

Estrazione di informazioni agronomiche tramite l'analisi dell'andamento temporale di indici vegetazionali per la mappatura automatica di aree agricole

Alberto Crema (*), Giacinto Manfron (*,**), Mirco Boschetti (*), Roberto Confalonieri (**)

(*) Consiglio Nazionale delle Ricerche (CNR), Istituto per il Rilevamento Elettromagnetico dell'Ambiente (IREA)
Via Bassini 15, 20133 Milano, crema.a@irea.cnr.it, manfron.g@irea.cnr.it, boschetti.m@irea.cnr.it

(**) Università degli Studi di Milano (UNIMI), Dipartimento di Scienze Agrarie e Ambientali, Produzione, Territorio,
Agroenergia (DISAA), Via Celoria 2, 20133 Milano, roberto.confalonieri@unimi.it

Riassunto

La produzione di informazioni agronomiche circa lo stato dei sistemi colturali è oggetto di indagine scientifica da parte di diverse discipline. In particolare il contributo offerto dal telerilevamento riguarda la possibilità di supportare il monitoraggio continuo dell'agroecosistema, fornendo su larga scala informazioni spazialmente distribuite inerenti la localizzazione delle colture, la dinamica stagionale ed il loro stato fitosanitario. L'approccio alla caratterizzazione delle aree agricole, volta all'identificazione delle differenti colture, richiede molto spesso informazioni esterne non sempre disponibili e presenti solamente in forma di statistiche aggregate per unità amministrative. Il telerilevamento satellitare, ed in particolare l'analisi multi-temporale, può rappresentare un utile strumento per l'estrazione di informazioni agronomiche di dettaglio ed aggiornate per sistemi con forte variabilità. In questo lavoro si è voluto approfondire la possibilità di mappare ed estrarre informazioni fenologiche di colture risicole in maniera automatica, proponendo un algoritmo in grado di ricavare questo tipo di informazioni valutando l'andamento temporale di indici spettrali ricavati da immagini satellitari MODIS. Sono state acquisite immagini del sensore relative ai principali distretti risicoli nel Nord Italia per gli anni 2005, 2006, 2007. I risultati hanno dimostrato come la dinamica di indici spettrali quali l'*Enhanced Vegetation Index* (EVI) e il *Land Surface Water Index* (LSWI), possano essere utilizzati per distinguere superfici interessate da risicoltura da aree caratterizzate da vegetazione spontanea o altre colture agricole. Un accurato lavoro di analisi della dinamica del segnale è stato inoltre in grado di identificare specifiche informazioni fenologiche della coltura. Il presente contributo fornisce un confronto tra i risultati di mappatura ottenuti con la metodologia proposta con una metodologie precedentemente applicata ed una già nota in letteratura, discute inoltre le potenzialità e i limiti dell'approccio proposto.

Abstract

The production of agricultural information about the status of cropping systems is being investigated by several scientific fields of research. In particular, the contribution of remote sensing can provide continuous monitoring of the agro-ecosystem, allowing the use of large-scale spatially distributed information regarding the location of crops, the seasonal dynamics and their state of health. The approach to the automatic mapping of agricultural areas, aimed at identifying the different crops, often needs external information not always available and when present often in the form of statistical averages. Satellite remote sensing, thanks to the multi-temporal monitoring, can be a useful tool for extracting update agricultural information. In this study we wanted to investigate the possibility of automatically extracting information of rice crop extension and phenology defining an algorithm that uses time series of satellite images. MODIS sensor data of the rice-growing areas of northern Italy were analyzed to derive time series of spectral indices EVI and

LSWI. The study highlighted how information on crop extension and phenology can be extrapolated from a dynamic analysis of spectral indices time series. We noticed how the dynamics of some indices, allow to differentiate rice crop from other natural and agronomical features. Secondly a careful analysis of the signal, is also able to specifically identify phenological crop information of the crop during the on-going season. This paper shows also an application of the methods proposed in the literature called analysis on the “Pareto Boundary”.

