

Dati satellitari per l'identificazione degli agroecosistemi

Emanuele Spirito - T18720

Introduzione e obiettivi

L'obiettivo è rispondere alle domande scientifiche:

Si possono usare serie temporali di dati satellitari per identificare sistemi colturali differenti?

Per che classi colturali vi sono omissioni e commissioni?

Quali possibili applicazioni?

Il presente report si pone, pertanto, lo scopo di identificare il piano colturale 2024 di un insieme di 188 appezzamenti appartenenti alla società agricola Bonifiche Ferraresi S.p.A. situati nei pressi del comune di Jolanda di Savoia (FE, Italia), utilizzando tecniche di telerilevamento. In particolare, saranno utilizzate immagini satellitari acquisite dalla costellazione di satelliti denominata Sentinel-2 del programma spaziale Copernicus dell'ESA. Sarà poi eseguita una discussione sul confronto tra il piano colturale dichiarato dall'azienda agricola e il piano colturale determinato dalla procedura sotto delineata.

Identificare il piano colturale viene tradotto in classificazione dei campi nelle categorie di cereali autunno vernini, estive, riso, cover crop/altre colture invernali, perenni, altro/unclassified, indicando eventualmente la presenza di una seconda coltura.

Materiali

Dati Sentinel-2

Come verrà descritto nel seguito, la determinazione del piano colturale verrà fatta attraverso l'utilizzo di una serie temporale di un indice spettrale, motivo per cui i dati satellitari utilizzati devono coprire la suddetta regione per un periodo temporale sufficiente a permettere l'identificazione della scansione temporale delle fasi fenologiche di ciascuna coltura in ognuno degli appezzamenti. Per tale ragione, dovendo concentrare l'attenzione sul 2024, si è ritenuto opportuno scegliere l'intervallo temporale dal 1 ottobre 2023 al 28 Novembre 2024 per un totale di 78 immagini satellitari in formato raster codificato GeoTIFF (.tif), di cui un esempio è mostrato in Figura 1.



Figura 1- [ritaglio di un tile Sentinel-2 che ricopre l'insieme degli appezzamenti visualizzato in RGB. Nome: S2B_MSIL2A_20231104T101119_N0509_R022_T32TQQ_20231104T113842.SAFE. Timestamp di acquisizione: 2023-11-04T10:11:19.024000Z. Credits: QGIS - ESA/Copernicus]

La data finale corrisponde al giorno in cui il presente progetto è iniziato e quindi il giorno in cui sono stati scaricati i dati satellitari. Il motivo della scelta di anticipare rispetto all'inizio del 2024 sta nel voler osservare almeno un anno solare intero. Inoltre, nell'elaborare una serie temporale è utile partire da un intervallo di tempo più largo da entrambi i lati rispetto a quello in studio per poter effettuare operazioni di "smoothing" (vedi "Analisi di serie temporali di indici spettrali") che intrinsecamente tagliano porzioni di testa e di coda nei dati.

Le bande scelte per scaricare i dati sono, nel formato **identificativo banda** (nome della zona dello spettro): B02 (blue), B03 (green), B04 (red), B05 (red edge 1), B06 (red edge 2), B07 (red edge 3), B08 (NIR), B8A (red edge 4), B11 (SWIR 1), B12 (SWIR 2). Le prime tre bande sono necessarie per la visualizzazione RGB. Oltre a queste, le altre bande servono, nel presente report, per costruire indici spettrali che sintetizzano ed evidenziano alcune caratteristiche della superficie riflettente. Gli indici spettrali da utilizzare sono stati forniti solamente dopo il download delle immagini satellitari quindi si è proceduto a scaricare tutte le bande disponibili che tipicamente vengono utilizzate nell'ambito del telerilevamento agricolo.

Oltre a tali bande è stato scelto anche di scaricare, in seconda battuta, anche la banda SCL (Scene Classification Map) che corrisponde ad una bitmask a 4 bit di classificazione del raster multibanda a cui appartiene la banda. Questo verrà usato in seguito per eliminare i pixel che non rientrano nell'analisi (vedi "Applicazione della maschera SCL").

Confini aziendali e piano colturale

I confini aziendali sono stati forniti dal docente in formato vettoriale codificato Shapefile e corrispondono all'insieme dei 188 poligoni necessari a delineare il profilo spaziale dei campi agricoli da analizzare. Gli attributi (non nulli e non mascherati) per ciascuna geometria sono: **azienda**, **tenimento**, **id_geom**, e **id_appez** dove i primi due

corrispondono per tutti gli appezzamenti a BONIFICHE e JOLANDA. `id_geom` e `id_appez` sono analoghi, a patto quando un appezzamento è stato suddiviso su più geometrie ed in tal caso `id_geom` acquista oltre al numero corrispondente ad `id_appez` anche una lettera che identifica la suddivisione dell'appezzamento. Per tale motivo l'unico attributo rilevante di tale Shapefile corrisponde a `id_geom`. Nel seguito, per una fissata data verrà effettuata una media/mediana del valore dell'indice spettrale considerato sui pixel all'interno dei confini di ciascun campo e si vorrà tenere conto al variare della data come varia il valore così calcolato, da cui l'importanza dell'avere il vettoriale riguardante i confini aziendali.

Il piano colturale è stato fornito dal docente in formato vettoriale codificato Shapefile ed aggiunge rispetto allo Shapefile dei confini aziendali gli attributi raccolto, anno, superficie, varietà, note, coltura. La sua utilità sta nel poter confrontare la previsione fatta con la procedura che verrà qui seguita e il dato dichiarato dall'azienda agricola.

