

**-ESERCITAZIONE 1-2****-Composizione RGB ed estrazione di un subset di dati**

Nell'esercitazione è richiesto in che modalità eseguire una composizione RGB e l'estrazione di un subset di dati dal dataset del Landsat 8 OLI .

**Estrazione di un subset**

Come primo passo ,aggiungo i layer e carico le 7 bande . I layer sono GeoTIFF cioè nell'header sono già scritte le caratteristiche di georeferenziazione , ho il sistema di coordinate che è WGS84 / UTM zone 33N , ed ho informazioni sul massimo , minimo , valore medio e deviazione standard . Carico lo shape file che sarebbe la mia area studio ,si sovrappone al file raster caricato in precedenza . Lo shape file lo utilizzo per "tagliare" il mio file raster ,in modo da trovarmi la mia area studio. Con il comando estrazione -> clip raster by mask layer , in input layer inserirò di volta in volta(per tutte le 7 bande ) il singolo raster , con la Mask layer la mia "Area\_Studio12" , infine salvo con nome il file temporaneo , darò il nome di 'Banda 1 fino a 7'.

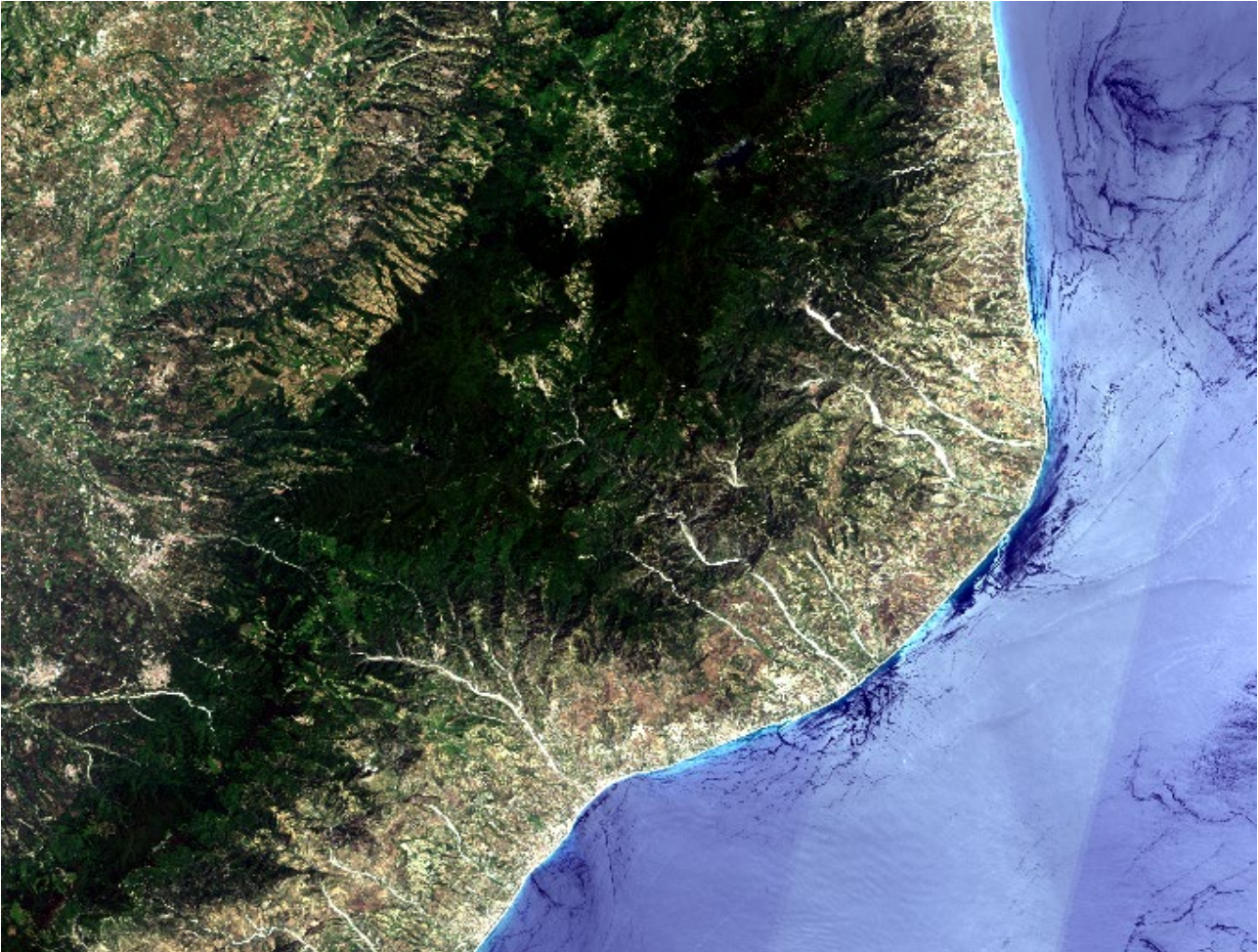
**Composizione RGB**

La composizione RGB è un modello di colori additivo composto da tre colori : rosso, verde e blu .

I colori RGB non si possono ottenere da nessun altro colore ma questi, possono generare tutti gli altri colori ottenuti dalla combinazione della luce con i tre colori .

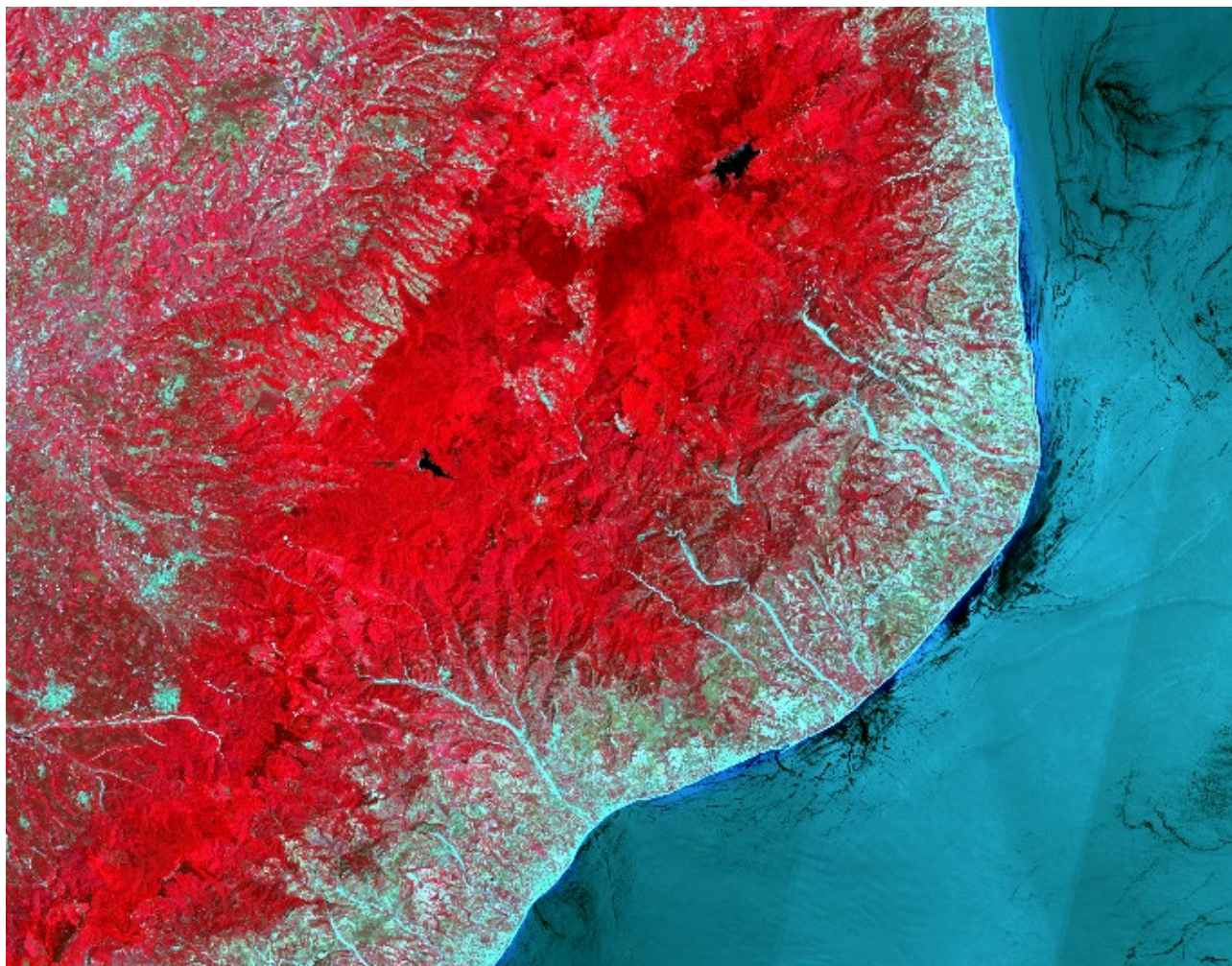
Procedo con la composizione a colori , Raster -> Miscellanea -> Crea raster Virtuale , mi consente di costruire la composizione a colori RGB . Scelgo le bande (che mi permettono una composizione a colori veri) : scelgo Banda 2 per il blu , Banda 3 per il verde e Banda 4 per il rosso . Creo un file in RGB , ho i colori non veritieri , allora cambio i colori alle bande da proprietà layer , così ho i colori veri .

Ecco il mio risultato :





Posso fare anche una composizione a falsi colori per avere un risalto e una visualizzazione migliore , utilizzando le bande 5 su rosso 4 sul verde e 3 sul blu .



2. Calcolo della correlazione tra le bande , matrice di varianza-covarianza , matrice di correlazione , indice OIF .

Con il ritaglio dei subset dell' immagine , posso ricavare i dati che riguardano il valore medio e la deviazione standard . Con questi dati posso costruirmi la matrice di varianza-covarianza , dove la varianza è definita alla singola banda , la covarianza è definita come la variazione congiunta di due variabili intorno alla loro media comune .

Ho estratto le medie e le deviazioni standard delle immagini e poi sono state tabellate .

Ecco il risultato :

	Valore Medio	Deviazione Standard
Banda 1	11190.36	1323.69
Banda 2	10298.36	1364.72
Banda 3	9593.39	1069.16
Banda 4	8932.95	1539.29
Banda 5	17649.99	5738.83
Banda 6	12222.32	2940.50
Banda 7	9176.99	1995.28

Mi servono per lo studio della correlazione fra le bande , identificando quali bande danno maggiori informazioni.

