

Metodi di classificazione su dati multitemporali a media risoluzione geometrica per l'analisi delle trasformazioni di uso dei suoli

William Cerull, Benedetto Figorito, Eufemia Tarantino

Politecnico di Bari, Via Orabona 4, 70125 Bari, Tel./Fax 0805963417
E-mail: william.cerull@gmail.com; figorito@ba.issia.cnr.it; e.tarantino@poliba.it

Riassunto

I dati satellitari costituiscono un insostituibile ausilio nel contesto del monitoraggio del territorio; effettuare analisi sempre più accurate, orientate alle trasformazioni di uso dei suoli, risulta di fondamentale importanza per ottenere risultati affidabili.

In questo studio sono stati sviluppati e confrontati differenti approcci di classificazione supervisionata di uso dei suoli (*Parallelepiped*, *Maximun Likelihood*, *Decision Tree* e *Feature Extraction*) applicati a dati Landsat TM5 dell'area del Parco Nazionale dell'Alta Murgia (Puglia), al fine di verificarne la riproducibilità nell'ambito di analisi di *change detection*.

Abstract

Satellite data are irreplaceable aid in the context of territorial monitoring; executing more accurate analysis for land use transformation is of fundamental importance to obtain reliable classification results. In this study different supervised land use classification approaches (*Parallelepiped*, *Maximun Likelihood*, *Decision Tree* e *Feature Extraction*) were developed and compared applied on Landsat TM5 data on the "Alta Murgia" national park in order to evaluate reproducibility for further change detection analysis.

Introduzione

In Italia la superficie sottoposta a tutela è circa il 5% del territorio nazionale: 772 aree protette inserite nell'elenco ufficiale del Ministero dell'Ambiente, di cui 22 parchi nazionali iscritti nell'Elenco Ufficiale delle Aree Protette (EUAP) e 105 parchi naturali regionali.

Nel corso degli anni abbiamo assistito a un'evoluzione del ruolo e delle funzioni delle aree protette: da una visione di rigida "conservazione e protezione dei valori ambientali", che limitava qualsiasi attività antropica, a una concezione cosiddetta di "tutela attiva" delle aree protette (conservazione e sviluppo sostenibile) (Roselli, 2005). Da sempre le campagne pugliesi offrono un paesaggio suggestivo con aree che usano ancora sistemi agricoli gestiti in maniera estensiva, caratterizzate da una forte naturalità e una ricca varietà floristica, e che fungono da ambienti ideale e fonte trofica per molte specie che vi trovano rifugio. Negli ultimi decenni molti di questi ambienti hanno subito notevoli cambiamenti a causa della diffusione di alcune discutibili pratiche agronomiche che hanno comportato notevoli cambiamenti sulla biodiversità, impoverendola, e sul paesaggio, semplificandolo e banalizzandolo (Angiuli et al, 2007). Lo spietramento e la macinazione del suolo, il diserbo chimico, l'impiego dell'acqua salmastra per l'irrigazione, la mancata rotazione colturale e il ricorso ad altre tecniche stanno avviando alcune aree alla desertificazione (Canora et al., 2003; Lopez et al., 2005).

Il dato sull'uso del suolo risulta essere tra le informazioni che meglio caratterizzano un territorio e allo stesso tempo fornisce una base per la formulazione di strategie di gestione sostenibile e di tutela dell'ambiente. Scaturisce pertanto l'esigenza di sperimentare tecniche di monitoraggio che consentano di ottenere dati attendibili sulle trasformazioni del territorio con il grado di precisione richiesto, e possano essere agevolmente replicabili per applicazioni in campo operativo. Il

telerilevamento offre quindi, oltre ai benefici di una grande copertura sinottica, la possibilità di usufruire dell'informazione multispettrale mediante tecniche automatiche e semi-automatiche per l'estrazione delle caratteristiche d'uso dei suoli (Caprioli et al., 2006).

