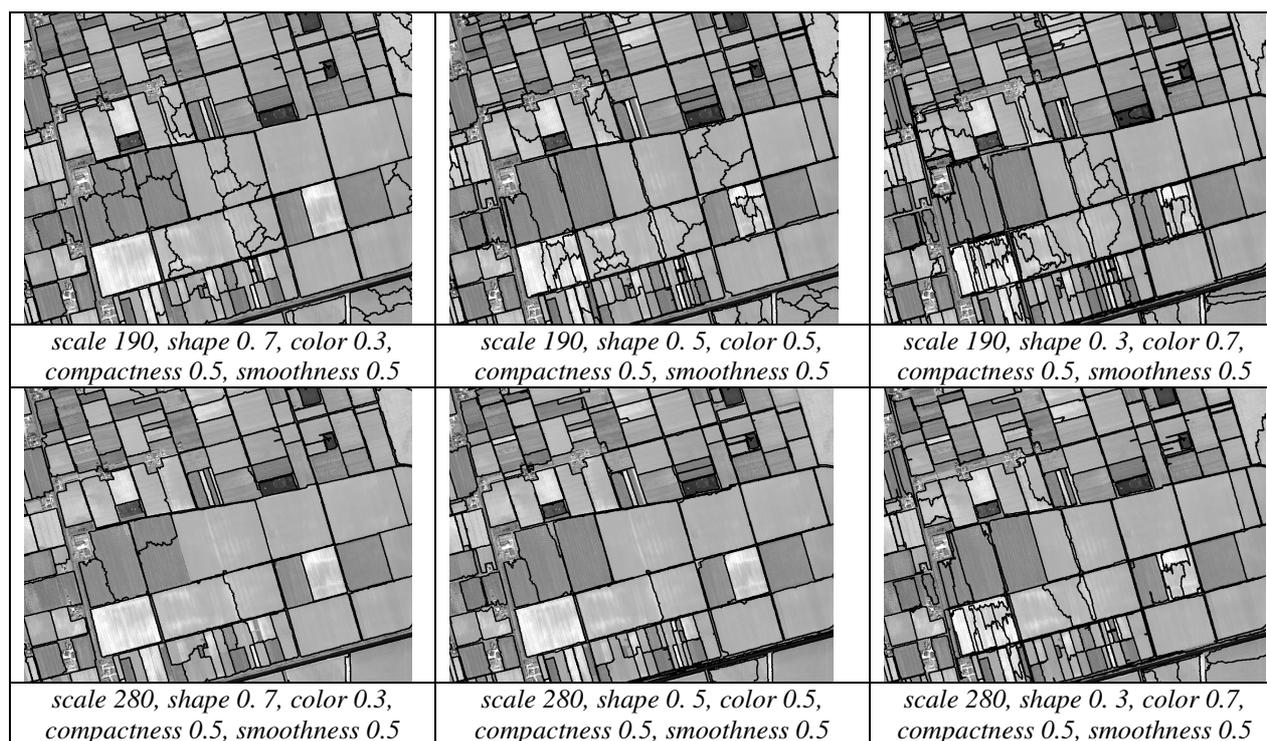


Figura 7.3 – Esempi di segmentazione: variazione nella forma, dimensione e numero degli oggetti creati utilizzando diversi valori dei parametri di color e shape per l'immagine QuickBird pancromatica

Sono stati testati diversi valori dei parametri di scala, forma e colore per individuare la combinazione ottimale degli stessi con cui estrarre, dalle immagini QuickBird, i perimetri dei terreni agricoli: non essendo disponibile una mappa di riferimento contenente gli appezzamenti, si è fatta una valutazione a video dei risultati, individuando quello che meglio si adatta agli obiettivi

preposti. In figura 7.3 sono riportati alcuni esempi di immagini segmentate secondo diversi valori dei parametri.

Il risultato della segmentazione diventa una mappa di supporto in cui è presente una demarcazione dei confini che, seppure “grossolana” in alcuni casi, permette di velocizzare il successivo lavoro del fotointerprete, il quale deve comunque intervenire per operazioni di aggregazione di poligoni e di eliminazione di quelli non significativi ai fini dell’indagine svolta.

Infine, in ambiente GIS è possibile migliorare i perimetri dei poligoni estratti mediante una generalizzazione delle linee con gli algoritmi di *Bend Simplify* e *Point Remove*.