Metodi

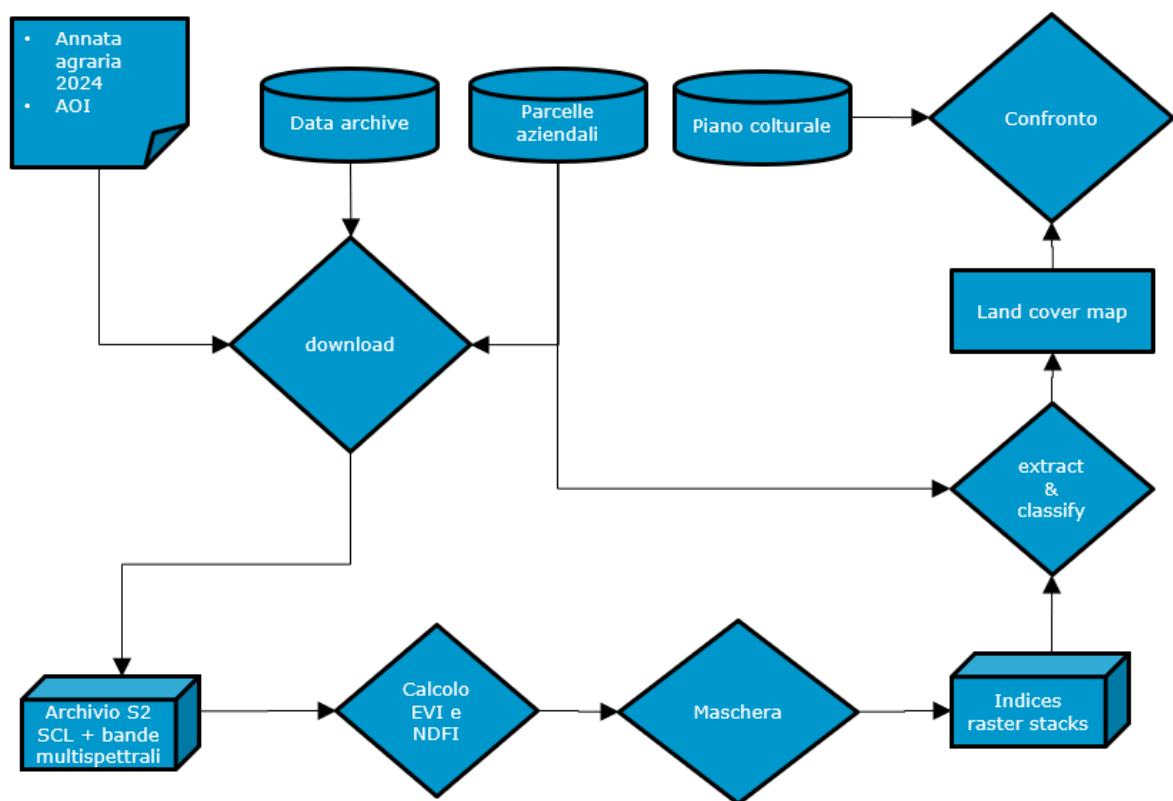


Figura 2 – [schema della metodologia seguita]

Download dei dati satellitari

Le immagini satellitari sono state scaricate attraverso il portale OpenEO¹ utilizzando uno script di R² che eseguisse l'accesso, la richiesta, il controllo della richiesta e il download, attraverso funzioni apposite del pacchetto `openeo`, quali `connect`, `login`, `processes`, `load_collection`, `save_result`, `create_job`, `start_job`, `list_jobs`, `download_results`. A tali funzioni sono stati forniti i parametri di estensione spaziale del rettangolo che contenesse i campi da analizzare (`west = 11.87754`, `south = 44.82105`, `east = 12.03337`, `north = 44.90196`), l'estensione temporale ("`2023-10-01`", "`2024-11-28`"), le bande ("`B02`", "`B03`", "`B04`", "`B05`", "`B06`", "`B07`", "`B08`", "`B8A`", "`B11`", "`B12`") e separatamente ("`SCL`"), il formato di download (GTiff).

Calcolo dell'indice vegetazionale e di allagamento

Nel presente report verranno utilizzati come indici spettrali i seguenti indici:

$$\begin{aligned} EVI_2 &= 2.5 * \left(\frac{B8 - B4}{B8 + 6 * B4 - 7.5 * B2 + 1} \right) = \\ &= 2.5 * \left(\frac{NIR - red}{NIR + 6 * red - 7.5 * blue + 1} \right) \\ NDFI_2 &= \frac{B4 - B12}{B4 + B12} = \frac{red - SWIR2}{red + SWIR2} \end{aligned}$$

Nel calcolo dell'indice vegetazionale EVI_2^3 (Enhanced Vegetation Index) è importante tenere in conto che i valori delle bande di Sentinel-2 vengono forniti moltiplicati per un fattore di scala di 10^4 in modo che il valore di riflettanza superficiale sia un intero compreso tra 0 e 10000 così che la codifica binaria sia di un intero unsigned a 16 bit⁴. Per evitare di effettuare troppi calcoli che comportano un errore di approssimazione, è stato scelto di sostituire nella formula espressa sopra al valore di 1 il valore di 10000 in modo da ottenere un risultato equivalente matematicamente ma più efficiente e preciso dal punto di vista computazionale. La vegetazione in uno stato di salute buona ha generalmente valori compresi tra 0.2 e 0.8, tenendo sempre conto che ciò è influenzato sia dal fenomeno di saturazione dell'indice che dalla variabilità del LAI (Leaf Area Index).

¹ <https://openeo.dataspace.copernicus.eu/>

² <https://cran.r-project.org/>

³ Glenn *et al.* (2008)

⁴ Il valore di riflettanza potrebbe teoricamente arrivare fino a $2^{16} - 1 = 65535$ ma la riflettanza superficiale è una percentuale quindi una volta normalizzata non può superare l'unità.

Per quanto riguarda l'indice NDFI2⁵ (Normalized Difference Flooding Index) non è necessaria la normalizzazione. In quanto differenza normalizzata l'intervallo dei valori di tale indice varia tra -1 e 1 . Un suolo allagato porta tale indice ad essere più vicino ad 1 rispetto ad un suolo secco che si porta verso l'altro estremo.

Questi indici sono stati calcolati tramite script R per ciascuna immagine satellitare scaricata nel punto precedente, ovvero per ciascuna data, producendo altrettanti raster monobanda sia di EVI2 che di NDFI2. Le funzioni prese dal pacchetto `terra` utilizzate sono `rast`, per importare il file raster dell'immagine satellitare, e `writeRaster`, per salvare il raster dell'indice spettrale.

Applicazione della maschera SCL

Sempre all'interno del pacchetto `terra`, è presente la funzione `mask` con cui sono stati mascherati i raster degli indici attraverso la maschera fornita da SCL. Sono stati mascherati tutti i pixel che hanno un valore di SCL pari a $(3, 8, 9, 10, 11)$, ovvero classificati rispettivamente come (ombra delle nuvole, media probabilità che sia nuvola, alta probabilità che sia nuvola, cirro, neve/ghiaccio).

Estrazione delle serie temporali mediate sui campi

Attraverso la funzione `extract` del pacchetto `terra` sono stati calcolati per ciascuna immagine satellitare, quindi per ciascuna data, la media dei valori dell'indice spettrale su ciascuna geometria. I parametri inseriti nella funzione `extract` sono (`index_stack_masked`, `vector_shapefile`, `fun = mean`, `na.rm = T`, `method = 'bilinear'`, `cells = TRUE`, `touches = TRUE`), ovvero nell'ordine l'insieme di tutti i raster dell'indice mascherato, lo shapefile vettoriale delle geometrie dei campi, la richiesta di effettuare la media su ogni campo, la rimozione dei valori non validi, l'utilizzo del metodo bilineare per interpolare il valore tra i cinque pixel più vicini al punto scelto all'interno della geometria (il pixel del punto e i quattro più vicini a croce), richiesta di restituzione del numero di celle, richiesta di prendere anche i pixel che sono toccati dalla geometria.

Il dataframe risultante contiene nelle righe le geometrie, identificate dal parametro `id_geom`, e in ogni colonna una data corrispondente alla data dell'immagine satellitare, estratta dal nome del raster dell'indice spettrale. Fare una serie temporale del campo identificato con `id_geom` corrisponde quindi a fissare la riga corrispondente al campo `id_geom` e ad osservare come varia il contenuto delle celle al variare delle colonne.

Dopo alcuni passaggi tecnici di manipolazione del dataframe descritto sopra, per ciascun campo viene eseguito il plot dell'andamento di EVI2 e di NDFI2 al variare del tempo nelle date considerate e poi anche posti tutti in una griglia per comodità di visualizzazione.