Gli algoritmi tuttora a disposizione per la classificazione di immagini digitali sono per la maggior parte di tipo convenzionale (Maximum Likelihood, Parallelepiped algorithm, Spectral Angle Mapper etc...). I classificatori che utilizzano altre informazioni oltre quelle spettrali puntuali possono essere divisi in due classi: contestuali (contextual classifiers) e basati sulle caratteristiche dell'oggetto (object-based classifiers) in contrapposizione con i classificatori convenzionali anche chiamati a pixel (pixel-based) (Tso et al., 2009).

Nel presente lavoro sono implementate analisi di *change detection* dell'uso dei suoli di una porzione di territorio dell'area dell'Alta Murgia, elaborando e confrontando differenti metodologie di classificazione di dati satellitari multi temporali a media risoluzione geometrica.

L'analisi è compiuta sulla base di due immagini telerilevate acquisite dal sensore LANDSAT 5 negli anni 2002 e 2009 su un'area di circa 6500 kmq, comprendente il Parco Nazionale dell'Alta Murgia. La scelta dei dati di riferimento ha tenuto conto della loro confrontabilità, disponibilità e riproducibilità del metodo.

Le scene LANDSAT sono state sottoposte nella prima fase ad un pre-processing di tipo geometrico e radiometrico allo scopo di ridurre gli errori dovuti alle distorsioni introdotte durante l'acquisizione e la trasmissione dei dati. Inoltre è stato applicato l'algoritmo MAD (*multivariate alteration detection*) per omogeneizzare le informazioni e permettere il confronto tra i dati multitemporali.

In seguito, sul dato LANDSAT del 2009 sono stati implementati i metodi di classificazione pixel-based (Parallelepiped algorithm e Maximum Likelihood), decision tree (Math method che utilizza una tecnica di splitting ortogonale) e object-based al fine di verificarne l'affidabilità e riproducibilità orientati alla successiva analisi di *change detection* rispetto al dato del 2002, per determinare l'entità e la qualità del cambiamento di uso dei suoli. Particolare attenzione è stata rivolta al confronto tra i risultati ottenuti mediante l'implementazione delle suddette tecniche sul dato di partenza, per stabilire, eventualmente, se l'una sia preferibile all'altra, in termini di accuratezza, costi, tempi, o per verificare l'indifferenza nella scelta di una specifica metodologia.

Infine, l'accuratezza dei metodi è stata verificata in modo empirico utilizzando una serie di campioni ottenuti da verità a terra riferiti a cartografia tecnica prodotta nei periodi di indagine e documentazione multisorgente relative al territorio indagato.

Metodologia

L'area di studio appartiene al territorio dell'Alta Murgia. Istituito nel marzo 2004, il Parco Nazionale dell'Alta Murgia si estende per circa 68077 ha, ripartiti in tre zone a tutela differenziata (Figura 1). Il Parco Comprende una zona boscata, 11000 ha di bosco ad alto fusto, e una zona a pascolo, una fascia della steppa a stipa, presente ad un'altitudine di circa 500 m, che caratterizza la Murgia Alta e rappresenta l'ultimo esempio di pseudo-steppa mediterranea presente nell'Italia peninsulare ed uno dei più importanti del Mediterraneo. Negli ultimi decenni molti di questi ambienti hanno subito notevoli cambiamenti a causa della diffusione di alcune discutibili pratiche agronomiche che hanno comportato drastici cambiamenti sulla biodiversità e sul paesaggio. Diffusa da sempre nell'area dell'Alta Murgia la pratica della "*spietatura*", ovvero della rimozione delle pietre affioranti dai campi coltivati alla fine di ogni ciclo produttivo, pietre che venivano poi riutilizzate per la costruzione di numerosi manufatti rurali che ancora oggi punteggiano il territorio (trulli, lamie, muretti); negli ultimi anni tale pratica è stata sostituita dallo "*spietramento*", che consiste nella trasformazione dei pascoli in seminativi attraverso la lavorazione profonda del terreno e la frantumazione meccanica della roccia presente. Tale pratica, legittimata dalla legge regionale 54/81 e incentivata dall'erogazione di finanziamenti pubblici, regionali ed europei, si è diffusa a macchia d'olio nell'area murgiana. Oggi i finanziamenti sono chiusi ma la pratica purtroppo, seppure in misura minore, continua e la possibilità di recuperare le aree spietrate richiede tempi molto lunghi ed esiti incerti.