E) Aggiornamento dei poligoni delle colture

Il passo più delicato dell’intero progetto è rappresentato dall’aggiornamento automatico del vettoriale contenente i poligoni degli appezzamenti. Le operazioni implementate con il tool Model Builder di Arcmap 9.1 sono di seguito descritte (Fig.7.4).

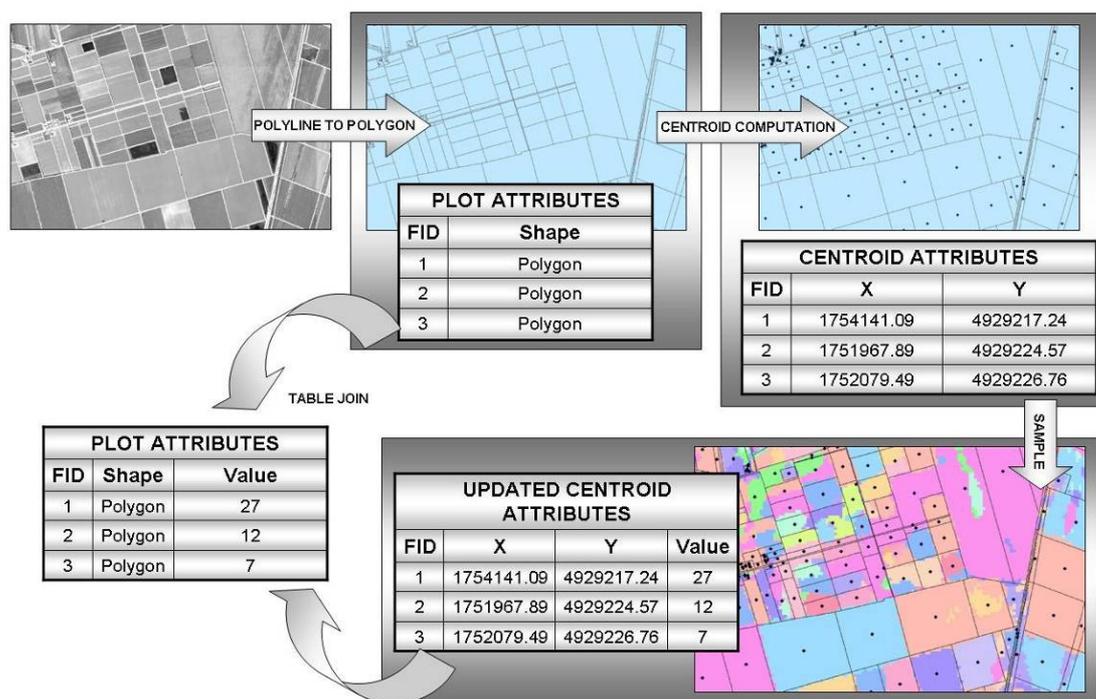


Figura 7.4 – Schema dell’aggiornamento automatico del vettoriale delle colture

La prima operazione è la conversione degli appezzamenti da polyline shapefile a polygon shapefile. Segue il calcolo delle coordinate del centroide di ciascun poligono e la creazione dello shape dei punti centroidi: il codice ID di ogni centroide, riportato nella tabella degli attributi, è lo stesso del corrispondente poligono. La funzione Sample permette quindi di associare ai singoli centroidi il

valore del raster sottostante, che è rappresentato dall'immagine Aster classificata: i valori riportati nella tabella che ne risulta (Updated Centroid Attributes) corrispondono alle diverse colture. Facendo un join tra la tabella e lo shape dei poligoni dei terreni i valori del raster possono essere assegnati a ciascun poligono. In tal modo l'attributo delle colture può essere associato all'appezzamento.

La procedura descritta si può ritenere adeguata per l'aggiornamento delle colture entro i limiti dell'incertezza di classificazione.