⁵ Boschetti, Nutini *et al.* (2014)

Classificazione della serie temporale

Le classi informative richieste sono **cereali autunno vernini, estive, riso, cover crop/Altre colture invernali, perenni (medica, prati), altro/unclassified**. Si richiede inoltre se è possibile, per ciascun campo, osservare la presenza di una seconda coltura.

La classificazione viene effettuata tramite interpretazione dei plot comparati di EVI2 e di NDFI2, utilizzando conoscenze agronomiche.

Risultati

Mappe di indici spettrali



Figura 3- [mappa di EVI2 (a) e di NDFI2 (b) per la data 12 Maggio 2024. Nella visualizzazione dell'EVI è stata utilizzata una visualizzazione a singola banda falso colore tra 0 e 1 con rampa di colori YlGn. Nella visualizzazione del NDFI è stata utilizzata una visualizzazione a singola banda falso colore tra -1 e 1 con rampa di colori Blues. Credits: QGIS]

Le mappe di indici spettrali sono state generate come descritto in “Calcolo dell’indice vegetazionale e di allagamento” e “Applicazione della maschera SCL” per un totale di 83 raster, uno per ogni data nell’intervallo di tempo considerato.

Il primo dato su cui concentrarsi è la presenza di nuvole sulla zona considerata nelle date considerate. In particolare, il 57% delle immagini satellitari è completamente corrotto dalla presenza di nuvole e il 6% delle immagini satellitari contiene solo qualche nuvola sparsa. Questo significa che nonostante la scelta di un intero anno di dati ci saranno periodi in cui la serie temporale sarà priva del dato. Laddove si verifici, la presenza di nuvole persistenti, ovvero per più immagini satellitari consecutive, potrebbe rendere di difficile identificazione una data, anche approssimativa, delle varie fasi fenologiche. Nell’analisi della serie temporale sarà visibile in dettaglio come la nuvolosità impatti sulla serie temporale complessiva. Per il presente studio la nuvolosità non ha impedito, se non in alcuni casi in cui interferisce con date specifiche come l’allagamento, la classificazione delle colture.

Nelle mappe mostrate in Figura 3 si osservano valori degli indici spettrali EVI2 ed NDFI2 relativi ad una singola immagine satellitare. Per capire se i valori degli indici sono nei

Telerilevamento per l'agricoltura – elaborato del laboratorio

range attesi vado ad osservare gli istogrammi di frequenza, riportati in Figura 4. Per quanto riguarda EVI2, il range effettivo è compreso tra 0 e 1 e corrisponde al range atteso quando si osservano campi vegetati da scarso a completamente vegetato. Si osserva nello stesso istogramma di frequenza anche la presenza di due picchi, uno attorno a 0.8 e uno poco sopra lo 0. Il primo potrebbe essere interpretato come vegetazione in salute verso l'apice della crescita, anche vista la data, mentre il secondo come campo non vegetato. Nella Figura 3 si vede come i campi da analizzare appaiano verso gli estremi della palette, ovvero alcuni campi sono visualizzati in giallo intenso (bassi valori di EVI2) e altri sono visualizzati in verde intenso (alti valori di EVI2), a conferma di quanto ipotizzato sopra. Si osservi come il range di NDFI2 sia compreso tra -0.8 e 0.2 . Valori negativi di indice NDFI2 indicano che tale campo non è allagato, o meglio che l'acqua eventualmente presente non ha contribuito significativamente rispetto al suolo e/o alla vegetazione soprastante. Valori positivi di indice NDFI2 indicano al contrario la presenza di un campo allagato. Dalla Figura 3 si vede che qualche campo ha un valore di NDFI2 molto superiore alla media dei campi, indicando che tali campi sono stati allagati. Questi corrispondono nell'istogramma di frequenza in Figura 4 alla coda oltre lo 0. Date le ridotte dimensioni della zona analizzata rispetto alla scala spaziale tipica degli eventi di pioggia, per cui la pioggia sarà con buona approssimazione uniformemente distribuita sulla zona interessata, una possibile spiegazione dei picchi di NDFI2, uno attorno a -0.4 e uno attorno a -0.2 , è che siano dovuti alla presenza di due diverse categorie di permeabilità di suolo a cui appartengono i suoli considerati.

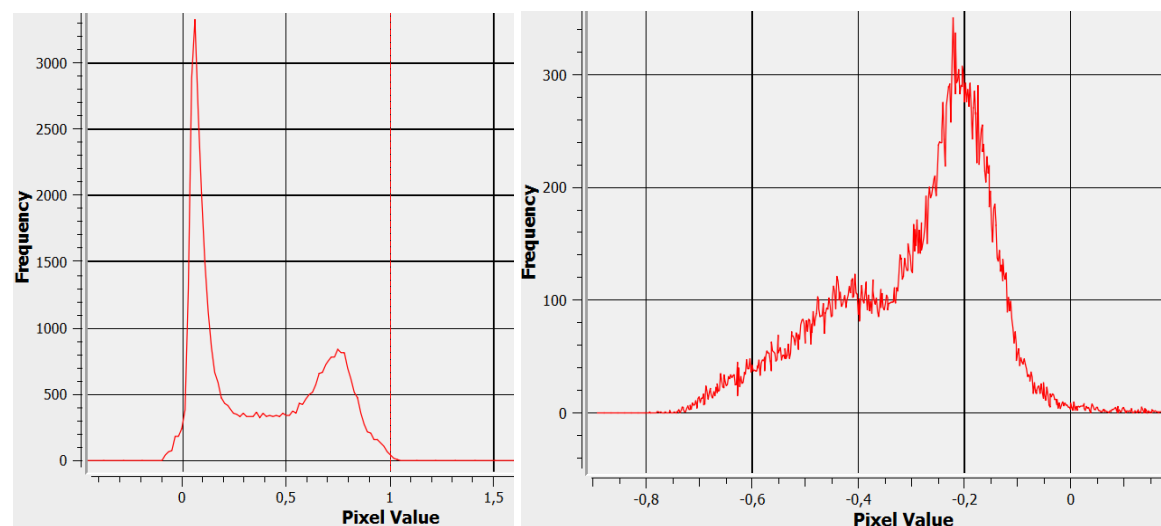


Figura 4 – [istogrammi di frequenza per le immagini della Figura 3 rispettivamente (a) per EVI2 e (b) per NDFI2. Credits: QGIS-ESA/Copernicus]