Introduzione

Il presente lavoro si concentra sulla elaborazione e sulla valutazione di un algoritmo per la mappatura di aree risicole e la restituzione di mappe fenologiche della coltura, tramite dati ottici a media risoluzione rilevati da sensore *Moderate Resolution Imaging Spectroradiometer* (MODIS) a bordo del satellite Terra. L'importanza delle colture risicole si testimonia infatti anche dall'elevato numero di metodi che sono stati proposti per la mappatura e il monitoraggio di questa risorsa sia a scala locale che globale. La conoscenza sin qui acquisita, ha permesso di focalizzare l'attenzione su un nuovo approccio alla base della metodologia qui proposta. Per la definizione del metodo si sono precedentemente valutate le performance di approcci per la mappatura automatica di colture risicole proposti in letteratura (Xiao et al., 2005; Peng et al. 2011) testandoli nelle diverse condizioni sperimentali delle risicoltura del nord Italia. In particolare, il riscontro di una limitata capacità di adattamento del metodo di Xiao et al 2005 (Crema et al, 2011) nel passaggio tra l'applicazione dell'algoritmo da zone tropicali a zone temperate caratterizzata dall'alta presenza di errori di commissione, ha fornito l'opportunità di definire un approccio alternativo. L'approccio qui proposto è stato pensato per analizzare l'andamento di indici spettrali basandosi sullo studio della dinamica di questi dati durante una stagione agronomica. Tale approccio, si è rivelato in grado di prestarsi bene per derivare informazioni di tipo fenologico, fornendo oltre che un metodo di riconoscimento delle colture anche uno strumento di monitoraggio (Eklund et al., 2004; Boschetti et al, 2009). Obiettivo di questo studio è proporre un approccio di mappatura alternativo e confrontare le sue performance rispetto a metodi precedentemente proposti comparandone i risultati con una cartografia tematica di riferimento.

Area di studio

In ambito mediterraneo l'Italia risulta essere il principale produttore di riso con la maggiore superficie investita alla coltivazione di questo cereale. Il riso in Italia è coltivato principalmente in un distretto agricolo compreso tra il Piemonte e la Lombardia (Bocchi et al., 2003). Come per gli altri cereali primaverile-estivi è presente un unico ciclo produttivo all'anno, che avviene durante un periodo che va dalla semina che si effettua in Aprile alla raccolta che generalmente si effettua a partire da Settembre ma che varia a seconda delle varietà. L'area di studio (Figura 1) abbraccia i principali distretti risicoli del Nord Italia e particolarmente intende focalizzarsi tra le province di Milano, Pavia, Vercelli e Novara dove sono investiti circa 250.000 ha e dove viene prodotto il 90% del riso coltivato in Italia, risultando il più grande distretto risicolo in Europa.

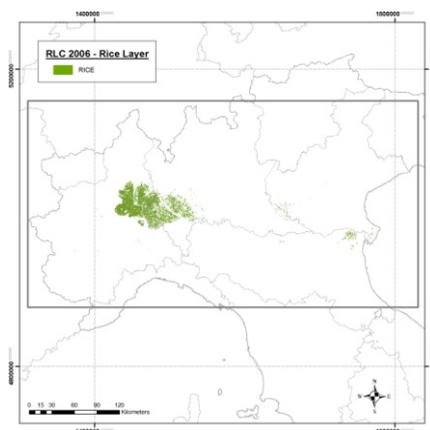


Figura 1. L'areale di studio si focalizza sui principali distretti risicoli del Nord Italia. In verde è evidenziata l'estensione risicola secondo il layer cartografico regionale, in grigio i confini regionali italiani.