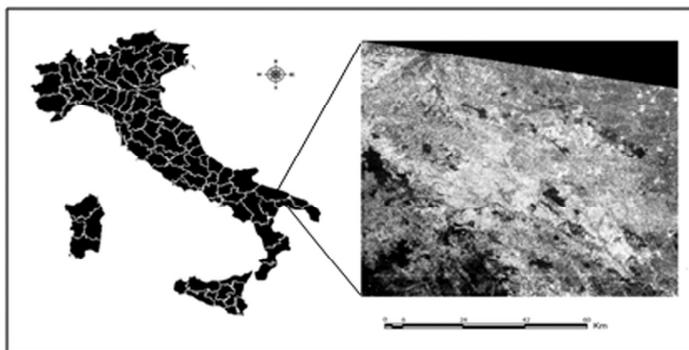


Figura 1. Area di studio relativa al territorio dell'Alta Murgia (Puglia).

L'analisi è stata condotta su due immagini multi spettrali TM5 (acquisition date: 22 june 2002; 27 july 2009) equivalente ad una superficie di circa 6500 km² e relativa ad un'area della Puglia caratterizzata da una forte eterogeneità, soprattutto dal punto di vista morfologico e topografico.

L'analisi spettrale delle 10 classi individuate è stata eseguita selezionando alcune regioni rappresentative che non presentassero significativi cambiamenti nelle due immagini dovute all'intervallo temporale intercorso tra esse. Nonostante tale gap temporale, poiché le due immagini sono state acquisite entrambe in estate, è possibile ritenere valida, con buona approssimazione, che la gran parte delle classi di vegetazione sia caratterizzata dallo stesso stato fenologico.

La strategia di analisi è stata articolata nelle seguenti fasi (Figura 2):

- 1) Correzione geometrica, calibrazione radiometrica, definizione della legenda per la classificazione e acquisizione della documentazione sulle aree di training e di verità a terra;
- 2) Classificazione con i metodi *Parallelepiped*, *Maximun Likelihood*, *Decision Tree* e *Feature Extraction*, implementati sul dato TM5-2009 mediante il software di image proccessing ENVI 4.7.
- 3) Calcolo della matrice di confusione e scelta dell'algoritmo con prestazioni migliori.
- 4) Analisi di change detection sui dati del 2002 e del 2009 e verifica dei risultati ottenuti.

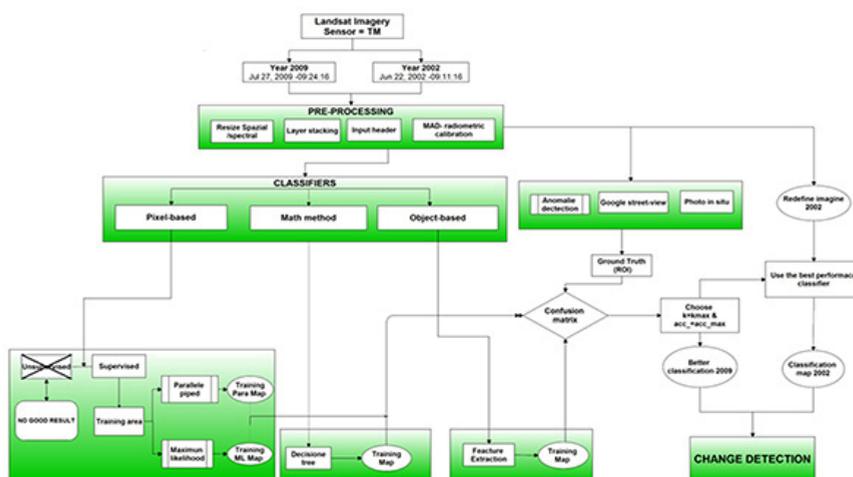


Figura 2. Diagramma di flusso della procedura implementata.