Analisi di serie temporali di indici spettrali

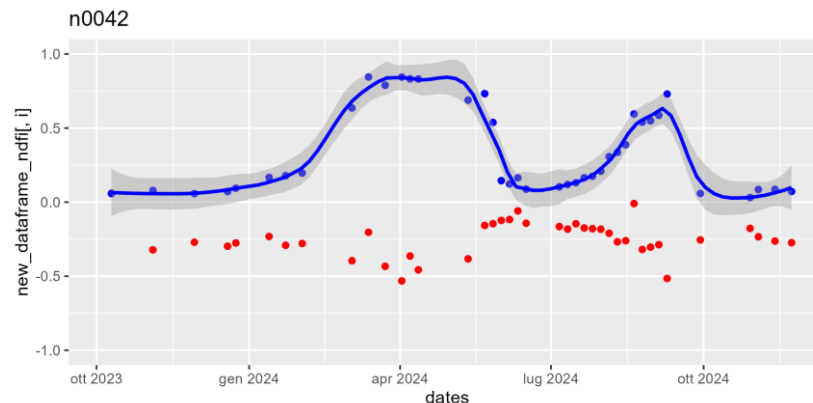
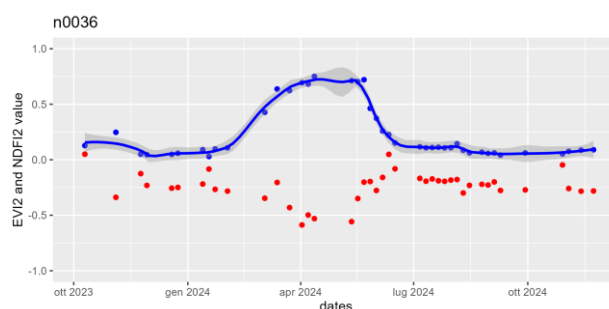


Figura 5- [serie temporale di EVI2 (blu) e di NDFI2 (rosso) per il campo con *id_geom* pari a 0042. Credits: R]

Come descritto in “Estrazione delle serie temporali mediate sui campi” si ottengono le serie temporali di entrambi gli indici per ciascun campo per un totale di 188 grafici come quello mostrato in Figura 5.

Si osservi innanzitutto che nei grafici è stato riportato oltre ai valori dei due indici anche un’interpolazione lineare con la funzione `geom_smooth` utilizzando il metodo `loess` con parametro di `span` pari a 0.2. L’interpolazione riguarda solo i valori di EVI2 perché sappiamo essere un parametro continuo, ovvero la massa vegetale fotosintetizzante non ha delle discontinuità di salto, mentre il valore di NDFI2 non è continuo a causa della rapidità con cui l’allagamento, sia esso artificiale o naturale, viene riassorbito dal terreno rispetto alla risoluzione temporale dei dati analizzati. Si ha infatti che un allagamento non coinvolge solitamente più di due date consecutive, corrispondenti ad un intervallo medio di 10 giorni.

Gli andamenti modello che andranno a costituire le *classi temporali* sono elencati in Figura 6.

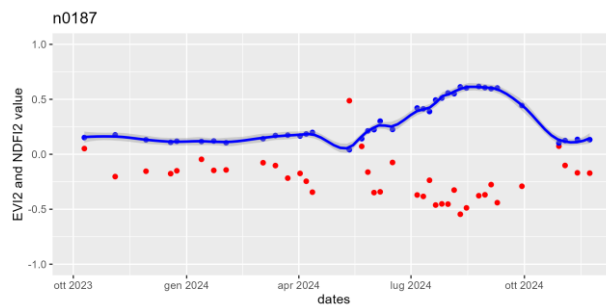


Classe temporale: A

Caratteristiche:

- risalita di EVI2 verso gennaio-febbraio;
- massimo di EVI2 verso aprile-maggio;
- discesa di EVI2 verso giugno.

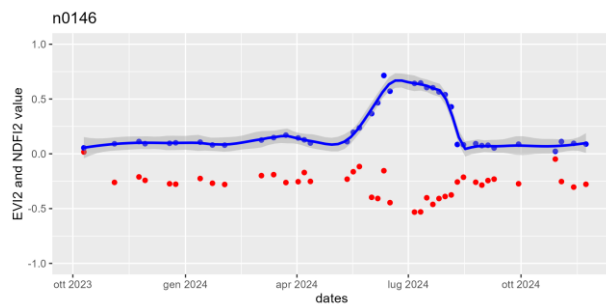
Telerilevamento per l'agricoltura – elaborato del laboratorio



Classe temporale: B

Caratteristiche:

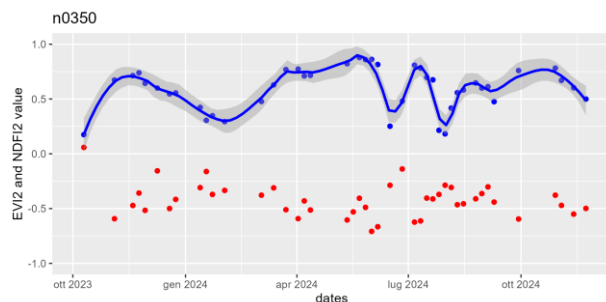
- valore di NDFI2 ben oltre lo 0 con discesa EVI2 intorno a maggio;
- salita di EVI2 da maggio ad agosto;
- picco di EVI2 attorno a agosto-settembre;
- discesa di EVI2 da settembre.



Classe temporale: C

Caratteristiche:

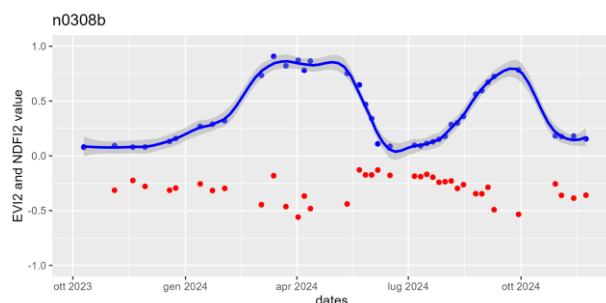
- valore di EVI2 che sale attorno a maggio;
- picco di EVI2 attorno a luglio;
- discesa di EVI2 verso settembre.



Classe temporale: D

Caratteristiche:

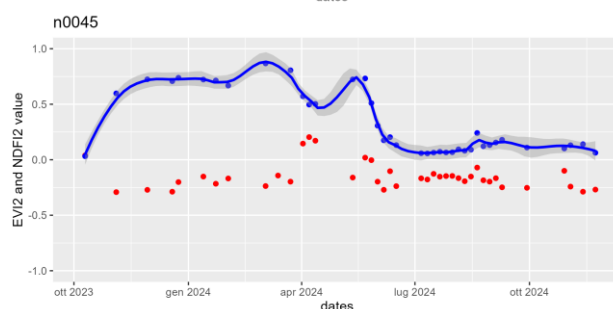
- presenza di più picchi di EVI2 in rapida successione;
- picco di EVI2 con valore massimo attorno a maggio.



Classe temporale: E

Caratteristiche:

- presenza di due picchi di EVI2 ben distinti, uno centrato in aprile-maggio, uno centrato in settembre;
- i due picchi di EVI2 hanno all'incirca la stessa altezza.

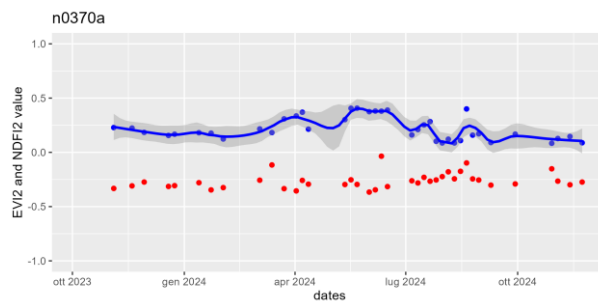


Classe temporale: F

Caratteristiche:

- andamento circa costante di EVI2 (si veda il dato originale piuttosto che quello interpolato) per il periodo invernale;
- picco a maggio di EVI2.