Ecco le formule usate per la costruzione della matrice :

VARIANZA

$$SS_k = \frac{\sum_{i=1}^N (BV_{ik} - \mu_k)^2}{N}$$

dove **SS<sub>k</sub>** si ottiene per 1=k; **BV<sub>ik</sub>** è la banda k-esima presa in considerazione; **μ<sub>k</sub>** è il valore medio della banda k-esima;

## COVARIANZA

Calcolo la matrice di covarianza per ciascuna immagine , di seguito la formula :

$$C_{kl} = \frac{\sum_{i=1}^N (BV_{ik} - \mu_k) \cdot (BV_{il} - \mu_l)}{N}$$

**BV<sub>ik</sub>** è la banda dell'immagine e **BV<sub>ij</sub>** è un'altra banda dell'immagine;

**μ<sub>k</sub>** e **μ<sub>j</sub>** sono i valori medii delle rispettive bande k e j.

In breve ecco la spiegazione della formula: elimino il valore medio dell'immagine 1 moltiplico per l'immagine 2 ed elimino il valore medio dell'immagine 2 il risultato sarà un immagine , dopo devo dividere la sommatoria dei valori per N cioè estraggo il valore medio . Così ottengo il confronto tra le bande e il valore medio.

Calcolata la covarianza rispetto a 2 bande, immetto il valore nella matrice di varianza-covarianza.

Avendo 7 bande spettrali avrò una matrice 7x7:

### Matrice di Varianza-Covarianza

	Matrice di Varianza-Covarianza						
	Banda1	Banda2	Banda3	Banda4	Banda5	Banda6	Banda7
Banda 1	1.752.155,21	1794231.94	1155938.82	1539257.26	-6430351.47	-1244882.48	453720.41
Banda 2	1794231.94	1.862.460,68	1274065.63	1723364.72	-6344512.67	-885319.33	738885.26
Banda 3	1155938.82	1274065.63	1.143.103,10	1596908.17	-2972354.25	783795.32	1401841.66
Banda 4	1539257.26	1723364.72	1596908.17	2.369.413,70	-4212483.20	1559026.80	2303983.34
Banda 5	-6430351.47	-6344512.67	-2972354.25	-4212483.20	32.934.169,76	9963479.48	1416291.75
Banda 6	-1244882.48	-885319.33	783795.32	1559026.80	9963479.48	8.646.540,25	5022552.78
Banda 7	453720.41	738885.26	1401841.66	2303983.34	1416291.75	5022552.78	3.981.142,27

### Matrice di correlazione

Per costruire la matrice di correlazione abbiamo bisogno dei valori ricavati dalla matrice di varianza-covarianza ,i valori di dev.standard e delle bande .

$$\rho_{kl} = \frac{C_{kl}}{\sigma_k \sigma_l} = \frac{C_{kl}}{\sqrt{\sigma_k^2 \sigma_l^2}}$$

La formula :

dove **C<sub>kl</sub>** è la covarianza; **σ<sub>k</sub>** e **σ<sub>l</sub>** sono i valori della dev.standard delle rispettive bande.

	Matrice Indice di Correlazione P						
	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 6	Banda 7
Banda 1	1	0,993227512	0,816781437	0,755447824	-0,846496299	-0,319831241	0,171790118
Banda 2	0,993227512	1	0,873183545	0,82037647	-0,810086423	-0,220615086	0,271349857
Banda 3	0,816781437	0,873183545	1	0,970324047	-0,484433877	0,249309473	0,65713169
Banda 4	0,755447824	0,82037647	0,970324047	1	-0,476863821	0,344438707	0,750161959
Banda 5	-0,846496299	-0,810086423	-0,484433877	-0,476863821	1	0,590427378	0,123687416
Banda 6	-0,319831241	-0,220615086	0,249309473	0,344438707	0,590427378	1	0,856050679
Banda 7	0,171790118	0,271349857	0,65713169	0,750161959	0,123687416	0,856050679	1

Il coefficiente di correlazione lineare può assumere valori compresi tra (-1,1) che determinano il grado di correlazione di una banda con un'altra .

Quando il valore dell'indice è prossimo a 1 allora le bande sono altamente correlate , produrranno immagini molto simili tra di loro .

Quando l'indice tende a -1 , allora le bande non sono correlate tra di loro .

### Calcolo OIF

OIF o indice di fattore ottimale (***Optimum Index Factor***), ovvero un indice statistico basato sull'utilizzo dei valori di dev.standard e di indici di correlazione per determinare combinazioni ottimali delle bande multispettrali di una stessa scena, così da costruire una composizione RGB ottimale.

L'intento è quello di esaltare la visibilità degli oggetti a terra presenti nell'area geograficamente ripresa. Per ottenere ciò, l'indice mette in relazione la somma dell'informazione (i valori della dev.standard) con l'ammontare delle duplicazioni (la somma dei valori assoluti degli indici di correlazione).

Si ricercano combinazioni mediante la formula del calcolo combinatorio .

L'indice OIF è dato da :

$$OIF = \frac{Std_j + Std_i + Std_k}{|Corr_{j,i}| + |Corr_{j,k}| + |Corr_{i,k}|}$$

Al numeratore è presente la somma delle dev.standard della combinazione di bande considerate , al denominatore la somma dei valori assoluti degli indici di correlazione delle bande considerate .

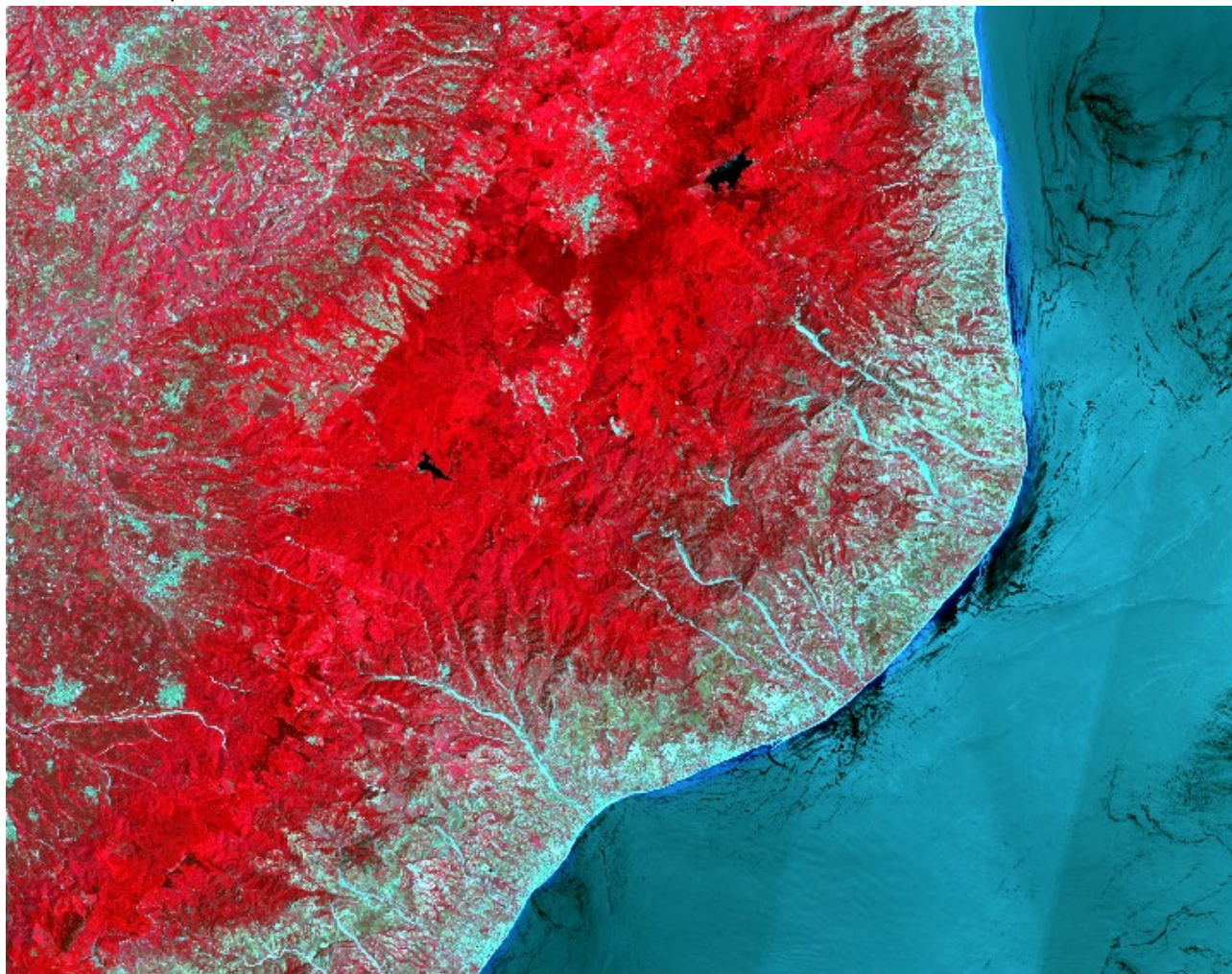
Di seguito è riportata la tabella con i valori dell'indice OIF relativo a ciascuna combinazione trovata :

OIF 123	1400,410149	OIF 246	6189,906204
OIF 124	1645,626604	OIF 247	2659,927877
OIF 125	-12703,96295	OIF 256	-22813,17321
OIF 126	12431,85491	OIF 257	-21922,29526
OIF 127	3260,788091	OIF 267	6948,170591
OIF 134	1546,53198	OIF 345	924768,1234
OIF 135	-15815,8124	OIF 346	3547,758159
OIF 136	7146,775084	OIF 347	1936,278489
OIF 137	2666,416326	OIF 356	27437,11908
OIF 145	-15146,37044	OIF 357	29702,1212
OIF 146	7439,831611	OIF 367	3407,073926
OIF 147	1706,796332	OIF 456	22311,28709
OIF 156	-17369,36479	OIF 457	23359,54015
OIF 157	-16438,27867	OIF 467	3319,439949
OIF 167	8840,939992	OIF 567	6798,39812
OIF 234	1491,49509		
OIF 235	-19397,09727		
OIF 236	5959,09913		
OIF 237	2458,370325		
OIF 245	-18524,05879		

