

Dottorato di ricerca in  
Metodi e Tecnologie per il Monitoraggio Ambientale  
Ciclo XII

Stima della concentrazione dei costituenti otticamente attivi  
dell'acqua marina mediante dati raccolti da sensori satellitari  
passivi multispettrali

Dottorando: Raffaele Grasso

Tutori:

Prof. Enzo Dalle Mese

Prof. Giovanni Corsini

---

# Introduzione

L'attività di ricerca che sarà presentata in questo lavoro di tesi riguarda lo studio del colore marino da dati acquisiti da sensori passivi multispettrali nella regione dello spettro elettromagnetico del visibile. Gli studi di colore marino tramite dati telerilevati hanno come scopo la stima delle concentrazioni dei costituenti otticamente attivi dell'acqua di mare che si trovano in essa sospesi od in soluzione. Tali costituenti fissano, interagendo con la radiazione luminosa proveniente dal sole, quelle proprietà ottiche che l'occhio umano percepisce come colore.

Il colore marino è una proprietà quantificabile attraverso la riflettanza spettrale sub-superficiale, grandezza che rappresenta, al variare della lunghezza d'onda, l'energia elettromagnetica che emerge dalla superficie marina, e che ha interagito con i costituenti otticamente attivi dell'acqua di mare, rapportata all'energia totale che giunge sulla superficie stessa dall'atmosfera (il rapporto è valutato immediatamente al di sotto della superficie di interfaccia aria-acqua da cui l'appellativo di sub-superficiale dato alla riflettanza). Lo scopo dell'intero lavoro è quello di definire uno schema d'inversione che dalla riflettanza sub-superficiale stimata dalle misure satellitari permetta di risalire alle concentrazioni dei costituenti otticamente attivi dell'acqua marina.

Nel presente studio si farà riferimento a due classi di costituenti che oltre all'acqua di mare pura contribuiscono a fissare le proprietà ottiche dell'acqua marina: le particelle in sospensione e la sostanza organica in soluzione. Le particelle in sospensione sono a loro volta classificabili in fitoplancton e particelle non clorofillacee. Il *Fitoplancton* è costituito da alghe unicellulari che interagiscono con la radiazione solare tramite pigmenti quali la clorofilla-a e la feofitina-a. Tali organismi sono localizzati all'interno della cosiddetta zona eufotica, lo strato in cui si concentra il 90% dell'energia solare che raggiunge la superficie d'interfaccia aria-acqua, e sono in grado di produrre carbonio organico tramite il processo di fotosintesi clorofilliana, assorbendo anidride carbonica dagli strati superficiali d'acqua ed assumendo nutrienti costituiti da fosfati e sostanze azotate provenienti dagli strati sottostanti. Le particelle non clorofillacee possono essere di natura organica od inorganica. Il particellato di natura organica è costituito in massima parte da microrganismi, batteri o virus, le cui dimensioni sono tali da rendere possibile l'interazione con la radiazione solare nel visibile da cui sono investiti. Le particelle di origine inorganica sono costituite principalmente da silicati e carbonati di origine sia biologica (ad esempio i coccoliti, scheletri carbonatici di alcune specie di alghe fitoplanctoniche) che terrestre (apporti in mare di detriti da parte di fiumi o scarichi industriali, erosione delle coste). La sostanza organica in soluzione o *sostanza gialla* è costituita dal prodotto di degradazione di sostanze organiche incluso il fitoplancton stesso.

Gli studi di colore marino trovano applicazione in vari campi. I più importanti riguardano il monitoraggio delle zone costiere, gli studi sulla circolazione oceanografica, la quantificazione del flusso di carbonio in atmosfera e l'osservazione di fenomeni climatici su larga scala. Il crescente interesse manifestato intorno a

tale disciplina e l'importanza che sta acquisendo negli anni, è dimostrato dall'alto numero di missioni organizzate dalle principali agenzie spaziali che prevedono, per gli anni a venire, la messa in orbita di sensori appositamente progettati per il monitoraggio del colore marino. Tra le varie missioni organizzate si ricordano la missione MERIS-ESA (su satellite ENVISAT) e le missioni della NASA MODIS-AM e MISR (su satellite EOS).

L'attività di ricerca è stata focalizzata sul problema della definizione degli algoritmi di stima delle concentrazioni dei costituenti otticamente attivi dalla riflettanza sub-superficiale, ricavata dai dati acquisiti dal sensore in ognuno dei suoi canali spettrali. L'approccio usato è di tipo *model based*: gli algoritmi di stima sono stati definiti tramite l'inversione di un modello diretto di riflettanza sub-superficiale che permette di calcolare le riflettanze medie sulle bande del sensore, dai valori delle concentrazioni dei parametri otticamente attivi. L'inversione viene effettuata scegliendo un modello inverso parametrico che accetta in ingresso il vettore delle riflettanze medie sulle bande del sensore e restituisce la stima della concentrazione del costituente di volta in volta preso in considerazione. I parametri del modello inverso sono stati stimati presentando in ingresso al modello stesso una serie di vettori di riflettanze e confrontando il valore dell'uscita desiderata con quella attuale. L'insieme di vettori di riflettanze, utilizzato per l'identificazione del modello inverso, è stato generato fornendo in ingresso al modello diretto una serie di valori della concentrazione dei costituenti otticamente attivi misurati nel bacino di interesse.

I modelli inversi parametrici che sono stati presi in considerazione in tale studio sono il modello multilineare ed il modello a rete neurale RBF (*Radial Basis Function*) il che costituisce una delle principali innovazioni presenti nel lavoro. I modelli inversi sono stati testati utilizzando una serie di dati simulati che sono rappresentativi di alcune delle condizioni che è possibile ritrovare nella realtà. A tale scopo si è reso necessario uno studio riguardante la variabilità statistica delle concentrazioni dei costituenti otticamente attivi. In particolare è stato presentato un modello di densità di probabilità delle concentrazioni che è stato validato con misurazioni *in situ* acquisite nel Mar Mediterraneo.

Le tecniche di inversione proposte sono state successivamente utilizzate per definire degli algoritmi di stima per un sito di sperimentazione situato nel Mediterraneo occidentale nei pressi dello Stretto di Gibilterra che è stato oggetto di studio nell'ambito del progetto OMEGA, progetto finanziato dall'Unione Europea il cui lo scopo è stato quello di studiare i fenomeni di circolazione che caratterizzano tale zona del Mediterraneo.

La tesi è organizzata in un primo capitolo in cui sono affrontati alcuni temi di carattere generale che riguardano il telerilevamento passivo applicato agli studi di colore marino. Segue un secondo capitolo in cui si affrontano le tematiche riguardanti la propagazione della radiazione solare attraverso l'atmosfera ed in cui si introduce il modello diretto di riflettanza sub-superficiale che è stato utilizzato per la definizione degli algoritmi di stima delle concentrazioni. Il terzo capitolo è dedicato ai metodi di inversione del modello diretto di riflettanza. Il quarto capitolo ha come argomento lo studio della variabilità statistica dei costituenti otticamente attivi mentre nel quinto ed ultimo capitolo sono riportati alcuni risultati sperimentali ottenuti applicando le tecniche di inversione ai dati simulati e misurati *in situ*.



---

# Capitolo 1

## *Il telerilevamento passivo e gli studi di colore marino*

Nel presente capitolo saranno affrontati alcuni temi di carattere generale ed introduttivo sul telerilevamento passivo applicato allo studio del colore marino. In particolare saranno presi in esame i seguenti aspetti:

- il problema generale del telerilevamento ed il suo uso nella determinazione di parametri d'interesse ambientale ai fini del monitoraggio della superficie terrestre, con particolare riferimento agli studi di colore marino;
- i costituenti otticamente attivi dell'acqua di mare che determinano le proprietà ottiche di quest'ultima e la cui stima delle concentrazioni massa su volume costituisce l'oggetto degli studi di colore marino;
- l'utilità degli studi di colore marino dal punto di vista del monitoraggio ambientale;
- l'interazione della radiazione luminosa con la materia e le principali grandezze fisiche misurabili dai sensori per il telerilevamento passivo e che sono utilizzate per la stima dei parametri d'interesse caratterizzanti la superficie osservata;
- le caratteristiche dei sensori passivi satellitari dedicati agli studi di colore marino con particolare enfasi rivolta ai sensori d'ultima generazione che saranno presi in considerazione in questo lavoro di tesi.

### **1.1 Il telerilevamento passivo applicato agli studi di colore marino**

Il telerilevamento ha lo scopo di rilevare a distanza le proprietà di un oggetto tramite sensori che sono in grado di misurare l'energia elettromagnetica che l'oggetto stesso irradia nello spazio ad esso circostante. I sensori per il telerilevamento possono essere classificati sia in base alla piattaforma su cui sono montati e quindi si avranno ad esempio sensori di tipo aviotrasportato e sensori di tipo satellitare, sia in base al tipo di sorgente d'energia elettromagnetica che serve ad illuminare l'oggetto sotto osservazione. In questo secondo caso si parlerà di sensori attivi, che hanno cioè la capacità di trasmettere un'onda elettromagnetica di determinate caratteristiche (frequenza, lunghezza d'onda ed intensità) verso l'oggetto da monitorare, e di sensori passivi che sfruttano cioè la radiazione solare come sorgente d'illuminazione.

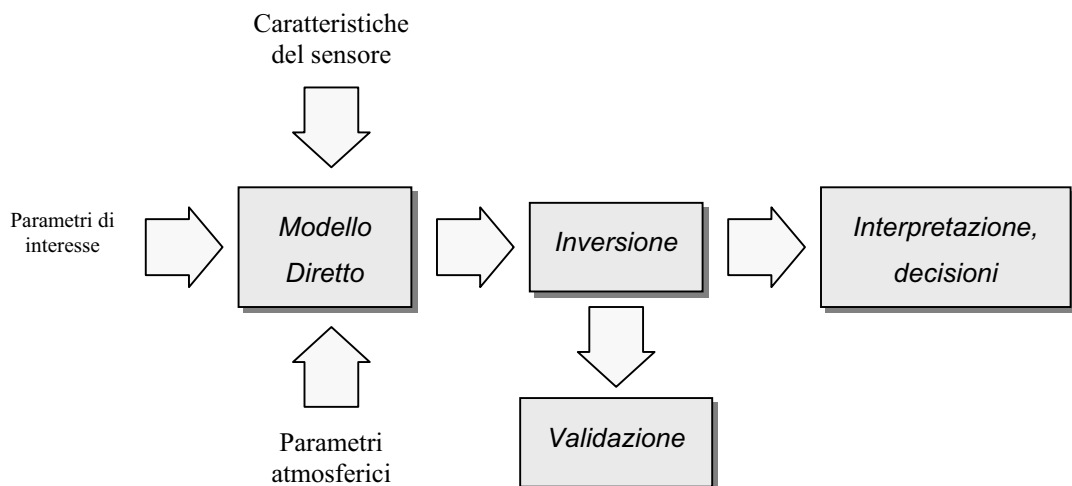
Nel presente lavoro saranno affrontati problemi inerenti al telerilevamento passivo da satellite applicato all'osservazione della superficie marina ai fini della stima di parametri che costituiscono dei veri e propri indici di qualità delle acque e più in generale dello stato dell'intero ecosistema marino.

Dal punto di vista sistemistico ogni problema di telerilevamento può essere inserito in uno schema generale di soluzione che comprende quattro fasi principali (Figura 1.1-1):

1. La definizione del problema diretto ovvero la definizione di un modello diretto che permette di calcolare il segnale di uscita al sensore fissato il valore di un determinato

numero di parametri di ingresso. Tali parametri specificano le proprietà dell'oggetto osservato, lo stato dell'atmosfera attraverso la quale avviene la propagazione, verso il sensore, della radiazione proveniente dalla superficie illuminata e le caratteristiche del sensore. Alla base del modello diretto stanno quei fenomeni fisici che regolano l'interazione della superficie illuminata con la radiazione elettromagnetica che incide su di essa. Fissato il particolare tipo di sensore e le caratteristiche di propagazione attraverso l'atmosfera, il modello permette di prevedere il comportamento dell'oggetto osservato in risposta a determinate condizioni di ingresso.

2. La stima dei parametri che determinano il comportamento del sistema osservato ovvero la risoluzione del problema inverso. Questa è la fase in cui il modello diretto è inserito in uno schema di inversione che permette di stimare dalle misure effettuate dal sensore i parametri di interesse.
3. La validazione degli algoritmi d'inversione tramite il confronto delle stime ottenute con dati di verità a terra.
4. Interpretazione ed uso dei risultati ottenuti. I parametri stimati consentono di analizzare il comportamento del sistema osservato nello spazio e nel tempo e permettono di prendere decisioni sulla gestione delle risorse naturali.



*Figura 1.1-1 Il problema del telerilevamento dal punto di vista sistemistico.*

Il telerilevamento passivo si avvale dell'uso di sensori che misurano l'intensità della radiazione solare reirradiata dalla superficie osservata. Tali sensori, montati in volo su piattaforme aeree o satellitari, effettuano una scansione della superficie mediante dispositivi meccanici od elettronici e compiono la misura in diversi canali spettrali centrati su determinate lunghezze d'onda della radiazione emessa. Le lunghezze d'onda sono scelte sulla base di alcune caratteristiche fisiche della superficie determinabili a priori. I diversi canali sono caratterizzati da una forma della risposta spettrale che ne fissa la larghezza di banda e sono posizionati all'interno di una particolare regione dello spettro elettromagnetico. Le applicazioni di tali sensori (detti multispettrali) dipendono ovviamente dalla particolare configurazione spettrale, si avranno in tal modo sensori adatti ad applicazioni di tipo terrestre o marino e sensori che permettono la misura di determinati parametri caratteristici della superficie considerata piuttosto che di altri. In questo lavoro di tesi l'applicazione che verrà presa in considerazione riguarda lo studio del colore marino da dati

acquisiti da sensori passivi nella regione dello spettro elettromagnetico del visibile. Gli studi di colore marino attraverso dati telerilevati hanno come scopo la stima delle concentrazioni dei costituenti otticamente attivi dell'acqua di mare che si trovano in essa sospesi o in soluzione. Tali costituenti fissano, interagendo con la radiazione luminosa proveniente dal sole, quelle proprietà ottiche che l'occhio umano percepisce come colore. Il colore marino è una proprietà quantificabile attraverso la riflettanza spettrale sub-superficiale, grandezza che rappresenta, al variare della lunghezza d'onda, l'energia elettromagnetica che emerge dalla superficie marina, e che ha interagito con i costituenti otticamente attivi dell'acqua di mare, rapportata all'energia totale che giunge sulla superficie stessa dall'atmosfera (il rapporto è valutato immediatamente al di sotto della superficie di interfaccia aria-acqua da cui l'appellativo di sub-superficiale dato alla riflettanza). Lo scopo dell'intero lavoro è quello di definire uno schema d'inversione che dalla riflettanza stimata dalle misure satellitari permetta di risalire alle concentrazioni dei costituenti otticamente attivi dell'acqua marina. Il problema si inserisce a livello dei primi due blocchi dello schema di Figura 1.1-1 relativi alla modellistica e alle procedure di inversione. In virtù di quanto su esposto la schematizzazione del problema può essere ampliata come mostrato in Figura 1.1-2 e Figura 1.1-3.

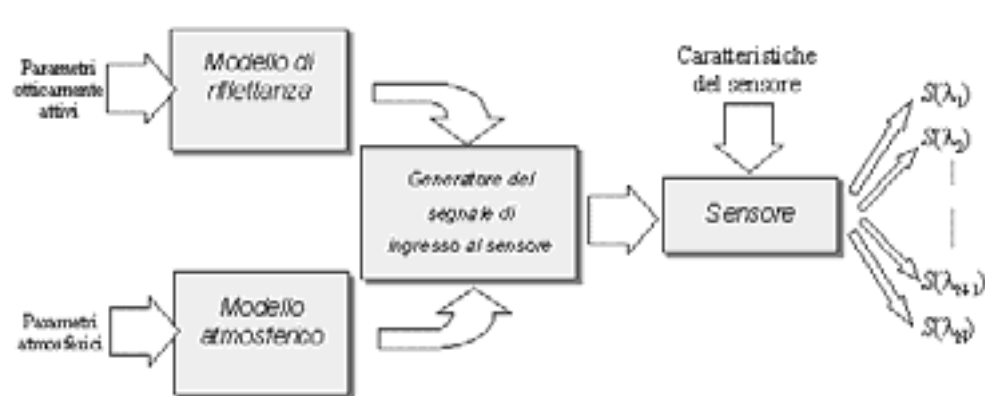


Figura 1.1-2: schematizzazione del problema della stima dei parametri otticamente attivi da dati satellitari: il problema diretto.

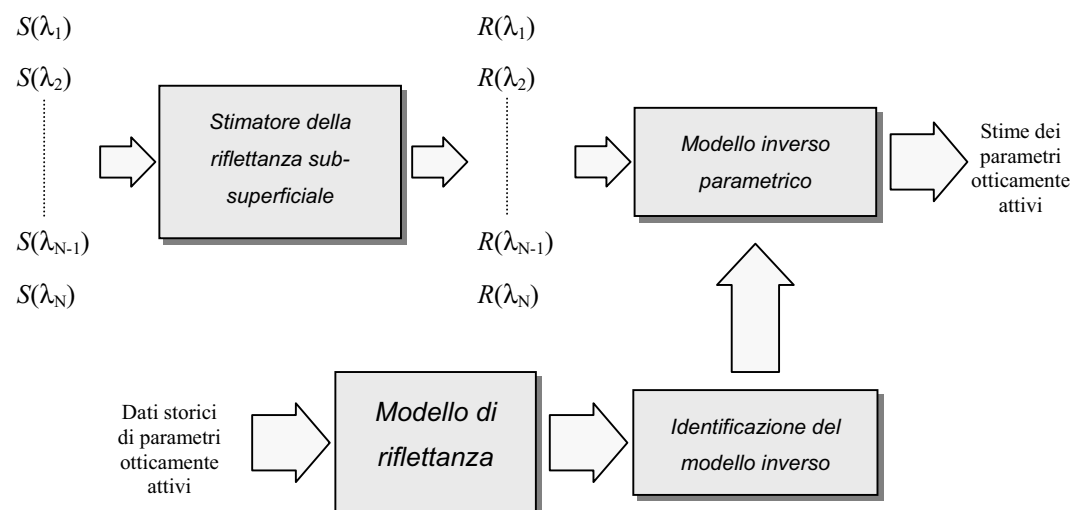


Figura 1.1-3: schematizzazione del problema della stima dei parametri otticamente attivi da dati satellitari: il problema inverso.

Il modello diretto, fissato il valore dei parametri otticamente attivi ed atmosferici, permette la generazione del segnale in ingresso al sensore. Specificate le caratteristiche di quest'ultimo il modello diretto è in grado di calcolare il segnale  $S(\lambda_i)$  in uscita in corrispondenza degli N canali del sensore centrati sulle

lunghezze d'onda  $\lambda_1, \dots, \lambda_N$ . Invertendo il modello atmosferico si è in grado di valutare la riflettanza sub-superficiale  $R(\lambda_i)$  per ogni canale del sensore e quindi l'informazione utile alla stima dei parametri otticamente attivi. La stima viene effettuata tramite l'uso di un modello inverso parametrico. I parametri del modello inverso sono stimati attraverso una procedura di identificazione che si avvale dell'uso congiunto di una serie di dati storici di concentrazioni di costituenti otticamente attivi e del modello diretto di riflettanza. L'algoritmo di inversione generato risulta avere validità limitata sia nello spazio che nel tempo in dipendenza della particolare regione geografica e del periodo dell'anno in cui è stata raccolta la serie di dati storici.

Il presente lavoro di tesi sarà in massima parte indirizzato alla soluzione del problema dell'identificazione del modello inverso di riflettanza tramite l'utilizzo di varie tecniche che saranno introdotte nei capitoli successivi.

## 1.2 I costituenti otticamente attivi dell'acqua marina

In questo paragrafo saranno introdotti e descritti brevemente i principali costituenti otticamente attivi dell'acqua di mare che saranno presi in considerazione nel proseguo. Non si farà in tale ambito cenno alcuno alle proprietà ottiche di tali costituenti dato che queste saranno oggetto di capitoli successivi.

Nel presente studio si farà riferimento a due classi di costituenti che oltre all'acqua di mare pura contribuiscono a fissare le proprietà ottiche dell'acqua marina: le particelle in sospensione e la sostanza organica in soluzione. Le particelle in sospensione sono a loro volta classificabili in fitoplancton e particelle non clorofillacee. L'approccio a tre componenti ha una validità riscontrabile in letteratura in numerosi lavori scientifici tra i quali, quello di *Sathyendranath et al.* [1] è da considerarsi senza dubbio fondamentale. Il modello descritto in tale articolo costituirà la base del modello di riflettanza utilizzato in questo lavoro di tesi.

In Figura 1.2-1 è descritta la natura dei costituenti otticamente attivi. Il *Fitoplancton* è costituito da alghe unicellulari che interagiscono con la radiazione solare tramite pigmenti quali la clorofilla-a e la feofitina-a. Tali organismi sono localizzati all'interno della cosiddetta zona eufotica, lo strato in cui si concentra il 90% dell'energia solare che raggiunge la superficie d'interfaccia aria-acqua, e sono in grado di produrre carbonio organico tramite il processo di fotosintesi clorofilliana, assorbendo anidride carbonica dagli strati superficiali d'acqua ed assumendo nutrienti costituiti da fosfati e sostanze azotate provenienti dagli strati sottostanti. Il fitoplancton assume importanza primaria nell'equilibrio dell'ecosistema marino: esso costituisce il primo anello della catena alimentare e su larga scala influisce in maniera determinante sul processo di scambio dell'anidride carbonica tra aria ed acqua, e quindi sul ciclo globale del carbonio, ragioni queste che da sole giustificano la necessità di monitorarne la presenza nei mari.

Le particelle non clorofillacee possono essere di natura organica od inorganica. Il particolato di natura organica è costituito in massima parte da microrganismi, batteri o virus, le cui dimensioni sono tali da rendere possibile l'interazione con la radiazione solare nel visibile da cui sono investiti. Le particelle di origine inorganica sono costituite principalmente da silicati e carbonati di origine sia biologica (ad esempio i coccoliti, scheletri carbonatici di alcune specie di alghe fitoplanctoniche) che terrestri (apporti in mare di detriti da parte di fiumi o scarichi industriali, erosione delle coste).

La sostanza organica in soluzione o *sostanza gialla* è costituita dal prodotto di degradazione di sostanze organiche incluso il fitoplancton stesso. Notevoli quantità di sostanza gialla sono presenti nelle zone



costiere fortemente antropizzate. Il telerilevamento su scala locale della concentrazione sia di sostanza gialla che di particelle non clorofillacee, nei pressi di tali aree, risulta un valido strumento di monitoraggio delle sorgenti di inquinamento e della qualità delle acque. Su scala oceanica la sostanza gialla rappresenta una delle più grandi riserve di carbonio in forma ridotta presenti sulla superficie terrestre e che risulta non ancora del tutto quantificata. Il telerilevamento della sostanza gialla quindi assume importanza non solo su scala locale, ma anche su scala globale per gli studi sul ciclo del carbonio.

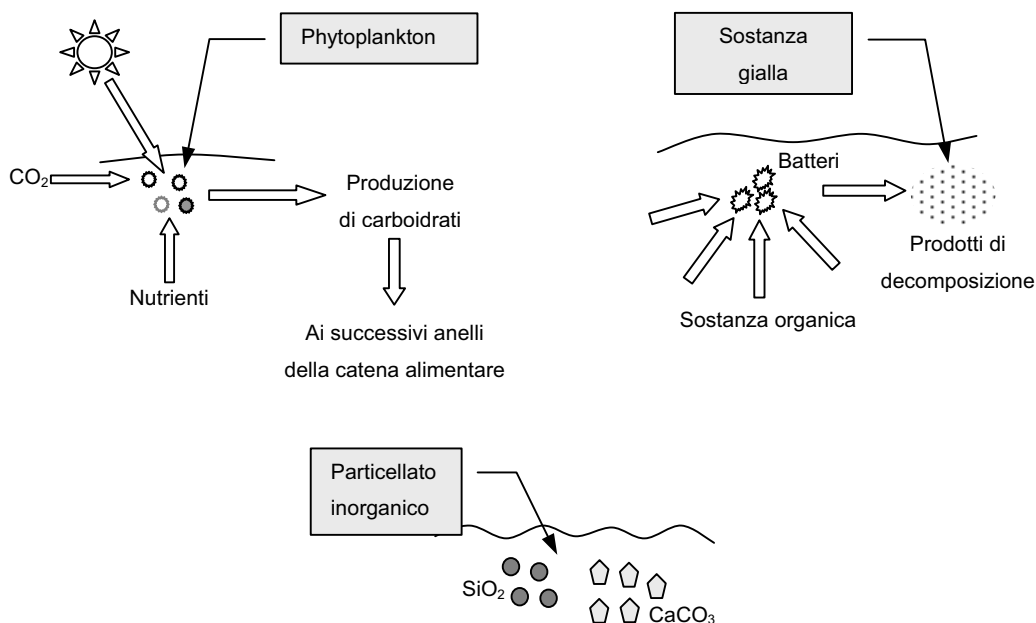


Figura 1.2-1: i principali costituenti otticamente attivi dell'acqua di mare.

### 1.3 L'importanza degli studi di colore marino

I dati di colore marino telerilevati da satellite possono trovare applicazione nell'ambito delle seguenti aree tematiche:

- quantificazione del flusso di carbonio in atmosfera e studio delle sue variazioni annuali;
- osservazione di fenomeni climatici su larga scala;
- studi di circolazione;
- monitoraggio delle zone costiere.

Lo scopo primario del telerilevamento del colore marino è quello di fornire indici della biomassa fitoplanctonica nei bacini oceanici. Il fitoplancton su scala globale consuma ogni anno circa cinquantamila milioni di tonnellate di anidride carbonica in un processo che prende il nome di produttività primaria. Il processo di scambio di carbonio tra atmosfera e mare ha quindi importanza notevole all'interno del ciclo globale del carbonio che influenza pesantemente l'andamento dei cambiamenti climatici terrestri (si pensi alle conseguenze dell'effetto serra). Gli studi di colore marino costituiscono un valido strumento di indagine scientifica in tale campo; i sensori satellitari per il colore marino sono in grado di fornire dati sinottici su larga scala e con elevata frequenza temporale. Gli indici della massa fitoplanctonica ricavati da questi dati possono essere quindi utilizzati con successo per stimare la portata del flusso di carbonio tra atmosfera e mare, per capire come tale flusso è controllato e come varia annualmente. Inoltre essendo la

massa fitoplanctonica sensibile ai cambiamenti climatici, i dati di colore marino sono utilizzati con notevole efficacia per monitorare cambiamenti climatici su larga scala sia spaziale che temporale permettendo di separare fenomeni di natura periodica (si pensi ad esempio al riscaldamento anomalo della superficie marina che si manifesta periodicamente nell'Oceano Pacifico, indicato con il nome di *El-Nino*), da fenomeni indotti da attività umane.

Lo studio del colore marino trova applicazione nel monitoraggio delle zone costiere caratterizzate da elevata densità di insediamenti industriali e di attività umane in genere. La stima delle concentrazioni dei costituenti otticamente attivi in tali zone permette il controllo delle varie sorgenti di inquinamento che provocano la variazione del carico di particelle in sospensione e di sostanze organiche in soluzione all'interno della colonna d'acqua. Il monitoraggio delle zone costiere permette di capire l'entità dell'impatto che le attività umane hanno su queste aree e quindi di predisporre dei piani di gestione integrata e coordinata di tali attività per favorirne lo sviluppo sostenibile. Gli insediamenti industriali hanno senz'altro il maggiore impatto ambientale sull'ecosistema marino, ma anche lo sviluppo incontrollato e malamente pianificato di attività quali il turismo e la pesca possono portare a gravi conseguenze. Le attività agricole inoltre hanno portato, con l'uso massiccio di fertilizzanti e pesticidi, ad un arricchimento di nutrienti nelle zone costiere. Tale apporto incontrollato di nutrienti stimola la crescita del fitoplancton favorendo la comparsa di fenomeni di crescita algale dannosi per l'ecosistema (fenomeni di eutrofizzazione, fioriture algali tossiche, mucillagini).

I costituenti otticamente attivi sono utilizzati infine come traccianti per gli studi di circolazione marina, per evidenziare vortici e correnti. Un esempio di tale applicazione sarà dato nei prossimi capitoli in cui si prenderà in considerazione un sito di sperimentazione oggetto di studio nell'ambito del progetto OMEGA, progetto finanziato dall'Unione Europea e che ha lo scopo di studiare i fenomeni di circolazione che caratterizzano il Mediterraneo occidentale nei pressi dello stretto di Gibilterra e la loro connessione con l'attività biologica.

## 1.4 Classificazione delle acque

Dal punto di vista delle caratteristiche ottiche è possibile distinguere due tipi principali di acqua marina [2]:

- acque di Caso I,
- acque di Caso II.

Le **acque di Caso I** sono quelle in cui il fitoplancton ha il ruolo dominante nel fissare le proprietà ottiche del corpo idrico ed in cui la presenza degli altri costituenti è fortemente correlata alla presenza del fitoplancton stesso. Le **acque di Caso II** sono quelle in cui gli altri costituenti, sostanza gialla e particellato solido in sospensione, variano indipendentemente dalla concentrazione di fitoplancton e hanno il ruolo fondamentale nel fissare le proprietà ottiche del corpo idrico.

Tale classificazione dal punto di vista matematico si traduce in differenti proprietà statistiche delle concentrazioni dei costituenti otticamente attivi in termini di valore medio, deviazione standard e correlazione statistica. Come sarà visto in seguito le proprietà statistiche delle concentrazioni dei costituenti risultano molto importanti ai fini della definizione degli algoritmi di stima. La forma di tali algoritmi viene a dipendere dal tipo di acqua contenuta nel bacino da monitorare. Le applicazioni del

monitoraggio dei costituenti otticamente attivi inoltre dipendono dal tipo di acqua considerata (Figura 1.4-1). Monitorare le acque di caso I significa monitorare, su scala globale o regionale, le acque oceaniche, ricche di fitoplancton, per cui le applicazioni che rivestono la maggiore importanza sono gli studi di circolazione, gli studi di produttività primaria sul ciclo del carbonio e l'osservazione dei cambiamenti climatici su larga scala. Il monitoraggio delle acque di caso II invece è il monitoraggio, su scala locale, delle zone costiere in cui, data l'alta densità di attività umane, la variabilità dei tre costituenti risulta più elevata. In tali zone dunque, le applicazioni che rivestono maggiore importanza sono quelle orientate alla valutazione della qualità delle acque e all'osservazione delle sorgenti di inquinamento ai fini della valutazione del loro impatto sull'ecosistema marino.