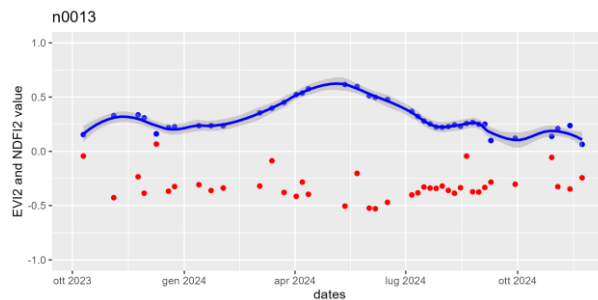
Telerilevamento per l'agricoltura – elaborato del laboratorio



Classe temporale: G

Caratteristiche:

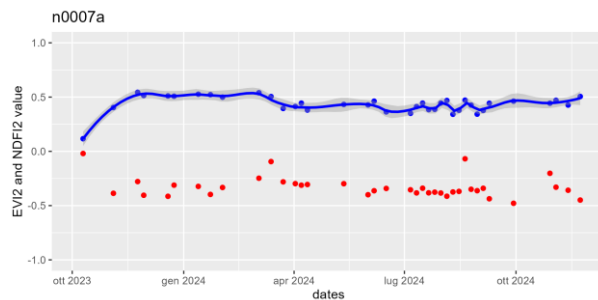
- non si osserva presenza di picchi di EVI2 marcato;
- si osserva un lieve aumento di EVI2 nella stagione estiva.



Classe temporale: H

Caratteristiche:

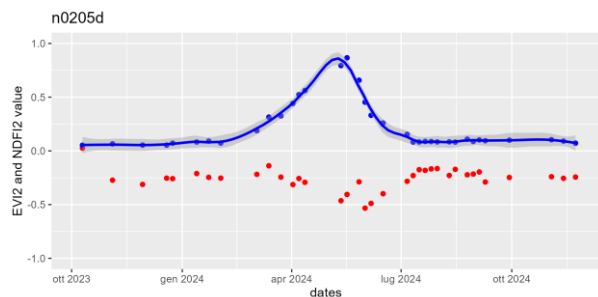
- si osserva un picco di EVI2 molto largo centrato in maggio.



Classe temporale: I

Caratteristiche:

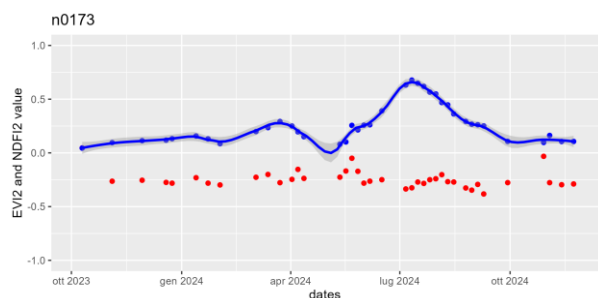
- si osserva un andamento costante di EVI2 sufficientemente superiore a zero.



Classe temporale: J

Caratteristiche:

- si osserva un picco di EVI2 più stretto rispetto alla classe A e centrato su maggio invece che su aprile-maggio.

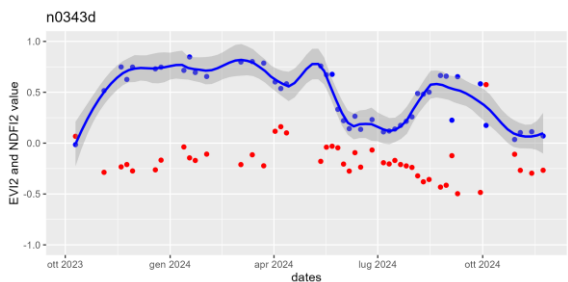


Classe temporale: K

Caratteristiche:

- picco basso di EVI2 attorno ad aprile;
- picco di EVI2 centrato su luglio più stretto rispetto alla classe temporale C.

Telerilevamento per l'agricoltura – elaborato del laboratorio



Classe temporale: L

Caratteristiche:

- presenza di un andamento costante (o in lieve salita) di EVI2 nella stagione invernale;
- presenza di un picco di EVI2 nella stagione estiva.

Figura 6 – [insieme degli andamenti modello delle serie temporali di indici EVI2 ed NDFI2. Credits: R]

Utilizzando conoscenze agronomiche si può stabilire quindi la seguente connessione tra classi temporali e classi informative, descritta in Tabella 1.

Classe temporale	Classe informativa
A	cereali autunno vernini
C, K	estive
B	riso
F, J	cover crop/altre colture invernali
E, L	coltura doppia
D, H, I	perenni (medica, prati)
G	altro/unclassified

Tabella 1 – [Collegamento tra classi temporali e classi informative. Le classi temporali sono prese dalla Figura 6]

Discussione degli errori di interpretazione

Prima di procedere con l’analisi delle serie temporali dei campi rispetto agli andamenti modello in Figura 6 si sono identificati ed eliminati dal dataset 4 campi di cui il dato colturale era assente o corrotto, vedi Figura 7. Una volta filtrati tali campi con dati non validi, si è effettuata la classificazione ed è stata confrontata l’ipotesi di piano colturale con la comunicazione aziendale. In Figura 8 è riportato il risultato di corrispondenza tra dato predetto e la comunicazione aziendale, mentre si veda la Figura 9 per il dettaglio del piano colturale predetto e la comunicazione aziendale.

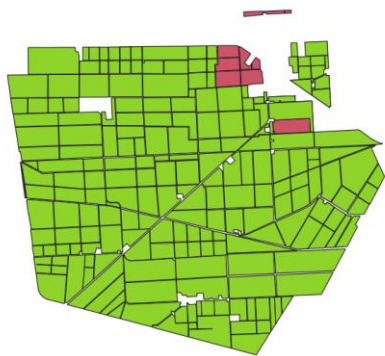


Figura 7 – [in marrone i campi di cui il dato è assente, in verde dove il dato è presente. Credits: QGIS]

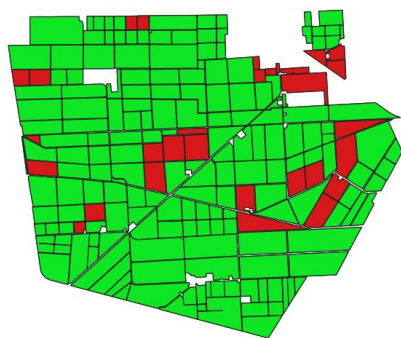


Figura 8 – [in verde i campi con dato valido in cui c’è corrispondenza tra l’ipotesi di piano colturale e la comunicazione aziendale. Credits: QGIS]