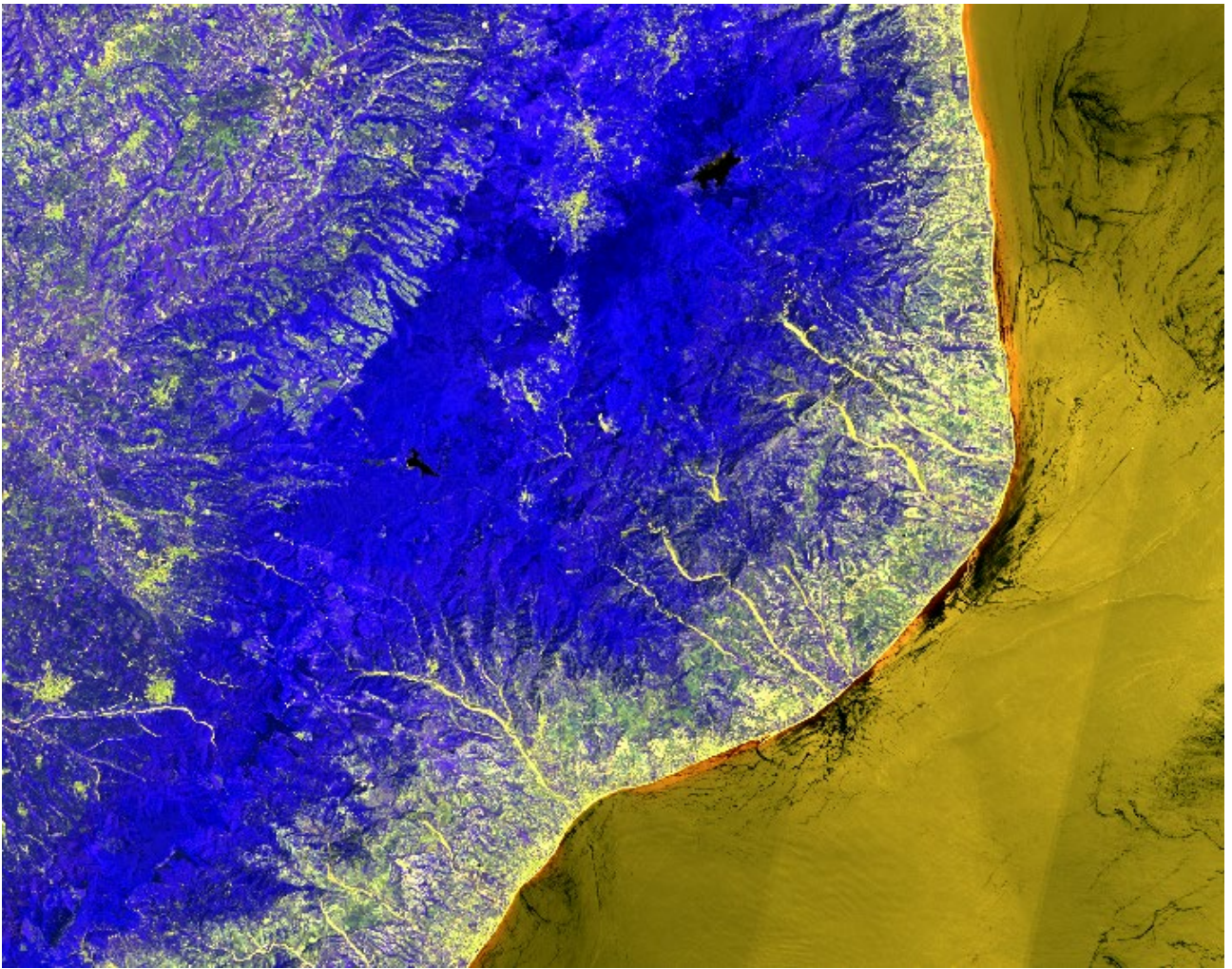
Più alto è il valore dell'indice OIF, più è performante l'utilizzo di quelle stesse bande per la composizione RGB.



Ecco un' esempio con bande 3-4-5 che hanno un indice OIF ottimale







### Esercitazione 3

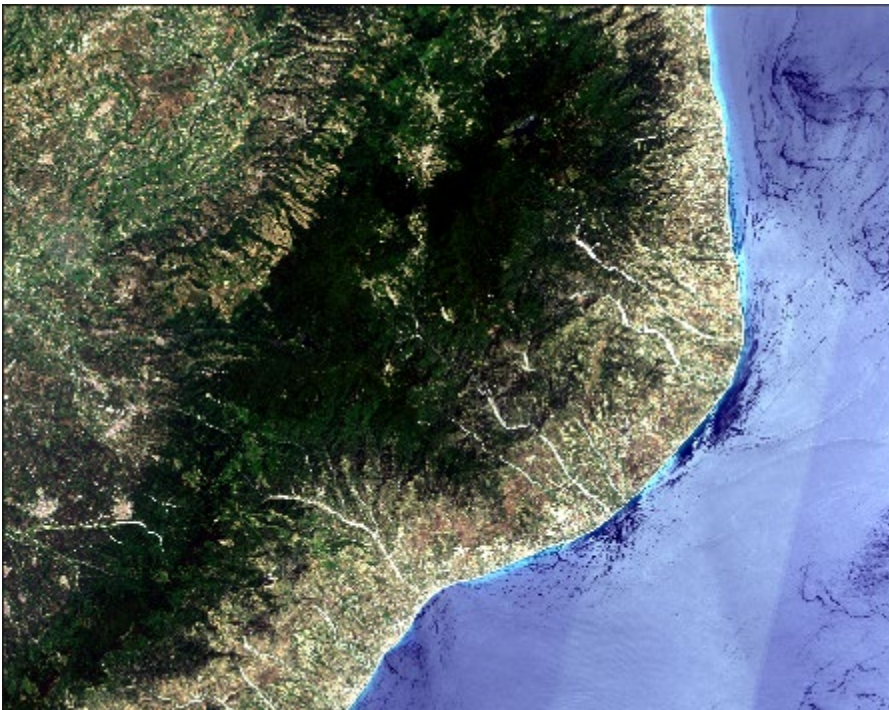
#### **Calcolo dell'indice NDVI e classificazione nelle tre classi acqua, suolo, vegetazione.**

Un indice di vegetazione molto utilizzato è l'indice NDVI(Normalized Difference Vegetation Index).

È basato sulla differenza normalizzata tra i BV della banda dell'infrarosso vicino e quelli della banda del rosso.

Come primo passo ho caricato su QGIS le bande 2(Blu)-3(Verde)-4(Rosso)-5(NIR vicino infrarosso) , mi servono principalmente la banda 4 e 5 perché sono le più canoniche , invece le bande 2-3-4 mi servono per fare una composizione in RGB .Ho già estratto la mia area studio con la funzione Clip Raster by Mask Layer in precedenza.

Continuo con la composizione RGB mi serve per ricercare con più facilità le classi di acqua , suolo e vegetazione .Prima di tutto vado in Miscellanea->virtual raster e inserisco le bande che utilizzo per la composizione in RGB , ovvero banda 2-3-4 .Siccome ho il layer virtuale e le bande sono invertite e i colori non sono veritieri , vado in proprietà e cambio il colore delle bande . Questo è il risultato :



Successivamente ottengo l'NDVI attraverso la formula :  $NDVI = (IR - Rosso) / (IR + Rosso)$

Per ottenere l'NDVI ho utilizzato il calcolatore Raster , inserendo  $(banda\ 5 - banda\ 4) / (banda\ 4 + banda\ 5)$  , lo salvo con il nome di NDVI1 e mi aspetto valore compresi tra -1 e 1 , il risultato calcolato è -0.05 e 0.65 , i colori scuri equivalgono ai valori più bassi e colori più chiari valori a più alti . L'indice mi dà con i valori bassi l'acqua , valori alti vegetazione e valori intermedi suolo . Adesso devo stabilire la soglia di separazione tra le tre classi in modo da riclassificare le immagini e ottenere 3 colori fondamentali , blu per l'acqua ,verde per la vegetazione e arancione per il suolo.

Utilizzo i Training Sites, ovvero dei raggruppamenti di pixel come esempi di classi nell'immagine. Questi possono essere identificati in 3 modi:

- 1)attraverso un campagna di 'verità al suolo' ;
- 2)attraverso dati raccolti in precedenza;
- 3)attraverso la fotointerpretazione.