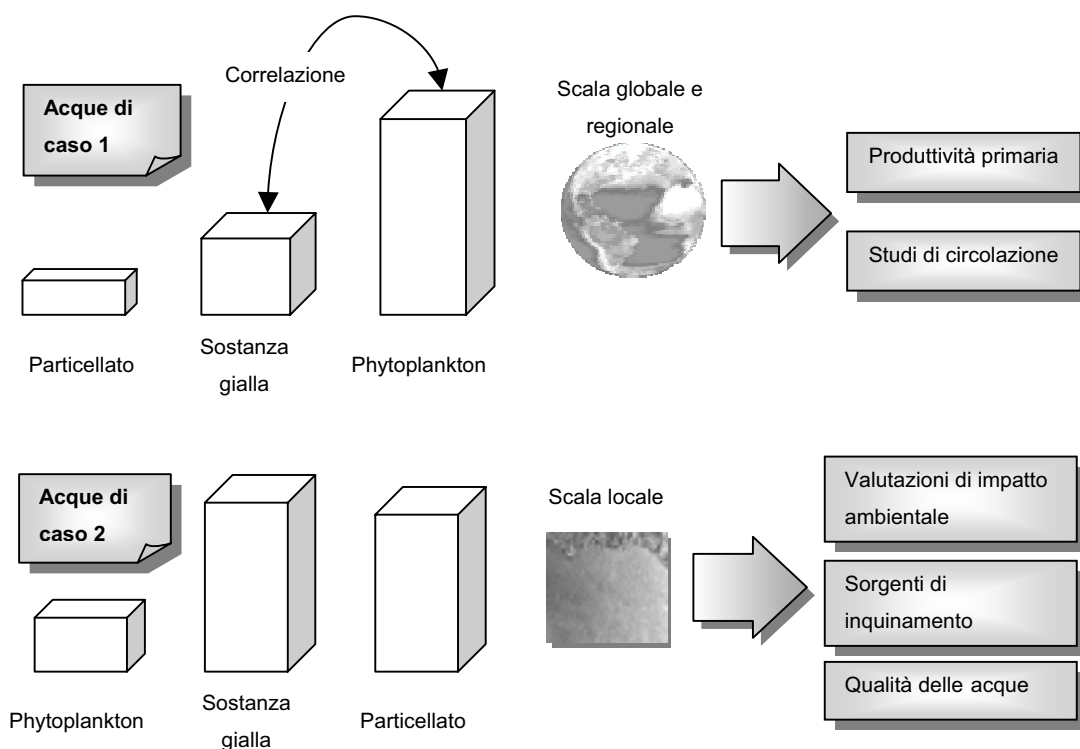


Figura 1.4-1: le acque di caso 1 e caso 2 e le applicazioni del telerilevamento dei costituenti otticamente attivi.

## 1.5 Grandezze radiometriche ed interazione della materia con la radiazione

In questo paragrafo saranno definite le principali grandezze radiometriche usate per la misura della radiazione elettromagnetica nella banda del visibile ed utilizzate nell'ambito del telerilevamento del colore marino; saranno inoltre introdotti i parametri che servono a quantificare i fenomeni di assorbimento e *scattering* che avvengono quando la materia interagisce con la radiazione elettromagnetica. Tali grandezze saranno usate in seguito per formalizzare il problema della stima dei costituenti otticamente attivi ovvero serviranno a definire il modello diretto schematizzato nel paragrafo 1.1.

### 1.5.1 Grandezze radiometriche.

Si indichi con  $\phi$  la potenza radiante misurata in watt, ovvero l'energia radiante  $Q$  (in joule) nell'unità di tempo, che emerge da o che incide su una generica superficie  $S$ .

Si definisce **irradianza** la densità di potenza radiante per unità di superficie che emerge o incide da o su un generico punto della superficie stessa:

$$E = \frac{\partial \phi}{\partial s_0}.$$

Eq. 1.5-1

L'irradianza è detta diffusa se nella sua definizione non si tiene conto della provenienza o della direzione verso cui fluisce la potenza radiante, il flusso totale è integrato cioè su un emisfero (Figura 1.5-1); l'irradianza è invece detta direzionale se si fa riferimento ad una precisa distribuzione direzionale del flusso radiante (ad esempio nel caso in cui si consideri un flusso di potenza radiante collimato proveniente da una fissata direzione).

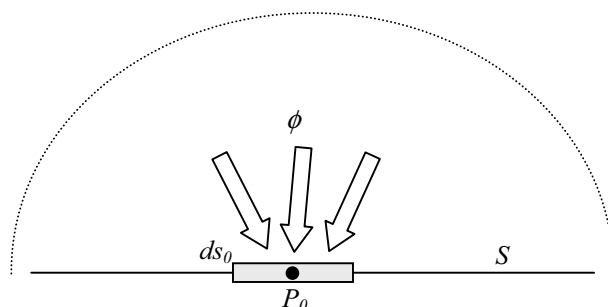


Figura 1.5-1: flusso di potenza radiante incidente sulla superficie S nel punto  $P_0$ .

Si definisce **radianza** la densità di potenza radiante per unità di superficie proiettata,  $ds_0 \cos(\theta)$ , e per unità di angolo solido  $\Omega$  che si propaga in direzione  $\theta$  da o verso un determinato punto della superficie considerata (Figura 1.5-2):

$$L = \frac{\partial^2 \phi}{\partial \Omega \cdot \partial s_0 \cdot \cos(\theta)}.$$

Eq. 1.5-2

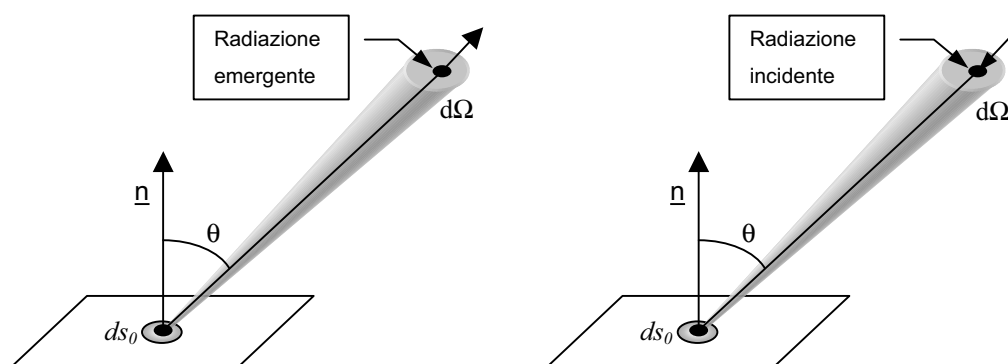


Figura 1.5-2: configurazione geometrica considerata nella definizione di radianza incidente ed emergente sulla o dalla superficie S.

Tale grandezza è fondamentale nel telerilevamento dato che i sensori durante la scansione della superficie osservano di volta in volta una porzione della superficie stessa lungo una fissata direzione all'interno di un determinato angolo solido. Ai fini del telerilevamento del colore marino la superficie marina può essere considerata con buona approssimazione una superficie lambertiana, cioè una superficie la cui radianza risulta costante al variare della direzione di propagazione. Per tali superfici la radianza in funzione dell'irradianza è pari a  $L=E/\pi$ .

Il flusso di potenza radiante, l'irradianza e la radianza per come sono state definite tengono conto dell'energia radiante integrata su tutto lo spettro del visibile. Per conoscere la frazione di energia radiante nell'intorno di una determinata lunghezza d'onda  $\lambda$  è necessario definire le grandezze spettrali corrispondenti:

$$\phi_\lambda = \frac{\partial \phi}{\partial \lambda}, \quad L_\lambda = \frac{\partial L}{\partial \lambda}, \quad E_\lambda = \frac{\partial E}{\partial \lambda}. \quad \text{Eq. 1.5-3}$$

cioè la densità di flusso, la radianza e l'irradianza per unità di lunghezza d'onda. Le grandezze spettrali saranno quelle usate in seguito per definire il modello di colore marino ed il modello atmosferico.

In Tabella 1.5-1 sono riassunte le definizioni delle grandezze radiometriche precedentemente introdotte e riportate le corrispondenti unità di misura.

Tabella 1.5-1

Grandezza	Definizione	Simbolo	Unità di misura (SI)
Energia radiante		$Q$	J (joule)
Flusso di potenza radiante	$dQ/dt$	$\phi$	W (watt)
Irradianza	$d\phi/ds_0$	$E$	W/m <sup>2</sup>
Radianza	$L = \frac{\partial^2 \phi}{\partial \Omega \cdot \partial s_0 \cdot \cos(\theta)}$	$L$	W/(sr·m <sup>2</sup> )
Flusso spettrale di potenza radiante	$\phi_\lambda = \partial \phi / \partial \lambda$	$\phi_\lambda$	W/nm
Irradianza spettrale	$E_\lambda = \partial E / \partial \lambda$	$E_\lambda$	W/(m <sup>2</sup> nm)
Radianza spettrale	$L_\lambda = \partial L / \partial \lambda$	$L_\lambda$	W/(m <sup>2</sup> sr nm)

Nel seguito, alla radianza e all'irradianza saranno associati gli aggettivi di incidente o emergente a seconda che esse descrivano un flusso di potenza radiante che incide o emerge su o da una determinata superficie.

### 1.5.2 La riflettanza sub-superficiale.

Il modello di colore marino come già anticipato sarà descritto in termini di riflettanza spettrale sub-superficiale. Si definisce riflettanza spettrale alla profondità  $z$  il rapporto tra l'irradianza spettrale diffusa *upwelling* che incide dal basso su un piano di riferimento posto a profondità  $z$  rispetto ad un determinato sistema di assi cartesiani, e l'irradianza spettrale diffusa *downwelling* che incide dall'alto sul piano stesso (Figura 1.5-3):

$$R(z, \lambda) = \frac{E_u(z, \lambda)}{E_d(z, \lambda)}. \quad \text{Eq. 1.5-4}$$

Negli studi di colore marino la riflettanza sub-superficiale è calcolata immediatamente al di sotto della superficie di interfaccia aria-acqua. Se l'origine dell'asse  $z$  di riferimento, orientato come in Figura 1.5-3, è posta in corrispondenza di tale superficie di separazione, la riflettanza sub-superficiale sarà indicata con:

$$R(0^-, \lambda) = \frac{E_u(0^-, \lambda)}{E_d(0^-, \lambda)}.$$

Eq. 1.5-5

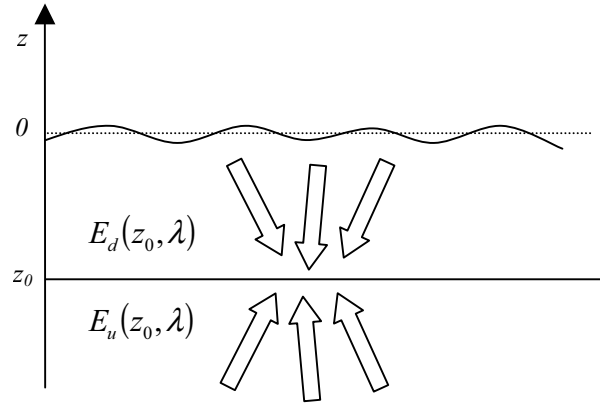


Figura 1.5-3: Irradianza down-welling ed up-welling usate nella definizione di riflettanza.

### 1.5.3 Assorbimento e scattering nella materia.

Si consideri uno strato di spessore infinitesimo a partire dalla coordinata  $z$  di un generico mezzo attraverso il quale viene fatto propagare un fascio di radiazione collimato di irradianza  $E(z, \lambda)$  (Figura 1.5-4). La radiazione subisce nell'attraversare il mezzo un'attenuazione dovuta ai fenomeni di assorbimento e *scattering*. L'assorbimento si ha quando il mezzo accresce la sua energia sottraendola al fascio di radiazione, mentre lo *scattering* o diffusione avviene quando i fotoni che compongono il fascio vengono deviati in tutte le direzioni dalle particelle di cui è costituito il mezzo. Si ha *backscattering* quando la deflessione del fotone è maggiore di  $90^\circ$  rispetto alla direzione di propagazione originaria del fascio mentre in caso contrario si parla di *forwardscattering*.

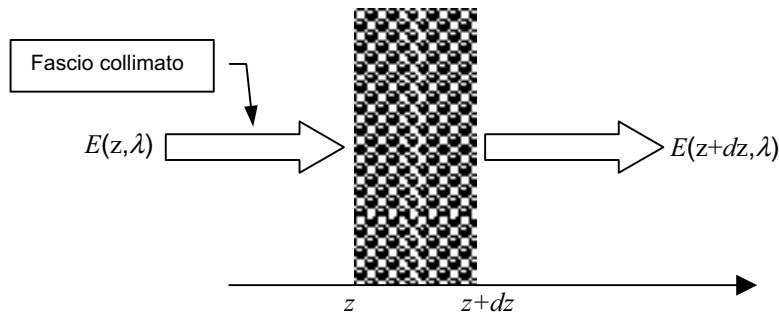


Figura 1.5-4: propagazione di un fascio collimato di irradianza  $E(z, \lambda)$  attraverso uno spessore infinitesimo  $dz$  di un generico mezzo.

Si definisce coefficiente di attenuazione alla lunghezza d'onda  $\lambda$  la variazione dell'irradianza nell'attraversare lo spessore  $dz$  relativa al valore dell'irradianza in  $z$ :

$$c_\lambda = -\frac{1}{E(z, \lambda)} \cdot \frac{\partial E(z, \lambda)}{\partial z}.$$

Eq. 1.5-6

Il coefficiente di attenuazione può essere scomposto in due contributi additivi che tengono conto dell'assorbimento e dello *scattering* che la radiazione subisce nell'attraversare il mezzo:

$$c_\lambda = a_\lambda + b_\lambda, \quad \text{Eq. 1.5-7}$$

in cui  $a_\lambda$  indica il coefficiente di assorbimento e  $b_\lambda$  il coefficiente di *scattering* entrambi alla lunghezza d'onda  $\lambda$ . Si possono definire quindi un coefficiente di assorbimento ed un coefficiente di *scattering* in maniera analoga a quanto fatto per il coefficiente di attenuazione supponendo presente alternativamente l'uno o l'altro dei due fenomeni:

$$a_\lambda = -\frac{1}{E(z, \lambda)} \cdot \frac{\partial E(z, \lambda)}{\partial z} \Big|_{\text{assorbimento}}, \quad \text{Eq. 1.5-8}$$

$$b_\lambda = -\frac{1}{E(z, \lambda)} \cdot \frac{\partial E(z, \lambda)}{\partial z} \Big|_{\text{scattering}}. \quad \text{Eq. 1.5-9}$$

A sua volta il coefficiente di *scattering* può essere scomposto nei contributi di *backscattering* e *forwardscattering*:

$$b_\lambda = b_{b\lambda} + b_{f\lambda}. \quad \text{Eq. 1.5-10}$$

I coefficienti di attenuazione così definiti costituiscono le proprietà intrinseche del mezzo dato che si fa riferimento ad una precisa distribuzione dell'irradianza incidente in  $z$ . Tali coefficienti serviranno a caratterizzare le proprietà intrinseche dei singoli componenti otticamente attivi dell'acqua marina e permetteranno di esprimere la riflettanza sub-superficiale in funzione delle loro concentrazioni.

## 1.6 Caratteristiche dei sensori per lo studio del colore marino

In questo paragrafo saranno esaminate brevemente le caratteristiche dei sensori multispettrali per il telerilevamento passivo dedicati agli studi di colore marino. In Figura 1.6-1 è schematizzata la geometria di ripresa di un sensore satellitare di questo tipo.

Il sensore osserva, da un'altezza  $h$  dal suolo, una porzione di superficie in direzione  $\theta$  con un determinato angolo di vista pari a IFOV (acronimo di *Instantaneous Field Of View*). La superficie viene scansionata lungo una linea perpendicolare alla direzione della traccia a terra del satellite. L'ampiezza del campo di vista del sensore è definita dall'angolo FOV (*Field Of View*) che insieme all'altezza fissa le dimensioni trasversali della linea di scansione ovvero della cosiddetta *swath* a terra. Il sensore acquisisce i dati lungo la direzione della traccia a terra e su un certo numero di linee di scansione adiacenti. Si è in grado in tal modo di comporre, per ogni canale spettrale, una mappa bidimensionale in cui ogni elemento (*pixel*) rappresenta una misura del flusso di potenza radiante medio sulla banda del generico canale che dalle singole superfici elementari si propaga in direzione del sensore. La dimensione a terra  $\Delta x$  dell'area elementare corrisponde alla risoluzione spaziale con cui il sensore osserva la superficie. Essa non è costante lungo la linea di scansione, dipendendo oltre che dall'altezza  $h$  e dall'angolo IFOV, anche dalla direzione di vista definita dall'angolo  $\theta$ , detto angolo zenitale. All'aumentare dell'angolo zenitale del sensore, aumenta la dimensione dell'area osservata con conseguente peggioramento della risoluzione. La risoluzione spaziale nominale lungo la linea di scansione che viene fornita nelle caratteristiche dei sensori

è riferita alla direzione di *nadir* ( $\theta=0$ ) e di solito risulta uguale alla risoluzione lungo la traccia a terra ( $\Delta x_{\theta=0}=\Delta y$ ). A seconda che la scansione della *swath* venga effettuata con dispositivi meccanici od elettronici è possibile distinguere i sensori in *whiskbroom* e *pushbroom*. I primi focalizzano il fascio di radiazione in ingresso su un unico sensore tramite un sistema di specchi che viene orientato opportunamente per scandire l'intera *swath*. I sensori del secondo tipo invece sono equipaggiati con dispositivi costituiti da *array* di sensori elementari (di tipo CCD Charge Coupled Device) su cui viene focalizzata la radiazione proveniente dalle diverse celle di risoluzione che costituiscono una linea di scansione. I dati corrispondenti ad ogni cella della *swath* vengono in questo modo acquisiti contemporaneamente.

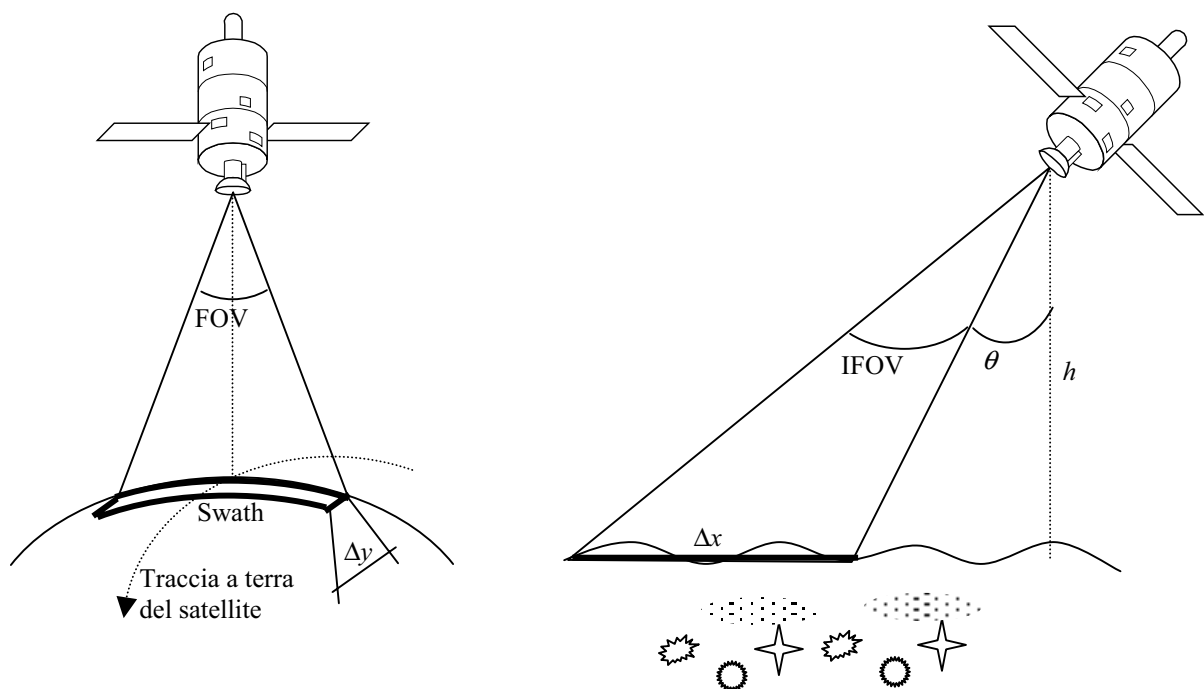


Figura 1.6-1: geometria di ripresa per un sensore passivo.

Le prestazioni di un sensore multispettrale sono definite facendo riferimento ad un certo numero di caratteristiche che possono essere raggruppate nelle seguenti categorie:

- **Caratteristiche orbitali** della piattaforma satellitare: inclinazione dell'orbita rispetto al piano equatoriale terrestre, altezza nominale, istante di attraversamento dell'equatore nella fase ascendente o discendente dell'orbita.
- **Risoluzione temporale** ovvero frequenza con cui uno stesso punto della superficie terrestre viene osservato.
- **Caratteristiche radiometriche** che comprendono gli standard di calibrazione del sensore, le prestazioni in termini di rumore, il *range* di dinamica e la quantizzazione del segnale in uscita.
- **Caratteristiche spettrali** che includono il numero di canali, la lunghezza d'onda a centro banda e la larghezza di banda.



- **Caratteristiche geometriche** in termini risoluzione spaziale, larghezza della *swath* a terra, possibilità di variare il piano di scansione (*tilting*).

Su tali caratteristiche si fonda la scelta del sensore più adatto all'applicazione che si vuol prendere in esame. In particolare le prestazioni richieste per l'osservazione su scala globale delle acque oceaniche, tipicamente di caso I, saranno diverse da quelle necessarie per monitorare zone costiere ricche di acque di caso II. Le specifiche richieste per il monitoraggio su scala globale ad esempio richiedono tipicamente una risoluzione spaziale di 4 od 8 Km ed una frequenza di passaggio di 3 o 5 giorni mentre, per l'osservazione su scala locale tali prestazioni dovranno essere migliori, con risoluzione sotto il Km e frequenza di passaggio giornaliera. Le caratteristiche spettrali inoltre influiscono sulla possibilità di stimare o meno le concentrazioni dei diversi costituenti otticamente attivi. Tipicamente per monitorare le acque di caso II, in cui le concentrazioni dei costituenti otticamente attivi variano indipendentemente le une dalle altre, è richiesto un maggior numero di canali ed una più alta risoluzione spettrale rispetto al numero di canali ed alla risoluzione spettrale necessari per il telerilevamento delle acque di caso I.

*Tabella 1.6-1: caratteristiche di alcuni sensori attualmente operativi o che hanno cessato di esserlo.*

Sensor	CZCS	OCTS	POLDER	MOS	SeaWiFS
Platform	Nimbus-7	ADEOS-1	ADEOS-1	IRS-P3	OrbView-2
Agency	NASA	NASDA	CNES	DLR	OSC/NASA
Country	USA	Japan	France	Germany/India	USA
Operation Start	Oct.1978	Aug.1996	Aug.1996	Mar.1996	Sep.1997
Operation End	Jun.1986	Jun.1997	Jun.1997	Mar.2001	Sep.2002
Orbital Inclination	99.3	98.6	98.6	98.7	98.2
Equatorial Crossing Time(h)	12:00	10:41	10:41	10:30	12:00
Altitude (km)	955	804.6	804.6	817	705
Resolution at Nadir (km)	0.825	0.7	6x7	0.5	1.1
Swath (km)	1566	1400	2400	200	2800
Tilt (degrees)	±20	±20	Variable	No	±20
Direct Link	No	UHF/X-band	X-band	S-band	L-band
Recorded	Yes	X-band	X-band	None	S-band
Solar Calibration	No	Yes	No	Yes	Yes
Lunar Calibration	No	No	No	No	Yes
Lamp Calibration	Yes	Yes	No	Yes	No

*Tabella 1.6-2: caratteristiche di alcuni sensori di nuova generazione.*

Sensor	MERIS	OSMI	MODIS-AM	MISR	OCI
Platform	Envisat	KOMPSAT	EOS-AM1	EOS-AM1	ROCSAT
Agency	ESA	KARI	NASA	NASA	Taiwan
Country	Europe	Korea	USA	USA	Taiwan
Operation Start	Mar.2000	Jul.1999	1999	1999	Feb.1999
Operation End	Feb.2004	Jul.2002	Jun.2003	Jun.2003	Feb.2003
Orbital Inclination	98.5	98.13	98.2	98.2	35
Eq. Crossing Time (h)	10:00	10:50	10:30	10:30	09:00/15:00
Altitude (km)	800	685	705	705	600
Resolution at Nadir (km)	1.2/0.3	0.85	1	0.25	0.8
Swath (km)	1150	800	2330	360	704
Tilt (degrees)	No	No	No	Variable	No
Direct Link	X-band	X-Band	X-band	No	S-band
Recorded	X-band	Yes	X-band	X-band	None
Solar Calibration	Yes	Yes	Yes	No	—
Lunar Calibration	No	No	Yes	Yes	—
Lamp Calibration	No	No	Yes	No	—

*Tabella 1.6-3: caratteristiche spettrali di alcuni sensori per il telerilevamento del colore marino.*

Center	FWHM	NE/L	Center	FWHM	NE/L	Center	FWHM	NE/L	Center	FWHM	NE/L	Center	FWHM	NE/L
443	20	0.078	412	20	0.186	443	20	0.138	438	10	0.132	412	20	0.092
620	20	0.173	443	20	0.133	460	20	0.147	443	10	0.124	443	20	0.077
650	20	0.166	460	20	0.088	525	20	0.088	460	10	0.148	460	20	0.096
670	20	0.094	520	20	0.121	670	20	0.063	520	10	0.147	510	20	0.049
730	10	0.045	565	20	0.091	763	10	0.090	570	10	0.150	558	20	0.043
			670	20	0.037	755	40	0.086	615	10	0.108	670	20	0.031
			755	40	0.037	865	40	0.034	685	10	0.141	765	40	0.016
			865	40	0.011	910	20	0.045	790	10	0.102	865	40	0.018
									815	10	0.034			
									870	10	0.033			

Center	FWHM	NE/L	Center	FWHM	NE/L	Center	FWHM	NE/L	Center	FWHM	NE/L	Center	FWHM	NE/L
443	20	0.093	412	20	0.048	412.5	10	0.026	443	20	—	443	20	—
657	20	0.081	460	10	0.032	462.5	10	0.025	460	20	—	460	20	—
672	10	0.068	488	10	0.025	490	10	0.022	510	20	—	510	20	—
687	10	0.031	521	10	0.018	510	10	0.016	565	20	—	565	20	—
			525	10	0.018	560	10	0.016	670	20	—	670	20	—
			667	10	0.008	625	10	0.014	865	40	—	865	40	—
			678	10	0.007	685	10	0.013						
			748	10	0.009	681	7.5	0.014						
			870	10	0.006	709	8	0.011						
						775	14	0.008						
						870	20	0.007						
						890	10	0.011						
						900	10	0.010						

In Tabella 1.6-1 sono riportate alcune caratteristiche di sensori multispettrali attualmente operativi o che hanno cessato di esserlo mentre in

Tabella 1.6-2 sono riportate le caratteristiche dei sensori di nuova generazione non ancora operativi. La Tabella 1.6-3 contiene invece le caratteristiche spettrali dei sensori presi in considerazione.

Il sensore CZCS è stato il primo sensore multispettrale dedicato all'osservazione del colore marino. Per gli anni a venire sono previste numerose missioni che prevedono la messa in orbita di sensori per l'osservazione del colore marino a dimostrazione dell'importanza che tale oggetto di studio ha assunto nel corso del tempo.

Nel presente lavoro di tesi saranno presi in considerazione i sensori MOS e SeaWiFS tra quelli attualmente operativi e il sensore MERIS tra quelli di prossima messa in orbita. In Figura 1.6-5 sono riportate le posizioni e la larghezza di banda dei canali dei tre sensori.



*Figura 1.6-2: particolare del sensore SeaWiFS.*

Il Sea Wide Field-of-view Sensor (SeaWiFS) è un sensore multispettrale di tipo *whiskbroom* con una risoluzione spaziale al nadir di circa 1 km ed una larghezza della *swath* di 2800 km all'equatore. Il sensore è equipaggiato con sei canali nel visibile (VIS), con una banda di 20 nm, specificatamente progettati per la stima delle concentrazioni dei parametri otticamente attivi, e due canali nell'infrarosso vicino (NIR), con una larghezza di banda di 40 nm che permettono la stima del contributo atmosferico al segnale acquisito. La risoluzione radiometrica è di 10 bit e la massima radianza misurabile è pari a  $136.3 \text{ Wm}^{-2}\text{sr}^{-1}\mu\text{m}^{-1}$  nel canale 1 e  $21.3 \text{ Wm}^{-2}\text{sr}^{-1}\mu\text{m}^{-1}$  nel canale 8. Le caratteristiche geometriche e la possibilità di avere una copertura globale in 2 giorni, fanno tale sensore particolarmente adatto al monitoraggio di bacini a media e larga scala con alta risoluzione temporale.



Figura 1.6-3: particolare del sensore MOS.

Il Modular Optoelectronic Scanner (MOS) è uno spettrometro di tipo *pushbroom* costituito da due distinti dispositivi: il MOS-A con quattro canali di 1.34 nm di larghezza di banda posizionati tra 755 e 768 nm disegnati per il telerilevamento dell'atmosfera, ed il MOS-B con tredici canali con banda di 10 nm e posizionato tra 408 e 1010 nm le cui caratteristiche spettrali sono state scelte in accordo alle caratteristiche spettrali delle acque oceaniche e costiere. Il sensore MOS-B è equipaggiato con otto canali nel VIS e tre canali nel NIR tra 700 e 900 nm. La risoluzione spaziale del sensore al nadir è di 520 m mentre la larghezza della *swath* è di circa 200 km. La risoluzione radiometrica è pari a 16 bit. La risoluzione spaziale e spettrale ed il numero di canali nel VIS del MOS-B rendono tale sensore particolarmente adatto al monitoraggio delle acque costiere tipicamente di caso II.



Figura 1.6-4: particolare del sensore MERIS

Il Medium Resolution Imaging Spectrometer (MERIS) di prossima messa in orbita da parte dell'ESA sulla piattaforma satellitare ENVISAT è un sensore di tipo *pushbroom* adatto per l'osservazione sia su scala globale che su scala regionale. La risoluzione spaziale del sensore in modalità *Full Resolution* è pari a 300 m al *nadir* requisito adatto al monitoraggio delle zone costiere. In modalità *Reduced Resolution*, usata per l'osservazione su larga scala, la risoluzione del sensore scende a 1200 m al nadir. Il campo di vista (FOV) di 68.5° combinato con le caratteristiche orbitali del satellite ENVISAT producono una larghezza di *swath* pari a 1150 Km permettendo la copertura globale del pianeta in tre giorni. Il MERIS ha quindici canali programmabili all'interno dello spettro del visibile e del vicino infrarosso tra 390 e 1040 nm. La larghezza

di banda dei canali va da 1.25 a 30 nm a seconda della posizione assunta nella regione di spettro coperta. I canali utilizzati per la stima dei costituenti otticamente attivi dell'acqua di mare, all'interno della regione del visibile tra 412.5 e 665 nm, hanno una larghezza di banda di 10 nm. Alla lunghezza d'onda di 681.5 nm è posizionato inoltre un canale per la detezione del segnale di fluorescenza della clorofilla con una larghezza di banda pari a 7.5 nm. La risoluzione radiometrica del sensore è pari a 16 bit.