F) Creazione Geodatabase

Tutti i dati disponibili inizialmente e quelli prodotti dalle diverse elaborazioni sono state collocate all'interno del Geodatabase creato con ArcCatalog 9.1, che rappresenta l'archivio di tutti i raster, i vettoriali e le tabelle necessari per la progettazione del GIS:

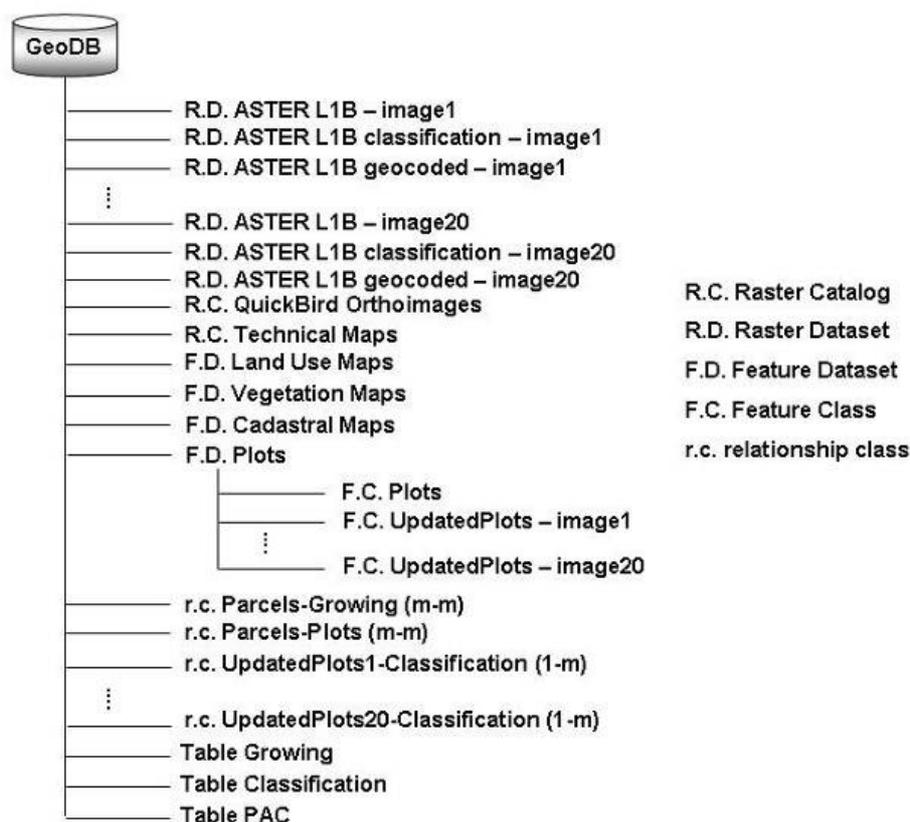


Fig. 7.5 – Struttura del Geodatabase

- i dati raster sono stati archiviati nei Raster Datasets (qualsiasi formato raster organizzato in bande) o in Raster Catalog (insieme di Raster Datasets definiti in formato tabella, nella quale ciascuna riga rappresenta un raster datasets)

- i dati vettoriali sono stati archiviati come Feature Classes in Feature Datasets (relativi alla stessa porzione di territorio e con lo stesso sistema di coordinate)
- le relazioni tra gli oggetti delle Feature Classes e le tabelle, sia le tabelle contenenti gli attributi e che le tabelle realizzate con Access, sono state create tramite Relationship Classes.

La struttura del Geodatabase è mostrata in figura 7.5

Considerazioni finali

Nel presente lavoro è stato messo a punto un processo di elaborazione e predisposizione dei dati necessari alla creazione di un GIS per la gestione delle colture. In particolare si è voluto definire una procedura di integrazione tra tecniche di classificazione e di segmentazione *object-oriented* per un aggiornamento veloce e continuo e quanto più possibile automatizzato delle informazioni relative alle coltivazioni.

In merito all'estrazione automatica degli appezzamenti, al fine di verificare l'attendibilità dei risultati, sarebbe necessaria una verifica dell'accuratezza degli stessi mediante confronto con una mappa di riferimento contenente l'esatta delimitazione degli appezzamenti individuabili nell'immagine. In ogni caso, la procedura di segmentazione descritta permette di estrarre in tempi rapidi (nel caso dell'intera scena dell'immagine QuickBird si parla di circa 10-15min) e con una buona approssimazione i perimetri dei terreni, semplificando in tal modo le successive operazioni di editing manuale da parte dell'utente.

Il principale limite della procedura testata è legato all'esito dell'accuratezza della classificazione dell'immagine Aster.