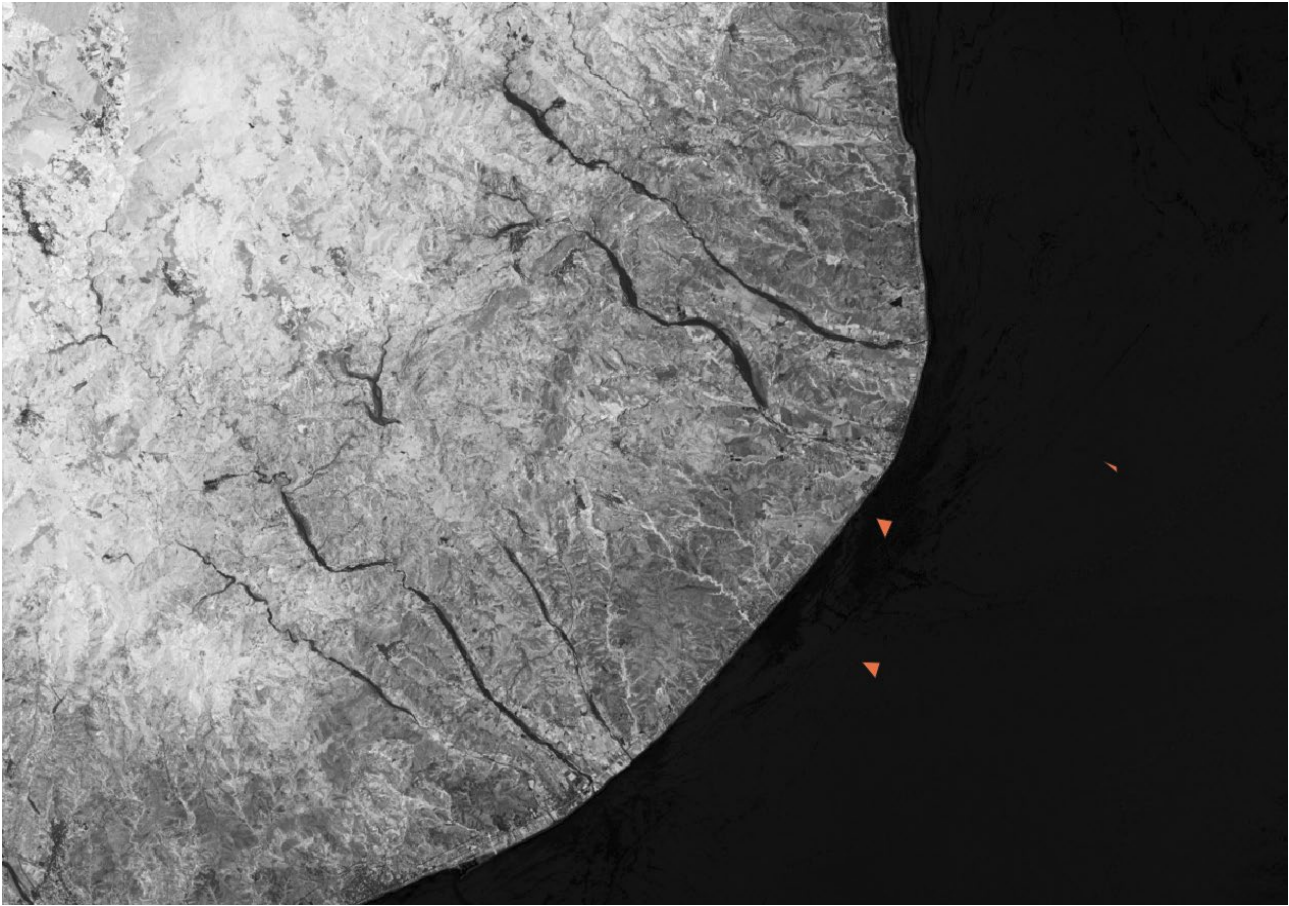
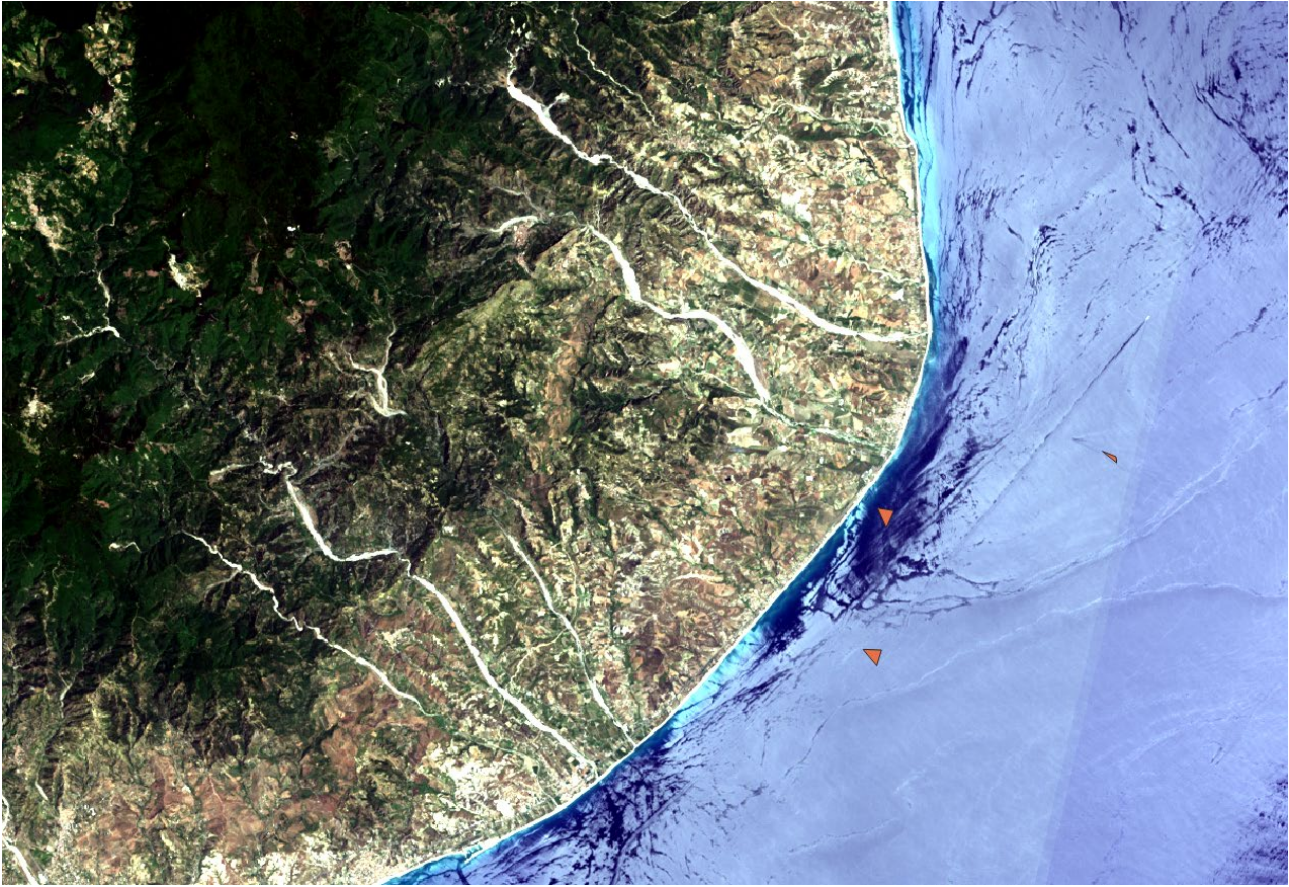
I Training Sites che ho creato tramite dei poligoni disegnati sull'immagine, ci forniscono un punto di riferimento per le classi di copertura. Inoltre, maggiore è il numero di Training Sites a nostra disposizione, più accurato sarà il contenuto informativo di quella determinata classe di copertura. Ad ogni Training Sites è stato associato anche un ID identificativo, per esempio alla classe acqua ho assegnato un ID=1, al suolo ID=2 e alla vegetazione ID=3.

Procedo nel disegnare i poligoni , creo un nuovo shape file , e lo chiamo acqua , ed incomincio a creare dei poligoni nell' acqua con id1 . La stessa procedura la uso anche per il suolo id2 e per la vegetazione con id 3.

Ecco i Training Sites raccolti per l' acqua , suolo e vegetazione.

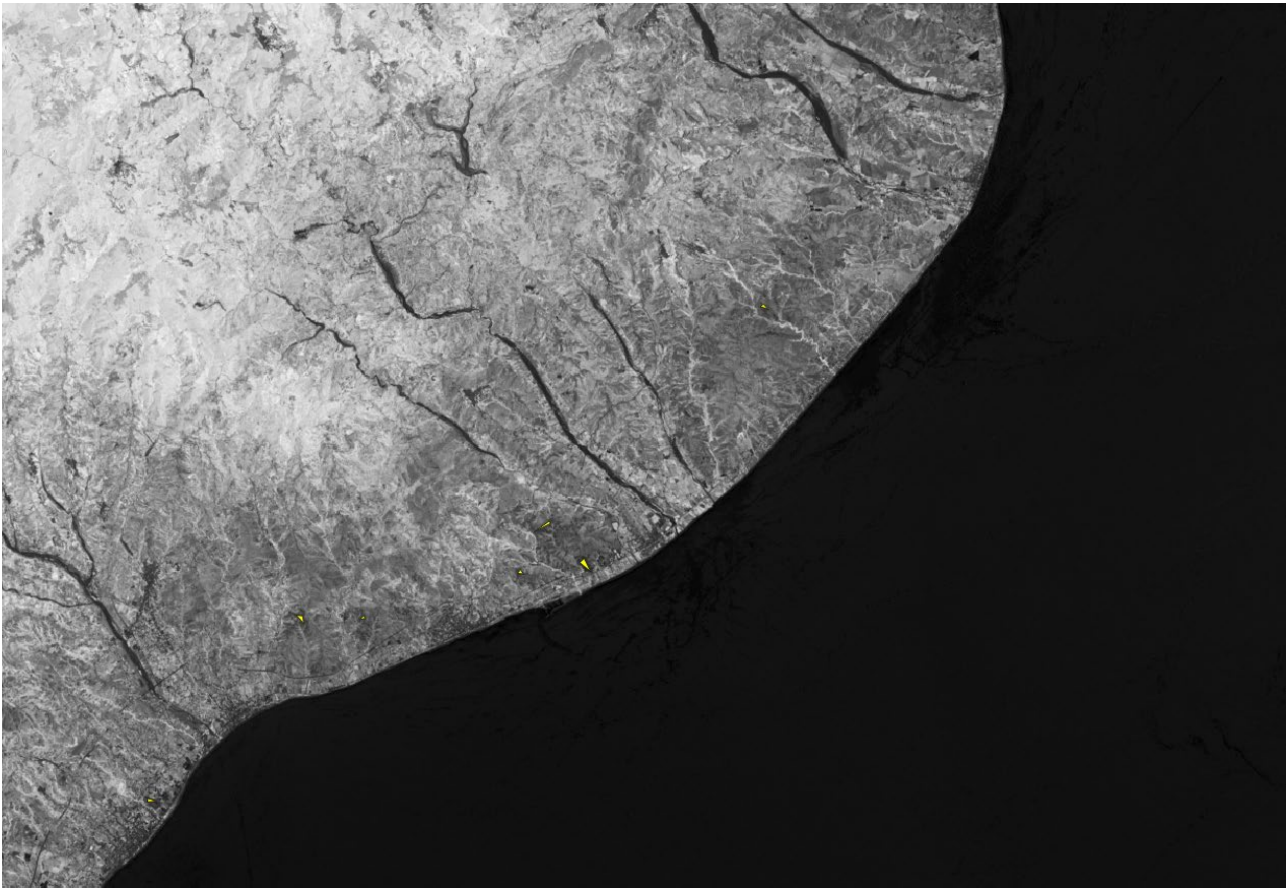


Acqua :