Materiali

Dati MODIS

Nel presente lavoro sono state impiegate immagini satellitari di riflettanza rilevate dal sensore *Moderate Resolution Imaging Spectroradiometer* (MODIS) presente a bordo del satellite Terra (EOS). Il sensore MODIS misura in 36 bande spettrali, le prime 7 sono state designate in particolare per lo studio della vegetazione e della superficie terrestre. Le prime due bande (rosso a 620-670 nm e NIR a 841-876 nm) sono acquisite a 250 m, mentre le successive 5, blu (459-479 nm), verde (545-565 nm), NIR (1230-1250 nm) e SWIR (1628-1652 nm, 2105-2155 nm) sono acquisite a 500 m. Tutte le altre bande hanno una risoluzione di 1000 m (Lillesand et al, 2000). Questi dati sono liberamente scaricabili dal portale *United State Geological Survey* (USGS), (<http://www.usgs.gov/>). Il prodotto MODIS considerato in questo studio prende il nome di MOD09A1 ed è il risultato di un composite temporale volto a selezionare la migliore osservazione di riflettanza registrata sulla data superficie in un arco temporale di 8 giorni.

Cartografia di riferimento

Per la valutazione dell'accuratezza delle mappe tematiche risultanti dall'algoritmo proposto è stato utilizzata una cartografia di copertura del suolo unendo 4 prodotti cartografici regionali, reperiti dai portali cartografici di riferimento delle regioni Piemonte, Lombardia, Veneto ed Emilia-Romagna. Questo prodotto, ad una maggiore risoluzione spaziale rispetto a quella dei dati MODIS, ha una scala di 1:10.000 e un dettaglio tematico articolato in 16 differenti classi tra cui quella riferita alle coltivazioni di riso.

Metodi

Indici spettrali

Per analizzare il comportamento temporale della vegetazione al suolo, è stato scelto di utilizzare l'indice *Enhanced Vegetation Index* (EVI) secondo l'equazione [1]. Tale indice, proposto da Huete et al. (2002), si caratterizza per essere fortemente correlato ad importanti parametri della vegetazione come ad esempio il *Leaf Area Index* (LAI). Altre due importanti caratteristiche di questo indice sono la capacità di subire meno gli effetti della distorsione atmosferica sulla radiazione riflessa e quella di riuscire ad attenuare la tendenza alla saturazione dei suoi valori all'infittirsi della copertura vegetativa, difetto caratteristico di altri indici spettrali come il *Normalized Difference Vegetation Index* (NDVI, Tucker C.J, 1974).

$$EVI = 2.5 \times \frac{\rho_{ntr} - \rho_{red}}{\rho_{ntr} + 6 \times \rho_{red} - 7.5 \times \rho_{blue} + 1} \quad [1]$$

L'indice utilizzato per caratterizzare la presenza di acqua (umidità/contenuto idrico) al suolo è il *Land Surface Water Index* (LSWI) [2], indice applicato da Xiao et al. (2004) ai dati forniti dai sensori MODIS. Questo indice utilizza le bande dell'infrarosso vicino e dell'infrarosso medio per essere sensibile nei confronti del contenuto d'acqua al suolo e dell'umidità.

$$LSWI = \frac{\rho_{ntr} - \rho_{swtr}}{\rho_{ntr} + \rho_{swtr}} \quad [2]$$

Processamento dati

In figura 2 è proposto un diagramma di flusso per sintetizzare la metodologia studiata durante la mappatura automatica delle aree risicole e delle caratteristiche fenologiche ad esse associate. L'algoritmo è stato implementato in ambiente di sviluppo *Interactive Data Language* (IDL) ed è in grado di gestire automaticamente l'intero flusso dati satellitari, sino all'ottenimento degli output. E' possibile descrivere l'intero processamento suddividendolo nelle sue quattro fasi principali: i) la fase di acquisizione dei dati, ii) l'elaborazione di dati per rilevare il grado di nuvolosità e per l'ottenimento di indici spettrali, iii) la fase di analisi temporale degli indici e, infine, iv) la fase di calcolo degli *output* inerente i criteri di ottenimento delle mappe di localizzazione e fenologia.