Conclusioni

Dopo la descrizione iniziale degli strumenti necessari per lo svolgimento dell'indagine preposta, il Telerilevamento Multispettrale, il sensore Aster e i metodi di classificazione, nei capitoli della Seconda e Terza parte della Tesi, sono state affrontate, discutendole e in buona parte risolvendole, le diverse problematiche riscontrate durante lo studio finalizzato al riconoscimento di colture agricole mediante dati satellitari e all'ideazione di una procedura semplice per la sua realizzazione. In particolare l'approfondimento relativo ai dati, alla loro natura e reperibilità, ha messo in luce alcune difficoltà, in termini di disponibilità delle informazioni, di tempi e di capacità professionali, non note al momento di iniziare il lavoro: la stesura dell'elaborato ha quindi rappresentato un'ottima occasione per analizzare nel dettaglio tutti quegli aspetti della ricerca che spesso in letteratura sono erroneamente ritenuti scontati o richiamati con superficialità. La metodologia via via impostata, e della quale si è riportata una schematizzazione semplificata, ha permesso infatti di valutare il problema secondo più punti di vista, appartenenti alle diverse materie di studio coinvolte, in modo da non trascurare quei particolari interdisciplinari che in passato sono stati un notevole limite nell'ambito della classificazione delle coperture del suolo.

E' proprio all'esame dettagliato di ogni passaggio della procedura prevista per questo tipo di indagini, che sono in gran parte da attribuire i soddisfacenti risultati ottenuti nella discriminazione delle colture presenti sul terreno. I principali problemi incontrati, le soluzioni proposte e le raccomandazioni per il loro superamento, sono di seguito riassunti.

Innanzitutto, indipendentemente dai metodi utilizzati, è necessario disporre di una documentazione relativa all'uso del suolo che sia aggiornata e con un livello di dettaglio compatibile con quello delle indagini effettuate. Questa documentazione è fondamentale sia per la fase di addestramento del classificatore utilizzato sia per la verifica della qualità dei risultati ottenuti. Si ricorda però che, mentre le diverse cartografie sono reperibili con una certa facilità perché realizzate da enti pubblici, i dati relativi alle colture agricole sono raccolti e archiviati presso associazioni private. Per poter consultare queste informazioni quindi è necessario prendere accordi con l'associazione stessa motivando anche gli obiettivi del lavoro e discutendo il loro utilizzo, per esempio nell'ambito di un contesto quale la ricerca scientifica e lo sviluppo di prototipi di interesse reciproco. I dati raccolti sono spesso di natura sensibili (privati e fiscali, sulle dichiarazioni dei coltivatori) e quindi deve essere garantita una certa riservatezza, a partire dal nome stesso dell'azienda agricola se questa non autorizza un ufficiale coinvolgimento. Per tutti questi motivi, il contatto e la collaborazione con un'agenzia che tutela gli interessi degli agricoltori possono costituire un grosso limite. Per chi opera

invece all'interno dell'ente, privato e proprietario dei dati, la visione di tali documenti non rappresenta un problema nella procedura ideata.

Altro aspetto complesso è quello relativo alla scelta dei dati da utilizzare per il riconoscimento dell'uso del suolo. Alcuni sensori possono essere maggiormente idonei di altri ma la loro accessibilità potrebbe essere limitata ad un arco temporale che è quello di vita media del satellite. Nel caso del sensore ASTER, scelto per le applicazioni descritte in questa tesi, va ricordato che la disponibilità dei dati potrebbe cessare da un momento all'altro, considerando che la vita prevista dai progettisti è già terminata. Andrebbe allora valutata anche l'esportabilità dei metodi visti ai dati di altri sensori, simili per risoluzione geometrica e spettrale. Questo aspetto rappresenta una ulteriore complicazione dell'indagine proposta. Va sottolineato però che una semplificazione del problema deriva dalla previsione che il futuro più prossimo delle indagini multispettrali è legato principalmente, anche se non esclusivamente, alla continuazione della missione Landsat: infatti, essendo i predecessori di ASTER, questi satelliti acquisiscono dati sui quali sono stati implementati e testati gran parte dei software di classificazione delle immagini, quindi non dovrebbero esserci delle difficoltà nella gestione dei loro formati. In ogni caso deve essere accertato che le bande delle nuove immagini, Landsat o meno, e il loro numero siano adeguate all'elaborazione richiesta.