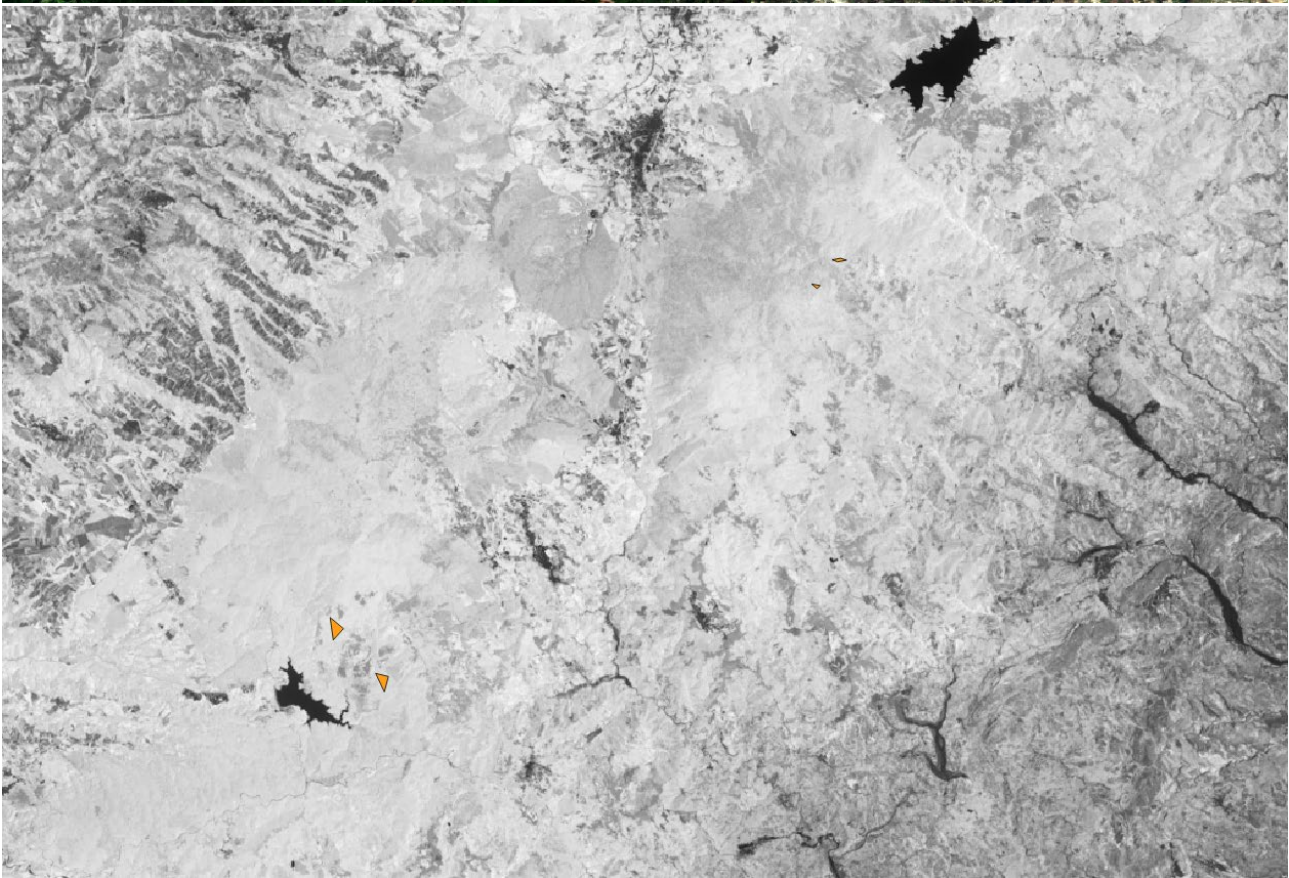


Suolo:





Vegetazione:



Devo dissolvere i Traing Sites attraverso gli strumenti di Geoprocessing . Consiste nel raggruppare i training sites in unica classe di riferimento. Dopo aver raggruppato tutti i training sites , utilizzo statistiche zonali per ricavarli i valori di ciascuna classe .

	Numero Pixel	Media	Dev.St
Acqua	287	-0,01279	0,012433663
Suolo	207	0,111388	0,032429753
Vegetazione	180	0,556877	0,01939436

Ho calcolato i punti di equilibrio  $s^*$  tra la dev.standard delle 2 classi mediante l'utilizzo della formula :

$$s^* \sigma_1 + s^* \sigma_2 = \mu_2 - \mu_1$$

$S^*$  è il punto di equilibrio ;

$\sigma_1$  e  $\sigma_2$  sono le dev.standard della classe 1 e classe 2 ;

Fatto per i valori della classe acqua-suolo e suolo-vegetazione .

S Acqua – Suolo 2,767948894

S Suolo – Vegetazione 8,596177115

Con questi valori posso calcolarmi la soglia di entrambe tramite la formula :

$$S_1 = \mu_1 + s^* \sigma_1$$

$S^*$  è il punto di equilibrio ;

$\mu_1$  è il valore medio della prima classe ;

$\sigma_1$  è la dev.standard della prima classe

Soglia Acqua – Suolo 0,021624148

Soglia Suolo-Vegetazione 0,390159949

Calcolate le soglie di separazione , trasformato l'NDVI in una carta tematica a 3 classi , dove le mie soglie sono 0,021 e 0,390 .



I calcoli li svolgo con il calcolatore raster , dove applico la formula :

$$(\text{"NDVI@1"} \leq 0.021624148) * 1 + (\text{"NDVI@1"} > 0.021624148 \text{ AND } \text{"NDVI@1"} \leq 0.390159949) * 2 + (\text{"NDVI@1"} > 0.390159949) * 3$$

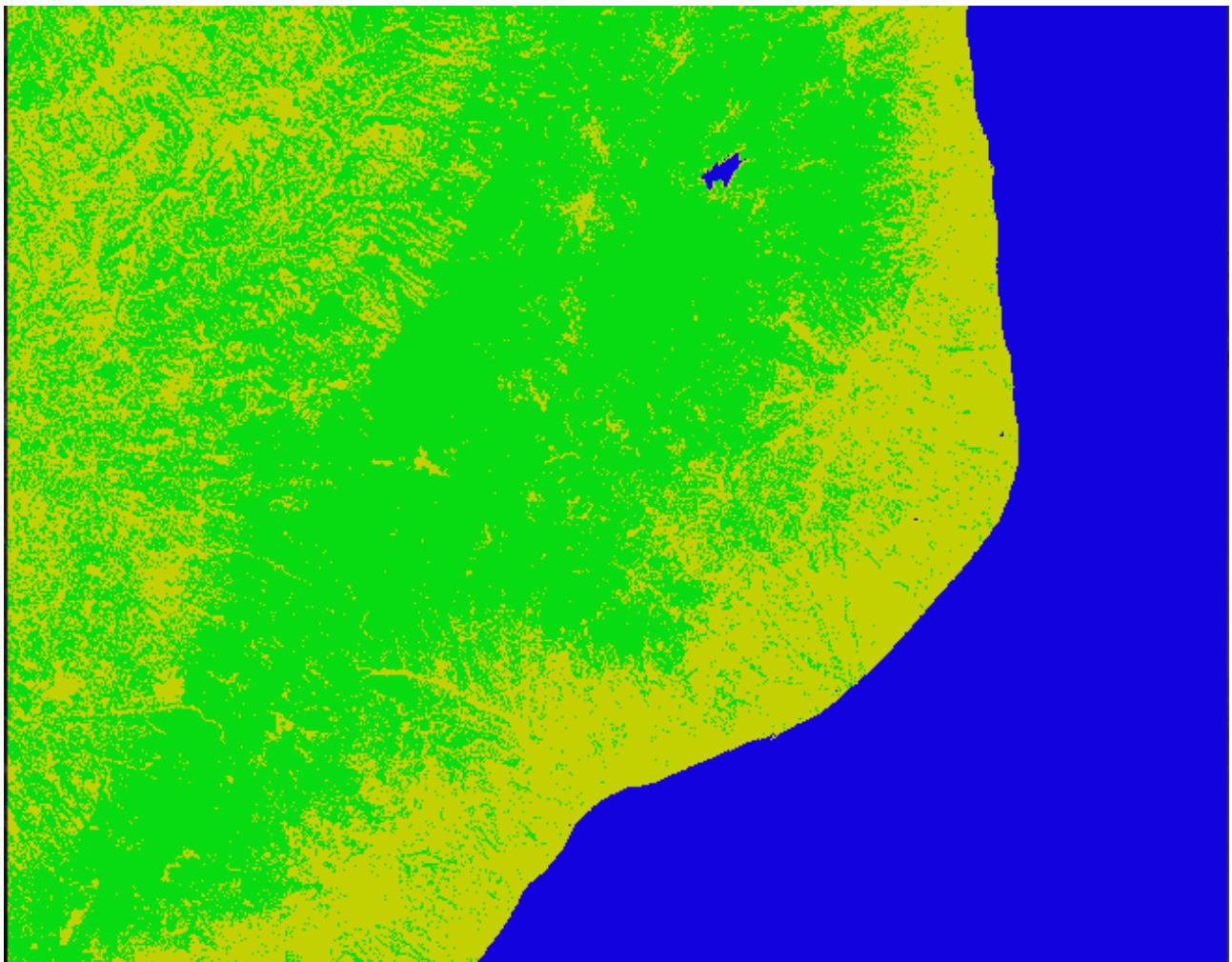
se l'ndvi è minore-uguale al valore limite che ho trovato, assegna il valore 1;

se l'ndvi è compreso tra i due valori limite, assegna il valore 2;

se l'ndvi supera il valore limite, assegna il valore 3.

Con questa formula classifichiamo l' ndvi e lo trasformiamo in una carta tematica a 3 colori.

Ecco il mio risultato :





## ESERCITAZIONE 4

### ACCURATEZZA TEMATICA

L'accuratezza tematica serve a stabilire i risultati ottenuti dai test sites, distinti dai training sites, ma come essi rappresentativi delle singole classi.