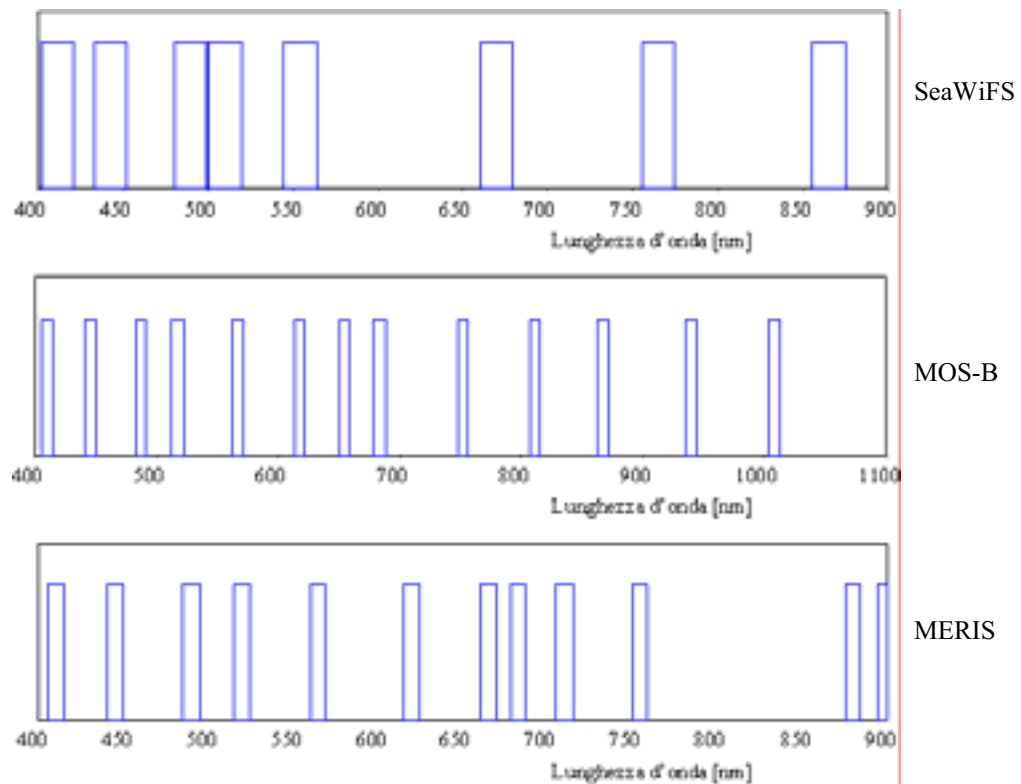


Figura 1.6-5: posizione dei canali spettrali dei sensori SeaWiFS, MOS-B e MERIS.

---

## 2 Capitolo 2

### *La stima dei parametri otticamente attivi dell'acqua marina da dati multispettrali: il problema diretto*

Questo capitolo sarà dedicato alla definizione del modello diretto che permette di calcolare la radianza che raggiunge il sensore in ciascuno dei suoi canali, una volta fissate le concentrazioni dei costituenti otticamente attivi, i parametri atmosferici e le caratteristiche spettrali del sensore stesso.

Il modello sarà costituito da due blocchi principali: il modello atmosferico ed il modello di colore marino definito attraverso la riflettanza spettrale sub-superficiale dell'acqua marina.

Del primo sarà fornita unicamente una versione semplificata, di tipo *single scattering*, in cui gli effetti atmosferici sulla radiazione proveniente dalla superficie marina sono dovuti a tre componenti atmosferici principali: molecole di gas permanenti, ozono stratosferico e particelle di aerosol. Una trattazione più accurata esula dagli scopi del presente lavoro di tesi essendo questo più orientato verso lo studio e l'applicazione di una metodologia per la definizione degli algoritmi di stima delle concentrazioni dei parametri otticamente attivi dai dati di riflettanza sub-superficiale. Il modello atmosferico semplificato consentirà comunque di definire una procedura per ridurre gli effetti atmosferici sul segnale al sensore permettendo così di stimare la riflettanza sub-superficiale e quindi applicare le tecniche di inversione del modello di colore marino a dati reali.

Il modello di colore marino che sarà preso in considerazione è basato sull'approccio di *Sathyendranath et al.* [1] di cui è possibile trovare una versione riveduta ed aggiornata in [3]. Il modello esprime la riflettanza spettrale sub-superficiale in funzione delle proprietà ottiche intrinseche dell'acqua di mare e cioè dei coefficienti di assorbimento e *backscattering* dell'acqua di mare pura, del fitoplancton, del particolato solido in sospensione e della sostanza gialla. Tali coefficienti sono parametrizzati in funzione delle concentrazioni dei tre costituenti otticamente attivi. Il modello che scaturisce da tale approccio è di tipo semi-analitico essendo esso un misto di relazioni teoriche, soluzione delle complicate equazioni del trasferimento radiativo che regolano la propagazione della radiazione attraverso la materia, e di relazioni empiriche frutto di intense campagne di raccolta di dati *in-situ* condotte su diverse aree geografiche e per diversi tipi di acqua.

Il capitolo è organizzato in un primo paragrafo in cui sono descritti i diversi contributi alla radiazione totale che giunge al sensore il quale osserva una porzione di superficie marina di dimensioni pari ad una cella di risoluzione a terra. Seguirà un secondo paragrafo in cui è definito il modello atmosferico semplificato. Il paragrafo conclusivo è dedicato alla definizione del modello diretto di colore marino.

## 2.1 La radianza al sensore

Saranno analizzati di seguito i vari contributi della radiazione totale che giunge al sensore. In Figura 2.1-1 è visualizzato il cammino che la radiazione solare compie prima di giungere sulla superficie del sensore. Il sensore osserva una porzione di superficie marina di dimensioni pari ad una cella di risoluzione a terra. La radiazione solare che si propaga in atmosfera, viene assorbita e diffusa in tutte le direzioni dalle particelle dei costituenti atmosferici (in massima parte gas permanenti, aerosol ed ozono) che agiscono sulla radiazione stessa da vero e proprio filtro alterandone la distribuzione spettrale (Figura 2.1-1-a). Una parte di questa, opportunamente attenuata, raggiunge la superficie marina per via diretta o per via diffusa ed una parte viene direttamente diffusa verso il sensore. Il fascio di radiazione che raggiunge la superficie marina (Figura 2.1-1-b) in parte si propaga verso gli strati sottostanti ed in parte viene diffuso nuovamente verso l'atmosfera dando luogo alla cosiddetta componente di *glint*. La radiazione che si propaga verso gli strati sottostanti d'acqua interagisce con i componenti otticamente attivi dell'acqua di mare (Figura 2.1-1-c). Una frazione di questa verrà diffusa verso l'interfaccia aria-acqua e si propagherà in atmosfera in direzione del sensore costituendo la componente che contiene le informazioni utili alla stima delle concentrazioni dei parametri otticamente attivi. Al sensore giungono quindi tre componenti principali (Figura 2.1-1-d): la componente di diffusione atmosferica costituita dalla radiazione solare che non ha raggiunto la superficie marina, il *glint* e la radiazione che emerge dalla superficie marina e che si propaga in atmosfera in direzione del sensore.

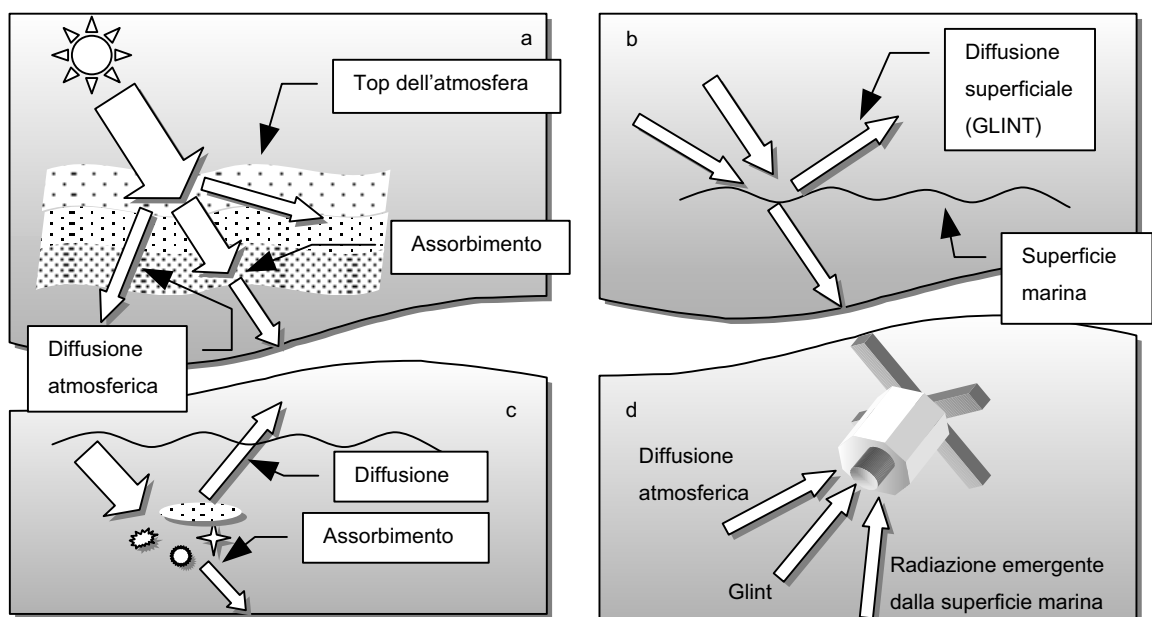


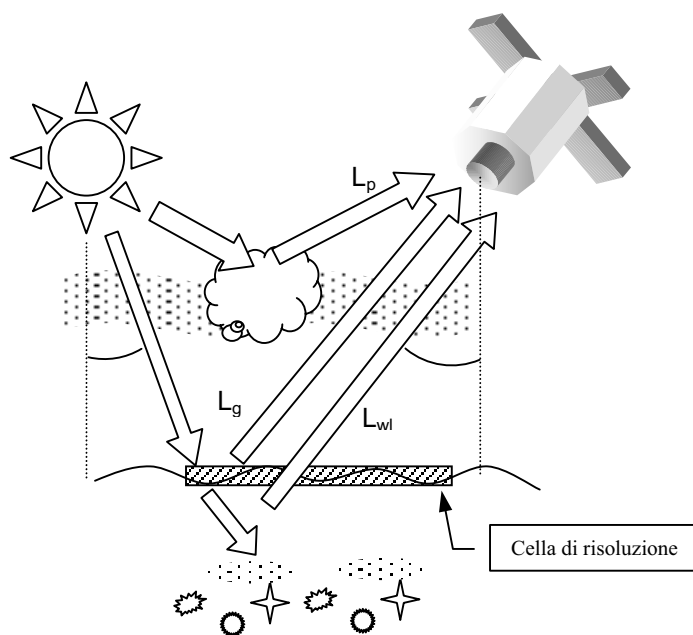
Figura 2.1-1: contributi alla radiazione incidente sul sensore e proveniente da una porzione di superficie marina di dimensioni pari ad una cella di risoluzione a terra.

La radianza spettrale totale  $L_t(\lambda)$  alla lunghezza d'onda  $\lambda$ , che incide sulla superficie del sensore, può quindi essere scomposta nella somma di tre contributi (Figura 2.1-2):

$$L_t(\lambda) = L_p(\lambda) + L_g(\lambda) \cdot T(\lambda) + L_{wl}(\lambda) \cdot T(\lambda). \quad \text{Eq. 2.1-1}$$

Il primo contributo  $L_p(\lambda)$ , detto di *path radiance*, tiene conto della frazione di radiazione diffusa dagli strati dell'atmosfera che giunge direttamente al sensore senza interagire con la superficie marina e con gli strati di acqua sottostanti. Il secondo contributo  $L_g(\lambda)$  rappresenta la radianza di *glint* moltiplicata per la

trasmissione atmosferica in salita  $T(\lambda)$ , coefficiente questo che tiene conto dell'attenuazione che la radiazione subisce lungo il cammino ottico superficie marina-sensore. Il terzo ed ultimo contributo rappresenta la radiazione che dagli strati sottostanti della superficie marina si propaga nuovamente in atmosfera in direzione del sensore. Tale termine è costituito dalla radianza emergente dalla superficie marina,  $L_{wl}(\lambda)$  (*Water Leaving Radiance*), moltiplicata anch'essa per la trasmissione atmosferica in salita. Il terzo contributo contiene informazioni sulle proprietà di assorbimento e diffusione dei componenti otticamente attivi dell'acqua di mare; è questo, quindi, il contributo utile ai fini della stima delle concentrazioni dei costituenti stessi. La catena di *processing* dei dati telerilevati, utilizzata per la stima delle concentrazioni dei costituenti otticamente attivi, dovrà prevedere una fase di elaborazione che servirà



ad eliminare dai dati stessi quei contributi che non portano informazioni utili. Il contributo di *glint* viene minimizzato predisponendo per il sensore un'opportuna geometria di ripresa che evita la riflessione diretta dalla superficie marina del fascio di radiazione solare (*tilt* del piano di scansione). La *path radiance* può essere ridotta utilizzando degli opportuni modelli atmosferici che permettono di stimarne il valore che può essere così sottratto dalla radianza totale misurata dal sensore.

Figura 2.1-2: i contributi della radiazione che giunge al sensore in termini di radianza.

Il segnale elettrico in uscita al sensore (che può essere indifferentemente una tensione o una corrente) viene amplificato, campionato e quantizzato su un certo numero di livelli. L'uscita del sensore, per il generico canale spettrale centrato sulla lunghezza d'onda  $\lambda$ , consiste in un dato digitale (DN: *Digital Number*) che rappresenta la radianza totale alla superficie del sensore mediata sulla banda del canale stesso. Il dato digitale in uscita al generico canale è legato alla radianza totale da una funzione di guadagno. Di tale guadagno si dovrà tenere conto nella catena di *processing* quando si andrà ad operare la calibrazione dei dati digitali per associarli a valori reali di radianza.

## 2.2 Modello atmosferico semplificato

In questo paragrafo sarà descritto il modello atmosferico a cui si farà riferimento per definire in seguito una procedura di correzione atmosferica per dati telerilevati relativi ad acque di caso I. Il modello dovrà



permettere di calcolare la *path radiance* in corrispondenza di una determinata lunghezza d'onda e per una fissata geometria di ripresa (Figura 2.2-1), definita dagli angoli zenitali  $\theta$  e  $\theta_0$  e dagli angoli azimutali  $\phi$  e  $\phi_0$  rispettivamente del sensore e del sole. Dovrà inoltre, nelle stesse condizioni, fornire il valore della trasmittanza in salita  $t(\lambda)$  che sarà rimossa dal contributo utile alla radianza totale per stimare la radianza *water leaving* a cui è legata la riflettanza spettrale sub-superficiale. Il modello che sarà considerato prevede l'atmosfera costituita da tre componenti principali:

- molecole di gas permanenti,
- particelle di aerosol atmosferico,
- ozono stratosferico.

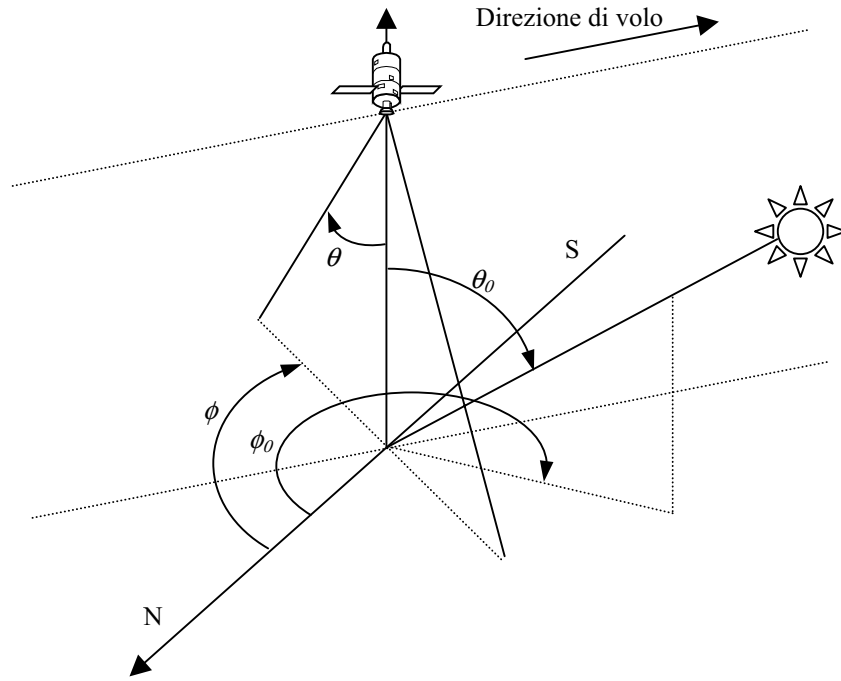


Figura 2.2-1: geometria di ripresa del sensore e definizione degli angoli zenitali ed azimutali del sole e del sensore.

L'ozono stratosferico è considerato un componente puramente assorbente. La *path radiance*, dovuta alla diffusione atmosferica della radiazione in direzione del sensore, è costituita dai seguenti contributi additivi [4]:

$$L_p(\lambda) = L_r(\lambda) + L_{ra}(\lambda) + L_a(\lambda), \quad \text{Eq. 2.2-1}$$

in cui si è indicato con  $L_r(\lambda)$  il contributo allo *scattering* da parte delle molecole di gas permanenti, con  $L_a(\lambda)$  quello dovuto alle particelle di aerosol e con  $L_{ra}(\lambda)$  il contributo misto che rappresenta la radiazione diffusa prima dai gas permanenti e successivamente diffusa dall'aerosol o viceversa. Nella Eq. 2.2-1 è possibile trascurare tale termine nel caso in cui venga ipotizzata l'assenza di *scattering* multiplo. Inoltre se la geometria di ripresa del sensore è tale da minimizzare il *glint* l'Eq. 2.1-1 diventa:

$$L_t(\lambda) = L_r(\lambda) + L_a(\lambda) + t(\lambda) \cdot L_{wl}(\lambda). \quad \text{Eq. 2.2-2}$$

Sotto l'ipotesi di *scattering* singolo le radianze  $L_r(\lambda)$  e  $L_a(\lambda)$  possono essere espresse come segue [5]:

$$L_x(\lambda) = \frac{\omega_x(\lambda) \cdot \tau_x(\lambda) \cdot F_0(\lambda) \cdot p_x(\theta, \theta_0, \lambda)}{4 \cdot \pi \cdot \cos(\theta)} \cdot T_{oz}(\theta, \theta_0, \lambda) \quad x = a, r, \quad \text{Eq. 2.2-3}$$

in cui:

a)  $\omega_x(\lambda)$  è l'albedo single *scattering* per il costituente  $x$  pari al rapporto tra energia diffusa ed energia complessivamente attenuata dal costituente, b)  $\tau_x(\lambda)$  lo spessore ottico per il costituente  $x$ , c)  $F_0(\lambda)$  l'irradianza solare al *top* dell'atmosfera, d)  $T_{oz}(\theta, \theta_0, \lambda)$  la trasmittanza dovuta all'attenuazione dello strato di ozono che la radiazione solare subisce nei due cammini discendenti ed ascendenti per giungere al sensore e pari a:

$$T_{oz}(\theta, \theta_0, \lambda) = \exp \left[ -\tau_{oz}(\lambda) \cdot \left( \frac{1}{\cos(\theta)} + \frac{1}{\cos(\theta_0)} \right) \right], \quad \text{Eq. 2.2-4}$$

in cui  $\tau_{oz}(\lambda)$  indica lo spessore ottico dell'ozono ricavabile, al variare della lunghezza d'onda, da apposite tabelle [6], e)  $p_x(\theta, \theta_0, \lambda)$  pari a:

$$p_x(\theta, \theta_0, \lambda) = \{P_x(\theta_+, \lambda) + [r(\theta) + r(\theta_0)] \cdot P_x(\theta_-, \lambda)\}, \quad \text{Eq. 2.2-5}$$

in cui si è indicato con  $P_x(\theta_{\pm}, \lambda)$  la funzione di fase del costituente  $x$  che rappresenta la probabilità che la radiazione sia diffusa in direzione  $\theta_{\pm}$  e con  $r(\theta)$  il coefficiente di riflessione di Fresnel all'interfaccia aria-acqua pari a:

$$r(\theta) = \frac{1}{2} \cdot \left[ \frac{\sin^2(\theta - \theta')}{\sin^2(\theta + \theta')} + \frac{\tan^2(\theta - \theta')}{\tan^2(\theta + \theta')} \right], \quad \text{Eq. 2.2-6}$$

$$\theta' = \arcsin \left[ \frac{\sin(\theta)}{n} \right],$$

dove con  $n$  si è indicato il coefficiente di rifrazione dell'acqua di mare. Gli angoli  $\theta_{\pm}$  sono detti di diffusione (o *scattering*) e sono pari a:

$$\cos(\theta_{\pm}) = \pm \cos(\theta_0) \cdot \cos(\theta) - \sin(\theta_0) \cdot \sin(\theta) \cdot \cos(\phi - \phi_0). \quad \text{Eq. 2.2-7}$$

La diffusione da particelle di gas permanenti può essere modellata tramite la teoria dello *scattering* di Rayleigh in cui la dimensione delle particelle del mezzo attraverso cui si propaga la radiazione, sono molto minori della lunghezza d'onda della radiazione stessa. Per tali mezzi la funzione di fase e lo spessore ottico hanno la seguente espressione generale:

$$P_r(\theta_{\pm}) = \frac{3}{4} \cdot [1 + \cos^2(\theta_{\pm})], \quad \text{Eq. 2.2-8}$$

$$\tau_r(\lambda) = 0.008735 \cdot \left[ \frac{P_{atm}}{1013.5} \right] \cdot \lambda^{-4.08}, \quad \lambda [\mu\text{m}], \quad \text{Eq. 2.2-9}$$

in cui  $P_{atm}$  è la pressione atmosferica a livello del mare espressa in Pascal. Per quanto riguarda il valore dell'albedo spettrale  $\omega_r(\lambda)$  questo può essere posto uguale ad 1 [7], ipotizzando in tal modo il comportamento delle molecole di gas permanenti puramente diffusivo.

Lo *scattering* da particelle di aerosol atmosferico dipende fortemente dalla distribuzione dimensionale delle particelle stesse e risponde alle leggi della diffusione di Mie, in cui le particelle del mezzo hanno dimensioni paragonabili (o maggiori) alla lunghezza d'onda della radiazione che lo attraversa. L'aerosol di

tipo marittimo nel VIS e nel NIR può essere considerato con buona approssimazione puramente diffusivo [7], il che comporta considerare  $\omega_a(\lambda)$  circa pari ad 1. Per quanto riguarda la funzione di fase e lo spessore ottico questi non hanno una forma generale al pari delle particelle di gas permanenti. Esistono diversi modelli parametrici di funzione di fase che dipendono dalla distribuzione dimensionale delle particelle. Un modello comunemente usato è la funzione di fase Two Terms Henyey-Greenstein (TTHG) [7]:

$$P_a(\theta_{\pm}, \lambda) = \frac{(1 - g_1^2) \cdot a}{[1 + g_1^2 - 2 \cdot g_1 \cdot \cos(\theta_{\pm})]^{1.5}} + \frac{(1 - g_2^2) \cdot (1 - a)}{[1 + g_2^2 + 2 \cdot g_2 \cdot \cos(\theta_{\pm})]^{1.5}}, \quad Eq. 2.2-10$$

in cui i parametri  $a$ ,  $g_1$  e  $g_2$  sono detti fattori di asimmetria e dipendono dalla lunghezza d'onda e dalle dimensioni delle particelle di aerosol. In Tabella 2.2-1 sono riportati i valori dei fattori di asimmetria per alcuni modelli di aerosol marittimo con umidità relativa (RH) fissata alla lunghezza d'onda di 550 nm e 670 nm [7].

*Tabella 2.2-1: valori dei fattori di asimmetria della funzione TTHG per alcuni modelli di aerosol marittimo*

Tipo di aerosol	$\lambda$ (nm)	$a$	$g_1$	$g_2$
Marittimo 70% RH	550	0.968	0.748	-0.682
Marittimo 90% RH	550	0.973	0.833	-0.671
Marittimo 99% RH	550	0.980	0.902	-0.717
Marittimo 80% RH	670	0.983	0.820	-0.550

La funzione TTHG è comunque considerata indipendente dalla lunghezza d'onda nel VIS e NIR. La mancanza di conoscenza della distribuzione dimensionale dell'aerosol può portare, nel processo di correzione atmosferica, ad errori maggiori rispetto a quelli introdotti trascurando la dipendenza della funzione di fase dalla lunghezza d'onda. Per quanto riguarda lo spessore ottico questo può essere modellato seguendo l'approccio di Angstrom:

$$\tau_a(\lambda) = \alpha \cdot \lambda^{-\beta}, \quad Eq. 2.2-11$$

in cui  $\alpha$  e  $\beta$  (detto esponente di Angstrom) dipendono ancora una volta dalla distribuzione dimensionale delle particelle di aerosol.

Dati gli spessori ottici dei singoli costituenti atmosferici è possibile calcolare la trasmittanza in salita che tiene conto dell'attenuazione che la radiazione emergente dalla superficie marina subisce lungo il cammino ottico tra superficie e sensore [7]:

$$t(\lambda) = \exp \left[ - \frac{\left( \frac{\tau_r}{2} + \tau_{oz} + \frac{\tau_a}{6} \right)}{\cos(\theta)} \right]. \quad Eq. 2.2-12$$

Per bassi valori dello spessore ottico dell'aerosol (per cui è valida l'ipotesi di *scattering* singolo) il termine  $\tau_a/6$  (il fattore 1/6 è un limite superiore [5]) può essere trascurato rispetto agli altri il che tornerà utile nel processo di correzione atmosferica usato nelle applicazioni del presente lavoro di tesi.

## 2.3 Il modello diretto di riflettanza sub-superficiale dell'acqua marina

Nel presente paragrafo sarà analizzato un modello di riflettanza basato sull'approccio di *Sathyendranath et al.* [1], di cui è possibile trovare una versione riveduta ed aggiornata in [3]. Numerosi studi sull'argomento hanno mostrato che è possibile esprimere la riflettanza spettrale sub-superficiale in funzione delle proprietà ottiche intrinseche dell'acqua di mare, cioè dei coefficienti di assorbimento e di *backscattering* [8]:

$$R(0^-, \lambda) = \frac{E_u(0^-, \lambda)}{E_d(0^-, \lambda)} = f(b_{b\lambda}, a_\lambda). \quad \text{Eq. 2.3-1}$$

Le relazioni proposte in tali studi possono essere ridotte alla seguente forma generale:

$$R(0^-, \lambda) = r \cdot \frac{b_{b\lambda}}{a_\lambda + b_{b\lambda}}, \quad \text{Eq. 2.3-2}$$

che per condizioni di mare aperto in cui  $b_{b\lambda} \ll a_\lambda$  diventa:

$$R(0^-, \lambda) = r \cdot \frac{b_{b\lambda}}{a_\lambda}, \quad \text{Eq. 2.3-3}$$

Il fattore  $r$  dipende dagli angoli zenitali del sole e del sensore e dalla funzione di fase delle particelle di cui è composto il corpo idrico. Nella pratica tale fattore viene ricavato per regressione risolvendo le complicate equazioni del trasferimento radiativo che regolano la propagazione del flusso radiante attraverso la materia. Le curve di riflettanza ed il rapporto  $b_{b\lambda}/a_\lambda$  sono simulati per diverse condizioni di propagazione della radiazione nel mezzo e cioè per diversi angoli di zenit del flusso incidente sulla superficie marina e diverse forme della funzione di fase delle particelle dell'acqua di mare. Si ottiene  $r$  effettuando la regressione tra i valori di  $R(0^-, \lambda)$  e  $b_{b\lambda}/a_\lambda$  in tal modo simulati. In [9] è stato ottenuto un valore per  $r$  pari a 0.3244 per incidenza normale del flusso di potenza radiante sulla superficie marina, ed un valore di 0.3687 per un flusso diffuso uniforme in tutte le direzioni. Risultati analoghi sono stati ottenuti in [10] in cui gli autori propongono la seguente relazione per la riflettanza:

$$R(0^-, \lambda) = 0.33 \cdot \frac{b_{b\lambda}}{a_\lambda} \cdot (1 + \Delta), \quad \text{Eq. 2.3-4}$$

Il termine  $\Delta$  dipende dalla distribuzione della radianza incidente e dalla funzione di fase delle particelle ed assume un valore, anche in condizioni estreme, che è sempre minore del 5%, risultando inoltre, nella regione dello spettro elettromagnetico del visibile, debolmente variabile al variare della lunghezza d'onda. Recentemente in [8] è stata proposta una relazione analitica per  $r$  in funzione delle condizioni di illuminazione e di ripresa della superficie marina. La relazione è stata ottenuta sotto l'ipotesi di *scattering* quasi-singolo ed è valida per una colonna d'acqua le cui proprietà ottiche sono omogenee lungo la verticale. I risultati conseguiti in questo lavoro, per i casi particolari di incidenza normale e di flusso radiante diffuso, sono in buon accordo con quelli citati pur riportando una variazione del termine  $\Delta$  maggiore rispetto a quella riportata in [10]. Sempre in tale lavoro si afferma l'importanza di  $r$  nel determinare la riflettanza sub-superficiale quando questa è utilizzata nelle applicazioni di telerilevamento. In particolare il valore di  $r$  influenza la precisione degli algoritmi di stima delle concentrazioni dei costituenti otticamente attivi che utilizzano il valore assoluto della riflettanza sub-superficiale. Il modello

analitico proposto introduce comunque delle complicazioni dal punto di vista pratico, rendendolo di non facile realizzazione quindi, pur riconoscendo l'importanza delle variazioni di  $r$  con le condizioni di illuminazione e di ripresa della superficie nel determinare il valore assoluto della riflettanza, nel seguito per l'implementazione del modello di colore marino, si farà riferimento alla Eq. 2.3-4 in cui si trascura il termine  $\Delta$ :

$$R(0^-, \lambda) = 0.33 \cdot \frac{b_{b\lambda}}{a_\lambda}, \quad \text{Eq. 2.3-5}$$

rimandando a lavori futuri una più accurata modellazione della dipendenza della riflettanza dai parametri geometrici.

I coefficienti di assorbimento e *backscattering* possono essere espressi come somma dei coefficienti di assorbimento e *backscattering* dei singoli componenti otticamente attivi dell'acqua di mare:

$$a_\lambda = a_{w\lambda} + a_{ph\lambda} + a_{yel\lambda} + a_{sed\lambda}, \quad \text{Eq. 2.3-6}$$

$$b_{b\lambda} = b_{bw\lambda} + b_{bph\lambda} + b_{bsed\lambda}, \quad \text{Eq. 2.3-7}$$

in cui si è indicato con:

- $a_{w\lambda}$  il coefficiente di assorbimento dell'acqua di mare pura [ $\text{m}^{-1}$ ],
- $a_{ph\lambda}$  il coefficiente di assorbimento del fitoplancton [ $\text{m}^{-1}$ ],
- $a_{yel\lambda}$  il coefficiente di assorbimento della sostanza gialla in soluzione [ $\text{m}^{-1}$ ],
- $a_{sed\lambda}$  il coefficiente di assorbimento delle particelle non clorofillacee [ $\text{m}^{-1}$ ],
- $b_{bw\lambda}$  il coefficiente di *backscattering* dell'acqua di mare pura [ $\text{m}^{-1}$ ],
- $b_{bph\lambda}$  il coefficiente di *backscattering* del fitoplancton [ $\text{m}^{-1}$ ],
- $b_{bsed\lambda}$  il coefficiente di *backscattering* delle particelle non clorofillacee [ $\text{m}^{-1}$ ];

la sostanza gialla è considerata normalmente un costituente puramente assorbente.