Una volta ammessa l'esistenza e l'idoneità delle immagini satellitari si è visto che, contemporaneamente al problema del recupero dei dati, si presenta quello della loro disponibilità in periodi dell'anno utili per gli studi sulle colture, in considerazione della stagionalità delle specie, del loro ciclo di maturazione, del momento in cui si effettuano le raccolte e di tutte le altre variabili che possono essere legate allo specifico territorio in esame. La situazione ideale sarebbe quella in cui vengono programmate una serie di acquisizioni, da parte del sensore, a copertura dell'area di interesse, nel momento di maggior efficacia dello studio. E' questa un'operazione che non ha costi molto elevati ma si possono presentare rischi legati alla concomitanza di altri ordini o di esigenze diverse nell'uso del satellite da parte dell'Agenzia proprietaria. Si è visto comunque, nell'ambito di lavoro della tesi, che questa programmazione non è sempre semplice, anche se tuttavia l'archivio dati ASTER nella zona di interesse è piuttosto ricco e i dati sono acquisiti in periodi interessanti per le analisi proposte.

Deve essere scelto poi il metodo migliore per classificazioni finalizzate al riconoscimento delle colture. I due classificatori testati, per pixel e object-oriented, evidenziano prestazioni diverse a livello globale dell'immagine, considerando cioè tutte le tipologie di coperture presenti a terra oltre le coltivazioni, ma molto più simili tra loro nella discriminazione delle firme spettrali delle sole colture. In ogni caso quella pixel-oriented sembra essere più semplice ai fini dell'ingegnerizzazione

del metodo. La classificazione ad oggetti, infatti, prevede una fase aggiuntiva di segmentazione che deve essere eseguita attentamente e con algoritmi rigorosi e ripetibili. Ovviamente un confronto tra metodi pixel-based e object-based ha un senso relativo e deve essere vista in funzione dei dati particolari che sono stati utilizzati: quelli del sensore ASTER. E' evidente, infatti, che la scelta di immagini satellitari a maggiore risoluzione geometrica e, normalmente, minore risoluzione spettrale, potrebbe spostare la scelta verso i metodi pixel-based. Entrambi hanno mostrato un buon livello di accuratezza nei risultati finali: il calcolo di parametri statistici e l'utilizzo di test statistici hanno permesso il confronto quantitativo tra le due metodologie.

Nell'arco di questo lavoro di ricerca, quindi, il percorso di analisi dei dati è stato affrontato interamente, dalla scelta dei dati alla valutazione dell'affidabilità raggiunta nell'analisi finale.

Un altro aspetto da considerare è quello relativo ai software che sono stati utilizzati nelle indagini sul monitoraggio del territorio. In particolare essi rientrano tra quelli commerciali e maggiormente utilizzati nella classificazione: Idrisi 32 ed eCognition. La scelta del primo è dovuta al numero consistente di algoritmi che prevede per la classificazione per pixel, suddivisi in hard e soft: in studi precedenti infatti, efficaci ricostruzioni dell'uso del suolo e caratterizzazioni delle acque interne e costiere, sono state ottenute con l'integrazione sinergica dei risultati provenienti dall'applicazione di tutte le tecniche disponibili. In sede di tesi di Dottorato poi, si è deciso di ricorrere, nell'ottica di costruire una procedura semplice e lineare, ad un unico algoritmo, quello della Massima Verosimiglianza che ha generato le classificazioni più affidabili. Il ricorso ad eCognition invece è stato quasi obbligato, perché attualmente è uno dei pochi software dedicato alla segmentazione e alla classificazione object-oriented. Il costo elevato delle licenze di entrambi i programmi, però, poco si presta alla realizzazione di una metodologia accessibile a qualsiasi tipo di utente: per tale motivo è stata condotta una ricerca, parallela alla sperimentazione, rivolta all'individuazione di software scaricabili gratuitamente da Internet, paragonabili per caratteristiche e capacità ai due commerciali sopra citati. Sono stati identificati come possibili sostituti Multispec (realizzato da School of Electrical and Computer Engineering, Purdue University – Indiana) e Spring 4.0 (National Institute for Space Research, San Jose dos Campos, Brasilia). Mentre Multispec prevede solo i tre tradizionali algoritmi presenti ormai da tempo in letteratura, ovvero Parallelepipedo, Minima Distanza dalla Media e Massima Verosimiglianza, il software Spring 4.0 permette di segmentare le immagini telerilevate, anche se sulla base di un numero inferiore di parametri, riducendo quindi la quantità di informazioni discriminanti i vari poligoni. E' tuttora in fase di studio l'applicazione dei due programmi ai dati ASTER, in particolare devono essere risolti alcuni problemi di compatibilità di formato all'interno di Spring 4.0.