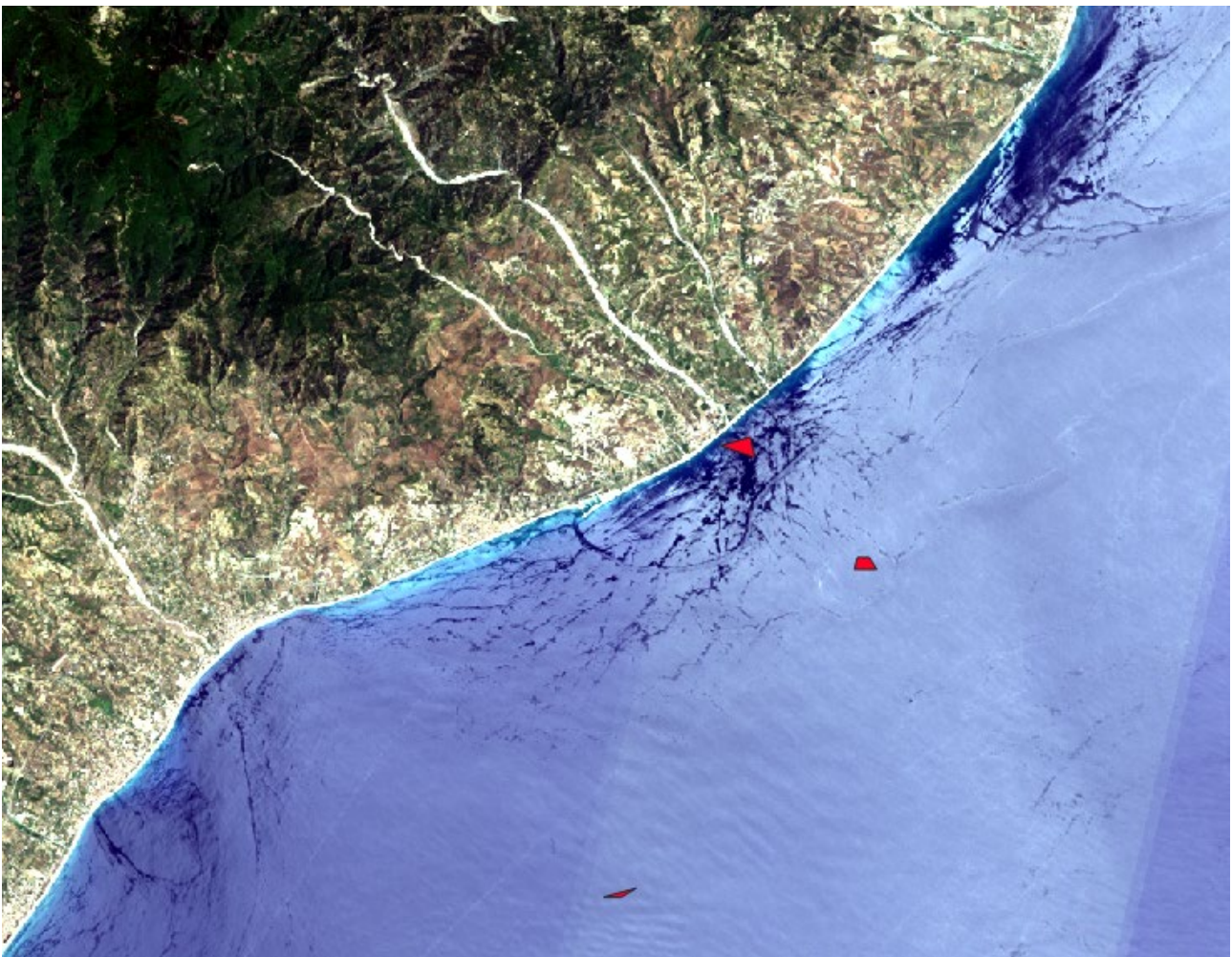
Il primo passo è quello di caricare le bande che mi interessano , carico anche l'indice NDVI che ho classificato nell'esercitazione 3 ed inizio a creare quelli che saranno i miei Test Sites .

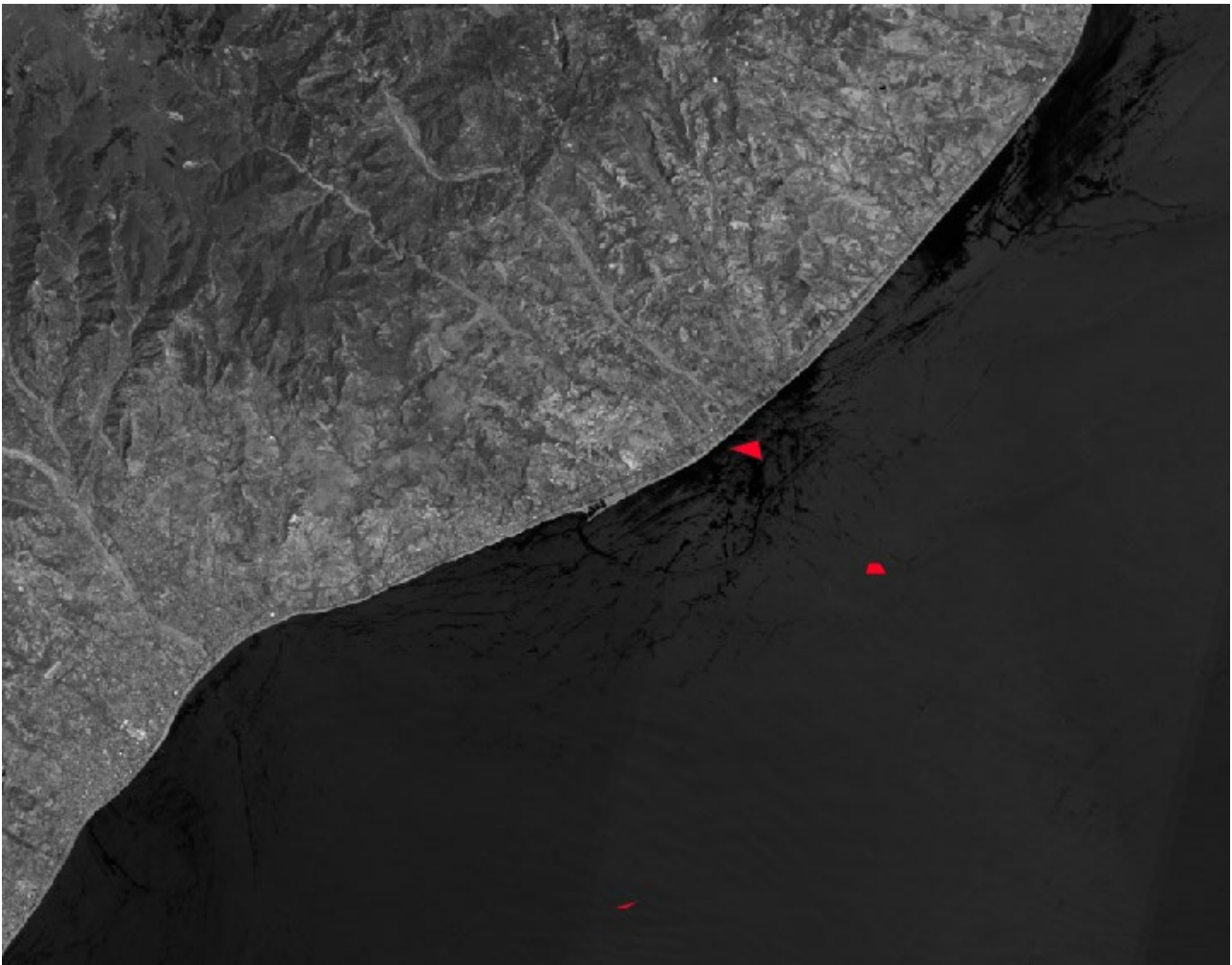
Ho 3 classi di copertura , quindi devo creare 3 test sites diversi che successivamente mi serviranno per creare la matrice di confusione (restituisce una rappresentazione dell'accuratezza di classificazione ).

La matrice di confusione è una matrice MxM alla quale andrò a classificare i pixel delle 3 classi ,per poter capire quanti pixel appartengono a quella classe . Raggruppati tutti i test sites nella stessa categoria (Acqua id =1 , Suolo id=2, Vegetazione id =3).

Ecco i miei test sites :