Il coefficiente di *backscattering* a sua volta può essere espresso in funzione del coefficiente di *scattering* totale:

$$b_{b\lambda} = b'_{bw} \cdot b_{w\lambda} + b'_{bph} \cdot b_{ph\lambda} + b'_{bsed} \cdot b_{sed\lambda}, \quad \text{Eq. 2.3-8}$$

in cui si indica con:

- $b_{w\lambda}$  il coefficiente di *scattering* totale dell'acqua di mare pura [ $\text{m}^{-1}$ ],
- $b_{ph\lambda}$  il coefficiente di *scattering* totale del fitoplancton [ $\text{m}^{-1}$ ],
- $b_{sed\lambda}$  il coefficiente di *scattering* totale delle particelle non clorofillacee [ $\text{m}^{-1}$ ],
- $b'_{bw}$  il rapporto tra il coefficiente di *backscattering* e di *scattering* dell'acqua di mare pura,
- $b'_{bph}$  il rapporto tra il coefficiente di *backscattering* e di *scattering* del fitoplancton,
- $b'_{bsed}$  il rapporto tra il coefficiente di *backscattering* e di *scattering* delle particelle non clorofillacee.

Il modello analizzato è a tre componenti cioè si ipotizza che le proprietà ottiche dell'acqua di mare siano dovute, oltre che all'acqua di mare pura, ai tre principali componenti otticamente attivi precedentemente introdotti: fitoplancton, sostanza gialla e particelle non clorofillacee. I coefficienti di assorbimento e

*backscattering* di tali costituenti, al variare della lunghezza d'onda, sono a sua volta parametrizzati in funzione delle concentrazioni massa su volume dei costituenti stessi o di alcuni parametri indice della concentrazione e che ad essa possono essere messi direttamente in relazione. In particolare i parametri a cui si fa normalmente riferimento negli studi di colore marino sono i seguenti:

- $C$ : concentrazione massa su volume di clorofilla- $a$  più feofitina- $a$  espressa in  $\text{mg/m}^3$ , indice della biomassa fitoplanctonica,
- $X$ : coefficiente di *scattering* delle particelle non clorofillacee a 550 nm di lunghezza d'onda espresso in  $\text{m}^{-1}$ , indice della concentrazione massa su volume delle particelle non clorofillacee,
- $Y$ : coefficiente di assorbimento della sostanza gialla a 440 nm di lunghezza d'onda espresso in  $\text{m}^{-1}$ , indice della concentrazione massa su volume di sostanza gialla disciolta in acqua.

Ai parametri su introdotti si dà il nome di parametri otticamente attivi dell'acqua di mare e saranno di qui in poi indicati con l'acronimo di OAPs (dall'inglese *Optically Active Parameters*). I prossimi capitoli saranno quindi dedicati al problema della stima di  $C$ ,  $X$ ,  $Y$  dai dati telerilevati di riflettanza. Nel paragrafo successivo saranno introdotte le espressioni per i coefficienti di assorbimento e di *backscattering* dei diversi costituenti in funzione della lunghezza d'onda e del valore degli OAPs.

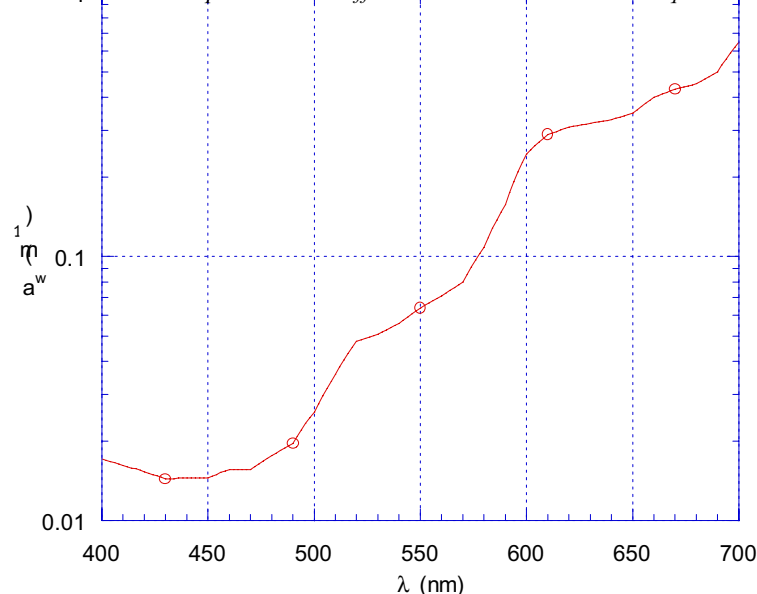
### 2.3.1 Il coefficiente di assorbimento.

Le relazioni che saranno di seguito fornite per il calcolo del coefficiente di assorbimento totale dell'acqua marina sono frutto di intense campagne di misura in mare condotte da diversi studiosi di ottica marina in varie aree geografiche.

#### 2.3.1.1 Il coefficiente di assorbimento dell'acqua di mare pura.

Il coefficiente dell'acqua di mare pura coincide praticamente con quello dell'acqua chimicamente pura dato che alle lunghezze d'onda del VIS gli ioni in soluzione contribuiscono solo allo *scattering*. I valori del coefficiente di assorbimento presi a riferimento sono quelli forniti da *Smith* e *Baker* in [11]. In Figura 2.3-1 è riportato l'andamento del coefficiente di assorbimento nello spettro del VIS tra 400 e 700 nm.

Figura 2.3-1: andamento spettrale del coefficiente di assorbimento dell'acqua di mare pura (Smith-Baker).



Come è possibile notare dalla figura, il minimo di assorbimento si ha nella regione del blu a circa 450 nm; si giustifica in tal modo la predominanza del colore blu assunto dall'acqua in condizioni di purezza. I valori del coefficiente di assorbimento sono forniti in forma tabellare con una risoluzione spettrale di 2 nm. La tabella di valori sarà usata direttamente nell'implementazione del modello finale di riflettanza.

### 2.3.1.2 Il coefficiente di assorbimento del fitoplancton.

L' Eq. 2.3-9 è l'espressione del coefficiente di assorbimento del fitoplancton parametrizzato secondo la concentrazione massa su volume  $C$  di clorofilla-a più feofitina-a.

$$a_{ph\lambda} = A(\lambda) \cdot C^{1-B(\lambda)} \quad Eq. 2.3-9$$

I valori dei parametri  $A(\lambda)$  e  $B(\lambda)$  in funzione della lunghezza d'onda  $\lambda$  sono riportati in un lavoro di Bricaud ed altri [12]. Gli autori hanno ricavato tali parametri tramite analisi regressiva, con risoluzione di 2 nm, analizzando un set di dati costituito da 815 campioni misurati in acque in cui i valori di  $C$  sono compresi tra 0.02 mg/m<sup>3</sup> e 25 mg/m<sup>3</sup> ed aventi bassa concentrazione relativa di feofitina, condizioni queste che fissano i limiti di validità del modello di riflettanza.

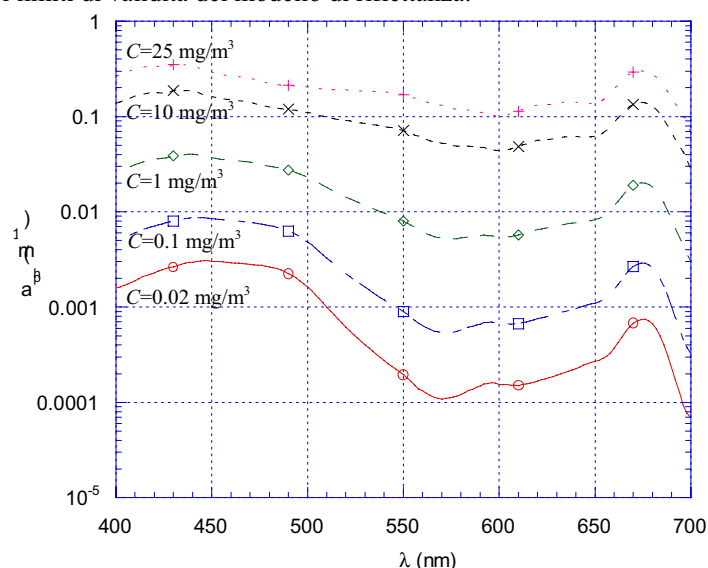


Figura 2.3-2: : coefficiente di assorbimento del fitoplancton (Bricaud) per diversi valori di  $C$ .

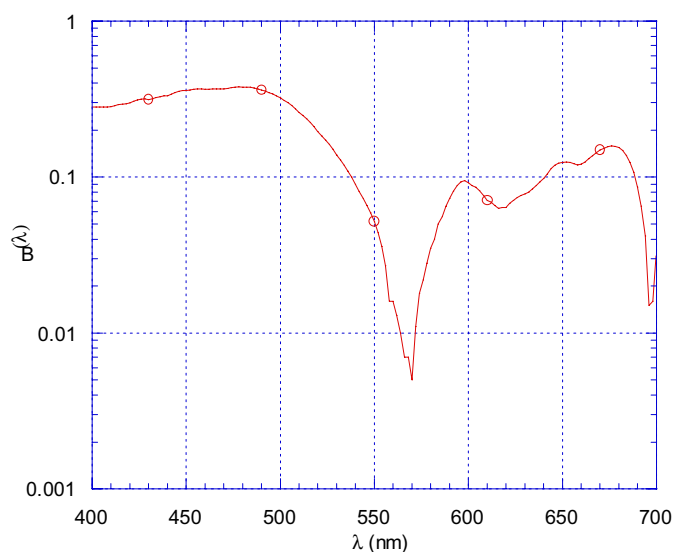


Figura 2.3-3: andamento spettrale del parametro  $B(\lambda)$ .

In Figura 2.3-2 sono riportate le curve del coefficiente di assorbimento per valori di  $C$  che coprono l'intervallo di validità della parametrizzazione. La curva a concentrazione unitaria corrisponde all'andamento della funzione  $A(\lambda)$ . La forma dello spettro di assorbimento al variare di  $C$  presenta un andamento caratteristico con un primo massimo relativo nella regione del blu tra 420 nm e 490 nm, un minimo nella regione del verde tra 450 nm e 600 nm ed un secondo massimo nella regione del rosso tra 650 nm e 700 nm. In Figura 2.3-3 è riportato l'andamento spettrale di  $B(\lambda)$ .

### 2.3.1.3 Il coefficiente di assorbimento delle particelle non clorofillacee.

La forma generale del coefficiente di assorbimento delle particelle non clorofillacee è riportata nell' Eq. 2.3-10 (in [3]) in cui si è indicato con  $a_{sed}^*(\lambda_0)$  il coefficiente di assorbimento alla lunghezza d'onda  $\lambda_0$  reso specifico rispetto a  $X$ . La lunghezza d'onda  $\lambda_0$  presa come riferimento è 440 nm per cui *Prieur* e *Sathyendranath* [13] hanno ricavato un valore di  $a_{sed}^*(\lambda_0)$  pari a 0.042 (adimensionale) analizzando una serie di misure *in situ*.

$$a_{sed\lambda} = X \cdot a_{sed}^*(\lambda_0) \cdot \exp[-s_x \cdot (\lambda - \lambda_0)] . \quad Eq. 2.3-10$$

Il parametro  $s_x$  regola la velocità con cui l'esponenziale tende a zero al crescere della lunghezza d'onda. In generale tale parametro non può essere considerato costante dipendendo esso dal particolare tipo di acqua preso in considerazione. In [14] è riportato un intervallo di variabilità pari a  $0.011 \pm 0.0025 \text{ nm}^{-1}$ . Nel modello di riflettanza sarà preso in considerazione il valore medio di tale intervallo, sarà posto cioè  $s_x = 0.011$ . L'espressione finale del coefficiente di assorbimento diventa quindi:

$$a_{sed\lambda} = 0.042 \cdot X \cdot \exp[-0.011 \cdot (\lambda - 440)] . \quad Eq. 2.3-11$$

In Figura 2.3-4 è riportato l'andamento del coefficiente di assorbimento specifico ricavato dividendo per  $X$  il coefficiente di assorbimento riportato nella Eq. 2.3-11 mentre in Figura 2.3-5 è riportato l'andamento del coefficiente di assorbimento per diversi valori di  $X$  compresi tra  $0.01 \text{ m}^{-1}$  e  $10 \text{ m}^{-1}$ .



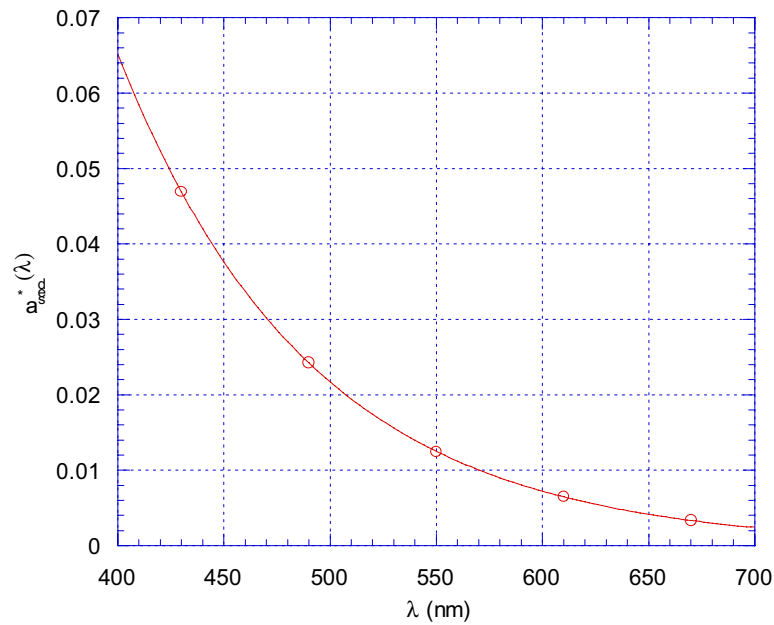


Figura 2.3-4: andamento spettrale del coefficiente di assorbimento specifico delle particelle non clorofillacee.

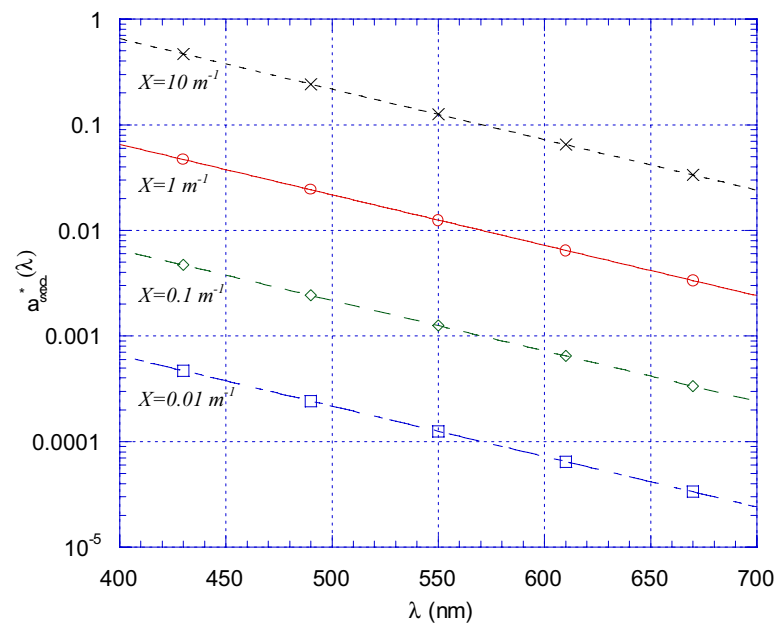


Figura 2.3-5: andamento del coefficiente di assorbimento delle particelle non clorofillacee per diversi valori di  $X$

#### 2.3.1.4 Il coefficiente di assorbimento della sostanza gialla.

La relazione comunemente accettata per modellare l'andamento spettrale del coefficiente di assorbimento della sostanza gialla è, analogamente al coefficiente di assorbimento delle particelle non clorofillacee, di tipo esponenziale:

$$a_{yel\lambda} = Y \cdot \exp[-s_y \cdot (\lambda - 440)]. \quad \text{Eq. 2.3-12}$$

*Morel e Prieur* [15] propongono un valore per  $s_y$  pari a  $0.014 \text{ nm}^{-1}$ , valore successivamente confermato in alcuni lavori scientifici sull'argomento ed assimilato nel modello di riflettanza usato per lo sviluppo degli algoritmi di stima degli OAPs del sensore SeaWiFS (in [16]). Non sono comunque da escludere valori diversi per  $s_y$ : sono vari gli studi, condotti su diversi tipi di acque sia in zone costiere che in mare aperto, che riportano valori del parametro leggermente maggiori e dipendenti dalla salinità delle acque il che suggerirebbe l'introduzione di tale parametro nel modello del coefficiente di assorbimento. Nel presente lavoro di tesi sarà comunque usato il valore di riferimento proposto da *Morel e Prieur* e ripreso in [3]. L'espressione del coefficiente di assorbimento utilizzata sarà quindi la seguente:

$$a_{yel\lambda} = Y \cdot \exp[-0.014 \cdot (\lambda - 440)] \quad \text{Eq. 2.3-13}$$

In Figura 2.3-6 è riportato l'andamento spettrale del coefficiente di assorbimento per diversi valori del parametro  $Y$  tra  $0.01 \text{ m}^{-1}$  e  $1 \text{ m}^{-1}$ , mentre in Figura 2.3-7 è riportato il confronto tra i coefficienti di assorbimento specifici, normalizzati al valore del coefficiente di assorbimento a  $\lambda=440 \text{ nm}$ , delle particelle non clorofillacee e della sostanza gialla in cui si può notare la maggiore selettività dell'assorbimento di quest'ultimo costituente.

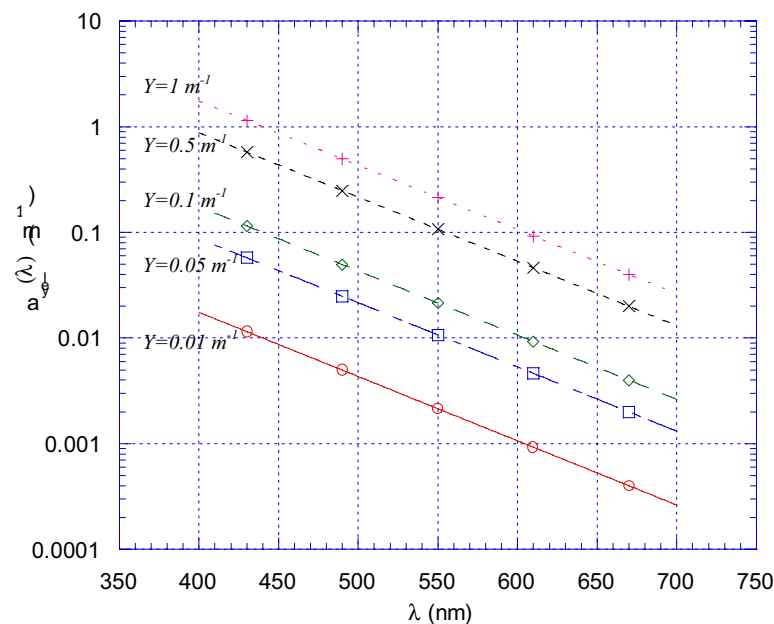


Figura 2.3-6: andamento spettrale del coefficiente di assorbimento della sostanza gialla per diversi valori del parametro  $Y$ .

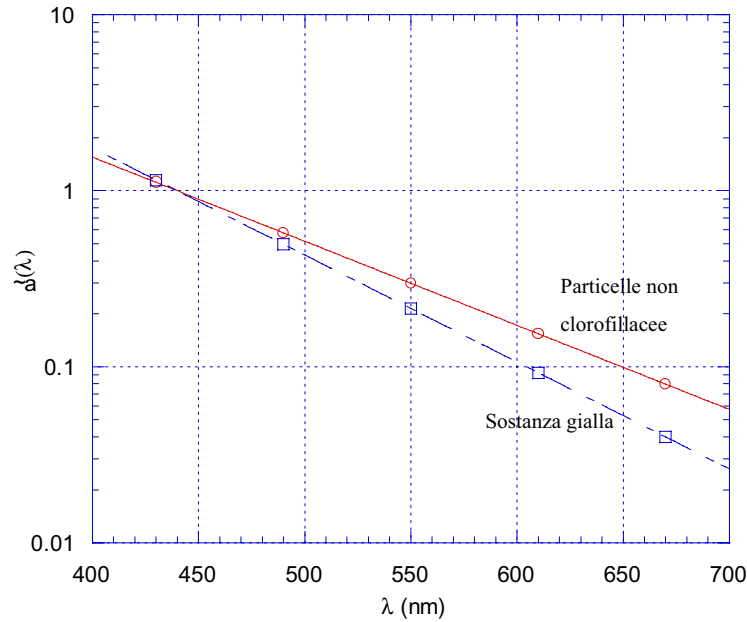


Figura 2.3-7: confronto tra i coefficienti di assorbimento specifici, normalizzati al valore del coefficiente di assorbimento a  $\lambda=440$  nm, della sostanza gialla e delle particelle non clorofillacee.

## 2.3.2 Il coefficiente di *backscattering*.

### 2.3.2.1 Il coefficiente di *backscattering* dell'acqua di mare pura.

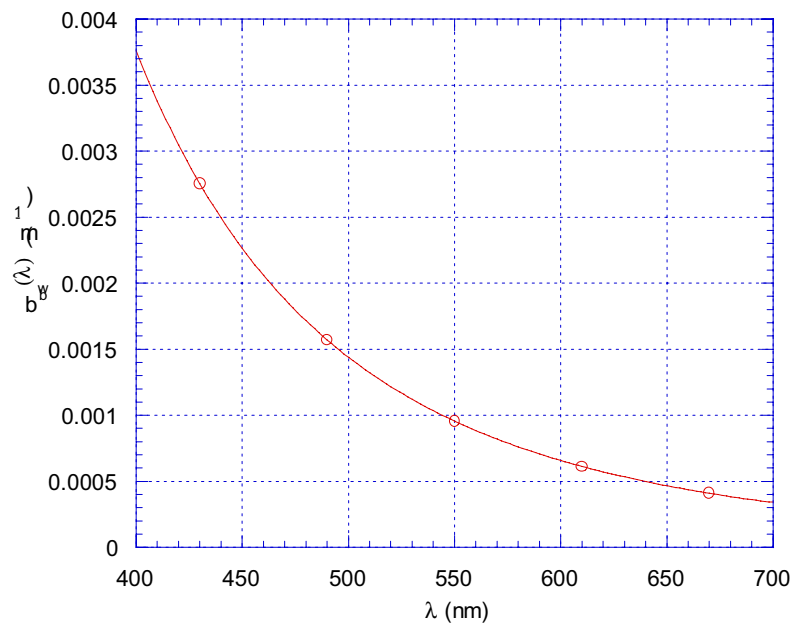
L'espressione diffusamente usata negli studi di colore marino per modellare il coefficiente di *backscattering* dell'acqua di mare pura è fornita da Morel in [17]:

$$b_{bw\lambda} = b'_{bw} \cdot b_{w\lambda} = b'_{bw} \cdot b_w(\lambda = 500 \text{ nm}) \cdot \left( \frac{\lambda}{500} \right)^{-4.3}, \quad \text{Eq. 2.3-14}$$

in cui si è indicato con  $b_w(\lambda=500 \text{ nm})$  il coefficiente di *scattering* a 500 nm pari a  $0.00288 \text{ m}^{-1}$ . Il valore del rapporto tra il coefficiente di *backscattering* e di *scattering*  $b'_{bw}$  è pari a 0.5 data la simmetria della funzione di fase delle molecole di acqua. Lo *scattering* è di tipo Rayleigh. L'espressione definitiva del coefficiente di *backscattering* è quindi la seguente:

$$b_{bw\lambda} = 0.5 \cdot 0.00288 \cdot \left( \frac{\lambda}{500} \right)^{-4.3} \quad \text{Eq. 2.3-15}$$

di cui è riportato l'andamento in Figura 2.3-8:



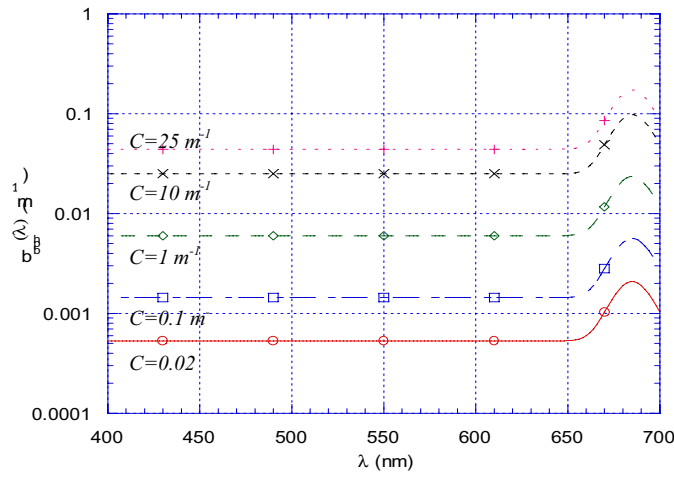


Figura 2.3-8: andamento spettrale del coefficiente di backscattering dell'acqua di mare pura.

### 2.3.2.2 Il coefficiente di backscattering del fitoplancton.

Il coefficiente di scattering del fitoplancton, legato al coefficiente di backscattering tramite il rapporto  $b'_{bph}$ , è, al pari del coefficiente di assorbimento, di difficile modellazione. Esso infatti, dipende fortemente dalle caratteristiche delle celle fitoplanctoniche in termini di stato fisiologico e di concentrazione relativa delle diverse specie algali presenti in acqua. Tutto ciò crea notevoli difficoltà quando si voglia cercare di modellare l'andamento spettrale del coefficiente di scattering del fitoplancton. In virtù delle indicazioni fornite in [3], la scelta più opportuna ai fini della modellazione della riflettanza sembra quella di usare un'espressione per il coefficiente di backscattering che sia indipendente dalla lunghezza d'onda:

$$b_{bph\lambda} = 0.002 \cdot 0.3 \cdot C^{0.62}, \quad \text{Eq. 2.3-16}$$

eccetto che per una zona spettrale, all'incirca compresa tra 680 nm e 690 nm, in cui il coefficiente misurato per varie specie di fitoplancton, presenta un picco, attribuibile a fenomeni di fluorescenza della clorofilla. Pur essendo tali fenomeni di difficile modellazione, in [3] è riportata un'espressione che prevede, nella zona spettrale considerata, un andamento di tipo gaussiano del coefficiente di backscattering. Il picco gaussiano che si somma al valore costante dato dalla Eq. 2.3-16 è centrato sulla lunghezza d'onda di 685 nm ed ha una deviazione standard pari a 10 nm. L'ampiezza della funzione gaussiana è scelta in modo tale che il coefficiente di backscattering complessivo a 685 nm risulti uguale a dieci volte il valore della Eq. 2.3-16. Il coefficiente di backscattering complessivo in funzione di  $C$  e della lunghezza d'onda  $\lambda$  è dato quindi dalla seguente relazione:

$$b_{bph\lambda} = 0.002 \cdot 0.3 \cdot C^{0.62} + 0.022 \cdot \frac{C^{0.62} \cdot \sqrt{2} \cdot \exp\left[-\frac{(\lambda - 685)^2}{200}\right]}{\sqrt{\pi}}. \quad \text{Eq. 2.3-17}$$

In Figura 2.3-9 è riportato l'andamento del coefficiente di backscattering totale per diversi valori di  $C$  compresi tra 0.02 mg/m<sup>3</sup> e 25 mg/m<sup>3</sup>.

Figura 2.3-9: andamento spettrale del coefficiente di backscattering del fitoplancton per diversi valori di  $C$ .

### 2.3.2.3 Il coefficiente di backscattering delle particelle non clorofillacee.

Le particelle non clorofillacee danno il maggior contributo al backscattering totale ed anche in acque di caso I, in cui la concentrazione di sedimento inorganico è normalmente minore di quella del fitoplancton, tale contributo risulta essere importante. La dipendenza dalla lunghezza d'onda del coefficiente di scattering delle particelle è modellata tramite una legge iperbolica del tipo  $\lambda^{-n}$ . Normalmente si assume per  $n$  un valore pari ad uno. L'espressione del coefficiente di backscattering in termini di coefficiente di scattering risulta essere:

$$b_{bsed\lambda} = b'_{bsed} \cdot b_{sed\lambda} = b'_{bsed} \cdot b_{sed}(\lambda = 550 \text{ nm}) \cdot \left(\frac{\lambda}{550}\right)^{-1}. \quad Eq. 2.3-18$$

Il coefficiente di *scattering* a 550 nm è, come visto precedentemente, il parametro  $X$  preso come indice della concentrazione delle particelle non clorofillacee. La fonte di maggiore variabilità nella Eq. 2.3-18 è rappresentata dal rapporto  $b'_{bsed}$  tra il coefficiente di *backscattering* e quello di *scattering*. In [3] è stato assunto un valore  $b'_{bsed}=0.02$  in accordo a quello riscontrabile in Morel [18]. L'espressione finale del coefficiente di *backscattering* quindi, risulta essere:

$$b_{bsed\lambda} = 0.02 \cdot X \cdot \left(\frac{\lambda}{550}\right)^{-1} \quad Eq. 2.3-19$$

di cui è riportato l'andamento per diversi valori di  $X$  nella Figura 2.3-10.

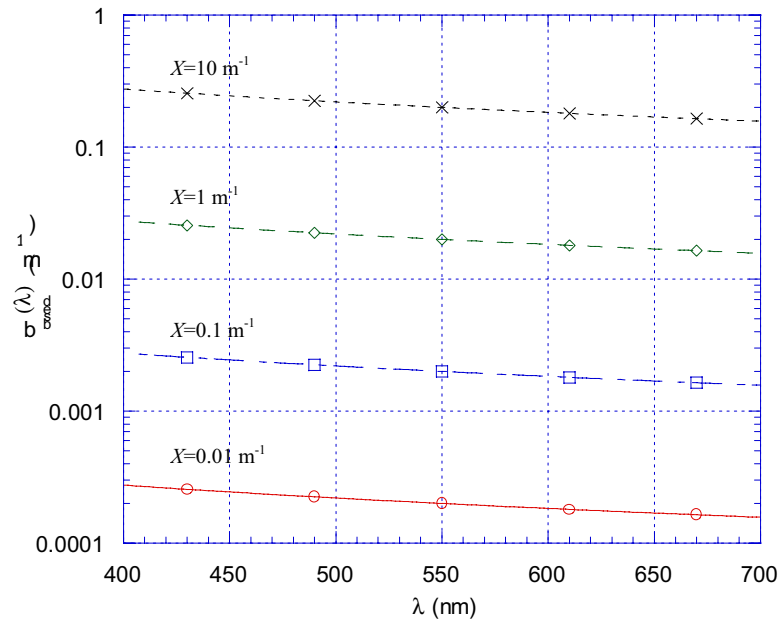


Figura 2.3-10: andamento spettrale del coefficiente di backscattering delle particelle non clorofillacee per diversi valori di  $X$ .