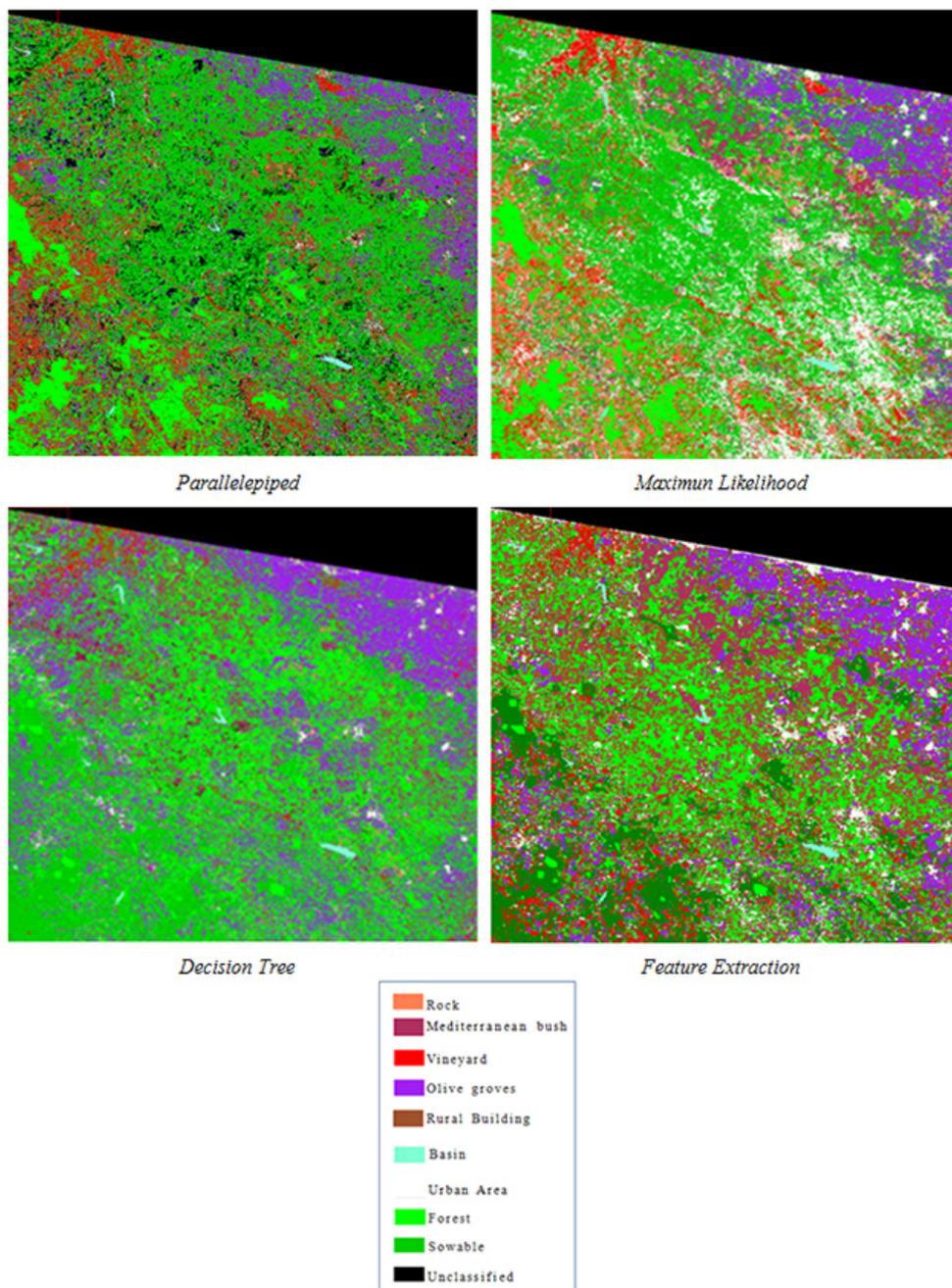


Figura 3. Risultati degli approcci di classificazione applicati al dato Landsat TM5 del 2009.