La fase di **acquisizione dati** ha previsto il reperimento di un totale di 46 immagini MODIS MOD09A1 per ogni anno di indagine (2005, 2006 e 2007) per un totale 138 immagini. Da queste, si sono successivamente estratte le informazioni relative alla contaminazione atmosferica e alla presenza delle nuvole. In dettaglio, sono stati estratti dati di riflettanza nella terza banda (blu, 459–479 nm) e i meta-dati presenti nello strato informativo HDF n° 12 relativo ai *Quality flags* (*500m state flags; internal cloud algorithm flag - MOD35 cloud*). L'integrazione di questi dati ha permesso di generare una maschera indicante il grado di contaminazione nuvolosa, utile nella successiva fase di *smoothing* del segnale EVI.

La fase di computo degli **indici spettrali** ha previsto l'elaborazione di 2 indici. L'indice EVI (strettamente correlato alla presenza di copertura vegetale al suolo), è stato sfruttato durante tutta la fase di analisi dei segnali, a partire dal suo iniziale filtraggio (*smoothing*) sino al definitivo ottenimento dei risultati di mappatura e fenologia del riso. L'indice EVI è stato inoltre utilizzato assieme all'indice LSWI (strettamente correlato alla presenza dell'acqua al suolo) per la rivelazione dello stato di "potenziale sommersione agronomica" delle superfici osservate, seguendo il criterio [3] adottato da Xiao et al. (2005):

$$LSWI+0.05 \geq EVI \quad [3]$$

Un'operazione di **filtraggio dati** precede l'analisi temporale degli indici. Questa elaborazione è stata applicata sull'indice EVI e si è resa necessaria per rimuovere il "rumore" che caratterizza le serie temporale e diminuire la possibilità di incorrere in errori di omissione / commissione durante la fase di analisi. Il filtraggio applicato prevede inizialmente un *gap-filling* sui dati della serie temporale e successivamente l'applicazione del filtro *Savitzky – Golay* (Chen et al. 2004). Questo filtro implementa una regressione polinomiale locale ed ha il vantaggio di agire sulle serie di dati preservando i punti di massimo e minimo relativo della funzione. Il filtro prende in considerazione la qualità del dato di origine, modulando la sua azione a seconda del grado di contaminazione nuvolosa rilevata durante in fase di acquisizione dei dati. Lo *smoothing* locale del segnale è stato effettuato pesando ogni singola osservazione della finestra di analisi (*temporal window*) in funzione della contaminazione nuvolosa presente.

Nella fase di **analisi temporale del dato**, si è operato calcolando la derivata prima dell'indice EVI (EVI') ed individuando successivamente i punti di massimo e minimo relativo della serie secondo la relazione [4] e [5].

$$\text{MINrel}_{(t)} = \text{EVI}'_{(t-1)} < 0 \quad \square \quad \text{EVI}'_{(t+1)} > 0 \quad [4]$$

$$\text{MAXrel}_{(t)} = \text{EVI}'_{(t-1)} > 0 \quad \square \quad \text{EVI}'_{(t+1)} < 0 \quad [5]$$

Una serie di criteri sono poi stati implementati per individuare tra tutti i massimi e minimi relativi quelli aventi un **significato agronomico** relativo all'andamento colturale risicola, questi punti sono stati definiti come MAX e MIN.

Sulla serie temporale di ciascun pixel, un punto di minimo relativo è individuato e considerato indicativo di una sommersione risicola (MIN) soltanto se occorre in corrispondenza della condizione [1]. Sulla serie temporale di ciascun pixel un punto di massimo relativo è stato individuato ed associato alla presenza di una coltura risicola (MAX) se occorre la concomitanza dei seguenti criteri: i) il punto di massimo relativo si presenta a seguito di un periodo di rapida crescita vegetativa, criterio implementato come una sequenza di almeno 3 punti con derivata prima positiva sui 5 punti che precedono il MAXrel analizzato; ii) al punto di MAXrel analizzato segue una diminuzione dell'andamento vegetativo identificato in almeno 3 punti con derivata prima negativa sui successivi 5; iii) il periodo di tempo che intercorre tra la sommersione e la fase di massima crescita vegetativa della coltura ha una logica coerente con la durata agronomica del ciclo vegetativo della coltura risicola, questo criterio è stato implementato considerando una finestra temporale compresa tra 56 e 120 giorni (7÷15 MODIS composite) tra l'individuazione di una sommersione, punto di MIN, ed il punto MAXrel considerato; iv) in successione al considerato MAXrel si accerta la presenza di una senescenza agronomica, descritta individuando lungo la serie temporale EVI il valore che presenta un decremento di almeno 1/3 del EVI, al punto MAXrel considerato, nei successivi 40 giorni (5 MODIS composite).