### 2.3.3 Espressione della riflettanza sub-superficiale.

Si ritiene utile riportare l'espressione finale della riflettanza sub-superficiale in funzione della lunghezza d'onda e degli OAPs. Sostituendo i singoli contributi di assorbimento e *backscattering* presenti nell'Eq. 2.3-13, Eq. 2.3-6 e Eq. 2.3-15, tale espressione risulta essere la seguente:

$$R(C, X, Y; \lambda) = \frac{b_w(500) \cdot b'_{bw} \cdot \left(\frac{\lambda}{500}\right)^{-4.3} + b'_{bph} \cdot 0.30 \cdot C^{0.62} + 0.022 \cdot \frac{C^{0.62} \cdot \sqrt{2} \cdot e^{\left[-\frac{(\lambda-685)^2}{200}\right]}}{\sqrt{\pi}} + b'_{bsed} \cdot X \cdot \left(\frac{\lambda}{550}\right)^{-1}}{a_w(\lambda) + A(\lambda) \cdot C^{1-B(\lambda)} + 0.042 \cdot X \cdot e^{\left[-s_x \cdot (\lambda-440)\right]} + Y \cdot e^{\left[-s_y \cdot (\lambda-440)\right]}} \quad Eq. 2.3-20$$

in cui il valore dei diversi parametri presenti è riportato nella Tabella 2.3-1. Tali parametri sono in evidenza perché costituiscono la maggiore fonte di variabilità all'interno del modello di riflettanza e

quindi possono essere oggetto di uno studio di sensibilità, in cui si va a valutare l'influenza che essi hanno sui parametri del modello inverso utilizzato per stimare gli OAPs.

L'implementazione del modello di colore marino [3] permette di calcolare la riflettanza spettrale nell'intervallo di lunghezze d'onda del visibile che va da 400 nm fino a 700 nm con un passo di 2 nm. Il campo di validità del modello è fissato dalle condizioni di validità della parametrizzazione del coefficiente di assorbimento del fitoplancton (*Eq. 2.3-9*) proposta da *Bricaud* in [12] e cioè:

- $0.02 \text{ mg/m}^3 < C < 25 \text{ mg/m}^3$ ,
- non valida per alte concentrazioni relative di feofitina,
- non valida per acque con caratteristiche bio-ottiche particolari quali certe acque antartiche ad elevata concentrazione di nutrienti.

*Tabella 2.3-1: valore dei parametri presenti nell'Eq. 2.3-20 del modello di riflettanza.*

$r$	$b_w(500) \text{ (m}^{-1}\text{)}$	$b'_{bw}$	$b'_{bph}$	$b'_{bsed}$	$S_x$	$S_y$
0.33	0.00288	0.5	0.002	0.02	0.011	0.014

#### 2.3.4 Relazione tra riflettanza sub-superficiale e *Water Leaving Radiance*.

Per la definizione completa del problema diretto è necessario valutare la radianza *water leaving* in termini della riflettanza sub-superficiale. Ciò è possibile esprimendo, nell'*Eq. 1.5-5*, l'irradianza *upwelling* in funzione della radianza *water leaving*:

$$E_u(0^-, \lambda) = \frac{Q \cdot n^2 \cdot L_{wl}(\lambda)}{(1 - \rho)}, \quad \text{Eq. 2.3-21}$$

in cui  $Q$  è un fattore che vale  $\pi$  per superfici lambertiane,  $n$  è il coefficiente di rifrazione dell'acqua di mare pari a 1.34 e  $\rho$  il coefficiente di riflessione di Fresnel per l'interfaccia acqua-aria considerato costante e pari a 0.021. Esprimendo infine l'irradianza *downwelling* in termini dell'irradianza solare al *top* dell'atmosfera  $F^0(\lambda)$ :

$$E_d(0^-, \lambda) = (1 - \rho') \cdot t(\lambda, \theta_0) \cdot F^0(\lambda) \cdot \cos(\theta_0), \quad \text{Eq. 2.3-22}$$

in cui  $\rho'$  è il coefficiente di riflessione di Fresnel per l'interfaccia aria-acqua,  $t(\lambda, \theta_0)$  la trasmittanza atmosferica *downward* che tiene conto dell'attenuazione che la radiazione solare subisce nel cammino ottico discendente (sole-superficie marina) e  $\theta_0$  l'angolo zenitale del sole, è possibile ricavare la relazione tra radianza e riflettanza sostituendo  $E_u(0^-, \lambda)$  e  $E_d(0^-, \lambda)$  nell'*Eq. 1.5-5*:

$$L_{wl}(\lambda) = \frac{(1 - \rho) \cdot (1 - \rho') \cdot T(\theta_0, \lambda) \cdot F^0(\lambda) \cdot \cos(\theta_0)}{Q \cdot n^2} \cdot R(\lambda). \quad \text{Eq. 2.3-23}$$

L'*Eq. 2.3-23* sarà utilizzata in seguito per stimare la riflettanza sub-superficiale dalla radianza *water leaving* valutata dai dati telerilevati corretti dal contributo atmosferico.

## 2.4 Schematizzazione del problema diretto.

In Figura 2.4-1 è schematizzato il modello diretto che permette il calcolo della radianza totale al sensore in ciascun canale spettrale. Tramite il modello di colore marino è possibile calcolare la riflettanza spettrale

sub-superficiale in corrispondenza dei parametri di ingresso  $C, X, Y$ . Il modello atmosferico una volta acquisiti i parametri relativi all'aerosol (tipo di funzione di fase e parametri della funzione di fase), il valore di pressione atmosferica  $P_{atm}$  e la configurazione della geometria di ripresa del sensore (angoli zenitali ed azimutali del sole e del sensore:  $\theta, \theta_0, \phi, \phi_0$ ), restituisce in uscita la trasmittanza in discesa  $T(\lambda)$ , la trasmittanza in salita  $t(\lambda)$  e la *path radiance* somma della radianza di *Rayleigh* e della radianza dell'aerosol  $L_r(\lambda)+L_a(\lambda)$ . Tramite l'Eq. 2.3-23, data la riflettanza  $R(\lambda)$ , la trasmittanza  $T(\lambda)$  e l'angolo zenitale del sole  $\theta_0$  è possibile calcolare la radianza *water leaving*  $L_w(\lambda)$ . Quest'ultima viene moltiplicata per la trasmittanza  $t(\lambda)$  e sommata alla *path radiance* per costituire la radianza totale al sensore  $L_t(\lambda)$ . La radianza totale viene infine integrata sulle bande del sensore per fornire in uscita la radianza totale media per i singoli canali spettrali:

$$L_t(\lambda_i) = \frac{\int H_i(\lambda) \cdot L_t(\lambda) \cdot d\lambda}{\int H_i(\lambda) \cdot d\lambda}, \quad \text{Eq. 2.4-1}$$

in cui  $H_i(\lambda)$  è la risposta spettrale per il canale  $i$ -esimo ed  $L_t(\lambda_i)$  la radianza totale media sul canale  $i$ -esimo.

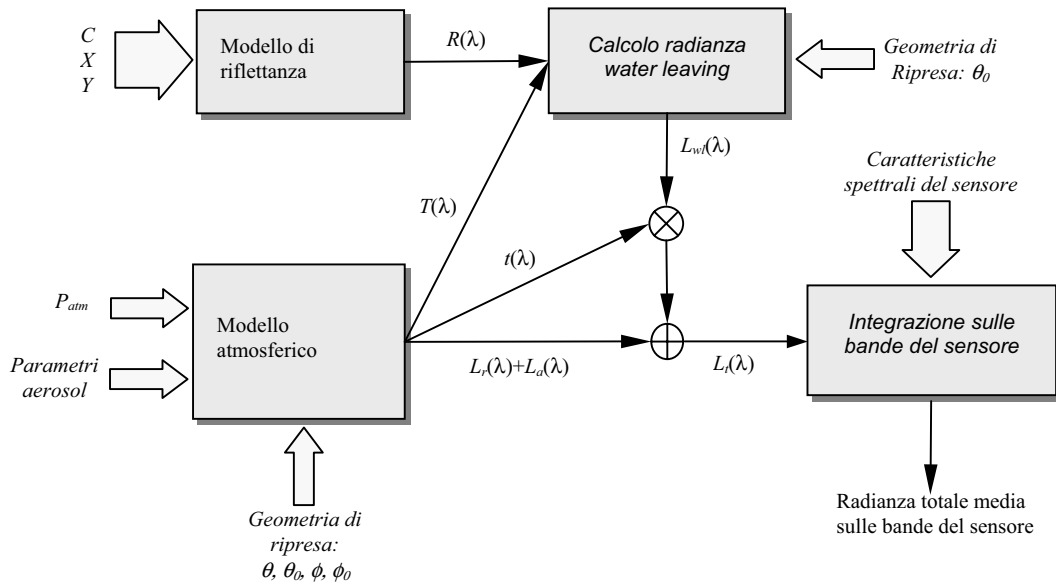


Figura 2.4-1: schema del modello diretto per il calcolo della radianza totale al sensore.

---

## 3 Capitolo 3

### 4 *La stima dei parametri otticamente attivi dell'acqua marina da dati multispettrali: inversione del modello di colore marino.*

Nel presente capitolo sarà affrontato il problema dell'inversione del modello diretto di colore marino e saranno fornite le basi teoriche per la definizione degli algoritmi di stima degli OAPs dai dati di riflettanza sub-superficiale. La relazione che lega la riflettanza agli OAPs risulta altamente non lineare come è evidente dall'osservazione dell'Eq. 2.3-21. Un'inversione diretta del modello di riflettanza non è perciò attuabile, si dovrà ricorrere quindi all'uso di metodi indiretti ed approssimazioni. Il capitolo comprende un primo paragrafo in cui è descritta, in termini generali, la procedura per la definizione degli algoritmi di stima degli OAPs, procedura che fa utilizzo del modello diretto di riflettanza sub-superficiale descritto nel precedente capitolo. Il secondo paragrafo è dedicato alle tecniche di inversione del modello diretto. Sono descritti, in un primo sotto paragrafo, gli algoritmi classici, a **rapporto di bande e mono-banda**, basati su certe caratteristiche peculiari della forma della curva di riflettanza spettrale, diffusamente usati nelle applicazioni e di cui è possibile trovare riferimento in svariati lavori scientifici sull'argomento [16]. Tali algoritmi non fanno uso di tutta l'informazione fornita dai dati relativi ai singoli canali del sensore e possono risultare quindi non adatti alla stima degli OAPs nelle acque di caso 2 [2]. Nei sotto paragrafi successivi sono prese in esame alcune tecniche, ancora sperimentali nell'ambito degli studi di colore marino, che fanno uso di tutta l'informazione contenuta nei dati del sensore. Gli algoritmi che saranno presi in considerazione in tale sezione comprendono:

- **algoritmi multi-banda**, definiti mediante regressione lineare effettuata tramite l'**analisi delle componenti principali** (PCA: *Principal Component Analysis*) dei dati di riflettanza,
- **algoritmi a rete neurale RBF** (RBF-NN: *Radial Basis Function-Neural Networks*).

I tre approcci possono rientrare in uno schema generale d'inversione in cui il problema della definizione degli algoritmi di stima degli OAPs è ricondotto all'identificazione di un sistema inverso parametrico che, accettando in ingresso i dati di riflettanza, restituisce il valore delle stime di  $C$ ,  $X$  e  $Y$ . I metodi di inversione proposti permettono di stimare i parametri del modello inverso mediante tecniche di regressione lineare in cui la minimizzazione di adeguate funzioni di costo avviene attraverso l'uso dei metodi dell'algebra lineare, rendendo la fase di definizione degli algoritmi di stima degli OAPs estremamente rapida ed efficiente.



#### 4.1 Uso del modello di riflettanza sub-superficiale per la soluzione del problema inverso

In questo paragrafo è descritta in termini generali la procedura per la definizione degli algoritmi di stima degli OAPs. Tale procedura fa uso del modello diretto di colore marino per simulare le riflettanze medie sulle bande del sensore a partire da un *data set* di OAPs costituito da un certo numero di triplette  $[C, X, Y]$ ; il metodo è schematizzato in Figura 4.1-1.

Sia  $R(0^-, \lambda)$  la riflettanza sub-superficiale dell'acqua di mare alla lunghezza d'onda  $\lambda$ . In accordo al modello di colore marino descritto nel capitolo 2 (Eq. 2.3-21) questa può essere espressa come funzione dei tre parametri otticamente attivi  $C, X, Y$ :

$$R(0^-, \lambda) = f(C, X, Y; \lambda). \quad \text{Eq. 4.1-1}$$

Le riflettanze sub-superficiali medie, stimate dai dati del sensore corretti dal contributo atmosferico in ognuno degli  $N_c$  canali spettrali, possono essere arrangiate in un vettore di dimensione  $N_c$ . L'Eq. 4.1-1 può essere quindi scritta in forma vettoriale nel seguente modo:

$$\underline{R} = \underline{f}(\underline{OAP}), \quad \text{Eq. 4.1-2}$$

in cui  $\underline{OAP} = [C, X, Y]^T$  e  $\underline{R} = [f(\underline{OAP}, \lambda_1), \dots, f(\underline{OAP}, \lambda_{N_c})]^T$ . La  $\underline{f}$  è la funzione vettoriale che mappa lo spazio tridimensionale degli OAPs nello spazio  $N_c$ -dimensionale delle riflettanze  $\underline{R}$  stimate dai dati del sensore. In tal modo  $\underline{f}$  è il modello diretto del problema inverso che deve essere risolto per stimare il vettore degli OAPs dalle riflettanze misurate. La soluzione di questo problema può essere raggiunta scegliendo un modello inverso parametrico:

$$\underline{OAP} = \underline{g}(\underline{R}; \underline{a}), \quad \text{Eq. 4.1-3}$$

dove  $\underline{a}$  è un vettore di parametri incogniti da stimare e  $\underline{g}(\cdot)$  una funzione di forma nota opportunamente scelta. Fornendo in ingresso al modello diretto di colore marino un *set* di  $N$  misure *in-situ* di OAPs, acquisite sull'area geografica di interesse, si simula il vettore delle riflettanze stimate dai dati del sensore  $\underline{R}$ . Si stimano quindi i parametri  $\underline{a}$  del modello inverso minimizzando l'errore quadratico medio (MSE: Mean Square Error) tra gli OAPs misurati e quelli ottenuti applicando la funzione  $\underline{g}(\cdot)$  ai dati di riflettanza simulati:

$$\underline{a} = \arg \left[ \min_{\underline{a}} \left( \frac{1}{N} \sum_{i=1}^N |\underline{OAP}_i - \underline{\hat{OAP}}_i|^2 \right) \right], \quad \text{Eq. 4.1-4}$$

in cui con  $\underline{OAP}_i$  si è indicata l' $i$ -esima tripletta  $[C, X, Y]_i$  appartenente al *data set* di misure *in-situ* di OAPs ovvero il valore osservato e con  $\underline{\hat{OAP}}_i$  l' $i$ -esima tripletta di parametri otticamente attivi valutati tramite l'Eq. 4.1-3 ovvero i cosiddetti valori predetti degli OAPs. La bontà del modello inverso può essere successivamente valutata calcolando il coefficiente di correlazione tra le misure di OAPs ed i valori stimati:

$$\underline{\rho} = [\rho_C, \rho_X, \rho_Y]^T, \quad \text{Eq. 4.1-5}$$

$$\rho_x = \frac{1}{N} \cdot \frac{\sum_{i=1}^N (x_i - \langle x \rangle) \cdot (\hat{x}_i - \langle \hat{x} \rangle)}{\sigma_x \cdot \sigma_{\hat{x}}}, \quad x=C, X, Y, \quad \text{Eq. 4.1-6}$$

in cui con  $\rho_x$  si è indicato il coefficiente di correlazione per il parametro  $x$ , con  $\sigma_x$  e  $\sigma_{\hat{x}}$  la deviazione standard del costituente  $x$  rispettivamente osservato e predetto e con  $\langle x \rangle$  e  $\langle \hat{x} \rangle$  il valor medio del parametro  $x$  rispettivamente osservato e predetto.

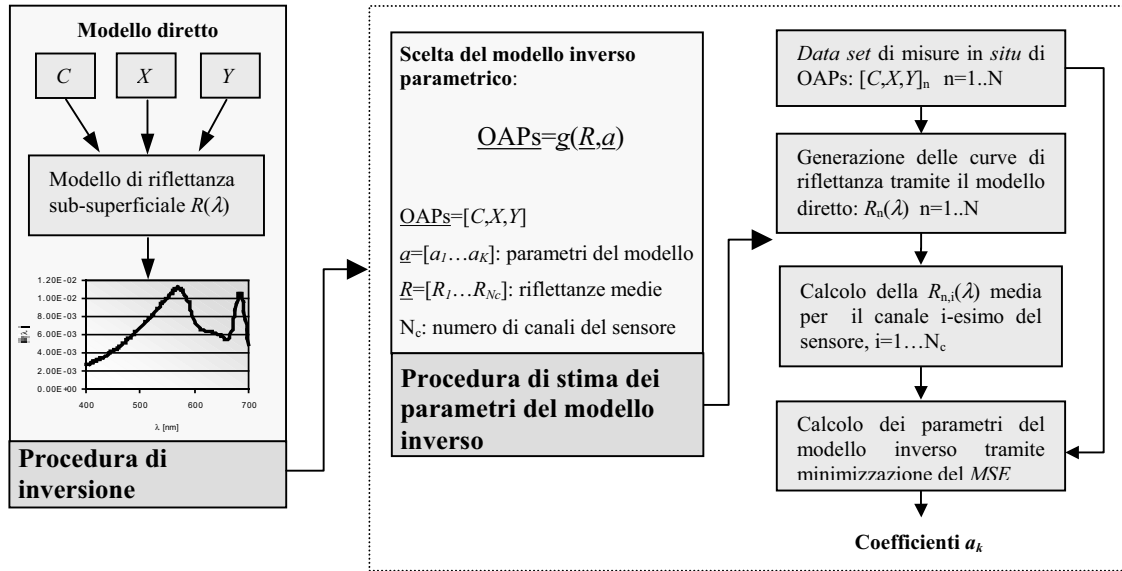


Figura 4.1-1: procedura di definizione degli algoritmi di stima degli OAPs da dati di riflettanza.

Classicamente, la definizione degli algoritmi per gli studi di colore marino viene effettuata ricorrendo ad un approccio di tipo empirico, vengono cioè stimati i coefficienti del modello inverso correlando misure *in-situ* di OAPs e di riflettanza. L'approccio seguito in questo lavoro di tesi è invece di tipo semi-analitico, si sfrutta cioè un modello analitico di colore marino per simulare la riflettanza sub-superficiale nelle bande del sensore in corrispondenza di un *set* di misure *in-situ* di OAPs. L'approccio *model based* ha il vantaggio di non richiedere la misura delle riflettanze durante le campagne di raccolta dati in mare con notevole risparmio in termini di costi di strumentazione. Di contro, la bontà o meno dei risultati che si ottengono con tali metodi, dipende fortemente dall'accuratezza con cui la riflettanza spettrale è modellata. Risulta perciò cruciale attuare un'intensa fase di validazione, mediante campagne di misura in mare contemporanee all'acquisizione dei dati satellitari, orientata a stabilire sia le capacità del modello di colore marino di rappresentare il più fedelmente possibile l'andamento della curva di riflettanza spettrale, sia la precisione con cui gli algoritmi sviluppati con l'approccio descritto stimano il valore degli OAPs.

## 4.2 Algoritmi per la stima dei parametri otticamente attivi

Nel presente paragrafo sono descritti in dettaglio gli algoritmi per la stima degli OAPs e le basi teoriche necessarie alla loro definizione. Saranno presi in considerazione i seguenti modelli inversi parametrici:

- modelli a rapporto di bande e mono-banda,
- modelli multibanda ed inversione tramite PCA dei dati di riflettanza,
- modelli multibanda ed inversione tramite RBF-NN.

Nel sotto paragrafo relativo agli algoritmi del primo tipo saranno introdotte alcune notazioni ed elementi teorici relativi ai metodi di regressione lineare. Ciò risulterà utile nel proseguo dato che gli aspetti teorici trattati potranno essere facilmente estesi agli algoritmi multibanda sia di tipo PCA che RBF-NN. Vale la pena ribadire il fatto che tutti gli algoritmi proposti possono essere ricondotti ad uno schema di regressione lineare in cui i parametri del modello inverso sono stimati ricorrendo ai metodi dell'algebra lineare. Tale scelta rende il processo di definizione degli algoritmi di stima degli OAPs estremamente più efficiente e semplice rispetto ad altri metodi di inversione non lineari.

#### 4.2.1 Algoritmi a rapporto di bande e mono-banda.

Gli algoritmi a rapporto di bande utilizzano il rapporto tra le riflettanze spettrali misurate in due bande distinte del sensore. Il modello che lega il valore del generico *OAP* ( $C, X$  o  $Y$ ) a tale rapporto è il seguente:

$$OAP = A_0 \cdot \left( \frac{R_i}{R_j} \right)^{a_1}, \quad \text{Eq. 4.2-1}$$

in cui si sono indicate con  $R_i$  ed  $R_j$  le riflettanze spettrali medie per i canali  $i$  e  $j$  del sensore;  $A_0$  ed  $a_1$  sono i parametri del modello che vanno stimati minimizzando l'MSE in accordo all'Eq. 4.1-4. Applicando il logaritmo in base dieci ad entrambe i membri dell'Eq. 4.2-1 si ottiene:

$$\text{Log}(OAP) = a_0 + a_1 \cdot \text{Log}\left(\frac{R_i}{R_j}\right), \quad \text{Eq. 4.2-2}$$

in cui  $a_0 = \text{Log}(A_0)$ . Il modello diventa un semplice modello di regressione lineare in cui la variabile indipendente è il  $\text{Log}(R_i/R_j)$  e la variabile dipendente il  $\text{Log}(OAP)$ . Valutando il secondo membro dell'Eq. 4.2-2 in corrispondenza delle riflettanze generate con il modello di colore marino a partire da un *data set* di  $N$  triplette  $[C, X, Y]$  di OAPs osservati ed eguagliandolo al valore osservato del  $\text{Log}(OAP)$  si ottiene una serie di  $N$  equazioni:

$$\begin{aligned} \text{Log}(OAP_1) &= a_0 + a_1 \cdot \text{Log}\left(\frac{R_{i1}}{R_{j1}}\right) + \varepsilon_1 \\ &\dots\dots\dots \\ &\dots\dots\dots \\ \text{Log}(OAP_N) &= a_0 + a_1 \cdot \text{Log}\left(\frac{R_{iN}}{R_{jN}}\right) + \varepsilon_N \end{aligned} \quad \text{Eq. 4.2-3}$$

in cui con  $\varepsilon_i$  si è indicato l'errore introdotto dal modello inverso nel valutare il  $\text{Log}(OAP)$ . Le Eq. 4.2-3 possono essere scritte in notazione matriciale come segue:

$$\underline{\psi}_{OAP} = \Gamma \cdot \underline{a} + \underline{\varepsilon}, \quad \text{Eq. 4.2-4}$$

in cui

$$\underline{\psi}_{OAP} = \begin{bmatrix} \text{Log}(OAP_1) \\ \dots \\ \text{Log}(OAP_N) \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 1 & \text{Log} \frac{R_{i1}}{R_{j1}} \\ \dots \\ 1 & \text{Log} \frac{R_{iN}}{R_{jN}} \end{bmatrix}, \quad \underline{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}, \quad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \dots \\ \varepsilon_N \end{bmatrix}, \quad \text{Eq. 4.2-5}$$

sono rispettivamente il vettore del logaritmo dell'OAP osservato, la matrice della variabile indipendente, detta matrice dei regressori, il vettore incognito dei parametri del modello ed il vettore degli errori. La stima del vettore dei coefficienti  $\underline{a}$  si effettua minimizzando l'errore quadratico (MSE a meno del fattore  $1/N$ ):

$$S = (\underline{\psi}_{OAP} - \Gamma \cdot \underline{a})^T \cdot (\underline{\psi}_{OAP} - \Gamma \cdot \underline{a}). \quad \text{Eq. 4.2-6}$$

Derivando l'Eq. 4.2-6 rispetto al vettore dei parametri incogniti ed uguagliando a zero si ottiene lo stimatore del vettore  $\underline{a}$  dei parametri incogniti:

$$\hat{\underline{a}} = (\Gamma^T \cdot \Gamma)^{-1} \cdot \Gamma^T \cdot \underline{\psi}_{OAP}, \quad \text{Eq. 4.2-7}$$

in cui con  $\hat{\underline{a}}$  si è indicato il vettore stimato dei coefficienti del modello. Lo stimatore così ottenuto è detto ai minimi quadrati. I valori predetti del logaritmo dell'OAP si ricavano moltiplicando la matrice dei regressori  $\Gamma$  per il vettore  $\hat{\underline{a}}$ :

$$\hat{\underline{\psi}}_{OAP} = \Gamma \cdot \hat{\underline{a}} = \Gamma \cdot (\Gamma^T \cdot \Gamma)^{-1} \cdot \Gamma^T \cdot \underline{\psi}_{OAP} = H \cdot \underline{\psi}_{OAP}, \quad \text{Eq. 4.2-8}$$

in cui  $H = \Gamma \cdot (\Gamma^T \cdot \Gamma)^{-1} \cdot \Gamma^T$ . Ponendo  $M = I - H$ , il residuo tra valori osservati e predetti è pari a:

$$\begin{aligned} \hat{\underline{\varepsilon}} &= \underline{\psi}_{OAP} - \hat{\underline{\psi}}_{OAP} = \underline{\psi}_{OAP} - H \cdot \underline{\psi}_{OAP} = (I - H) \cdot \underline{\psi}_{OAP} = \\ &= M \cdot \underline{\psi}_{OAP} = M \cdot \Gamma \cdot \underline{a} + M \cdot \underline{\varepsilon} = M \cdot \underline{\varepsilon} \end{aligned}, \quad \text{Eq. 4.2-9}$$

che risulta ortogonale ai valori predetti:

$$\hat{\underline{\psi}}_{OAP}^T \cdot \hat{\underline{\varepsilon}} = 0. \quad \text{Eq. 4.2-10}$$

Nel caso in cui siano soddisfatte le seguenti condizioni:

$$E(\underline{\varepsilon}) = \underline{0}, \quad E(\underline{\varepsilon} \cdot \underline{\varepsilon}^T) = \sigma^2 \cdot I, \quad \text{Eq. 4.2-11}$$

dette di *Gauss-Markov* (*G-M*), si dimostra [19] che lo stimatore del vettore dei parametri  $\underline{a}$  è non polarizzato, cioè  $E(\hat{\underline{a}}) = \underline{a}$ , ed inoltre:

$$\text{cov}(\hat{\underline{a}}) = \sigma^2 \cdot (\Gamma^T \cdot \Gamma)^{-1}. \quad \text{Eq. 4.2-12}$$

Se inoltre la traccia della matrice  $(\Gamma^T \cdot \Gamma)^{-1}$  tende a zero al tendere di  $N$  all'infinito lo stimatore di  $\underline{a}$  gode della proprietà di consistenza ovvero  $\text{cov}(\hat{\underline{a}}) \rightarrow 0$  per  $N \rightarrow \infty$ . Sempre sotto le condizioni di *G-M* è possibile dimostrare che:

$$E(\underline{\psi}) = \Gamma \cdot \underline{a}, \quad \text{cov}(\underline{\psi}) = \sigma^2 \cdot I, \quad \text{Eq. 4.2-13}$$

ed inoltre

$$\text{cov}(\hat{\underline{\epsilon}}) = \sigma^2 \cdot \mathbf{M}. \quad \text{Eq. 4.2-14}$$

Le proprietà introdotte, valide sotto le condizioni di *G-M*, possono risultare utili nella valutazione della bontà del modello di regressione. Ciò è possibile a patto di riuscire a stimare la varianza dell'errore  $\sigma^2$ . Lo stimatore per tale parametro che sotto le condizioni di *G-M* risulta essere non polarizzato e consistente, è il seguente [19]:

$$s^2 = \frac{1}{N-2} \cdot \hat{\underline{\epsilon}}^T \cdot \hat{\underline{\epsilon}}. \quad \text{Eq. 4.2-15}$$

Ai fini di determinare la bontà del *fitting* ottenuto, assume particolare importanza la valutazione della varianza della stima dei coefficienti  $\underline{a}$  (Eq. 4.2-12). Lo stimatore non polarizzato e consistente per la covarianza di  $\underline{a}$  risulta essere [19]:

$$\text{cov}(\hat{\underline{a}}) = s^2 \cdot (\Gamma^T \cdot \Gamma)^{-1}. \quad \text{Eq. 4.2-16}$$

Vedremo nel prossimo sotto paragrafo, relativo agli algoritmi multibanda, come la dipendenza lineare tra riflettanze nei diversi canali del sensore produca delle stime dei coefficienti con elevati valori di varianza. L'analisi delle componenti principali è un metodo con cui si cerca di ridurre tale effetto che può risultare dannoso per la qualità delle stime degli OAPs da dati reali.

Tutto quanto descritto precedentemente resta valido nel caso in cui si voglia usare anziché il rapporto delle riflettanze in due bande distinte, la sola riflettanza misurata in uno degli  $N_c$  canali del sensore; l'algoritmo così realizzato è di tipo monobanda ed ha la seguente espressione:

$$\text{Log}(OAP) = a_0 + a_1 \cdot \text{Log}(R_i). \quad \text{Eq. 4.2-17}$$

In particolare resta valido il modello di regressione lineare dell'Eq. 4.2-4 (e quindi l'Eq. 4.2-7 dello stimatore ai minimi quadrati del vettore dei coefficienti  $\underline{a}$ ) a patto di sostituire nella matrice dei regressori  $\Gamma$ , il  $\text{Log}(R_i/R_j)$  con il  $\text{Log}(R_i)$ .

La scelta delle bande e del tipo di algoritmo più adatti alla stima di  $C$ ,  $X$  e  $Y$ , viene effettuata sulla base di considerazioni di tipo fisico ovvero sulla base di conoscenze a priori dell'effetto che i singoli OAPs hanno sull'andamento delle curve di riflettanza spettrale. In particolare, per quanto riguarda la stima di  $C$  nelle acque di caso I, si opta per un algoritmo a rapporto in cui si usano due bande in cui l'effetto di  $C$  è rispettivamente massimo e minimo ovvero in corrispondenza di lunghezze d'onda in cui è massimo e minimo il coefficiente di assorbimento del fitoplancton. L'effetto dell'aumento di  $C$  è quello di portare verso il basso i valori di riflettanza nella banda del blu. Nella banda del verde le variazioni della riflettanza sono di solito minime; in tale regione dello spettro esiste un punto nodale in cui praticamente convergono tutte le curve relative ai diversi valori di  $C$  (Figura 4.2-1). Quindi, il rapporto di riflettanze per la stima di  $C$  può essere costruito utilizzando la riflettanza in un canale prossimo a tale punto nodale come riferimento e la riflettanza in un canale nella banda del blu sensibile alle variazioni di  $C$ . La scelta delle bande dovrà essere operata effettuando un compromesso tra tali esigenze e la possibilità di estrarre facilmente la riflettanza sub-superficiale dai dati telerilevati. Da questo punto di vista può risultare problematico l'uso di una banda nella regione del blu nella zona tra 400 nm e 420 nm in cui il segnale utile (la radianza *water leaving*) risulta per alti valori di  $C$  estremamente basso rispetto al contributo atmosferico. L'algoritmo a rapporto di bande è basato sulla forma tipica dello spettro di riflettanza e non sul valore assoluto della stessa, caratteristica questa che rende tali algoritmi particolarmente robusti ad

errori di stima della riflettanza sub-superficiale. Analoghe considerazioni possono essere effettuate per quanto riguarda la stima della sostanza gialla ovvero di  $Y$ . Come è possibile notare dalla Figura 4.2-3, il comportamento della riflettanza al variare di  $Y$ , per valori fissati di  $C$  ed  $X$ , è analogo a quello relativo alle variazioni di  $C$ . La stima di  $Y$  quindi, può essere effettuata ricorrendo ad un algoritmo a rapporto di bande in cui i canali spettrali usati sono scelti in base ai criteri usati per l'algoritmo di  $C$ . Comportamento del tutto differente si ha quando consideriamo acque in cui dominano le variazioni di  $X$ . Le curve di riflettanza spettrale al crescere di  $X$  traslano verso l'alto come è evidente dall'osservazione della Figura 4.2-2. L'algoritmo più adatto a tale situazione è quindi di tipo mono-banda in cui il canale spettrale usato è posizionato nella regione del verde, intorno a 550 nm, in cui l'effetto degli altri parametri otticamente attivi è minimo.