Risultati e conclusioni

Nelle Tabelle 1a-1b-1c-1d sono raccolti i risultati in termini di accuratezza della classificazione delle quattro metodologie applicate al dato Landsat TM5 del 2009.

L'analisi quantitativa, in congiunzione con quella qualitativa evidenziata nella Figura 3, mostra che il risultato migliore riguarda la classificazione ottenuta con l'approccio *Feature Extraction*, che è stato il metodo prescelto per sviluppare la classificazione del dato del 2002.

La Tabella 3 mostra i cambiamenti individuati durante l'arco di tempo preso come riferimento (Giugno 2002 - Luglio 2009). I risultati sono riportati secondo la superficie di ogni singola classe in kmq. Con l'analisi di *Change detection* è emerso un costante incremento della superficie delle classi agricole (seminativo, uliveto, vigneto). Per lo stesso periodo si ha anche una piccola percentuale di conversione da urbano ad altre coperture, dovute probabilmente a false classificazioni sulla prima data.

La metodologia adottata ha permesso di aggiornare in tempi ridotti anche vaste superfici di territorio. Eventuali false classificazioni, essendo una percentuale molto esigua rispetto al totale, non creano limitazioni all'analisi e a livello di macro-aree non pregiudicano la qualità e localizzazione dei risultati.

Classified map	Ground Truth													
	Basin	Urban area	Forest	Sowable	Rock	Mediterr. bush	Vineyard	Olive groves	Rural building	Total	Prod. Acc.	User Acc.	Commission	Omission
Unclassified	3.39	0.08	5.84	0.59	1.07	0.09	0.00	0.12	0.84	2.93				
Basin	94.30	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	18.74	94.30	99.97	0.00	0.12
Urban area	0.26	97.97	0.00	77.67	17.15	86.15	14.91	98.44	3.38	37.64	97.97	45.72	0.00	0.00
Forest	0.02	0.00	93.08	0.00	0.00	0.00	0.00	0.00	0.00	32.07	93.08	99.99	14.91	90.44
Sowable	0.01	0.20	0.00	20.65	0.11	13.76	4.65	9.39	0.00	4.24	20.65	71.07	0.00	0.00
Rock	0.33	1.72	0.00	0.02	81.46	0.00	56.71	0.00	95.78	3.01	81.46	29.22	4.65	9.39
Mediterranean bush	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vineyard	1.08	0.03	1.06	1.04	0.11	0.00	23.73	0.04	0.00	1.37	23.73	45.77	56.71	0.00
Olive groves	0.01	0.00	0.00	0.03	0.11	0.00	0.00	0.00	0.00	0.01	0.00	0.00	23.73	0.04
Rural building	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Overall Accuracy = 72.53% k = 0.6475														

Tabella 1a. Matrice di confusione ottenuta con l'algoritmo Parallelepiped.

Classified map	Ground Truth													
	Basin	Urban area	Forest	Sowable	Rock	Mediterr. bush	Vineyard	Olive groves	Rural building	Total	Prod. Acc.	User Acc.	Commission	Omission
Unclassified	0.12	0.91	0.06	0.00	0.25	0.00	0.00	0.00	0.99	0.12				
Basin	99.79	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	9.53	99.79	52.86	47.14	0.21
Urban area	0.00	90.93	0.02	2.78	0.25	0.72	0.86	1.00	8.91	6.90	90.93	91.66	6.34	9.07
Forest	0.00	0.00	97.41	0.01	0.00	1.23	1.33	0.00	0.00	41.73	97.41	99.71	0.29	12.59
Sowable	0.00	1.44	0.06	66.66	0.37	2.19	0.56	2.04	0.99	14.39	66.66	95.60	4.40	33.34
Rock	0.09	1.14	0.05	1.40	96.63	0.05	0.73	0.00	9.90	1.66	96.63	72.00	28.00	3.37
Mediterranean bush	0.00	0.50	0.08	23.03	1.50	94.51	0.04	0.27	0.00	10.35	94.51	94.51	47.10	5.49
Vineyard	0.00	1.45	2.70	5.25	0.25	1.20	95.76	1.91	51.49	6.24	95.76	55.38	44.62	4.24
Olive groves	0.00	1.43	0.18	0.85	0.00	0.11	0.73	94.78	5.94	7.02	94.78	93.86	6.14	5.22
Rural building	0.00	0.21	0.00	0.00	0.75	0.00	0.00	0.00	21.78	0.06	21.78	55.00	45.00	78.22
Overall Accuracy = 85.31% k = 0.8030														