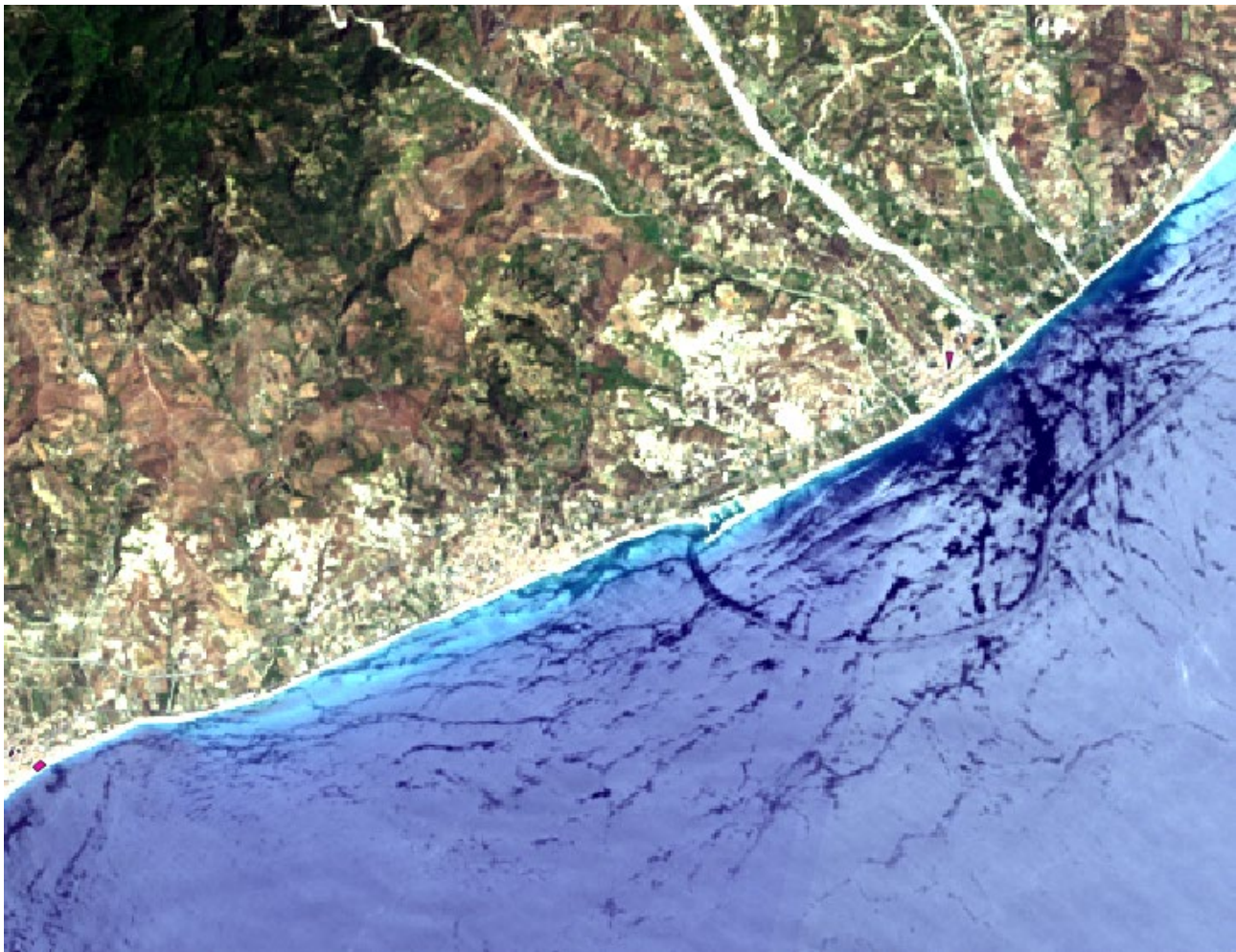
Per l'acqua







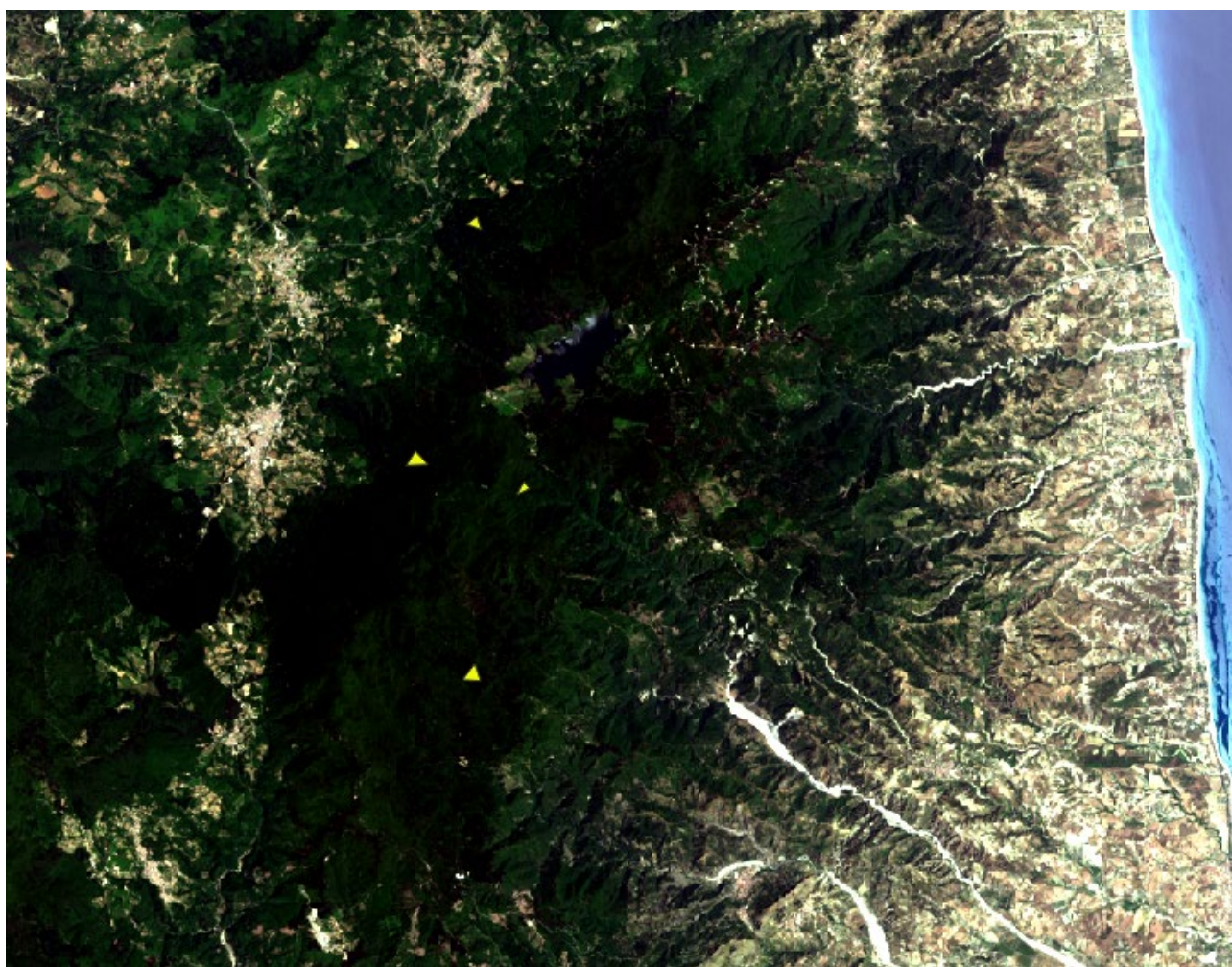
Per il suolo

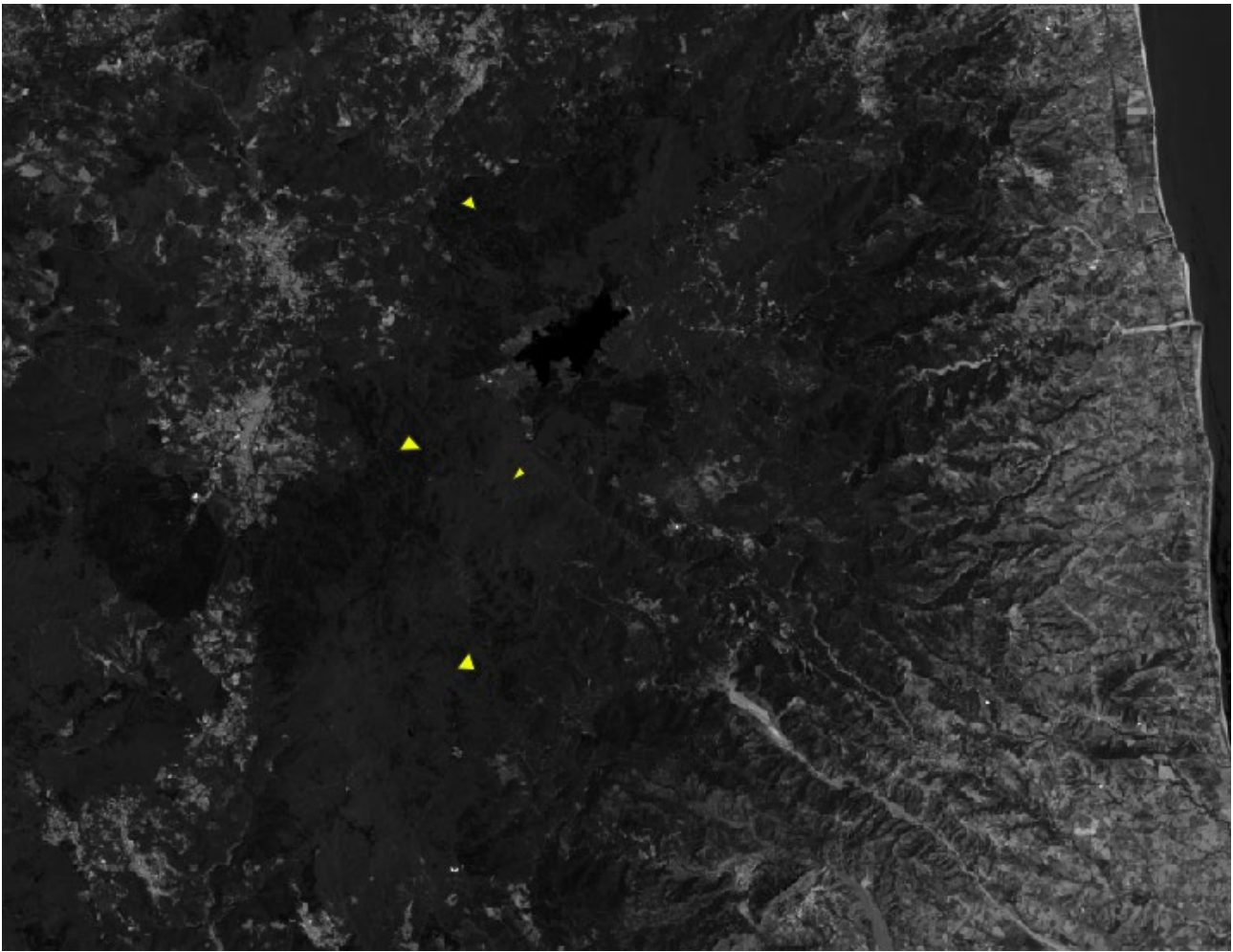






Per la Vegetazione:





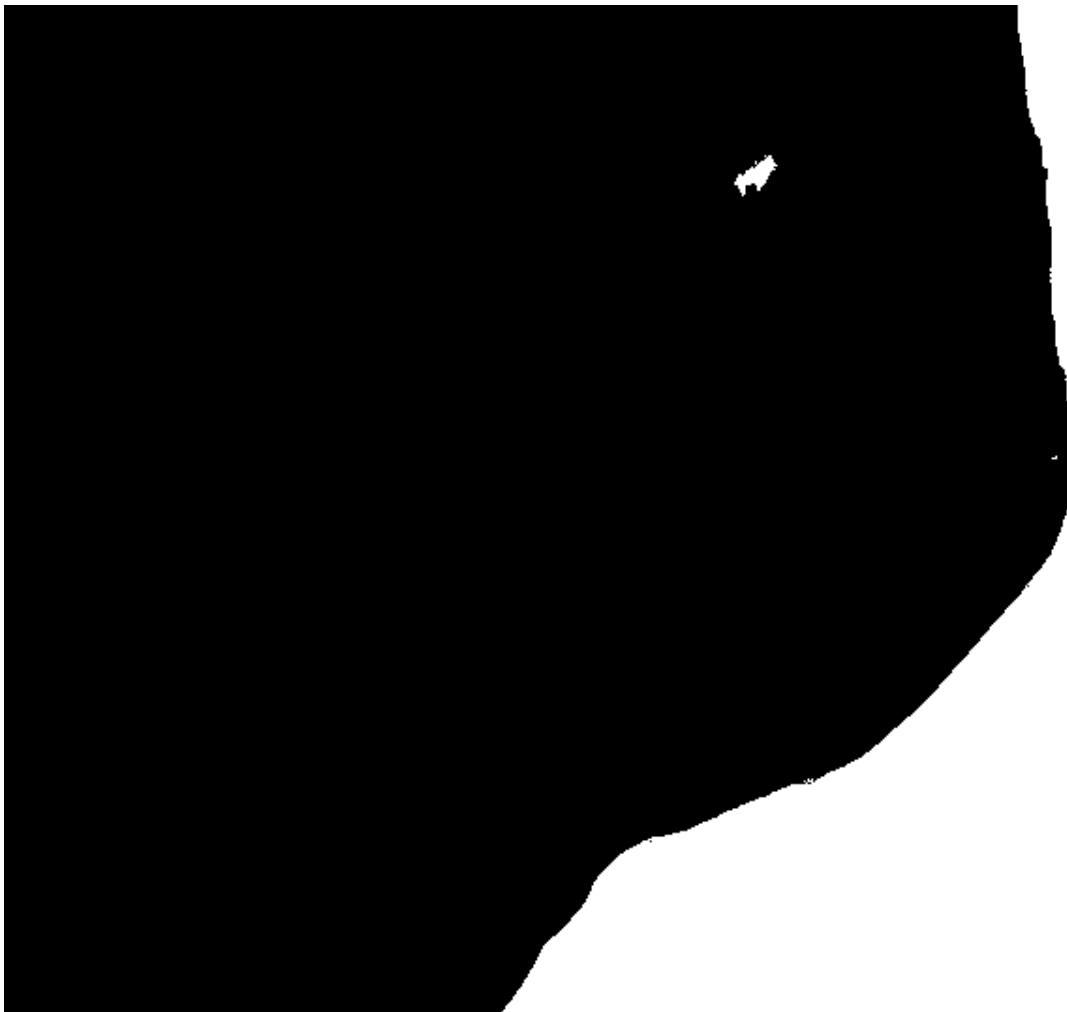
Riclassifico l'NDVI per creare 3 file raster , il primo file raster con Acqua pari a 1 e suolo e vegetazione pari a 0 , il secondo file raster con Suolo pari a 1 ed acqua e vegetazione pari a 0 , il terzo file raster con Vegetazione pari a 1 ed acqua e suolo pari a 0 .