Gli algoritmi a rapporto di bande e mono-banda hanno il grosso vantaggio di essere estremamente semplici e di possedere una maggiore robustezza nei confronti degli errori introdotti dalle procedure di correzione atmosferica. L'uso di tali algoritmi però può non essere adatto per acque di caso II in cui gli OAPs variano all'interno di ampi intervalli ed in maniera tra essi indipendente, al contrario delle acque di caso I in cui le proprietà ottiche sono praticamente fissate dal fitoplancton. Per discriminare gli OAPs in tali condizioni è necessario dotare i sensori di un adeguato numero di canali (molto maggiore del numero di parametri da stimare), in modo che coprano il più possibile lo spettro del visibile, e di un'alta risoluzione spettrale. La progettazione dei sensori di nuova generazione è orientata a dotare gli strumenti di tali requisiti (MOS, MERIS). Per il futuro quindi sarà sempre più sentita la necessità di mettere a punto algoritmi che sfruttino l'intera informazione contenuta nei dati multispettrali in modo da rendere più facile la discriminazione dei singoli OAPs. I prossimi sotto paragrafi saranno dedicati allo studio di tale approccio.

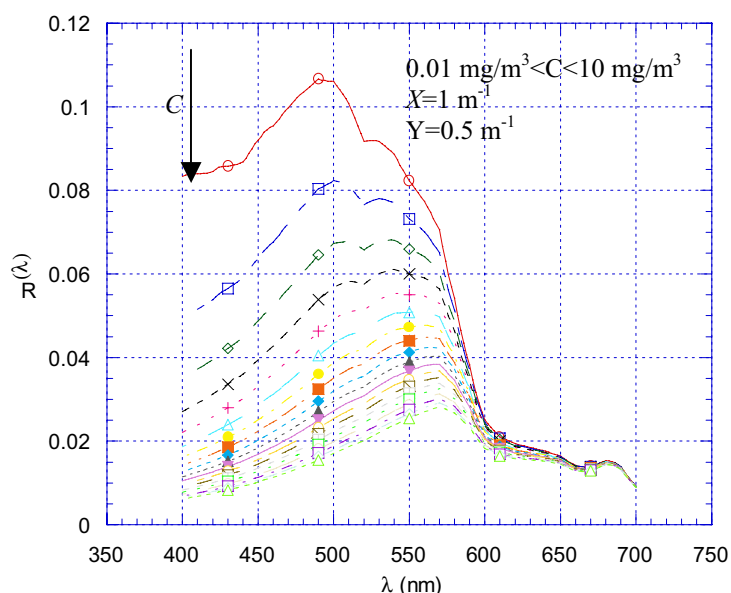


Figura 4.2-1: curve di riflettanza per acque in cui dominano le variazioni di  $C$ .

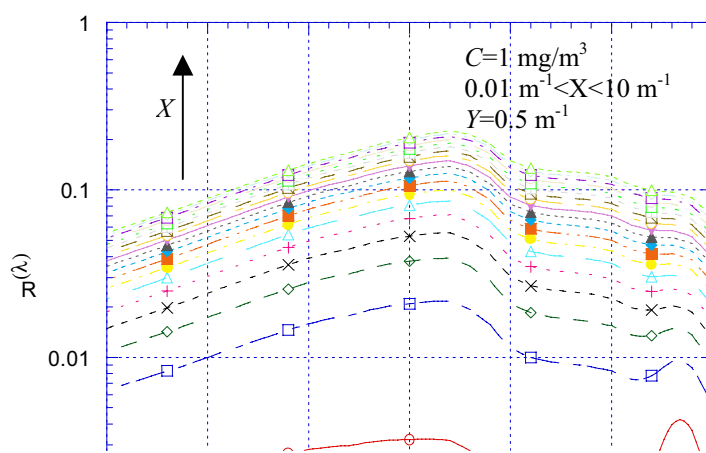


Figura 4.2-2: curve di riflettanza per acque in cui sono dominanti le variazioni di  $X$ .

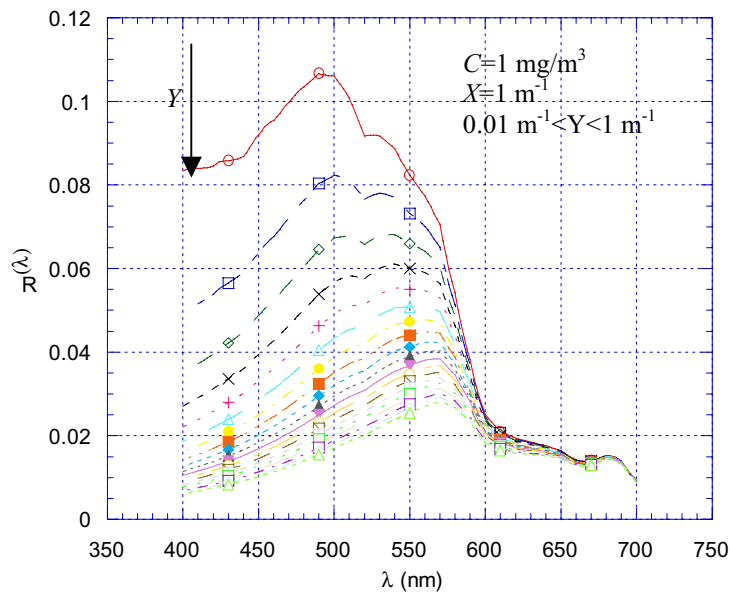


Figura 4.2-3: curve di riflettanza per acque in cui dominano le variazioni di  $Y$ .

#### 4.2.2 Algoritmi multibanda.

L'Eq. 4.2-1 relativa all'algoritmo a rapporto di bande descritto nel precedente sotto paragrafo può assumere una forma più generale includendo le riflettanze nei diversi canali del sensore:

$$OAP = A_0 \cdot R_1^{a_1} \cdot R_2^{a_2} \cdot \dots \cdot R_{N_c}^{a_{N_c}} = A_0 \cdot \prod_{i=1}^{N_c} R_i^{a_i} . \quad Eq. 4.2-18$$

Il modello inverso multibanda dell'Eq. 4.2-18 può essere ricondotto ad un modello di regressione multi-lineare operando una trasformazione logaritmica:

$$Log(OAP) = a_0 + a_1 \cdot Log(R_1) + \dots + a_{N_c} \cdot Log(R_{N_c}) = a_0 + \sum_{i=1}^{N_c} Log(R_i) . \quad Eq. 4.2-19$$

Per tale modello valgono tutte le considerazioni e le espressioni introdotte nel paragrafo relativo agli algoritmi a rapporto e monobanda. In particolare resta valida l'Eq. 4.2-4 del modello di regressione lineare in forma matriciale a patto di modificare la matrice dei regressori ed il vettore dei parametri incogniti come segue:

$$\Gamma = \begin{bmatrix} 1 & \text{Log}(R_{11}) & \text{Log}(R_{12}) & \dots & \text{Log}(R_{1N_c}) \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \text{Log}(R_{N1}) & \text{Log}(R_{N2}) & \dots & \text{Log}(R_{N,N_c}) \end{bmatrix}, \quad \underline{a} = \begin{bmatrix} a_0 \\ \dots \\ \dots \\ a_{N_c} \end{bmatrix}, \quad \text{Eq. 4.2-20}$$

inserendo cioè nella matrice  $\Gamma$  le colonne corrispondenti alle riflettanze nei diversi canali del sensore ed aggiungendo i relativi coefficienti al vettore  $\underline{a}$ . Si ottiene in tal modo una matrice  $\Gamma$  dei regressori di dimensioni  $N \times (N_c+1)$  ed un vettore dei parametri incogniti a  $N_c+1$  dimensioni. La stima a minimo errore quadratico dei coefficienti si effettua applicando l'Eq. 4.2-7.

L'algoritmo multi-banda sfrutta tutta l'informazione contenuta nei dati di riflettanza dei diversi canali del sensore permettendo, nei casi in cui sia possibile applicare un modello multilineare, di stimare con maggior precisione il valore dell'OAP. Il difetto di tali algoritmi è quello di non considerare la correlazione tra le riflettanze nei vari canali, è possibile cioè usare informazioni che non sono necessarie alla stima dell'OAP. La correlazione tra i canali ha l'effetto di produrre una stima dei coefficienti  $a_i$  con elevata varianza e di esaltare le imprecisioni presenti nella matrice  $\Gamma$  quando questa è costruita a partire da dati reali rumorosi [19] con ripercussioni negative sull'errore di stima del valore dell'OAP.

La dipendenza lineare tra colonne della matrice  $\Gamma$  fa sì che la matrice  $\Gamma^T \Gamma$  risulti singolare cosa che rende il problema mal condizionato. La stima dei coefficienti  $\underline{a}$  secondo l'Eq. 4.2-7 può quindi non essere unica. La singolarità della matrice  $\Gamma$  può essere definita in termini di esistenza di un vettore  $\underline{c}$  unitario, cioè tale che  $\underline{c}^T \underline{c} = 1$ , per cui il vettore ottenuto effettuando il prodotto riga per colonna della matrice  $\Gamma$  per il vettore  $\underline{c}$  sia uguale al vettore nullo:

$$\Gamma \cdot \underline{c} = \underline{0}, \quad \text{Eq. 4.2-21}$$

od in maniera equivalente che la norma di tale vettore risulti pari a zero:

$$\|\Gamma \cdot \underline{c}\|^2 = \underline{c}^T \cdot (\Gamma^T \cdot \Gamma) \cdot \underline{c} = 0. \quad \text{Eq. 4.2-22}$$

In pratica la condizione di singolarità non è mai raggiunta, ma possono esistere casi in cui a tale condizione ci si possa avvicinare ovvero per cui la norma di  $\Gamma \underline{c}$  assuma un valore  $\delta$  piccolo:

$$\|\Gamma \cdot \underline{c}\|^2 = \underline{c}^T \cdot (\Gamma^T \cdot \Gamma) \cdot \underline{c} = \delta. \quad \text{Eq. 4.2-23}$$

Tale condizione di quasi-singolarità può provocare l'aumento della varianza della stima del vettore  $\underline{a}$  al tendere di  $\delta$  a zero ovvero man mano che la condizione stessa tende alla singolarità. Dalla disuguaglianza di Cauchy-Schwartz infatti si ha che

$$\begin{aligned} 1 &= \underline{c}^T \cdot \underline{c} = \underline{c}^T \cdot (\Gamma^T \cdot \Gamma)^{1/2} \cdot (\Gamma^T \cdot \Gamma)^{-1/2} \cdot \underline{c} \\ &\leq \underline{c}^T \cdot (\Gamma^T \cdot \Gamma) \cdot (\Gamma^T \cdot \Gamma)^{-1} \cdot \underline{c} = \delta \cdot \underline{c}^T \cdot (\Gamma^T \cdot \Gamma)^{-1} \cdot \underline{c} \end{aligned} \quad \text{Eq. 4.2-24}$$

Da quest'ultima relazione e dalla Eq. 4.2-12 segue:

$$\text{var}(\underline{c}^T \cdot \hat{\underline{a}}) = \sigma^2 \cdot \underline{c}^T \cdot (\Gamma^T \cdot \Gamma)^{-1} \cdot \underline{c} \geq \frac{\sigma^2}{\delta}, \quad \text{Eq. 4.2-25}$$

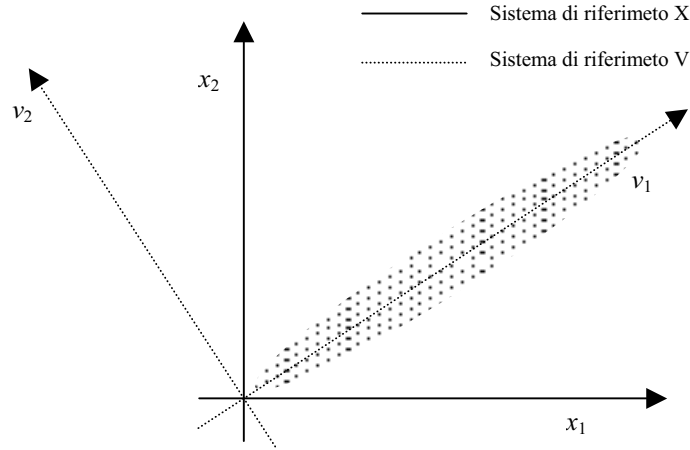
in cui il rapporto  $\sigma^2/\delta$  è grande quando  $\delta$  è piccolo nei confronti di  $\sigma^2$ , il che può comportare per alcuni degli  $a_i$  l'avere un elevato valore di varianza.



Dalle precedenti considerazioni nasce quindi la necessità di dotarsi di uno strumento matematico che permetta la rivelazione dei casi di quasi-singularità e quindi la possibilità di usare l'informazione effettivamente necessaria per stimare gli OAPs. Il prossimo sotto paragrafo sarà dedicato quindi alla trattazione matematica degli algoritmi di regressione lineare che fanno uso dell'analisi delle componenti principali strumento questo che si rivela di grande utilità nel ridurre gli effetti della dipendenza lineare tra le colonne della matrice dei regressori  $\Gamma$ .

#### 4.2.2.1 Inversione tramite l'analisi delle componenti principali (PCA) dei dati di riflettanza.

L'analisi delle componenti principali permette di analizzare in dettaglio la struttura della varianza totale spiegata da un *set* di dati multivariabile. Essa si basa sulla proiezione degli  $N$  vettori delle riflettanze nei singoli canali del sensore su un nuovo sistema di riferimento definito da una base ortonormale di vettori opportunamente scelta. In tale sistema di riferimento risulta possibile stabilire qual è la percentuale di varianza totale che ciascuna variabile indipendente spiega permettendo di considerare, ai fini dell'inversione del modello diretto di riflettanza, solo quelle variabili che contribuiscono alla maggior parte del contenuto informativo del *data set* di partenza. La PCA ha una facile ed intuitiva interpretazione



geometrica se si fa riferimento alla Figura 4.2-4 in cui è rappresentato un generico *data set* bidimensionale. I dati rispetto al sistema di riferimento X sono rappresentati dalla coppia di variabili  $x_1, x_2$ . Nel sistema di riferimento V l'intera variabilità del *data set* può essere catturata dalla sola coordinata  $v_1$  posizionata lungo la direzione principale del *data set* stesso trascurando quindi la variabilità in direzione ortogonale, attribuibile ad esempio ad errori di misura, e giungendo in tal modo ad una riduzione della dimensionalità dei dati.

Figura 4.2-4: interpretazione geometrica dell'analisi delle componenti principali di un *data set* multivariabile.

La PCA sarà utilizzata per definire un algoritmo di stima degli OAPs che faccia uso dell'intera informazione fornita dai dati di riflettanza del sensore nei diversi canali e che eviti allo stesso tempo gli inconvenienti che possono presentarsi a causa della dipendenza lineare tra le colonne della matrice dei regressori  $\Gamma$ . Si consideri a tale scopo il cosiddetto modello di regressione multi-lineare centrato:

$$\underline{\psi}_{OAP-0} = \Gamma_0 \cdot \underline{a}_0 + \underline{\varepsilon}_0, \quad \text{Eq. 4.2-26}$$

in cui

$$\underline{\psi}_{OAP-0} = \underline{\psi}_{OAP} - \underline{1} \cdot \overline{\psi}_{OAP} , \quad \text{Eq. 4.2-27}$$

è il vettore colonna della variabile dipendente (del logaritmo dell'OAP) privato del valore medio aritmetico  $\overline{\psi}_{OAP}$  pari a:

$$\overline{\psi}_{OAP} = \frac{1}{N} \cdot \sum_{k=1}^N \psi_{OAP,k} , \quad \text{Eq. 4.2-28}$$

e

$$\Gamma_0 = \Gamma - \overline{\Gamma} , \quad \text{Eq. 4.2-29}$$

$$\overline{\Gamma} = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \dots & \Gamma_{N_c} \\ \Gamma_1 & \Gamma_2 & \dots & \Gamma_{N_c} \\ \dots & \dots & \dots & \dots \\ \Gamma_1 & \Gamma_2 & \dots & \Gamma_{N_c} \end{bmatrix}, \quad \Gamma_i = \frac{1}{N} \cdot \sum_{k=1}^N \Gamma_{k,i} , \quad i=1, \dots, N_c, \quad \text{Eq. 4.2-30}$$

è la matrice di dimensioni  $N \times N_c$  le cui colonne sono i logaritmi delle riflettanze nei singoli canali del sensore private del corrispondente valor medio aritmetico  $\Gamma_i$ .

Sia  $V$  la matrice ortogonale (cioè  $V^T V = I$ ) di dimensioni  $N_c \times N_c$  tale che:

$$V^T \cdot (\Gamma_0^T \cdot \Gamma_0) \cdot V = \Lambda , \quad \text{Eq. 4.2-31}$$

con  $\Lambda$  matrice diagonale:

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{N_c}), \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N_c} . \quad \text{Eq. 4.2-32}$$

I  $\lambda_i$  sono gli autovalori della matrice  $\Gamma_0^T \cdot \Gamma_0$  e le colonne della matrice  $V$  sono i relativi autovettori che costituiscono la base ortonormale su cui sarà proiettato il *data set* di riflettanze. Data l'ortogonalità della matrice  $V$  il modello di regressione lineare centrato può essere espresso nella seguente forma:

$$\underline{\psi}_{OAP-0} = \Gamma_0 \cdot \underline{a}_0 + \underline{\varepsilon}_0 = \Gamma_0 \cdot V \cdot V^T \cdot \underline{a}_0 + \underline{\varepsilon}_0 = U \cdot \underline{\eta} + \underline{\varepsilon}_0 , \quad \text{Eq. 4.2-33}$$

in cui si è indicato con  $U$  la matrice di dimensioni  $N \times N_c$ , pari a

$$U = \Gamma_0 \cdot V , \quad \text{Eq. 4.2-34}$$

detta delle componenti principali e con  $\underline{\eta}$

$$\underline{\eta} = V^T \cdot \underline{a}_0 , \quad \text{Eq. 4.2-35}$$

il vettore dei nuovi coefficienti del modello di regressione. Le righe della matrice  $U$  rappresentano i vettori di dimensione  $N_c$  dei logaritmi delle riflettanze spettrali proiettati sul sistema di riferimento degli autovettori. Le  $N_c$  colonne della matrice  $U$  sono le componenti principali e rappresentano le nuove variabili indipendenti del modello di regressione multi-lineare. La stima a minimo errore quadratico del vettore dei coefficienti  $\underline{\eta}$  secondo l'Eq. 4.2-7 è pari a:

$$\begin{aligned} \hat{\underline{\eta}} &= (U^T \cdot U)^{-1} \cdot U^T \cdot \underline{\psi}_{OAP-0} = (V^T \cdot \Gamma_0^T \cdot \Gamma_0 \cdot V)^{-1} \cdot U^T \cdot \underline{\psi}_{OAP-0} = \\ &= \Lambda^{-1} \cdot U^T \cdot \underline{\psi}_{OAP-0} \end{aligned} , \quad \text{Eq. 4.2-36}$$

e quindi la covarianza della stima dei coefficienti risulta uguale a:

$$\text{cov}(\hat{\underline{\eta}}) = \sigma^2 \cdot \Lambda^{-1}. \quad \text{Eq. 4.2-37}$$

La varianza dei singoli coefficienti quindi è pari a:

$$\text{var}(\hat{\eta}_i) = \frac{\sigma^2}{\lambda_i}, i=1, \dots, N_c, \quad \text{Eq. 4.2-38}$$

che è tanto più grande quanto più piccolo è l'autovalore corrispondente rispetto alla varianza  $\sigma^2$  dell'errore  $\underline{\varepsilon}_0$ . Gli autovalori piccoli rispetto alla varianza dell'errore segnalano di solito la presenza di dipendenza lineare tra le colonne della matrice  $\Gamma_0$  del modello di regressione multi-lineare. Si dimostra che l'autovalore  $i$ -esimo rappresenta la frazione di varianza spiegata dalla corrispondente componente principale ovvero dalla colonna  $i$ -esima della matrice  $U$ . Sulla base degli autovalori quindi è possibile selezionare le componenti principali che contribuiscono alla maggior parte dell'energia totale del *data set* di riflettanze e considerare solo i coefficienti  $\eta_i$  con la più bassa varianza, riducendo in tal modo gli effetti della presenza di dipendenza lineare tra le colonne della matrice dei regressori  $\Gamma_0$ . Si noti che gli autovalori nella matrice  $\Lambda$  sono ordinati in ordine decrescente e quindi le corrispondenti componenti principali spiegano frazioni decrescenti di varianza totale. Facendo quindi riferimento agli autovalori si possono selezionare le prime  $r$  componenti principali (corrispondenti alle prime  $r$  colonne della matrice  $U$ ) che spiegano in totale una frazione fissata (di solito si fa riferimento al 90%, 95% o al 99%) dell'intera varianza spiegata dal *data set*. Il modello di regressione nell'Eq. 4.2-33 diventa quindi:

$$\underline{\psi}_{OAP-0} = U_r \cdot \underline{\eta}_r + \underline{\varepsilon}', \quad \text{Eq. 4.2-39}$$

in cui con  $U_r$  si indica la matrice costituita dalle prime  $r$  colonne della matrice delle componenti principali  $U$  e con  $\underline{\eta}_r$  il vettore dei coefficienti costituito dalle prime  $r$  componenti del vettore  $\underline{\eta}$ . I valori predetti del logaritmo dell'OAP si ottengono moltiplicando la matrice  $U_r$  per il vettore stimato dei coefficienti  $\hat{\underline{\eta}}_r$  ottenuto considerando le prime  $r$  componenti del vettore  $\hat{\underline{\eta}}$  dato dall'Eq. 4.2-36:

$$\hat{\underline{\psi}}_{OAP-0} = U_r \cdot \hat{\underline{\eta}}_r. \quad \text{Eq. 4.2-40}$$

L'Eq. 4.2-39 costituisce quindi la forma finale dell'algoritmo multi-lineare che lega il valore del logaritmo dell'OAP considerato alle prime  $r$  componenti principali ottenute effettuando la PCA sulla matrice  $\Gamma_0$  dei logaritmi delle riflettanze. In fase di definizione di tale algoritmo si dovrà effettuare in primo luogo la stima dei valori medi del logaritmo dell'OAP considerato e del logaritmo delle riflettanze e centrare con tali valori il modello di regressione lineare ottenendo l'Eq. 4.2-26. Si procede successivamente ad effettuare la PCA sulla matrice  $\Gamma_0$  ed a stimare il vettore dei coefficienti  $\underline{\eta}$  tramite l'Eq. 4.2-36.

In fase di applicazione dell'algoritmo a dati telerilevati si procederà a costruire la matrice  $\Gamma_0$  e ad applicare alla stessa la PCA. Si effettua successivamente l'analisi degli autovalori (ad esempio graficandone la somma cumulata rapportata alla somma totale) per individuare le componenti principali che spiegano una frazione fissata della varianza totale. Si stima quindi il valore del logaritmo dell'OAP tramite l'Eq. 4.2-40 a cui si somma il valore medio del logaritmo dell'OAP valutato nella fase di definizione.

L'applicazione di tali algoritmi per la stima degli OAPs è ovviamente possibile solo nei casi in cui sia lecito assumere come modello inverso un modello che lega linearmente il logaritmo degli OAPs al

logaritmo delle singole riflettanze spettrali. Nel caso in cui siano presenti forti non linearità si dovrà ricorrere a metodi differenti, più generali, ma che presentano una maggiore complessità.

#### 4.2.3 L'uso di reti neurali RBF per la soluzione del problema inverso.

I metodi di inversione analizzati precedentemente presuppongono che il modello inverso legghi il valore del logaritmo dell'OAP ad una combinazione lineare dei logaritmi delle riflettanze nei differenti canali del sensore a cui si somma un termine costante  $a_0$  (si noti che anche l'algoritmo a rapporto di bande può essere ricondotto ad una tale forma esprimendo il logaritmo del rapporto come differenza di logaritmi; i coefficienti della combinazione lineare tra i logaritmi delle riflettanze nelle due bande scelte saranno in tal modo uguali in valore assoluto, ma opposti in segno). Nei casi in cui non sia possibile fare tale assunzione, si dovrà ricorrere ad un modello più flessibile che permetta cioè di ricostruire la relazione tra il logaritmo degli OAPs e quello delle riflettanze qualunque forma essa assuma inclusa quella multilineare. Tale modello in generale dovrà mappare un vettore  $\underline{X}^{N_c}$  dimensionale appartenente ad un sottospazio vettoriale di  $\mathfrak{R}^{N_c}$  in un valore  $y$  appartenente ad un sottospazio di  $\mathfrak{R}$  qualunque sia la relazione  $y = F(\underline{X})$  che lega i due sottospazi. Tale problema può essere risolto efficacemente ricorrendo all'utilizzo di reti neurali di tipo RBF (RBF-NN: *Radial Basis Function-Neural Network*).

In Figura 4.2-5 è riportato lo schema di principio di una RBF-NN.

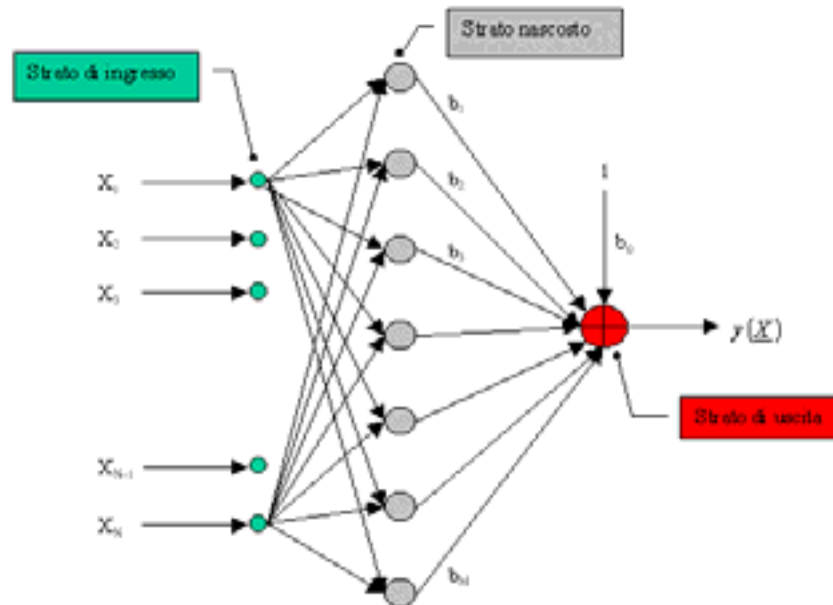


Figura 4.2-5: schema di principio di una RBF-NN a tre strati.

La rete è costituita da elementi detti neuroni (Figura 4.2-6) organizzati in strati. Nel presente lavoro di tesi si farà riferimento a reti a tre strati. Il primo strato o strato di ingresso è costituito da  $N_c$  neuroni. Questi hanno la funzione di fornire le  $N_c$  componenti  $x_k$  del vettore di ingresso  $\underline{X}$  ad ogni neurone del secondo strato detto strato nascosto. I neuroni dello strato nascosto, in numero pari ad  $M$ , effettuano operazioni non lineari sul vettore di ingresso  $\underline{X}$ . Lo strato di uscita ha il compito di combinare linearmente le uscite dei neuroni dello strato nascosto secondo i coefficienti  $b_i$  e di sommare un valore costante  $b_0$  per costituire l'uscita  $y$ . Ogni neurone dello strato di ingresso è collegato ad ogni neurone dello strato nascosto così come tutti i neuroni dello strato nascosto sono collegati al neurone dello strato di uscita.

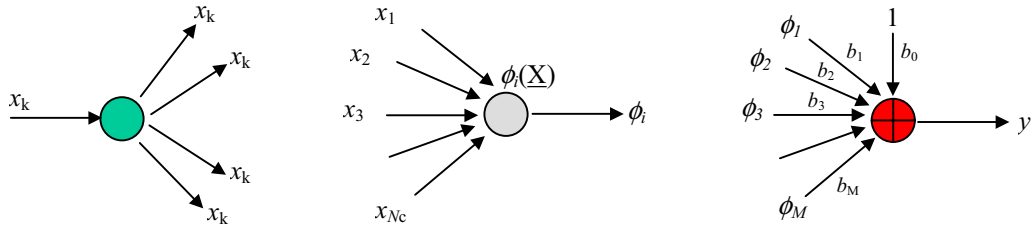


Figura 4.2-6: operazioni effettuate dai neuroni di ingresso, dello strato nascosto e dello strato di uscita.

Se  $\phi_i(\underline{X})$  è la relazione non lineare che lega gli ingressi e l'uscita dell' $i$ -esimo neurone dello strato nascosto allora la relazione ingresso uscita complessiva della RBF-NN può essere espressa come segue:

$$y(\underline{X}) = b_0 + \sum_{i=1}^M b_i \cdot \phi_i(\underline{X}), \quad \text{Eq. 4.2-41}$$

ovvero come sviluppo in serie di funzioni di base  $\phi_i(\underline{X})$  con coefficienti  $b_i$ . In una RBF-NN le funzioni di base sono a simmetria radiale nello spazio dei vettori di ingresso  $\underline{X}$ . Ogni funzione di base è centrata in un punto  $\underline{c}_i$  dello spazio di ingresso ed è caratterizzata da un parametro di forma che ne regola la dispersione intorno al centro. Se la forma delle funzioni di base è la stessa per ogni neurone  $i$  dello strato nascosto allora la Eq. 4.2-41 diventa:

$$y(\underline{X}) = b_0 + \sum_{i=1}^M b_i \cdot \phi(\|\underline{X} - \underline{c}_i\|), \quad \text{Eq. 4.2-42}$$

in cui con  $\|\underline{X} - \underline{c}_i\|$  si è indicata la distanza tra il punto  $\underline{X}$  ed il centro  $\underline{c}_i$  dell' $i$ -esimo neurone dello strato nascosto. La scelta della forma della funzione radiale di base non è determinante ai fini delle prestazioni della rete in termini di errore di ricostruzione della relazione tra spazio di ingresso e spazio di uscita [Aiken]. Un esempio di funzione di base radiale largamente usata nelle applicazioni [20] e a cui si farà riferimento nel proseguo del lavoro è la funzione gaussiana (Figura 4.2-7):

$$\phi(r) = \exp\left(-\frac{r^2}{\beta^2}\right), \quad \text{Eq. 4.2-43}$$

in cui  $r = \|\underline{X} - \underline{c}_i\|$  e  $\beta$  è il parametro che regola la dispersione della funzione intorno al centro.