Telerilevamento per l'agricoltura – elaborato del laboratorio

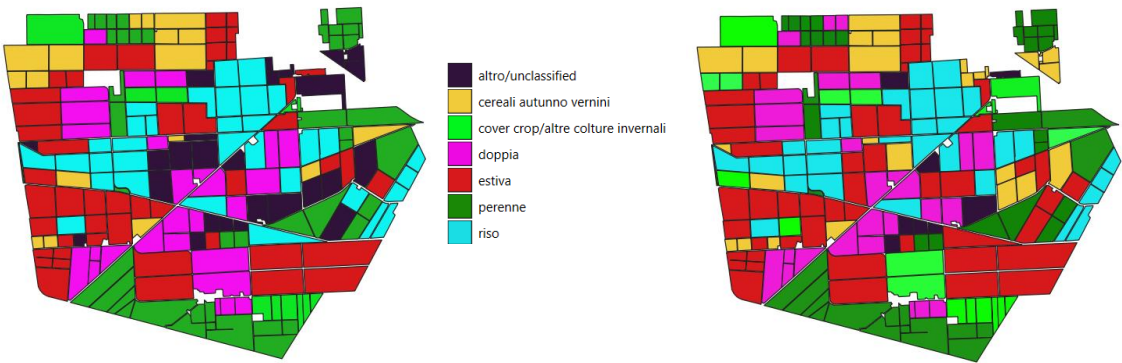


Figura 9 - [confronto tra ipotesi di piano colturale (a) e confronto con la comunicazione aziendale (b) per i campi con dato valido. Credits: QGIS]

Per un dettaglio maggiore sui campi che non sono stati identificati correttamente, si va a filtrare la Figura 9 con la Figura 8, ottenendo la Figura 10.

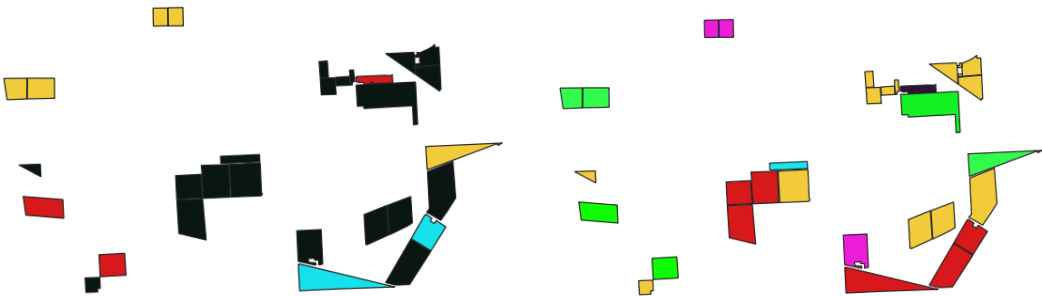


Figura 10 – [campi con dati validi che non corrispondono tra ipotesi di piano colturale (a) e comunicazione aziendale (b). Credits: QGIS]

Per poter quantificare la classificazione viene eseguito il calcolo della matrice di confusione, riportata in Tabella 2, ovvero una tabella a doppia entrata per classi tra il dato predetto e il dato a terra.

	cereale	riso	estiva	perenne	doppia	cover crop	altro/ unclassified	somma	PA %
cereale	10	0	0	0	0	0	12	22	45,5
riso	0	29	0	0	0	0	1	30	96,7
estiva	0	2	37	0	0	0	4	43	86,0
perenne	0	0	0	38	0	0	0	38	100
doppia	2	0	0	0	26	0	1	29	89,7
cover crop	3	0	2	0	0	8	1	14	57,1
altro/ unclassified	0	0	1	0	0	0	7	8	87,5
somma	15	31	40	38	26	8	26	184	
UA %	66,7	93,5	92,5	100	100	100	26,9		84,2

Tabella 2 – [matrice di confusione per la classificazione effettuata. Le classi che definiscono le colonne sono le classi predette mentre le classi che definiscono le righe sono le classi della comunicazione aziendale. Un esempio di lettura della tabella è che leggendo la colonna estiva dei 40 campi che sono stati predetti essere della classe estiva, nella

Telerilevamento per l'agricoltura – elaborato del laboratorio

comunicazione aziendale 37 erano della classe estiva, mentre 2 erano cover crop e 1 era altro/unclassified. A/U sta per altro/unclassified. UA sta per User Accuracy mentre PA sta per Producer Accuracy. Nella cella in basso a destra è descritta la Overall Accuracy]

Una prima osservazione che si può fare in Figura 9 è che le risaie, per la maggior parte identificate correttamente come descritto in Tabella 2, sono campi adiacenti. Questo ha come motivazione l'efficienza nella costruzione di canali di irrigazione per allagare tali campi.

Una seconda osservazione viene dalla Figura 10 e dalla Tabella 2 dove si vede come la confusione maggiore è tra la classe **altro/unclassified** dell'ipotesi di piano colturale e la classe **cereali autunno vernini** della comunicazione aziendale. Il motivo è mostrato in Figura 11 dove si vede come, a differenza della classe temporale A della Figura 6, è presente un picco nella stagione estiva (più basso rispetto al primo picco altrimenti sarebbe stato classificato come **doppia**), per cui si è scelto di non classificarlo nella categoria **cereali autunno vernini**.

Altri errori di classificazione che posso vedere in Tabella 2 sono:

- (ipotesi) **cereali autunno vernini** confusa con (comunicazione) **cover crop/altre colture invernali**. Nella comunicazione aziendali tali campi sono di *brassica carinata*. Dalla Figura 12 si vede come ci sia un andamento molto simile alla classe temporale A della Figura 6. In quanto cover crop è possibile che sia confusa con una coltura invernale di cereale dato che le cover crop vengono utilizzate tipicamente nel periodo invernale tra una coltura estiva e un'altra;
- (ipotesi) **estiva** confusa con (comunicazione) **cover crop/altre colture invernali**. Nella comunicazione aziendale corrispondono a due campi di *fagiolo secco* piantati non come coltura invernale ma come coltura estiva. Questo può succedere quando la cover crop viene utilizzata come unica coltura dell'annata agraria, anziché nella stagione invernale in previsione di una coltura estiva. Si osservi in Figura 13 come il fagiolo secco mostri un andamento simile alla classe temporale C della Figura 6;
- (ipotesi) **riso** confusa con (comunicazione) **estiva**. Nella comunicazione aziendale corrispondono a due campi di *pomodoro*. In questo caso non si è osservato un incremento di NDFI2, quindi si sarebbe potuto escludere **riso**, se non fosse che le date di maggio erano state tolte in quanto nuvolose quindi l'allagamento ci sarebbe potuto essere senza che sia visualizzato. Si osservi la Figura 14 per un esempio di campo di *pomodoro* e come esso sia simile alla classe temporale B della Figura 6.