Infine, come già riportato nei diversi capitoli, l'ottimizzazione dei risultati si può ottenere con la collaborazione tra due figure professionali, l'esperto di telerilevamento e l'agronomo, che, in funzione della propria esperienza, sono in grado insieme di studiare tutti i problemi descritti, valutandone i diversi aspetti interdisciplinari che hanno rappresentato un certo limite durante la stesura dell'elaborato di Tesi. Quindi, se la procedura, una volta rifinita nei suoi punti critici, dovesse essere adottata da un ente o una associazione il cui personale non è esperto nel trattamento di immagini satellitari, deve essere prevista una consulenza tecnica continua o almeno una fase di addestramento.

Nonostante tutte le difficoltà messe in luce, si può affermare che esistono le basi per ottenere una classificazione delle colture agricole soddisfacente, ma non si può prescindere da tutte le variabili in gioco: una procedura veloce ed efficace a disposizione di enti come l'Unione Agricoltori di Ravenna, può essere realizzata, e questa Tesi vuole esserne un suggerimento concreto, ma non si può ridurre di certo l'insieme di tutti gli aspetti analizzati ad una elementare successione di passaggi automatizzati che non richiedano l'intervento e l'esperienza dell'operatore.

Bibliografia

- Abrams M. 2007. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) on NASA's Terra platform. *Rivista Italiana di Telerilevamento*, 39, 9-18.
- Barbarella M., De Giglio M., Mancini F. 2006. Monitoraggio del territorio rurale nell'area emiliano-romagnola mediante dati Aster. CITIMAP: il Telerilevamento per un'agricoltura sostenibile. Piacenza 20 aprile 2006, ATTI su Cd-Rom. ISBN 88-902430-1-5.
- 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. Definiens Imaging GmbH Trappentreustrasse 1, München, Germany.
- Baatz M., Schäpe A. 1999. Object-oriented and multi-scale image analysis in semantic networks. Proceedings of 2nd International Symposium on operationalization of Remote Sensing. ITC, Enschede.
- Bins L.S.A., Fonseca L.M., Erthal G.J., Ii F.M. 1996. Satellite Image Segmentation: a region growing approach. VIII Simposio Brasileiro de Sensoramento Remoto, Salvador, Bahía, Brasil.
- Bishop Y.M.M., Fienberg S.E., Holland P.W. 1975. *Discrete Multivariate Analysis, Theory and Practice*, The MIT press, 555 pp.
- Bitelli G. 2007. ASTER, dati geometrici e generazione di DEM. *Rivista Italiana di Telerilevamento*, 39, 45-58.
- Blaschke T., Land S., Lorup E., Strobl J., Zeil P. 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. *Environmental Information for Planning, Politics and the Public*, Metropolis Verlag, Marburg, Volume II.
- Bonciarelli F. 2001. *Coltivazioni erbacee da pieno campo*. Calderoni Ed agricole.
- Brivio P.A., Lechi G., Zilioli E. 2006. *Principi e metodi di Telerilevamento*, CittàStudi Edizioni, Torino.
- Burrough P.A., 1986. "Principles of Geographical Information Systems for Land Resources Assessment (Monographs on Soil & Resources Survey)", Oxford University Press.
- Burrough P.A., Rachael A., 1998. "Principles of Geographical Information Systems" Second Edition, Oxford University Press.