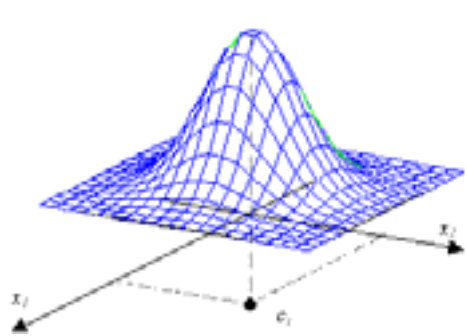


Figura 4.2-7: esempio di funzione radiale di base gaussiana a due dimensioni centrata in  $\underline{c}_i$ .

La fase di valutazione dei parametri della rete ovvero del numero  $M$  di neuroni dello strato nascosto, dei centri  $\underline{c}_i$ , del parametro  $\beta$  e dei pesi  $b_i$ , è detta di addestramento e consiste nel presentare di volta in volta alla rete stessa una serie di esempi costituiti da coppie di ingresso-uscita appartenenti ad un insieme detto *training set*:

$$T \equiv \{\underline{X}_j, y_j\}_{j=1}^{N_t} . \quad \text{Eq. 4.2-44}$$

La stima dei parametri viene effettuata minimizzando l'errore quadratico medio tra le uscite della rete e i le uscite  $y_j$  del *training set*. Fissato un valore iniziale dei parametri della rete, per ogni  $\underline{X} \in T$  si calcola il corrispondente valore  $\hat{y}_j$  di uscita tramite l'Eq. 4.2-42. Si valuta successivamente l'errore quadratico medio tra i valori di uscita  $y_j \in T$  ed i valori predetti  $\hat{y}_j$ :

$$\xi = \frac{1}{N_t} \cdot \sum_{j=1}^{N_t} (y_j - \hat{y}_j)^2 \quad \text{Eq. 4.2-45}$$

e si aggiornano i parametri della rete in maniera opportuna. Si ripete la procedura ripresentando alla rete gli esempi del *training set* sino a quando l'errore non risulta minore di un fissato valore. Tale metodo richiede l'uso di procedure di minimizzazione dell'errore di tipo non lineare di solito basate sulle derivate parziali dell'errore quadratico medio rispetto ai parametri della rete. Tali procedure hanno l'inconveniente, oltre che di introdurre una maggiore complessità di calcolo nella fase di identificazione del modello inverso, di bloccarsi su eventuali minimi relativi dell'errore quadratico medio. Supponendo però il numero di neuroni  $M$  dello strato nascosto, i centri  $\underline{c}_i$  ed il parametro di forma  $\beta$  fissati, l'Eq. 4.2-42 può essere ricondotta ad un modello di regressione lineare:

$$y_j = b_0 + \sum_{i=1}^M b_i \cdot f_i(\underline{X}_j) + \varepsilon_j , \quad \text{Eq. 4.2-46}$$

in cui

$$f_i(\underline{X}) = \phi(\|\underline{X} - \underline{c}_i\|) \quad \text{Eq. 4.2-47}$$

dipende solo da  $\underline{X}$ . Valutando il secondo membro dell'Eq. 4.2-46 in corrispondenza di ogni elemento  $\underline{X}_j$  appartenente all'insieme  $T$  è possibile ottenere un sistema di equazione che in forma matriciale è esprimibile come:

$$\underline{Y} = F \cdot \underline{b} + \underline{\varepsilon} , \quad \text{Eq. 4.2-48}$$

in cui

$$F = \begin{bmatrix} f_{1,1} & \dots & \dots & f_{1,M} \\ f_{2,1} & \dots & \dots & f_{2,M} \\ \vdots & & & \vdots \\ f_{N_t,1} & \dots & \dots & f_{N_t,M} \end{bmatrix} , \quad \text{Eq. 4.2-49}$$

è la matrice  $N_t \times M$  dei regressori i cui elementi sono pari a:

$$f_{j,i} = \phi(\|\underline{X}_j - \underline{c}_i\|) \quad j = 1 \dots N_t, i = 1 \dots M , \quad \text{Eq. 4.2-50}$$

$\underline{Y}$  il vettore colonna dei valori di uscita  $y_j \in T$ ,  $\underline{b} = [b_1, b_2, \dots, b_M]^T$  il vettore incognito dei coefficienti e  $\underline{\varepsilon} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{N_t}]^T$  il vettore dell'errore tra i valori predetti ed i valori  $y_j$  del *training set*. Dall'Eq. 4.2-7 discende quindi la stima ai minimi quadrati del vettore  $\underline{b}$  dei coefficienti incogniti:

$$\hat{\underline{b}} = (F^T \cdot F)^{-1} \cdot F^T \cdot \underline{Y} \quad \text{Eq. 4.2-51}$$

Per tale modello di regressione valgono tutte le relazioni dall'Eq. 4.2-7 all'Eq. 4.2-16, introdotte nel sotto paragrafo 4.2.1, a patto di sostituire la matrice  $\Gamma$  con la matrice  $F$ . Definendo la matrice di covarianza  $A = F^T \cdot F$  l'Eq. 4.2-51 diventa:

$$\hat{\underline{b}} = A^{-1} \cdot F^T \cdot \underline{Y}, \quad \text{Eq. 4.2-52}$$

e la Eq. 4.2-12 relativa alla covarianza del vettore  $\underline{b}$ :

$$\text{cov}(\hat{\underline{b}}) = \sigma^2 \cdot A^{-1}. \quad \text{Eq. 4.2-53}$$

Definendo inoltre la matrice di proiezione come  $P = I - F \cdot A^{-1} \cdot F^T$  è possibile esprimere la somma dei residui quadratici tra  $y_j \in T$  ed i valori predetti  $\hat{y}_j$  come:

$$S = \underline{Y}^T \cdot P^2 \cdot \underline{Y}. \quad \text{Eq. 4.2-54}$$

La definizione di matrice di proiezione  $P$  e l'Eq. 4.2-54 saranno utili nel prossimo sotto paragrafo relativo alla selezione dei centri della RBF-NN tramite procedure di *forward selection*.

#### 4.2.3.1 La procedura di *forward selection* per la scelta dei nodi dello strato nascosto.

La scelta dei centri  $\underline{c}_i$ , fissato il parametro di forma  $\beta$ , è una fase cruciale affinché la rete abbia elevate prestazioni in termini di errore di stima. Nella pratica i centri sono scelti direttamente dal *training set*  $T$ . Il problema è di come scegliere nella maniera più appropriata possibile i centri dal *training set*. La scelta di usare tutti i punti dell'insieme  $T$  non è praticabile nei casi in cui il *data set* contenga un elevato numero di elementi. In tale caso, infatti, la dimensione dello strato nascosto sarebbe eccessiva ed inoltre il problema della stima dei coefficienti  $\underline{b}$  risulterebbe mal condizionato per effetto della dipendenza lineare tra le colonne della matrice dei regressori  $F$  introdotta da funzioni di base con centri troppo vicini tra loro. Per risolvere il problema della scelta dei centri sono stati ideati metodi cosiddetti di *model selection* che consistono nel testare vari modelli di regressione costruiti con differenti combinazioni delle variabili indipendenti ovvero, nel caso di RBF-NN, con differenti sottoinsiemi di funzioni di base scelte da un insieme fissato di possibili funzioni di base. Si sceglie quindi il modello migliore secondo un fissato criterio di ottimalità. Risulta chiaro che una procedura del genere è proibitiva dal punto di vista computazionale nel caso in cui si preveda di testare in maniera esaustiva tutti i possibili modelli di regressione. Esistono infatti  $2^L - 1$  possibili sottoinsiemi in un insieme di  $L$  funzioni di base e quindi altrettanti modelli di regressione da esaminare. Per evitare l'uso del metodo esaustivo è possibile ricorrere ad una procedura di scelta del modello detta di *forward selection* [21]. La procedura parte da un modello vuoto a cui si aggiunge di volta in volta la funzione di base, prelevata dal set di  $L$  funzioni candidate, che massimizza il decremento della somma dei residui quadratici (Eq. 4.2-54). L'accrescimento del modello procede fino a quando non si raggiunge il minimo di un fissato criterio di ottimo. La procedura pur essendo non lineare ha il vantaggio di non richiedere la scelta a priori del numero  $M$  di centri del modello e di essere poco dispendiosa dal punto di vista computazionale.

L'inserimento di una nuova funzione di base nel modello si traduce nell'aggiunta di una colonna alla matrice dei regressori  $F$  proveniente dalla matrice dei regressori  $F^*$  costruita considerando tutte le  $L$  funzioni di base candidate. Consideriamo tale matrice partizionata in vettori colonna di dimensione  $N_j$ :  $F^* = [\underline{f}_1, \underline{f}_2, \dots, \underline{f}_L]$  in cui le componenti dei vettori sono pari a  $f_{j,i} = \phi(\|\underline{X}_j - \underline{c}_i\|)$  con  $i=1 \dots L$  e  $j=1 \dots N_j$ . L'insieme di vettori  $\underline{c}_i$  può essere costituito da tutti i vettori  $\underline{X} \in T$  o da vettori  $\underline{X}$  scelti in

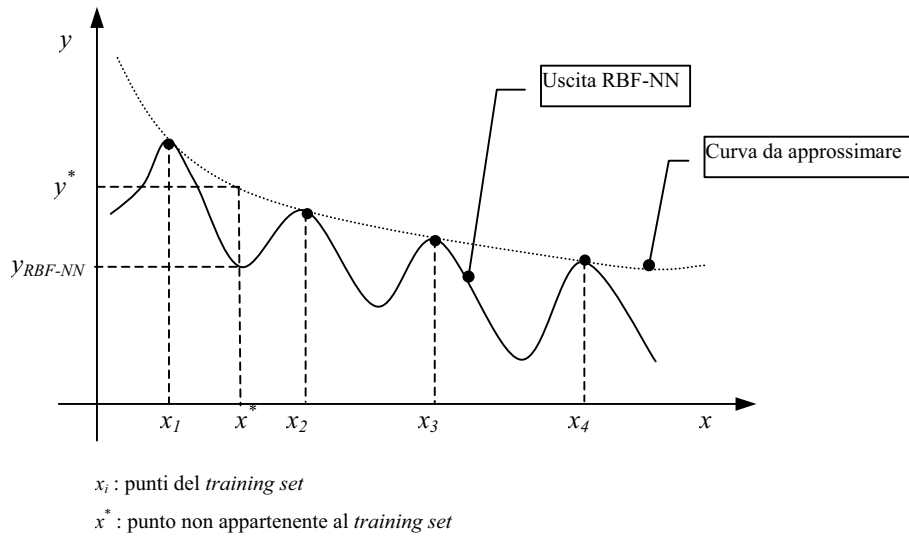
maniera casuale dal *training set*; nel primo caso sarà  $L=N_t$ . Sia  $F_m$  la matrice dei regressori del modello al passo  $m$  (notare che  $m$  indica anche il numero di colonne della matrice) e  $F_{m+1}$  la stessa matrice a cui è stata aggiunta una colonna  $\underline{f}_i$  della matrice  $F^*$ ; è possibile dimostrare [21] che la matrice di proiezione  $P_{m+1}$  del nuovo modello può essere espressa in funzione della  $P_m$  come segue:

$$P_{m+1} = P_m - \frac{P_m \cdot \underline{f}_i \cdot \underline{f}_i^T \cdot P_m}{\underline{f}_i^T \cdot P_m \cdot \underline{f}_i} \quad \text{Eq. 4.2-55}$$

e che il decremento della somma dei residui quadratici ottenuto aggiungendo la nuova funzione di base risulta essere:

$$S_m - S_{m+1} = \frac{(\underline{y}^T \cdot P_m \cdot \underline{f}_i)^2}{\underline{f}_i^T \cdot P_m \cdot \underline{f}_i}. \quad \text{Eq. 4.2-56}$$

Sull'Eq. 4.2-56 è basata la scelta della nuova funzione di base da includere nel modello. Ad ogni passo si valuterà la differenza  $S_m - S_{m+1}$  per ogni colonna della matrice  $F^*$  e si sceglierà come nuova colonna da includere nella matrice  $F$  quella che massimizza tale differenza. Il criterio di ottimalità scelto per arrestare la procedura di selezione dovrà quantificare le proprietà di generalizzazione della rete addestrata cioè dovrà fornire un valore numerico che rappresenti l'errore di predizione della RBF-NN ovvero l'errore nelle stime effettuate dalla RBF-NN stessa da dati di ingresso non appartenenti al *training set*. Tali criteri si rivelano utili al fine di evitare l'*overfitting* durante la fase di addestramento della rete. L'*overfitting* è quel fenomeno per cui, pur possedendo la rete un basso valore dei residui in corrispondenza di ogni punto



del *training set*, la stima effettuata in corrispondenza di vettori di ingresso  $\underline{X}$  che si discostano da tali punti risulta essere affetta da errore elevato (Figura 4.2-8).

Figura 4.2-8: esempio di overfitting in una rete RBF ad un ingresso ed un'uscita.

Esistono un certo numero di criteri per la selezione dei modelli usati comunemente nella pratica ed esprimibili, nel caso di modelli di regressione lineare, in funzione della matrice di proiezione  $P_m$ ; tra questi ricordiamo [21]:

il *generalised cross validation* (GCV):



$$GCV = \frac{N_t \cdot \underline{Y}^T \cdot P_m^2 \cdot \underline{Y}}{(N_t - m)^2}, \quad \text{Eq. 4.2-57}$$

l'unbiased estimate of variance (UEV):

$$UEV = \frac{N_t \cdot \underline{Y}^T \cdot P_m^2 \cdot \underline{Y}}{N_t - m}, \quad \text{Eq. 4.2-58}$$

il final prediction error (FPE):

$$FPE = \frac{N_t + m}{N_t - m} \cdot \frac{\underline{Y}^T \cdot P_m^2 \cdot \underline{Y}}{N_t}, \quad \text{Eq. 4.2-59}$$

ed il bayesian information criterion (BIC):

$$BIC = \frac{N_t + m \cdot [\ln(N_t) - 1]}{N_t - m} \cdot \frac{\underline{Y}^T \cdot P_m^2 \cdot \underline{Y}}{N_t}. \quad \text{Eq. 4.2-60}$$

Il processo di selezione di nuove funzioni di base procederà fino a quando il criterio scelto cesserà di decrescere segnalando in tal modo la comparsa di *overfitting*.

#### 4.2.3.2 Inversione del modello di riflettanza.

In Figura 4.2-9 è riportato lo schema di una RBF-NN usata per la stima degli OAPs. Il vettore degli ingressi è costituito dai logaritmi delle riflettanze nei singoli canali del sensore mentre l'uscita è costituita dal logaritmo del OAP considerato.

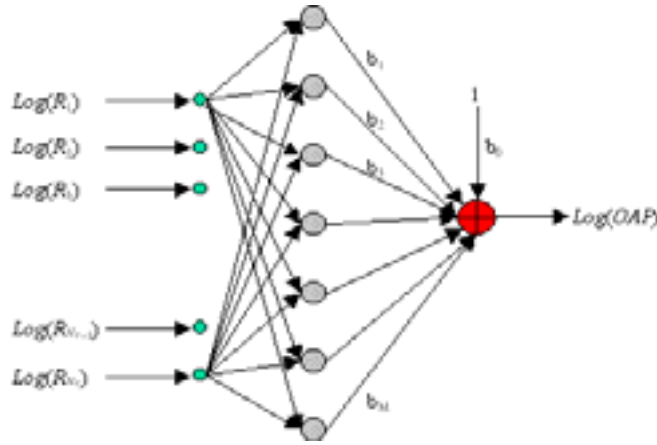


Figura 4.2-9: rete RBF usata per l'inversione del modello di riflettanza.

Il *training set* è quindi definito come segue:

$$T \equiv \{\underline{\Gamma}_j, \text{Log}(\text{OAP}_j)\}_{j=1}^{N_t} \quad \text{Eq. 4.2-61}$$

in cui  $\underline{\Gamma}_j = [\text{Log}(R_{1,j}), \text{Log}(R_{2,j}), \dots, \text{Log}(R_{N_c,j})]$  è il  $j$ -esimo vettore dei logaritmi delle riflettanze negli  $N_c$  canali del sensore. La procedura di addestramento fissato il parametro di forma  $\beta$  si articola nei seguenti passi (Figura 4.2-10):

1. generazione di un *data set* di OAPs e di riflettanze  $D \equiv \{\underline{R}_j, \text{OAP}_j\}_{j=1}^N$  mediante acquisizione di  $N$  triplette  $\text{OAPs}_j = [C, X, Y]_j$  di misurazioni *in-situ* e calcolo di  $N$  vettori  $\underline{R}_j = [R_{1,j}, R_{2,j}, \dots, R_{N_c,j}]$  tramite applicazione del modello diretto di riflettanza,

2. generazione del *data set*  $\Delta \equiv \{\underline{\Gamma}_j, \text{Log}(OAP_j)\}_{j=1}^N$  tramite calcolo dei logaritmi delle riflettanze e degli OAPs,
3. generazione del *training set*  $T$  tramite scelta casuale di  $N_t$  coppie  $\{\underline{\Gamma}, \text{Log}(OAP)\} \in \Delta$ ,
4. calcolo della matrice  $F^*$  dal *training set*  $T$  (si considerano come centri possibili tutti i punti del *training set* quindi  $L=N_t$ ),
5. scelta del criterio di selezione (GCV, UEV, FPE, BIC), applicazione della procedura di *forward selection* e generazione della matrice  $F$  dei regressori,
6. stima ai minimi quadrati dei coefficienti della rete  $\underline{b}$ ,
7. ulteriore verifica delle proprietà di generalizzazione della rete tramite calcolo dell'MSE tra l'uscita della rete in corrispondenza di ogni  $\underline{\Gamma}_j \in \Delta$  ed il  $\text{Log}(OAP_j) \in \Delta$ .

Il parametro  $\beta$  sarà fissato in maniera empirica. Per ottimizzare il valore di tale parametro, come già anticipato, non si farà uso in questo lavoro di tesi di algoritmi di ottimizzazione non lineari basati sul gradiente, rispetto ai parametri incogniti, dell'Eq. 4.2-45. Tramite la procedura di *forward selection* è comunque possibile effettuare una sorta di ottimizzazione del parametro di forma costruendo una matrice  $F^*$  in cui le colonne corrispondono a funzioni di base con differenti valori di  $\beta$  fissati. La procedura di selezione provvederà a scegliere le colonne della matrice  $F^*$  più adatte ad approssimare il legame tra le riflettanze nei diversi canali del sensore e l'OAP.

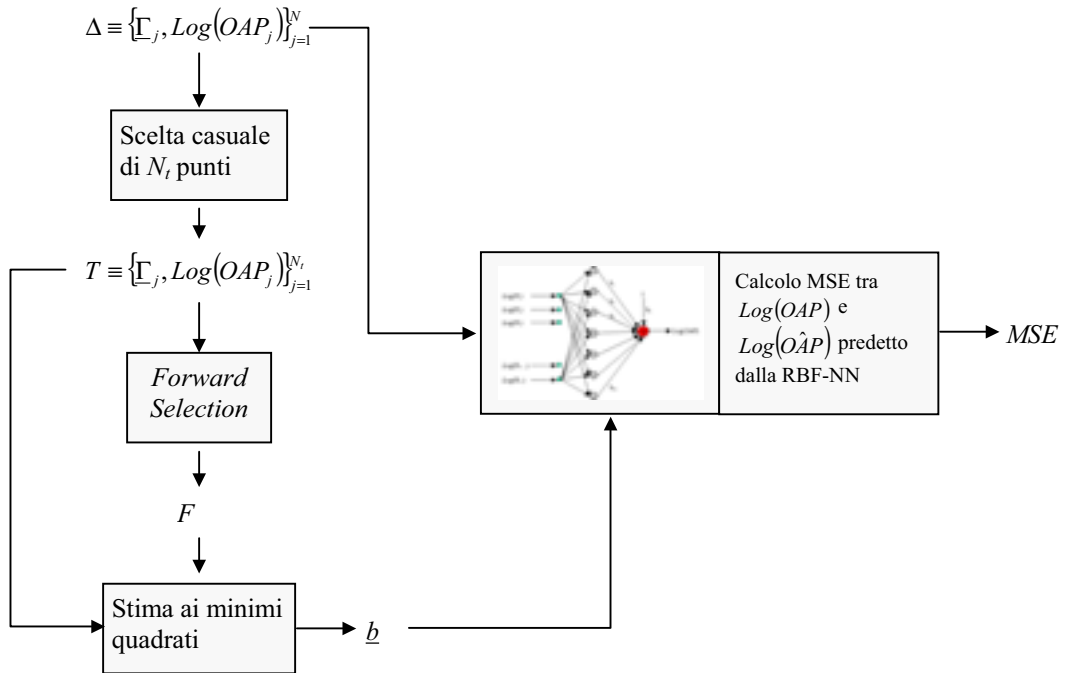


Figura 4.2-10: schema della procedura di addestramento di una RBF-NN per la stima degli OAPs.

---

## 5 Capitolo 4

### 6 *La distribuzione statistica dei parametri otticamente attivi dell'acqua marina*

Nel capitolo 2 è stato definito il modello diretto che permette di calcolare la riflettanza sub-superficiale fissati i tre parametri otticamente attivi dell'acqua marina. Nel capitolo 3 è stato affrontato il problema dell'inversione di tale modello al fine di stimare i tre parametri  $C$ ,  $X$ , ed  $Y$  dalla riflettanza sub-superficiale media ricavata dai dati misurati dai sensori nei vari canali di cui sono dotati. Sono state proposte diverse metodologie di inversione. Quelle più innovative cercano di sfruttare l'intera informazione spettrale fornita dai sensori e richiedono una più larga sperimentazione.

L'elemento che lega le differenti tecniche di inversione, le quali possono tra l'altro rientrare tutte in uno schema comune come già evidenziato nel capitolo 3, è il fatto di poter disporre di un numero statisticamente rilevante di triplette  $[C, X, Y]$  che possa essere usato come *data set* di *training* per la stima dei parametri del modello inverso. Il presente capitolo sarà dedicato quindi alla caratterizzazione statistica dei parametri otticamente attivi dell'acqua marina. In particolare sarà introdotto il modello di densità di probabilità log-normale che trova evidenza sperimentale in natura come messo in risalto da numerosi lavori sull'argomento tra cui quello fondamentale di Janet Campbell [22]. In tale lavoro è stata data una giustificazione teorica che porta ad assumere la distribuzione log-normale come modello valido per descrivere la variabilità statistica dei parametri otticamente attivi e sono state fornite numerose verifiche sperimentali che avallano tale ipotesi. Nel presente capitolo verranno ripresi alcuni concetti riportati nel lavoro di J. Campbell e saranno analizzati alcuni dati provenienti da campagne di misura *in situ* per effettuare alcune verifiche sperimentali. Il fine di tale trattazione preliminare è quello di fornire un modello statistico che permetta, attraverso l'uso di opportuni generatori di variabili casuali, di creare una serie di *data set* di parametri otticamente attivi che rappresenti il più fedelmente possibile le caratteristiche riscontrabili in natura dei parametri stessi. La generazione di tali *data set* simulati per i vari tipi di acque (di caso I e caso II) permetterà di testare le tecniche di inversione proposte in diverse condizioni che possono manifestarsi nei bacini sottoposti a monitoraggio. Le prestazioni dei vari algoritmi di inversione potranno in tal modo essere confrontate al variare delle diverse condizioni considerate e per i differenti sensori presi in esame permettendo la verifica della capacità che questi ultimi hanno di discriminare i tre principali parametri otticamente attivi dell'acqua marina.

Il capitolo è articolato in un primo paragrafo in cui si richiamano alcune proprietà della densità di probabilità log-normale. In tale paragrafo si esprimono le motivazioni teoriche che portano ad assumere la densità log-normale come modello fondamentale di variabilità statistica dei parametri otticamente attivi e si riportano le statistiche di una serie di dati *in situ* di tali parametri per cercare conferme sperimentale alle

ipotesi formulate. Nel secondo paragrafo si descrivono una serie di *data set* simulati di parametri otticamente attivi e di riflettanze medie sulle bande dei sensori rappresentativi dei differenti tipi di acque (caso I, caso II e casi intermedi). Nel terzo ed ultimo paragrafo si riporta la descrizione di un *data set* di parametri otticamente attivi misurati durante una campagna oceanografica condotta nel Mare di Alboran nel Mediterraneo occidentale condotta nell'ambito del progetto OMEGA (*Observing and Modelling of Eddy scale Geostrophic and Ageostrophic circulation*) finanziato dall'Unione Europea con lo scopo di studiare la variabilità fisica e biologica di tale bacino. Tale *data set* sarà utilizzato per sviluppare gli algoritmi di stima degli OAPs per il sensore SeaWiFS che sono stati usati per studiare la variabilità temporale e spaziale del bacino del Mare di Alboran caratterizzato da intensi fenomeni di circolazione a mesoscala.

## 6.1 Modelli di distribuzione statistica dei parametri otticamente attivi

### 6.1.1 La densità di probabilità log-normale.

Una variabile aleatoria  $X$  ha densità di probabilità log-normale quando il suo logaritmo  $\ln(X)$  è distribuito normalmente con valore medio  $\mu$  e varianza  $\sigma^2$ . La funzione di densità di probabilità log-normale viene espressa nella seguente forma:

$$f(X) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma \cdot X} \cdot e^{-\left[\frac{(\ln X - \mu)^2}{2 \cdot \sigma^2}\right]} \quad \text{Eq. 6.1-62}$$

La distribuzione di  $X$  è completamente determinata dai due parametri adimensionali  $\mu$  e  $\sigma^2$  da cui dipendono gli indici caratteristici della distribuzione stessa. In particolare fissati  $\mu$  e  $\sigma^2$  la media di  $X$  risulta pari a:

$$E[X] = e^{\left(\mu + \frac{\sigma^2}{2}\right)}, \quad \text{Eq. 6.1-63}$$

mentre la mediana e la moda sono uguali a rispettivamente:

$$\text{mediana}[X] = e^{\mu}, \quad \text{Eq. 6.1-64}$$

$$\text{moda}[X] = e^{(\mu - \sigma^2)}, \quad \text{Eq. 6.1-65}$$

ed in generale il momento di ordine  $n$  della distribuzione è dato dalla seguente espressione:

$$E[X^n] = e^{\left(n \cdot \mu + \frac{n^2 \cdot \sigma^2}{2}\right)}. \quad \text{Eq. 6.1-66}$$

In Figura 6.1-11 è visualizzata la curva della densità di probabilità log-normale con evidenziate le posizioni degli indici caratteristici.

I tre indici caratteristici di posizione possono essere significativamente diversi tra loro (a differenza della densità di probabilità gaussiana in cui essi coincidono); il loro rapporto risulta pari a:

$$\text{moda}(X) : \text{mediana}(X) : E[X] = e^{-\sigma^2} : 1 : e^{\frac{\sigma^2}{2}}. \quad \text{Eq. 6.1-67}$$

Come è possibile notare dalla Eq. 6.1-67 la moda della distribuzione è sempre minore della mediana qualunque sia il valore della varianza  $\sigma^2$ . Al contrario la media risulta essere sempre maggiore della mediana della distribuzione.

Il coefficiente di variazione  $\delta$  della variabile  $X$ , pari al rapporto tra la deviazione standard e la media di  $X$ , risulta essere uguale a:

$$\delta = \sqrt{e^{\sigma^2} - 1}, \quad \text{Eq. 6.1-68}$$

che permette di esprimere la Eq. 6.1-67 come:

$$\text{moda}(X) : \text{mediana}(X) : E[X] = \frac{1}{M} : 1 : \sqrt{M}, \quad \text{Eq. 6.1-69}$$

in cui si è indicata con  $M$  la seguente quantità:

$$M = e^{\sigma^2} = \delta^2 + 1. \quad \text{Eq. 6.1-70}$$

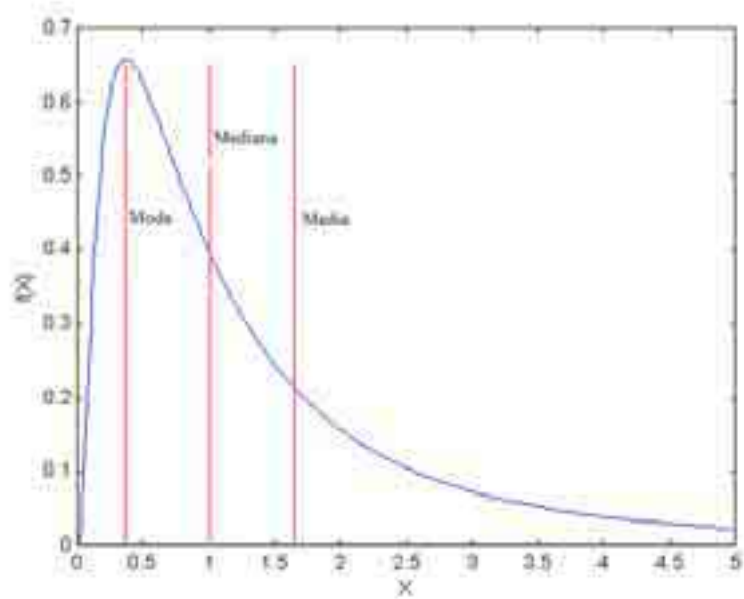


Figura 6.1-11: curva di densità di probabilità log-normale;  $\mu=0$ ,  $\sigma^2=1$ ;  $\text{moda}=e^{-1}$ ,  $\text{mediana}=1$ ,  $\text{media}=e^{1/2}$ .

In Tabella 6.1-2 è riportato il valore del rapporto *Moda:Mediana:Media* al variare del coefficiente di variazione  $\delta$ .

Tabella 6.1-2: rapporto *Moda:Mediana:Media* di una variabile log-normale in funzione di  $\delta$ .

$\delta$	Moda	Mediana	Media
0.0000	1.0000	1.0000	1.0000
0.10000	0.99010	1.0000	1.0050
0.20000	0.96154	1.0000	1.0198
0.30000	0.91743	1.0000	1.0440
0.40000	0.86207	1.0000	1.0770
0.50000	0.80000	1.0000	1.1180
0.60000	0.73529	1.0000	1.1662
0.70000	0.67114	1.0000	1.2207
0.80000	0.60976	1.0000	1.2806
0.90000	0.55249	1.0000	1.3454
1.0000	0.50000	1.0000	1.4142

La stima a massima verosimiglianza della media e del coefficiente di variazione (o equivalentemente della deviazione standard) di una variabile log-normale da un campione  $\{X_1, \dots, X_n\}$  di  $n$  elementi può essere effettuata sostituendo nella Eq. 6.1-63 e nella Eq. 6.1-68 la stima a massima verosimiglianza della media  $\mu$  e della varianza  $\sigma^2$  di  $\ln(X)$  ovvero della media aritmetica e dello scarto quadratico medio degli elementi del campione:

$$m = \frac{1}{n} \cdot \sum_{i=1}^n \ln(X_i), \quad \text{Eq. 6.1-71}$$

$$s^2 = \frac{1}{n} \cdot \sum_{i=1}^n [\ln(X_i) - m]^2. \quad \text{Eq. 6.1-72}$$

La stima a massima verosimiglianza della media secondo la Eq. 6.1-63 risulta più efficiente della stima effettuata direttamente dalla media aritmetica degli elementi  $X_i$  del campione. Infatti per piccoli campioni la media aritmetica tende a sottostimare la media vera della popolazione dato che i valori più alti di  $X$  hanno meno probabilità di presentarsi (pur pesando parecchio nel calcolo della media). Per quanto riguarda la stima della mediana e della moda in maniera analoga i rispettivi stimatori a massima verosimiglianza si ottengono sostituendo nella Eq. 6.1-64 e nella Eq. 6.1-65 la Eq. 6.1-71 e la Eq. 6.1-72.