Infine, in accordo con quanto proposto in precedenti studi da Jönsson et al. (2002) e da Boschetti et al. (2009), i segnali filtrati dell'indice EVI sono stati analizzati per derivarne i punti di inizio della stagione vegetativa risicola (Start of Season – SoS) e di fine (End of Season - EoS). Questi due punti, così come i punti MIN e MAX, sono in grado di restituire importanti interpretazioni sulla fenologia del riso nella stagione agronomica passata e sono stati rispettivamente calcolati nell' algoritmo come il momento in cui si localizza un incremento del 10% del valore di EVI dal punto di MIN (SoS) ed un decremento del 50% del valore di EVI dal punto di MAX (EoS).

Gli output dell'algoritmo hanno riguardato 5 differenti mappe: una mappa binaria, relativa alla presenza/assenza di colture nell'area di studio e quattro mappe indicanti l'occorrenza dei momenti fenologici rilevati (MIN, MAX, SoS, EoS).

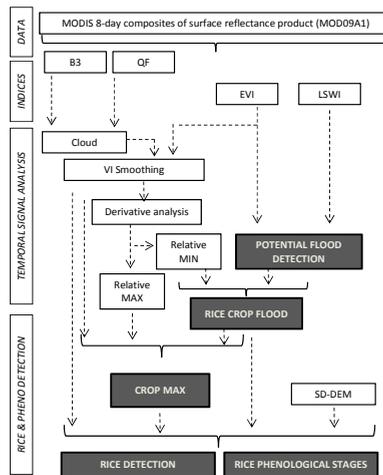


Figura 2. Rappresentazione schematica della metodologia proposta.

Valutazione della classificazione

Le mappe tematiche ottenute a seguito dell'applicazione dell'algorithm descritto sono state valutate con dati cartografici regionali a migliore risoluzione spaziale. La valutazione è stata condotta mediante il calcolo delle matrici d'errore (matrici di confusione), metodologia spesso utilizzata nella valutazione delle mappe ottenute da dato telerilevato (Brivio et al., 2006), in quanto consente di confrontare pixel a pixel la mappa ottenuta con quella di riferimento, ricavando l'accuratezza globale della mappa, l'accuratezza di classificazione della specifica classe e l'ammontare degli errori di commissione/omissione.

I risultati ottenuti, sono stati inoltre valutati utilizzando la metodologia di analisi suggerita da Boschetti et al. (2004). Questo metodo mette nelle condizioni di poter comparare le performance di classificazione dell'algorithm qui proposto, con quelle di classificatori utilizzati precedentemente sulla medesima area di studio (Crema et al., 2011; Manfron et al., 2012). La determinazione del confine di Pareto (*Pareto boundary*) permette di individuare inoltre la soglia di errore imputabile unicamente dalla condizione sperimentale funzione della dimensione e grado di eterogeneità della classe di interesse e non dalle caratteristiche dei classificatori (*bias error*).