- Câmara G., Souza R.C.M., Freitas U.M., Garrido J. 1996. SPRING: Integrating remote sensing and GIS by object-oriented data modelling, *Computers & Graphics* 20: (3)395-403.
- Chirici G, Corona P. 2006. Immagini satellitari ad alta risoluzione nel rilevamento delle risorse forestali. Aracne Editrice S.r.l, Roma.
- Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurements*, 20, 37-46.
- Cohen J. 1968. Weighted Kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4), 213-220.
- Congalton R.G., Oderwald R.G., Mead R.A. 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing*, 49(12), 1671-1678.
- De Giglio M. 2004. Utilizzo del sensore satellitare Aster per l'analisi dell'uso del suolo nell'area emiliano-romagnola. Tesi di laurea, Ingegneria per l'Ambiente e il Territorio, Università degli Studi di Bologna.
- Donner A., Shoukri M.M., Klar N., Bartfay E. 2000. Testing the equality of two dependent kappa statistics. *Statistics in Medicine* 19, 393-387.
- Eastman J.R., 1999, IDRISI 32 – Guide to GIS and Image Processing, Volume 1 and Volume 2, Clark Labs, Clark University, Worcester (MA).
- Espindola G. M., Camara G., Reis I.A., Bins L.S., Monteiro A.M. 2006. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *International Journal of Remote Sensing*, 27, 3035-3040.
- Everitt B.S. 1968. Moments of the statistics kappa and weighted kappa. *British Journal of Mathematical and Statistical Psychology*, 21, 97-103.
- Fleiss J.L., Cohen J., Everitt B.S. 1969. Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 72, 323-327.
- Foody G.M. 2004. Thematic Map Comparison: Evaluating the Statistical Significance of differences in Classification Accuracy. *Photogrammetric Engineering & Remote Sensing* 70, 627-633.
- Gao Y. Mas J.F., Maathuis B.H.P., Xiangmin Z., Van Dijk P.M. 2006. Comparison of pixel-based and object-oriented image classification approaches – a case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing*, 27(18), 4039-4055.

- Gao Y., Mas J.F., Navarrete J.A., Palacio J.L., García-Mora T.J. 2007. Image classification accuracy improvement with a region based segmentation. 32nd International Symposium on Remote Sensing of Environment, San José, Costa Rica, 25-29 giugno 2007.
- Gao Y., Mas J.F., Navarrete J.A., Palacio J.L., García-Mora T.J. 2007. Image segmentation optimization and the effect on classification accuracy. 32nd International Symposium on Remote Sensing of Environment, San José, Costa Rica, 25-29 giugno 2007.
- Gomasca M.A. 1997. Introduzione a telerilevamento e GIS per la gestione delle risorse agricole ambientali. Edizioni AIT, Firenze.
- Gusella, L., Bitelli, G., Mognol, A., 2005. Determinazione dei parametri di segmentazione in immagini aeree e telerilevate: un approccio semantico all'ottimizzazione. Atti Convegno SIFET, "Integrazione tra le tecniche innovative del rilievo del territorio e dei beni culturali", 2005, Palermo; CD-Rom, 8 pagg., ISBN 88-901939-1-3.
- Herold M., Mueller A., Guenter s., Scepan J. 2002. Object oriented mapping and analysis of urban use/cover using Ikonos data. Proceedings of the 22nd EARSEL Symposium Geoinformation for European wide Integrating, Prague.
- Hudson W.D., Ramm C.W. 1987. Correct formulation of the kappa coefficient of agreement. Photogrammetric Engineering and Remote Sensing, 53, 421-422.
- Italeco S.p.A Gruppo IRI. 1989. Se il bosco muore. Compagnia Generale Riprese aeree/Editore.
- Landgrebe D.A. 2003. Signal theory methods in multispectral remote sensing. Wiley, 508.
- Lillesand T.M., Kiefer R.W., Chipman J.W. 2004. Remote sensing and image interpretation (fifth edition). Wiley, 763 pp.
- Mognol A. 2007. Analisi di immagini aerofotogrammetriche e telerilevate per la caratterizzazione qualitativa e quantitative di ecosistemi forestali e della loro evoluzione. Tesi di Dottorato in Scienze Geodetiche e Topografiche. Università di Bologna.
- Piccoli F., Corticelli S., Dell'Acquila L., Pellizzari M. 1996. Vegetation map of the Regional Park of the Po Delta (Emilia Romagna Region). Allionia, 34, 325-331.
- Rossiter D.G. 2004. Technical note: statistical methods for accuracy of classified thematic maps. Department of earth system analysis, International Institute for geo-information science & Earth observation (ITC), Enschede (NL), 1-46.
- Verrucci E. 2005. Uso di immagini multispettrali ASTER per la classificazione delle colture presenti nell'area emiliano-romagnola. Tesi di Laurea. Università di Bologna.