I 3 file raster sono calcolati mediante calcolatore raster attraverso la formula (nel mio caso il mio NDVI è rinominato Classificazione ) :

Per Acqua e salvato come Ndzi Acqua1 :

```
( "Classificazione@1" =1 ) *1 + ("Classificazione@1">1) *0
```

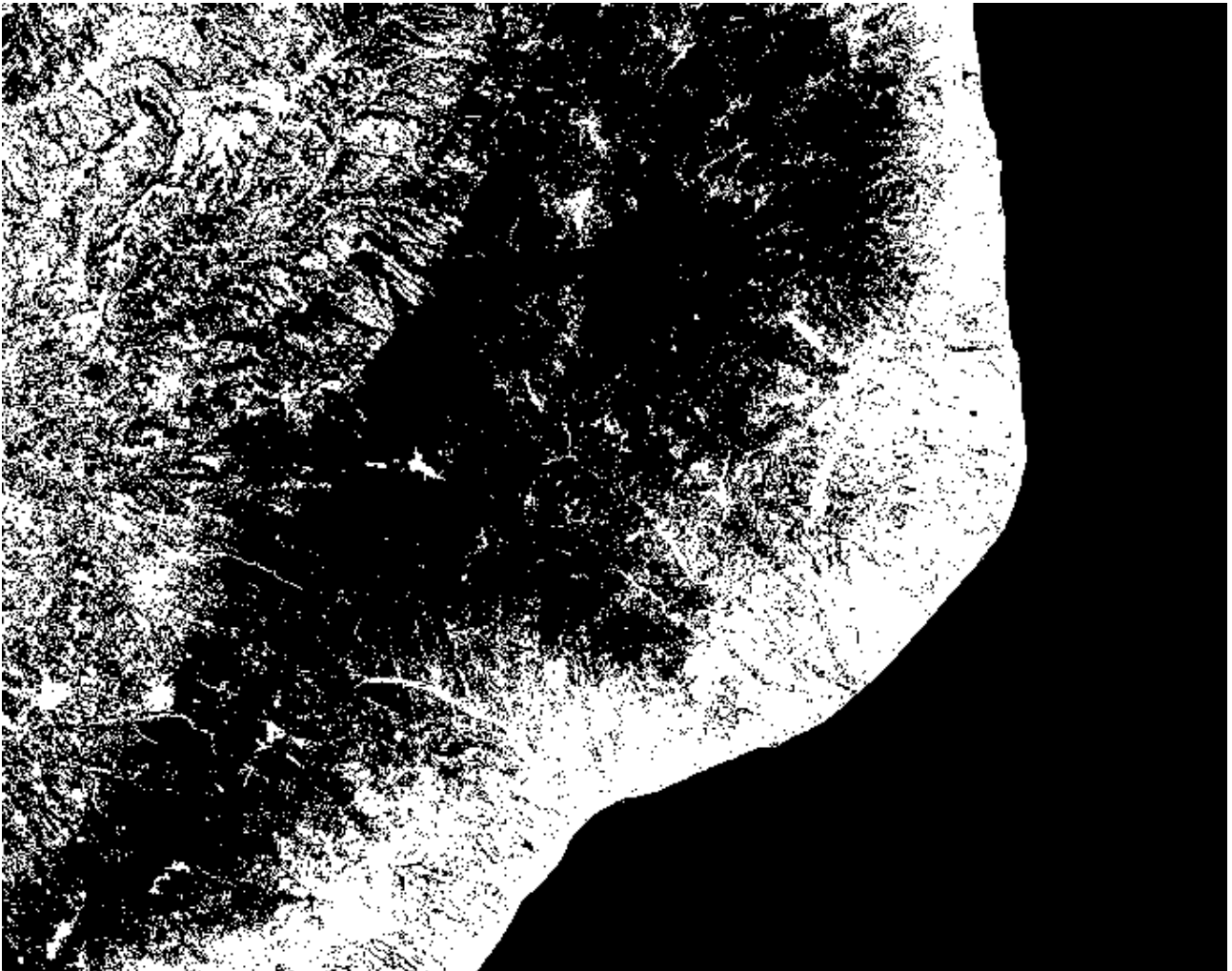
Questo è il risultato :



Per Suolo e salvato come Ndzi Suolo1:

$( \text{"Classificazione@1"}=1) * 0 + (\text{"Classificazione@1"}=2 ) * 1 + ( \text{"Classificazione@1"}=3) * 0$

Questo è il risultato , in nero sia acqua che vegetazione e bianco tutto ciò che è suolo :

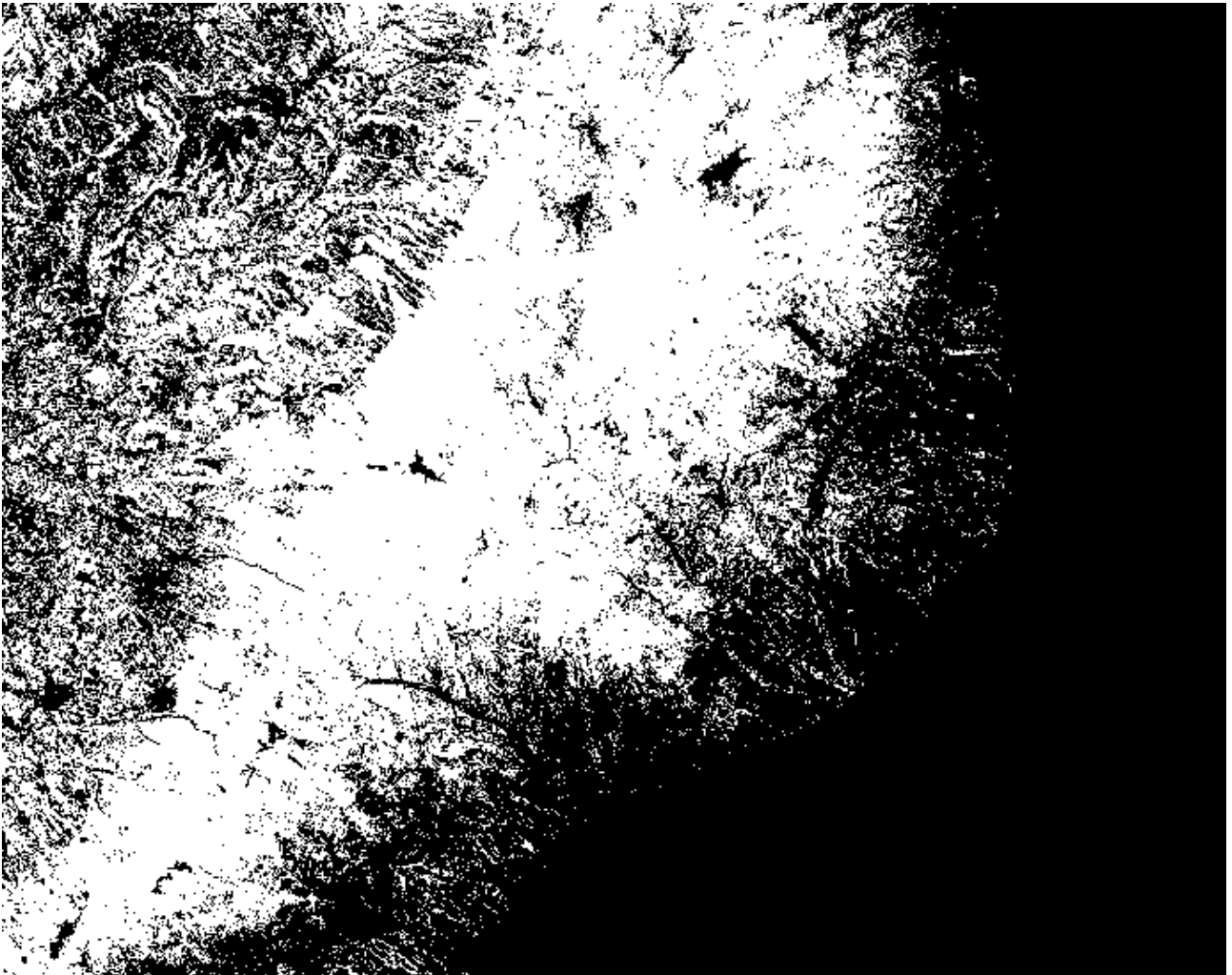




Per Vegetazione e salvato come Ndpi Vegetazione:

```
( "Classificazione@1" < 3)*0+("Classificazione@1"=3)*1
```

Questo è il risultato ,tutto ciò che è bianco è vegetazione :



I primi 3 file raster permettono di capire velocemente il grado di accuratezza dei pixel di una classe .

Per ogni raster riclassificato si utilizza lo strumento statistica zonale , usando funzioni di somma e conteggio .

Si procede con la statistica zonale sul vector dei test site applicata al raster riclassificato dell'acqua che ci darà, con la somma, i pixel di acqua classificati come acqua; la statistica zonale, sullo stesso vector, per il raster riclassificato del suolo ci darà i pixel di acqua erroneamente classificati come suolo; la differenza tra il numero totale dei pixel dei test site dell'acqua e la somma dei pixel delle precedenti due classi già considerate (acqua e suolo) fornisce i pixel erroneamente classificati come vegetazione.

Si ripete il tutto per i vector restanti e sempre per gli stessi raster riclassificati.

Una volta trovati i valori di somma e conteggio delle 3 classi, posso iniziare a costruire la matrice di confusione.

	Acqua	Suolo	Vegetazione	
Acqua	442	0	0	442
Suolo	18	135	0	153
Vegetazione	0	1	274	275
	460	136	274	870 Totale Pixel dei miei TestSites

Calcolo (Producer Accuracy , User Accuracy , Overall Accuracy)

1)Producer Accuracy, ovvero la frazione di pixel correttamente classificati e fra i pixel totali di test di una data classe;

2)User Accuracy, definisce la frazione dei pixel classificati correttamente fra i pixel assegnati ad una data classe;

3)Overall Accuracy, definisce la percentuale di pixel classificati correttamente sull'intero test sites.

Producer Accuracy	1	0,882353	0,996364
User Accuracy	0,96087	0,992647	1
Overall Accuracy	0,978161		