Tabella 1b. Matrice di confusione ottenuta con l'algoritmo Maximum Likelihood.

Classified map	Ground Truth													
	Basin	Urban area	Forest	Sowable	Rock	Mediterr. bush	Vineyard	Olive groves	Rural building	Total	Prod. Acc.	User Acc.	Commission	Omission
Unclassified	0.03	5.51	3.29	0.00	0.00	0.13	0.00	0.99	4.00	1.23				
Basin	99.45	0.00	0.01	0.00	0.00	0.00	0.00	7.92	0.00	5.03	99.45	99.72	0.28	0.55
Urban area	0.21	58.28	0.41	0.00	0.04	0.04	0.00	16.83	0.62	5.36	58.28	97.50	2.50	41.72
Forest	0.00	0.62	92.54	1.63	0.00	4.37	9.20	2.93	0.00	45.23	92.54	97.39	2.61	7.46
Sowable	0.00	14.32	6.57	57.72	4.87	75.13	9.71	6.75	3.96	21.56	57.72	55.24	44.76	42.28
Rock	0.00	1.55	0.00	1.25	72.91	0.00	0.21	0.00	13.86	1.33	72.91	67.91	32.09	27.09
Mediterranean bush	0.18	6.02	0.01	23.52	0.00	19.54	0.39	2.42	2.97	6.73	19.54	16.83	83.17	80.46
Vineyard	0.00	0.19	0.73	3.38	0.00	0.11	74.58	0.18	49.50	3.85	74.58	69.97	30.03	25.42
Olive groves	0.12	13.47	0.16	8.79	0.87	0.85	5.73	87.67	3.96	9.47	87.67	64.34	35.66	12.33
Rural building	0.00	0.05	0.00	0.00	16.73	0.00	0.00	0.00	0.00	0.21	0.00	0.00	100.00	100.00
Overall Accuracy = 77.00% k = 0.6813														

Tabella 1c. Matrice di confusione ottenuta con l'algoritmo Decision Tree.

Classified map	Ground Truth													
	Basin	Urban area	Forest	Sowable	Rock	Mediterr. bush	Vineyard	Olive groves	Rural building	Total	Prod. Acc.	User Acc.	Commission	Omission
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Basin	99.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.03	99.63	100.00	0.00	0.37
Urban area	0.12	77.91	0.08	1.58	1.87	0.00	0.34	0.20	0.00	7.41	77.91	94.35	5.65	22.09
Forest	0.09	0.29	99.78	0.03	0.00	4.37	3.47	2.42	0.00	48.08	99.78	98.79	1.21	0.22
Sowable	0.03	13.55	0.01	78.11	3.00	10.35	3.25	10.56	3.95	19.43	78.11	82.94	17.06	23.89
Rock	0.00	0.65	0.00	0.00	89.51	0.00	0.39	0.00	31.68	1.23	89.51	90.08	9.92	10.49
Mediterranean bush	0.12	0.55	0.12	7.23	0.00	69.65	0.09	3.15	0.00	5.52	69.65	66.85	33.15	36.35
Vineyard	0.00	0.19	0.01	5.14	0.00	0.64	91.66	18.17	49.50	5.77	91.66	57.35	42.65	8.34
Olive groves	0.00	0.00	0.00	0.22	4.12	20.42	0.81	65.19	14.85	5.86	65.19	77.28	22.72	34.81
Rural building	0.00	0.83	0.00	7.69	1.50	0.00	0.00	0.00	0.00	1.68	0.00	0.00	100.00	100.00
Overall Accuracy = 88.30% ; k = 0.8352														

Tabella 1d. Matrice di confusione ottenuta con l'algoritmo Feature extraction.