Risultati

In figura 3a è mostrato il risultato della mappatura automatica di aree risicole rilevate nell'areale di studio dall'algorithm proposto. La valutazione della mappa ha mostrato accuratezza globale (OA) del 98% ed una accuratezza dell'utente (UA) pari all'87%. Questi sono risultati molto soddisfacenti e trovano conferma ulteriore nel dato inerente gli errori di commissione (CE) pari al solo il 13%. Questi dati indicano bassa percentuale di errore da parte delle predizioni dell'algorithm nei confronti della classe riso. Ciononostante, l'analisi ha anche evidenziato che l'approccio dimostra essere ancora molto conservativo. Ha infatti prodotto un significativo errore di omissione (OE), pari al 68%. L'analisi condotta con la metodologia della "frontiera di Pareto" (Figura 3b) ha permesso di capire che la distribuzione della classe riso è molto frammentata sull'areale di studio e pertanto genera alla scala di rappresentazione MODIS (risoluzione spaziale 500m) un potenziale alto errore di omissione (*low resolution bias*). Questo errore non è imputabile alla bontà del classificatore ma dipende unicamente dalle caratteristiche dell'area di studio. Il confronto tra le metodologie di classificazione precedentemente applicate e quella proposta in questo lavoro indica che un approccio di mappatura basato sullo studio della dinamica degli indici spettrali nel tempo è comparabile in termini di performance di classificazione al altri metodi, localmente adattati, basati essenzialmente sul confronto diretto tra indici.



Figura 3. a) Mappatura dell'estensione risicola sull'areale di studio utilizzando il metodo proposto. In verde le aree risicole stimate, in grigio quelle indicate dalla cartografia di riferimento. b) Analisi di confronto tra metodologie di mappatura precedentemente applicate e quella qui proposta.

In figura 4 sono invece riportati i risultati riferiti alla mappatura delle caratteristiche fenologiche delle superfici risicole per gli anni 2005, 2006 e 2007. Questi risultati seppur affetti da alte omissioni e pertanto ancora conservativi nella mappatura, risultano essere in pieno accordo con i calendari agronomici dell'area di studio e rappresentativi delle seguenti fasi fenologiche della coltura risicola: **MIN**: fase di sommersione della risaia, prossima all'emersione della coltura. Fine Aprile; **SoS**: interpretabile come fase di rapido accrescimento vegetativo della coltura. Inizio Maggio; **MAX**: picco di crescita vegetativa della coltura, prossima alla fioritura del riso; Metà Luglio; **EoS**: fase prossima alla maturità fisiologica della coltura. Metà Settembre.

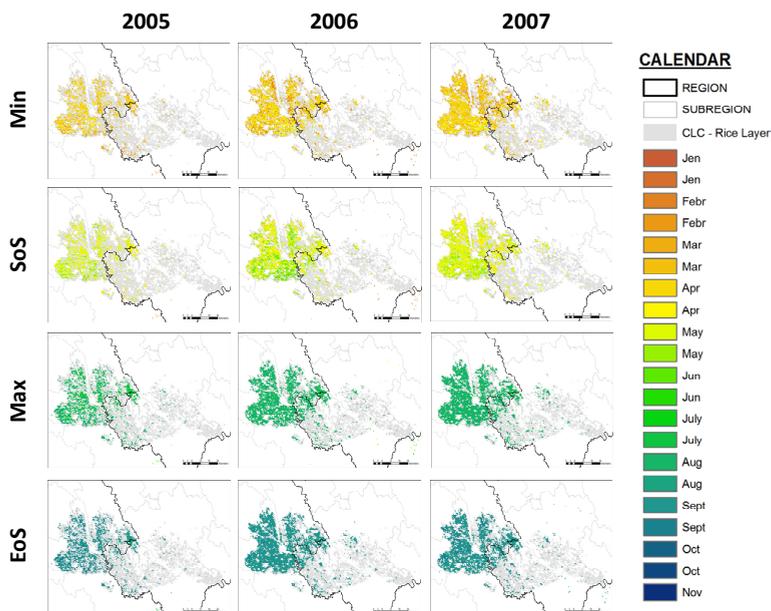


Figura 4. Mappatura delle caratteristiche fenologiche risicole, per il triennio 2005, 2006, 2007.