Telerilevamento per l'agricoltura – elaborato del laboratorio

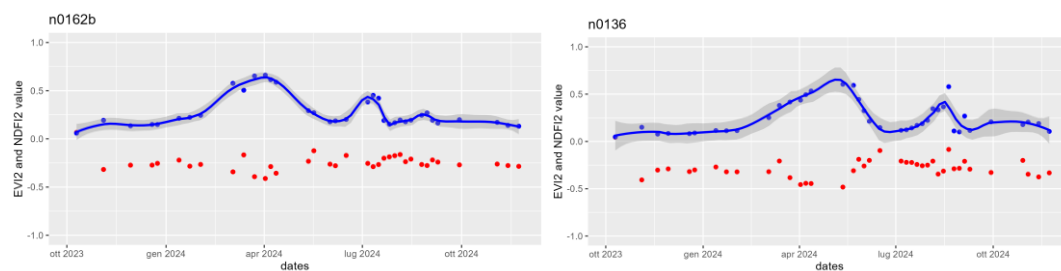


Figura 11 – [serie temporali degli indici spettrali per due campi classificati nell'ipotesi di piano colturale come *altro/unclassified* ma che nella comunicazione aziendale corrispondono alla classe *cereali autunno vernini*, in particolare orzo (a) e grano tenero (b). Credits: R]

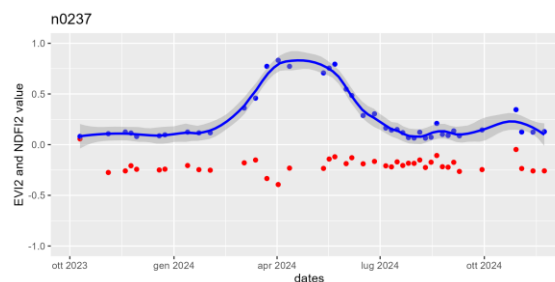


Figura 12 – [serie temporale degli indici spettrali per un campo che nella comunicazione aziendale è di *brassica carinata*. Credits: R]

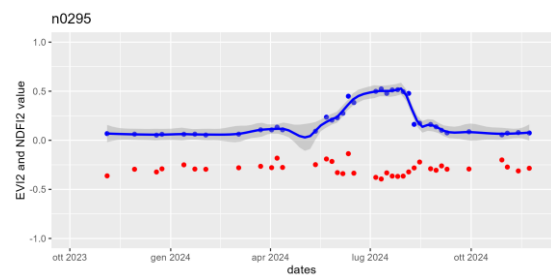


Figura 13 – [serie temporale degli indici spettrali per un campo che nella comunicazione aziendale è di *fa-giolo secco*. Credits: R]

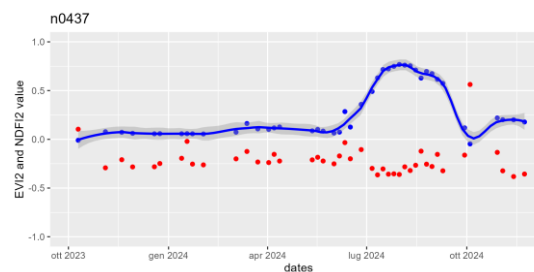


Figura 14 – [serie temporale degli indici spettrali per un campo che nella comunicazione aziendale è di *pomodoro*. Credits: R]

Conclusioni

L'obiettivo del lavoro era di rispondere alle domande scientifiche:

Domanda: *Si possono usare serie temporali di dati satellitari per identificare sistemi colturali differenti?*

Risposta: per i 184 campi di cui si avevano dati validi (vedi Figura 7) nel periodo dal 1 ottobre 2023 al 28 novembre 2024 è stata effettuata l'analisi delle serie temporali di indici spettrali (EVI2 e NDFI2) ricavati dai dati satellitari ESA/Copernicus/Sentinel-2, filtrati sulla copertura nuvolosa e mediati sulle geometrie dei campi. Sono stati identificati visivamente 12 andamenti modello riportati in Figura 6 e tramite conoscenze agronomiche in

Telerilevamento per l'agricoltura – elaborato del laboratorio

Tabella 1 sono stati assegnati tali andamenti alle classi con cui era stato chiesto di suddividere i campi. Nella Figura 8 e in Figura 9 sono mostrati i confronti tra l'ipotesi di piano colturale e la comunicazione aziendale escludendo i campi con dati non validi. A partire da tale figura si è ricavata la matrice di confusione riportata in Tabella 2 per quantificare l'accuratezza della classificazione, ottenendo una Overall Accuracy del 84.2%, valore compreso nel range ottimale di classificazione che è tra 80% e 95%, con cui è possibile rispondere affermativamente alla domanda.

Domanda: *Per che classi colturali vi sono omissioni e commissioni?*

Risposta: dalla Tabella 2 e dalla Figura 10 si vede che i campi in cui c'è maggior confusione è tra **altro/unclassified** e **cereali autunno vernini**. La User Accuracy della classe **altro/unclassified** si riduce al 26.9% e la Producer Accuracy della classe **cereali autunno vernini** si riduce al 45.5%. Il motivo è evidenziato nella Figura 11 confrontata con la Figura 6 ed è motivato dalla presenza di un secondo picco in estate che in questo lavoro non era contemplato in una coltura invernale.

Domanda: *Quali possibili applicazioni?*

Risposta: la prima applicazione è stata la crescita di conoscenza agronomica, ovvero grazie a questo lavoro è stato possibile incrementare la conoscenza sugli andamenti delle colture agrarie.

Inoltre, nella comunicazione aziendale è presente oltre che alla classificazione finora descritta anche il dettaglio della varietà coltivata. Utilizzando gli andamenti modello in Figura 6 è possibile determinare nuove classi, sottoclassi delle precedenti, con cui identificare le colture in modo più specifico. Con lo scopo di correggere gli errori di classificazione, è possibile aggiungere una classe temporale **M** che avrà come andamento modello quello in Figura 11. Si arriva così a generare una corrispondenza tra classi temporali e nuove classi informative come riportato in Tabella 3. Come punti di forza per l'osservazione di un agro-ecosistema:

- la capacità di osservare il terreno dalla scala locale, regionale, nazionale e continentale utilizzando le stesse procedure, senza doversi recare sul posto;
- i dati satellitari inoltre sono sempre disponibili integralmente e gratuitamente e quindi è possibile effettuare queste analisi per qualunque zona geografica;
- il rapporto costo/beneficio. Il dato satellitare ha un costo molto diluito sia per popolazione, es. Europa, che nel tempo, es. un programma spaziale ha una vita media di 10 anni. La stessa immagine satellitare può inoltre essere utilizzata per più scopi da più professionisti mantenendo cioè lo stesso costo e incrementando i benefici;
- rispetto alla classificazione pixel-based e object-based, l'analisi della serie temporale sfrutta la stagionalità della massa fotosintetizzante invece che il valore della riflettività superficiale o degli indici spettrali. Questo permette di eliminare i problemi di saturazione nel caso in cui la stagionalità sia un elemento distintivo della

Telerilevamento per l'agricoltura – elaborato del laboratorio

pianta, ad esempio per le colture. Non ci si basa infatti sul valore assoluto ma sull'andamento di tale valore assoluto nel tempo.