	Parallelepiped	Mahimum Likelihood	Decision Tree	Feature Extraction
Overall Accuracy	72.53	85.31	77.00	88.30
K	0.65	0.80	0.68	0.86

Tabella 2. Sintesi delle accuratze ottenute mediante l'implementazione dei quattro metodi di classificazione sul dato TMS – 2009.

		Initial State 2002										
		Basin	Urban area	Forest	Sowable	Olive groves	Vineyard	Rock	Mediterr. bush	Unclassified	Row Total	Class Total
Final State 2009	Unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	790.28	790.28	790.28
	Basin	12.95	0.06	0.00	0.22	0.02	0.06	0.00	0.02	0.00	13.33	13.33
	Urban area	0.22	83.90	0.53	238.95	37.67	9.29	4.37	13.59	0.00	388.50	388.50
	Forest	0.24	9.60	616.27	168.41	124.80	491.82	0.07	186.07	0.00	1597.27	1597.27
	Sowable	0.06	212.68	6.00	2075.42	432.64	55.52	9.94	258.26	0.00	3050.52	3050.52
	Olive groves	0.04	54.30	10.55	385.01	513.23	48.23	0.07	164.97	0.00	1176.40	1176.40
	Vineyard	0.06	9.21	54.36	115.97	17.86	312.70	2.83	39.46	0.00	552.46	552.46
	Rock	0.00	2.90	0.00	2.69	0.07	0.04	10.96	0.13	0.00	16.79	16.79
	Mediterr. bush	0.00	32.26	0.11	210.06	112.24	1.44	0.35	133.34	0.00	489.80	489.80
	Class Total	13.58	404.91	687.81	3196.73	1238.53	919.09	28.58	795.83	790.28	0.00	0.00
	Class Changes	0.63	321.02	71.54	1121.30	725.30	606.39	17.62	662.50	0.00	0.00	0.00
	Image Difference	-0.25	-16.41	909.45	-146.21	-62.14	-366.63	-11.79	-306.03	0.00	0.00	0.00

Tabella 3. Matrice di change detection calcolata in termini di kmq sui dati Landsat TMS del 2002 e del 2009.

Riferimenti bibliografici

- Angiuli, A., Brunetti G., Ciola G., Iannone R. F. (2007), "Il riutilizzo dei fanghi e delle acque reflue in agricoltura". Mario Adda.
- Canora F., Ferrigno L., Fidelibus M.D., Spilotro G., Straziuso K. (2003), "Modificazioni antropiche delle tessiture dei suoli carsici dell'Alta Murgia e finalizzazione agricola: implicazioni idrogeologiche". *Riassunti FIST*, Bellaria 2003, 476-478.
- Caprioli M., Tarantino E. (2006), "Identification of Land Cover Alterations in the Alta Murgia National Park (Italy) with VHR Satellite Imagery", *Int. J. of Sustainable Development and Planning*, 1(3), 261-270.
- Lopez R., Miano T.M. (2005), "Dissodamenti e spietramenti di suoli ricadenti nel Parco nazionale dell'Alta Murgia: elaborazioni geostatiche ed indicatori di stato". *Atti Società italiana della Scienza e del Suolo*, Bari 21-24 giugno 2005, 419-425.
- Roselli, L. (2005), "Sviluppo rurale e agricoltura sostenibile nelle aree protette. Il caso del Parco Nazionale dell'Alta Murgia", *Tesi di dottorato*, Dipartimento di economia e politica agraria, Università degli Studi di Napoli "Federico II", 2005.
- Tso B., Mather P (2009), "Classification methods for remotely sensed data", CRC Press.