Conclusioni

Il presente lavoro dimostra come sia possibile effettuare una mappatura automatica di aree risicole utilizzando un approccio basato sull'interpretazione della dinamica temporale degli indici spettrali. Questo approccio, applicato in maniera totalmente automatica, ha permesso di ottenere risultati

comparabili a quelli ottenuti con metodi specificatamente adattati alle condizioni locali (vedi Manfron et al., 2012). Tale nuovo approccio si presta ad essere indipendente da costanti, soglie e maschere di esclusione, risultando quindi più adatto ad essere esportato ed applicato in differenti contesti. Ciò nonostante questo metodo necessita di ulteriori studi e approfondimenti finalizzati a ridurre l'alta quota di omissione riconducibile anche ad una scarsa capacità predittiva della fase di sommersione. Quest'ultimo aspetto richiede perciò ulteriori indagini anche sfruttando diverse combinazioni spettrali delle bande del sensore MODIS.

Bibliografia

- Bocchi S., Confalonieri R., Genovese G., Mariani L., Martin S., Orlandi S., (2003), "An integrated rice yield forecasting system in Europe", CD of proceedings of the 3rd International Temperate Rice Conference. 10-13 March, Punta del Este – Uruguay.
- Boschetti L., Flasse S. P., & Brivio P. A., (2004), "Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary". *Remote Sensing of Environment*, 91(3-4): 280-292.
- Boschetti, M., Stroppiana, D., Brivio, P. A. and Bocchi, S., (2009), "Multi-year monitoring of rice crop phenology through time series analysis of MODIS images", *International Journal of Remote Sensing*, 30: 18, 4643 – 4662.
- Brivio P.A.; Lechi G.M., Zilioli E., (2006). *Principi e metodi di Telerilevamento. 1a ed*, UTET Università.
- Chen, J., Johnsson, P., Tamura, M., Gu, Z., Matsushita, B. and Eklundh, L., (2004). "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter". *Remote Sensing of Environment*, 91, pp. 332–344.
- Crema A., Boschetti M., Manfron G., Pepe M., Musanti M., (2011), "Individuazione automatica di aree coltivate a riso da serie temporale MODIS: primi risultati in ambiente mediterraneo", 15° Conf. Naz. ASITA, Colorno 15-18 Novembre, pp.779-789.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G., (2002), "Overview of the radiometric and biophysical performance of the MODIS vegetation indices", *Remote Sensing of Environment*, 83: 195–213.
- Jönsson and Lars Eklundh, (2002), "Seasonality Extraction by Function Fitting to Time-Series of Satellite Sensor Data", *Geoscience and remote sensing*, 40: 1824-1832.
- Lillesand T. M., Kiefer R. W., (2000), *Remote Sensing and Image Interpretation*, Chapter 7 – Wiley & Sons.
- Manfron G., Crema A., Boschetti M., Confalonieri, R. (2012), "Automatic procedures definition and validation to extract mapping and phenological rice fields informations through time series analysis of remote sensed MODIS data", *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIV*, Proc. of SPIE 2012 paper number 8531-56, in press.
- Peng D., Heute A.R., Huang J., Wang F., Sun H., (2011), "Detection and estimation of mixed paddy rice cropping patterns with MODIS data", *International journal of applied earth observation and geo-information*, 13: 13-23.
- Tucker, C.J., (1974), "Red and photographic infrared linear combinations for monitoring vegetation", *Remote Sensing of Environment*, 8: 127-150
- Xiao, X., He, L., Salas, W., Li, C., Moore, B., Zhao, R., et al., (2002), "Quantitative relationships between field-measured leaf area index and vegetation index derived from VEGETATION images for paddy rice fields", *International Journal of Remote Sensing*, 23: 3595– 3604.
- Xiao X., Boles S., Liu J., Zhuang D., Froking S., Li C., Salas W., Berrien M., (2005), "Mapping paddy rice agriculture in southern china using multi-temporal MODIS images", *Rem. Sen. Env.*, 95: 480-492.