Tra i punti di debolezza per l'osservazione dell'agro-ecosistema:

- la copertura nuvolosa. Essa ha l'effetto di ridurre la frequenza temporale del dato osservato, rischiando quindi di perdere dei dettagli sulle fasi fenologiche della coltura. Per questo motivo in zone dove la copertura nuvolosa è persistente è necessario utilizzare tecniche differenti, es. utilizzo dei dati satellitari radar oppure della fotografia aerea oppure della fotografia da drone;
- la risoluzione spaziale. Per riconoscere alcune colture potrebbe essere necessaria una risoluzione spaziale maggiore, ad esempio l'identificazione dei filari richiede un dato ad alta risoluzione per poter rimuovere il contributo del suolo andando a filtrare il dato con un DEM oppure tramite classificazione dei pixel;
- l'insensibilità degli indici spettrali alle variazioni forestali. In campo forestale, le piante non vengono piantate e poi riseminate, motivo per cui non è possibile analizzare le metriche fenologiche per stabilire la tipologia di pianta. Si può lo stesso utilizzare questa tecnica per distinguere le piante sempreverdi da quelle caducifoglie. In questa tipologia di ambiente è più efficace un approccio pixel-based oppure object-based, nel primo caso mettendo assieme le bande di riflettività superficiale della pianta in diversi periodi dell'anno, nel secondo caso stabilendo la forma e dimensioni di una pianta in un certo periodo dell'anno. Per quest'ultimo passaggio è necessaria una risoluzione spaziale più alta di quella del dato Sentinel-2 che si raggiunge tipicamente con satelliti commercial ma anche con droni.

Una miglioria che si può fare a questa procedura è quella di automatizzare la classificazione delle serie temporali, in analogia a quanto viene fatto con la classificazione pixel-based e object-based, con la differenza che l'impronta spettrale è qui un'impronta temporale. Per una classificazione unsupervised si può quindi pensare di passare la serie temporale ad un clusterer e poi una volta creati i cluster andare ad identificare, aprendo l'andamento medio del cluster, a quale classe informativa appartiene. Per una classificazione supervised si può passare un 10% dei campi della comunicazione aziendale scelti in modo casuale come insieme di training su cui far allenare un algoritmo di classificazione da usare poi sul restante 90% dei campi della comunicazione aziendale ed usare poi un 10% dei campi della comunicazione aziendale, differenti da quelli di training, come insieme di validazione della procedura per creare la matrice di confusione.

Si può inoltre pensare di eseguire un mascheramento di qualità dei pixel non solo basato sulla classificazione SCL ma anche sulle altre bitmask presenti nel dato Sentinel-2, es. MSK_CLDPRB che indica la percentuale di probabilità che quel pixel contenga una nuvola. Sempre sui fenomeni meteo, si può pensare di inserire un layer che tenga conto della precipitazione giornaliera registrata in modo da evitare di confondere un'alluvione con

Telerilevamento per l'agricoltura – elaborato del laboratorio

un allagamento. Questo è utile specialmente nei territori dove le alluvioni sono stagionali, ad esempio nel sudest asiatico e nel continente indiano.

La scelta dell'indice spettrale determina in modo significativo quali caratteristiche spettrali della superficie si vanno ad analizzare. Dall'utilizzo di un solo indice non è possibile ricavare direttamente le caratteristiche della pianta, ma combinando più indici è possibile incrociare ulteriori dati per arrivare ad una migliore⁶ conoscenza dell'oggetto di studio. Un esempio è quello mostrato in questo lavoro, ovvero l'indice NDFI2 combinato con l'EVI2 permette al 96.7% di Producer Accuracy di individuare il riso. Allo stesso modo utilizzare un altro indice che sia per esempio sensibile al contenuto di acqua della foglia oppure al contenuto di azoto della foglia potrebbe mostrare ulteriori dettagli sull'andamento stagionale della coltura e quindi la sua identificazione.

Si aggiunge che la computazione qui eseguita ha utilizzato metodi di calcolo sequenziale su R. Scrivere l'algoritmo con tecniche di calcolo parallelo permette di eseguire il codice sui server in cloud, es. Google Earth Engine, incrementando la scalabilità della procedura, ovvero potendo gestire una quantità di dati superiore incrementando meno che linearmente il tempo di computazione.

Come interesse personale, uno studio di questo tipo su un'area differente come può essere un paese in via di sviluppo può essere utilizzato come base scientifica per prendere decisioni sociali, ambientali e governative sulla gestione dell'ecosistema, sia esso agricolo o forestale, per esempio per monitorare e prevenire carestie, desertificazioni, deforestazioni. Questo è ancora più importante su territori vasti, poco monitorati in situ e privi di infrastrutture come il sud America e l'Africa.

Classe temporale	Nuova classe informativa
A, M	<i>cereale autunno vernino (orzo, frumento, grano)</i>
B	<i>riso</i>
C	<i>mais</i>
D	<i>erba medica</i>
E	<i>doppia: mix foraggero</i>
F	<i>colza⁷</i>
G	<i>set aside</i>
H, I	<i>arboreto</i>
J	<i>legume</i>
K	<i>girasole</i>

⁶ In un approccio ancora più completo si potrebbe simulare come da una configurazione della pianta (foglia, canopy e geometria) si ottenga una determinata riflettività superficiale. Si può quindi allenare un algoritmo per invertire questa relazione e riuscire così a ricavare dalla riflettività superficiale la configurazione della pianta.

⁷ Nel riconoscere che questa classe è colza è stata utilizzata anche la visualizzazione in QGIS in Figura 1 e in Figura 15.

L	<i>doppia: cover crop + coltura estiva</i>
altro/unclassified	<i>altro/unclassified</i>

Tabella 3 – [associazione tra classi temporali e nuove classi informative. Le classi temporali sono prese dalla Figura 6]

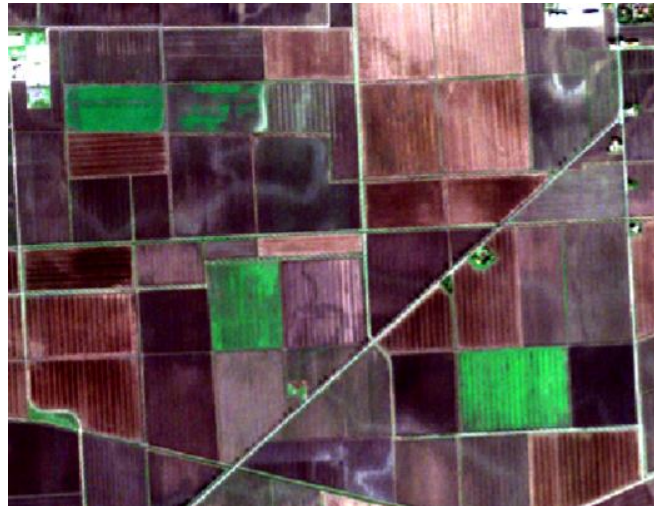


Figura 15 – [particolare dell'immagine satellitare in RGB del 19 dicembre 2023. La colza è caratterizzata da un colore verde fluorescente. Credits: QGIS – ESA/Copernicus]

Bibliografia

Boschetti, M., Nutini, F., Manfron, G., Brivio, P. A., & Nelson, A. (2014). Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems. *PLoS ONE*, 9(2). <https://doi.org/10.1371/journal.pone.0088741>

Glenn, E. P., Huete, A. R., Nagler, P. L., & Nelson, S. G. (2008). Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors*, 8, 2136–2160. <https://doi.org/10.3390/s8042136>

QGIS version 3.34.13-Prizren

RStudio version 2023.09.1 Build 494 "Desert Sunflower" Release

R version 4.3.1 "Beagle Scouts"