### 6.1.2 La densità log-normale come modello di variabilità statistica degli OAPs.

Come evidenziato in [Campbell 1995], il modello log-normale è preso come base per descrivere la variabilità statistica delle concentrazioni delle bio-masse, come ad esempio la concentrazione di clorofilla, e per la variabilità bio-ottica in generale. Tale assunto viene giustificato in base al teorema del limite centrale e assumendo che il tasso di accrescimento nel tempo della concentrazione della bio-massa sia proporzionale alla concentrazione stessa considerata al generico istante  $t$  ovvero detta  $X(t)$  la concentrazione all'istante  $t$  il suo tasso di accrescimento nel tempo risulterà pari a:

$$\frac{dX(t)}{dt} = K(t) \cdot X(t), \quad \text{Eq. 6.1-73}$$

in cui  $K(t)$  è il tasso di accrescimento specifico che può assumere valori positivi o negativi. La

Eq. 6.1-73 ha come soluzione:

$$X(t) = X(0) \cdot e^{\int_0^t K(t) \cdot dt}, \quad \text{Eq. 6.1-74}$$

dove  $X(0)$  rappresenta la concentrazione all'istante iniziale. Si consideri  $K(t)$  un processo stocastico e si divida l'intervallo di tempo  $[0, t]$  in  $n$  sottointervalli. Calcolando l'integrale all'esponente nei singoli sottointervalli temporali ed estraendo il logaritmo di ambo i membri, la Eq. 6.1-74 può essere espressa come somma di variabili aleatorie:

$$\ln[X(t)] = \ln[X(0)] + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} K(t) \cdot dt. \quad \text{Eq. 6.1-75}$$

Se si assume che gli integrali di  $K(t)$  nei differenti sottointervalli di tempo siano variabili aleatorie indipendenti (ad esempio per il phytoplankton, data la natura intermittente dei controlli che l'ambiente esercita sulla sua crescita, è ragionevole pensare che i  $K(t)$  in diversi istanti di tempo siano indipendenti) e se si considera  $n$  elevato, per il teorema del limite centrale la somma nella

Eq. 6.1-75 ha densità di probabilità gaussiana. Se anche  $\ln[X(0)]$  è normalmente distribuito allora  $\ln[X(t)]$  risulterà normalmente distribuito e quindi, la concentrazione  $X(t)$  avrà densità di probabilità log-normale. Leggi del tipo della Eq. 6.1-74 sono comuni anche in ottica ad esempio l'irradianza in un mezzo tra la superficie del mezzo stesso e la profondità  $z$  è espressa dalla legge di Beer:

$$E(z) = E(0) \cdot e^{-\int_0^z K(z) dz}, \quad \text{Eq. 6.1-76}$$

in cui  $K(z)$  è il coefficiente di attenuazione del mezzo. Si può arrivare a dimostrare, in maniera analoga per quanto fatto per  $X(t)$ , che anche  $E(z)$  è log-normale. In generale quindi, la densità log-normale può essere assunta come modello di riferimento della variazione statistica sia delle concentrazioni massa su volume dei costituenti otticamente attivi dell'acqua marina sia delle grandezze ottiche in uso nel telerilevamento passivo quali la radianza e l'irradianza.

Per fornire alcune evidenze sperimentali a conferma di quanto precedentemente detto, nei prossimi sotto paragrafi saranno esaminati alcuni *data set* di misure *in situ* prelevate durante specifiche campagne oceanografiche condotte nel Mar Mediterraneo da differenti enti nell'ambito di specifici progetti di ricerca o attività di monitoraggio in continua di determinati siti. Le misurazioni effettuate sono relative a differenti parametri otticamente attivi dell'acqua di mare. In particolare l'analisi statistica mostrata riguarderà misure di concentrazione di clorofilla-a, feofitina-a e sedimento solido in sospensione. Le aree di interesse sono la costa e l'arcipelago della Toscana ed il Mare di Alboran nel Mediterraneo Occidentale nei pressi dello stretto di Gibilterra.

#### 6.1.2.1 Il *data set* ARPAT.

In Tabella 6.1-3 sono riportate le statistiche di un *data set* di misure *in situ* di concentrazione di solidi in sospensione effettuate dall'Agenzia Regionale per la Protezione dell'Ambiente della Toscana (ARPAT). Le misurazioni sono relative ad una serie di punti di campionamento fissi posizionati lungo la costa della Toscana e dell'Isola d'Elba che sono monitorati periodicamente durante l'anno. La Tabella 6.1-3 riporta le statistiche della concentrazione di solido in sospensione  $S$ , espressa in  $\text{g/m}^3$ , e del logaritmo in base 10 della concentrazione. I dati sono relativi al periodo Aprile-Agosto 1998. Il numero di misure contenute nel *data set* è pari a 481.

Tabella 6.1-3: statistiche della concentrazione di sedimenti del *data set* ARPAT.

	$S [\text{g/m}^3]$	$\text{Log}(S)$
Minimo	0.50500	-0.29671
Massimo	27.805	1.4441
Somma	2283.0	305.32
Numero punti	481.00	481.00
Media	4.7464	0.63476
Mediana	4.3500	0.63849
RMS	5.4448	0.66055
Deviazione Standard	2.6706	0.18297
Varianza	7.1321	0.033478
Skewness	4.8383	0.090003

Curtosi	35.024	4.6597
---------	--------	--------

I valori di media e mediana del  $\text{Log}(S)$  risultano praticamente coincidenti fatto questo che fa supporre che la distribuzione del logaritmo della concentrazione  $S$  sia gaussiana ovvero che quella di  $S$  sia log-normale (in una curva gaussiana i valori di moda, mediana e media coincidono). Il valore del parametro di asimmetria (o *Skewness*, che è definito come il momento del terzo ordine della distribuzione normalizzato alla deviazione standard al cubo; per una curva gaussiana risulta pari a 0) è uguale a 0.09 mentre il parametro di curtosi (pari al momento del quarto ordine della distribuzione normalizzato alla deviazione standard elevata alla quarta potenza meno 3; in una curva gaussiana è uguale a 0) risulta uguale a circa 4.66 che indicano rispettivamente una leggera asimmetria ed una maggiore “piccosità” della curva di densità di probabilità empirica del  $\text{Log}(S)$  (in Figura 6.1-13 è mostrato l’istogramma dei valori del  $\text{Log}(S)$ ) rispetto ad una curva gaussiana con uguale valore medio e varianza.

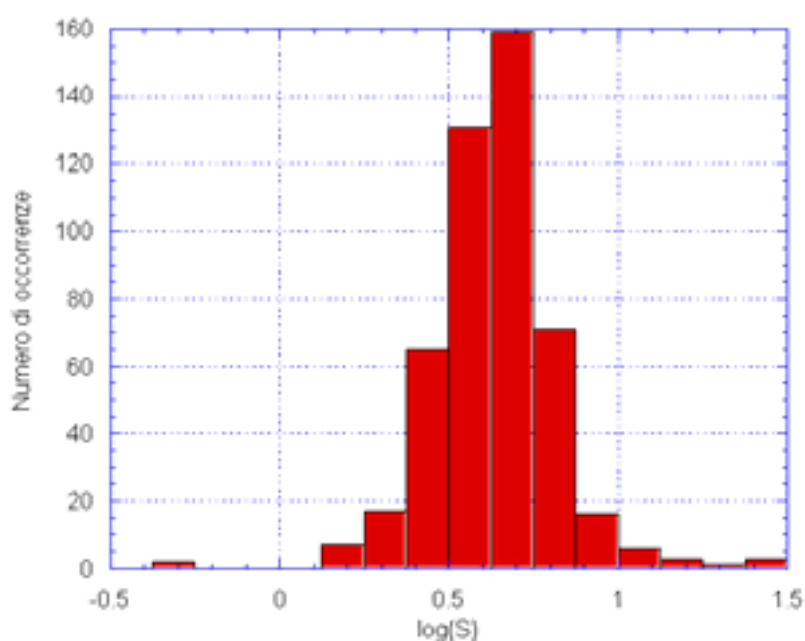




Figura 6.1-13: istogramma della concentrazione di sedimenti del data set ARPAT.

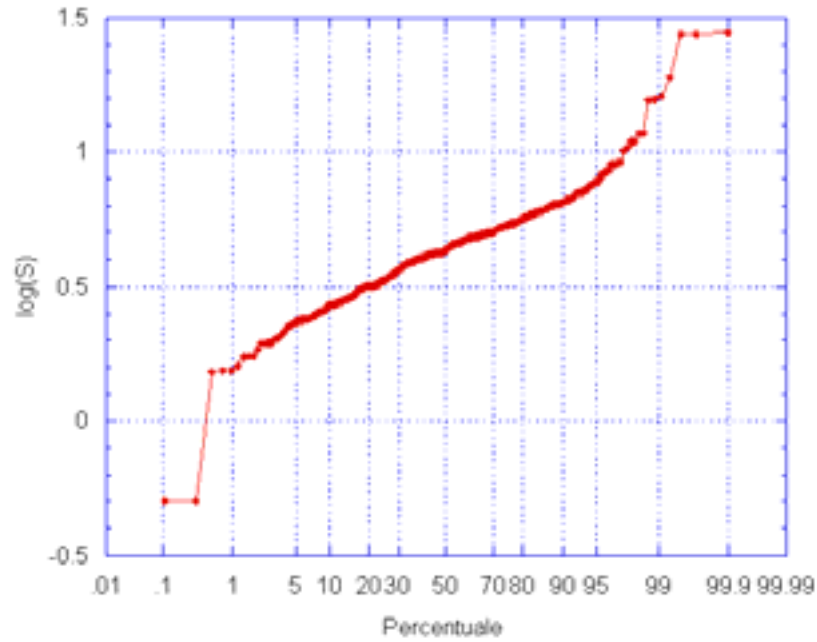


Figura 6.1-15: carta normale di probabilità del Log(S) relativo al data set ARPAT.

Un esame più approfondito del *data set* mostra un buon adattamento del  $\text{Log}(S)$  alla distribuzione normale. La Figura 6.1-15 mostra la carta di probabilità normale per i dati in oggetto. Tale carta rappresenta la distribuzione cumulata dei dati su scala logaritmica. Se i dati fossero distribuiti normalmente la curva in Figura 6.1-15 dovrebbe tendere ad assumere un andamento rettilineo. Come è possibile osservare dalla figura stessa l'andamento del grafico di probabilità è ben approssimato da una retta per valori di  $\text{Log}(S)$  compresi tra circa 0.2 ed 0.9.

Sul campione di dati log-trasformati è stato effettuato il test di verifica delle ipotesi del  $\chi^2$ . Le ipotesi da verificare sono la  $H_0$ : il campione proviene da una popolazione distribuita normalmente, contro l'ipotesi  $H_1$ : il campione non proviene da una popolazione distribuita normalmente. La statistica che viene usata nel test è la seguente:

$$\chi^2 = \sum_{i=1}^k \frac{(f_{oss} - f_{teor})^2}{f_{teor}} \quad \text{Eq. 6.1-77}$$

dove  $f_{oss}$  è la frequenza relativa osservata ed  $f_{teor}$  la frequenza relativa teorica. La statistica  $\chi^2$  è distribuita con densità di probabilità del  $\chi^2$  con  $\nu=k-1-m$  gradi di libertà. Il numero di gradi di libertà del test è noto una volta fissato il numero di classi in cui vengono raggruppati i dati del campione. Per ogni classe viene calcolata una frequenza relativa di presentazione che viene confrontata poi con la frequenza teorica della classe, calcolata considerando la distribuzione teorica i cui parametri vengono stimati dai dati stessi del campione. I gradi di libertà del test sono quindi pari a  $\nu=k-1-m$  dove  $k$  è il numero di classi e  $m$  il numero di parametri necessari al calcolo delle frequenze teoriche. Nel caso di distribuzione normale, completamente determinata nota la media e la varianza,  $m=2$ . Fissato  $\nu$  ed il livello di confidenza  $\alpha$  risulta fissato il cosiddetto valore di soglia critico  $\chi^2_{crit}$  pari al valore di  $\chi^2$  tale che  $\Pr\{\chi^2 < \chi^2_{crit}\} = 1 - \alpha$ . Tale

valore viene confrontato con la statistica  $\chi^2$  calcolato tramite la Eq. 6.1-77; se  $\chi^2$  è al di sotto della soglia critica allora si accetta l'ipotesi  $H_0$  secondo cui il campione è stato estratto da una popolazione distribuita secondo la distribuzione teorica considerata.

Il test del  $\chi^2$  sui dati log-trasformati del *data set* ARPAT (Tabella 6.1-4) porta a non rifiutare l'ipotesi di distribuzione gaussiana ad un livello di confidenza pari a 0.05. Il test è stato effettuato considerando un numero di gradi di libertà è  $\nu=17$  il che corrisponde a considerare un numero di classi pari a 20. Il valore critico è uguale a  $\chi^2_{crit}=27.5871$  mentre il valore della statistica risulta pari a  $\chi^2=10.9128$ .

Tabella 6.1-4: risultati del test del  $\chi^2$  effettuato sul Log(S).

Livello di confidenza $\alpha$	Soglia critica $\chi^2_{crit}$	Valore statistica $\chi^2$	Gradi di libertà $\nu$
0.05	27.5871	10.9128	17

#### 6.1.2.2 Il *data set* ICM.

La Tabella 6.1-5 mostra le statistiche di un *data set* di misure *in situ* di concentrazione di clorofilla-a (indicata con C ed espressa in  $\text{mg/m}^3$ ) e di concentrazione di feofitina-a (indicata con Ph ed espressa anch'essa in  $\text{mg/m}^3$ ) raccolte durante una campagna oceanografica condotta sul mare di Alboran nel Mediterraneo Occidentale tra il 4 Ottobre ed il 13 Ottobre 1996 dall'Istituto di Scienza del Mare di Barcellona (ICM) nell'ambito del progetto europeo OMEGA. Il campione risulta di piccole dimensioni, 44 punti di campionamento per la concentrazione di clorofilla e 35 per quella di feofitina. Media e mediana del campione del  $\text{Log}(C)$  si discostano leggermente tra loro assumendo il valore di -0.49 e -0.58 rispettivamente. Il parametro di asimmetria stimato risulta pari a circa 0.54 mentre il parametro di curtosi a -0.3. Per quanto riguarda il campione del  $\text{Log}(Ph)$  si ha un valore stimato di media e mediana pari a -1.83 e -1.92 rispettivamente. L'asimmetria ha un valore di 0.89 mentre la curtosi di 0.82. Nella Figura 6.1-16 e Figura 6.1-18 sono riportate le curve normali di probabilità per il  $\text{Log}(C)$  ed il  $\text{Log}(Ph)$  rispettivamente. I grafici mostrano un andamento oscillante intorno ad una linea retta.

Sui due campioni è stato effettuato il test del  $\chi^2$  i cui risultati sono riportati in

Tabella 6.1-7, per quanto riguarda il  $\text{Log}(C)$ , ed in Tabella 6.1-8 per quanto riguarda il  $\text{Log}(Ph)$ . I risultati del test portano a non respingere l'ipotesi di distribuzione gaussiana dei dati ad un livello di confidenza pari a  $\alpha=0.05$  (il test è stato effettuato con 6 gradi di libertà).

Tabella 6.1-5: statistiche della concentrazione di clorofilla-a e feofitina-a del *data set* ICM.

	C [ $\text{mg/m}^3$ ]	Ph [ $\text{mg/m}^3$ ]	Log(C)	Log(Ph)
Minimo	0.064000003	0	-1.19382	-2.3979399
Massimo	3.9749999	0.145	0.5993371	-0.83863199
Somma	23.156	0.72899999	-21.648657	-64.017494
Numero punti	44	35	44	35
Media	0.52627273	0.02025	-0.49201493	-1.8290712
Mediana	0.26300001	0.012	-0.58004421	-1.9208188
RMS	0.85300948	0.031693409	0.63786216	1.8570737

Dev. St.	0.67907489	0.024726359	0.41063579	0.32597098
Varianza	0.4611427	0.00061139283	0.16862175	0.10625708
Skewness	3.3046095	3.773486	0.54429653	0.89147326
Curtosi	13.318689	16.212325	-0.30078839	0.82304605

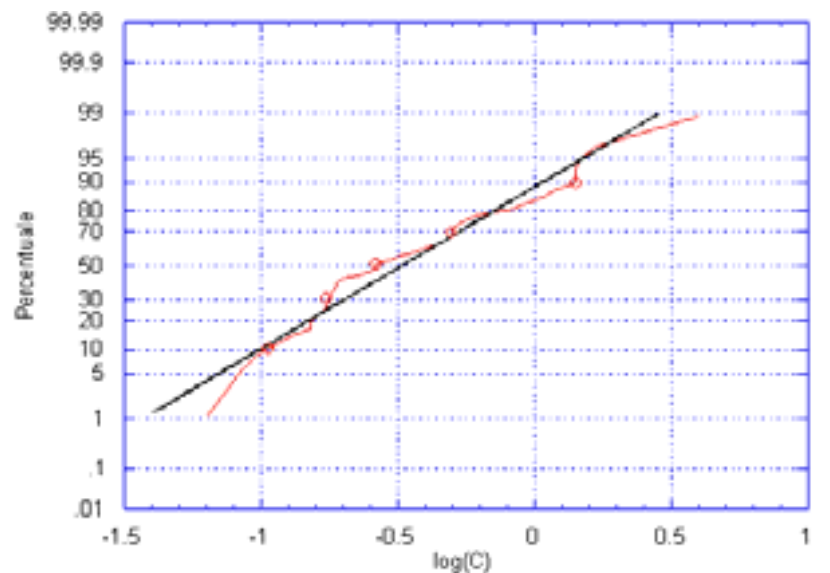


Figura 6.1-16: carta normale di probabilità del Log(C) relativo al data set ICM.

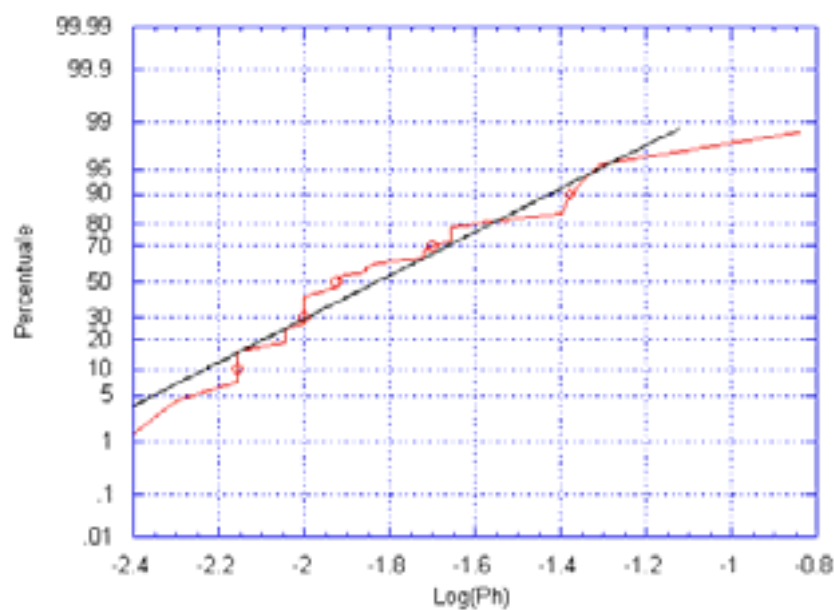


Figura 6.1-18: carta normale di probabilità del Log(Ph) relativo al data set ICM.

Tabella 6.1-7: risultati del test del  $\chi^2$  effettuato sul Log(C) data set ICM.

Livello di confidenza $\alpha$	Soglia critica $\chi^2_{crit}$	Valore statistica $\chi^2$	Gradi di libertà $\nu$
-----------------------------------	-----------------------------------	-------------------------------	---------------------------

0.05	12.5916	0.1573	6
------	---------	--------	---

Tabella 6.1-8: risultati del test del  $\chi^2$  effettuato sul Log(Ph) data set ICM.

Livello di confidenza $\alpha$	Soglia critica $\chi^2_{crit}$	Valore statistica $\chi^2$	Gradi di libertà $\nu$
0.05	12.5916	0.3349	6

### 6.1.2.3 Il data set SOC.

In Tabella 6.1-9 sono mostrate le statistiche di un *data set* di misure *in situ* di concentrazione di clorofilla-a (indicata con C ed espressa in  $\text{mg/m}^3$ ) raccolte durante una seconda campagna oceanografica condotta nel mare di Alboran nella prima metà del Dicembre 1996 dal *Southampton Oceanography Centre* (SOC) sempre nell'ambito del progetto europeo OMEGA. Il *data set* contiene 169 punti di campionamento distribuiti sul sito di interesse come mostrato in Figura 6.1-20 in cui sono riportate le tracce della crociera oceanografica. Media e mediana del campione del  $\text{Log}(C)$  sono pari a -0.29 e -0.30 rispettivamente. Il parametro di asimmetria stimato è uguale a circa 0.35 mentre il parametro di curtosi a 0.73.

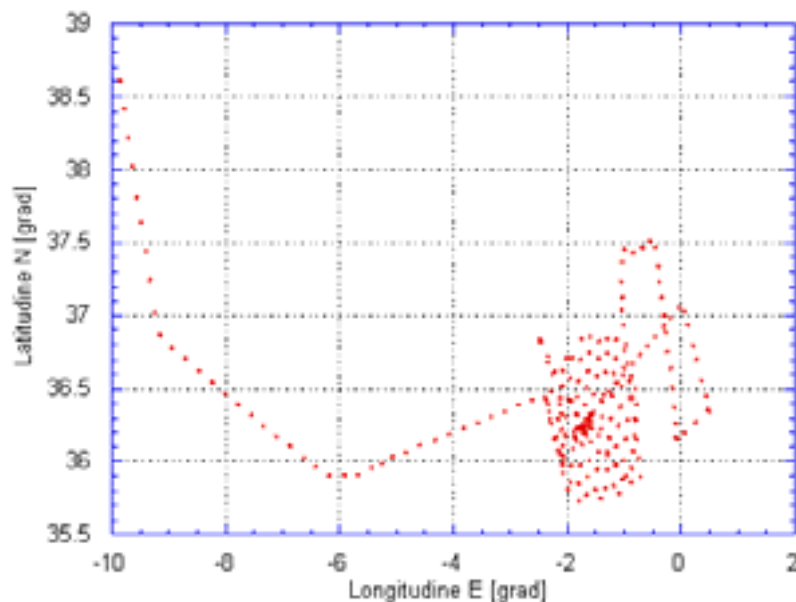


Figura 6.1-20: tracce della crociera condotta dal SOC nel Mare di Alboran. durante il Dicembre 96.

Tabella 6.1-9: statistiche della concentrazione di clorofilla-a e feofitina-a del data set SOC..

	$C [\text{mg/m}^3]$	$\text{Log}(C)$
Minimum	0.15369999	-0.81332999
Maximum	2.8074	0.4483
Sum	99.0509	-48.74288
Points	169	169
Mean	0.5861	-0.28841941
Median	0.495	-0.30539
RMS	0.68379586	0.35890173

Std Deviation	0.35327332	0.21423415
Variance	0.12480204	0.045896272
Skewness	2.8134079	0.35150332
Kurtosis	11.886689	0.73359161

Nella Figura 6.1-22 è riportato l'istogramma del  $\text{Log}(C)$  mentre in Figura 6.1-24 la relativa curva normale di probabilità che mostra un andamento sostanzialmente rettilineo tra  $-0.7$  e  $0.05$ ; al di fuori di tale intervallo la distribuzione si discosta maggiormente dalla normalità. Il test del  $\chi^2$ , effettuato con 17 gradi di libertà e i cui risultati sono mostrati in Tabella 6.1-11, porta ad accettare l'ipotesi di distribuzione normale del  $\text{Log}(C)$  sempre con un livello di confidenza pari a  $\alpha=0.05$  ( $\chi^2=0.1646 < \chi^2_{crit}=27.5871$ ).

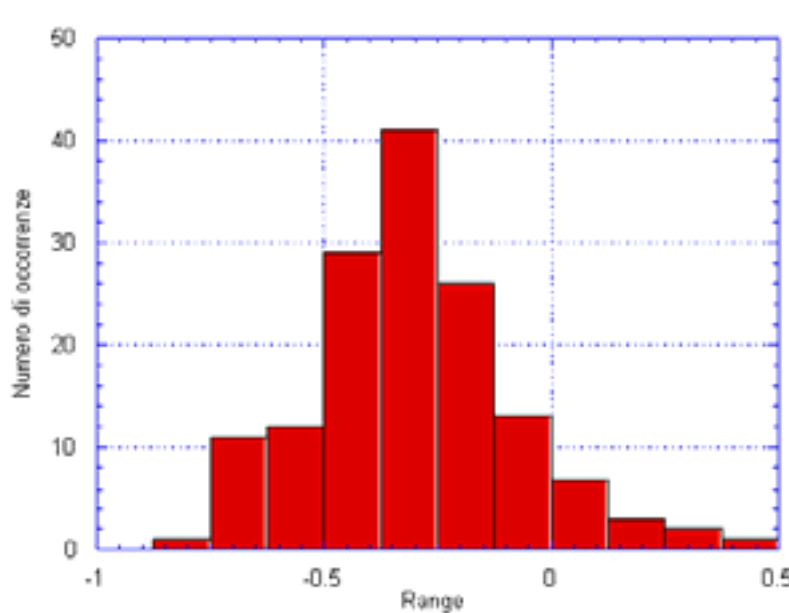


Figura 6.1-22: istogramma della concentrazione di sedimenti del data set SOC.

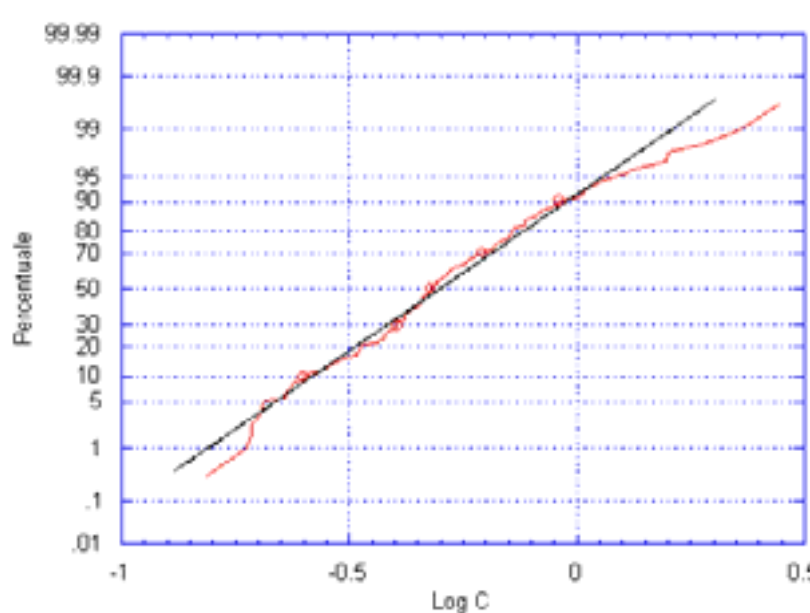


Figura 6.1-24: carta normale di probabilità del  $\text{Log}(C)$  relativo al data set SOC.

Tabella 6.1-11: risultati del test del  $\chi^2$  effettuato sul  $\text{Log}(C)$  data set SOC.

Livello di confidenza $\alpha$	Soglia critica $\chi^2_{crit}$	Valore statistica $\chi^2$	Gradi di libertà $\nu$
0.05	27.5871	0.1646	17

## 6.2 Generazione di *data sets* simulati di OAPs da distribuzioni note

Il modello di variabilità statistica log-normale è stato considerato come modello di base per la generazione di *data set* simulati di OAPs e di riflettanze medie sulle bande dei sensori. I *data set* generati saranno usati per testare le tecniche di inversione illustrate nel capitolo 3. Al fine di effettuare un'analisi il più completa possibile delle prestazioni di tali algoritmi, le tecniche di inversione proposte saranno sperimentate su *data set* che simulano il comportamento ottico del corpo idrico in diverse condizioni che possono essere incontrate nei casi pratici e che risultano di particolare importanza ai fini del monitoraggio. In particolare, tramite generatori di variabili casuali log-normali, sono stati generati tre *data set* di parametri otticamente attivi dell'acqua marina: il primo è tipico delle acque oceaniche in cui abbondano le acque di caso I, il secondo è rappresentativo per le acque di caso II mentre il terzo *data set* rappresenta un caso intermedio tra i primi due. Tramite il modello di riflettanza sub-superficiale i tre *data set* di OAPs sono usati per generare altrettanti *data set* di riflettanze medie sulle bande di ciascun sensore preso in esame.

La generazione dei *data set* di OAPs richiede di fissare, a seconda del tipo di acqua considerato, i parametri della distribuzione log-normale, valor medio e varianza del logaritmo di  $C$ ,  $X$  ed  $Y$ , ed i coefficienti di correlazione sempre tra i logaritmi di  $C$ ,  $X$  ed  $Y$ . In natura i tre OAPs possono assumere valori compresi nei seguenti intervalli [1]:

- $0.01 \text{ mg/m}^3 < C < 100 \text{ mg/m}^3$ ,
- $0.01 \text{ m}^{-1} < X < 10 \text{ m}^{-1}$ ,
- $0.01 \text{ m}^{-1} < Y < 1 \text{ m}^{-1}$ .

Per quanto riguarda l'intervallo di variazione di  $C$  si dovrà tenere conto dei limiti di applicabilità del modello di riflettanza che fissano un valore minimo di  $0.02 \text{ mg/m}^3$  ed un valore massimo pari a  $25 \text{ mg/m}^3$ . All'interno dei limiti riportati si collocano i tre casi citati precedentemente che saranno considerati nel proseguo del lavoro. Le statistiche degli OAPs nei vari casi saranno fissate in base a risultati riportati nella letteratura relativi a ricerche condotte in diversi bacini. Una più ampia trattazione riguardante le statistiche tipiche delle acque di caso I e II ed intermedie è fornita in [3].

### 6.2.1 *Data set* per acque di caso I.

In [23] e [24] sono stati suggeriti due criteri che permettono di stabilire se un'acqua è da ritenersi di caso I. Secondo tali autori le acque di caso I soddisfano le due seguenti condizioni:

- l'assorbimento dovuto alla sostanza gialla a  $380 \text{ nm}$  è minore di  $0.1 \text{ m}^{-1}$ ,
- la concentrazione totale di particelle (TSM: *Total Suspended Matter*), clorofillacee e non, è minore di  $0.5 \text{ g/m}^3$ .

La prima condizione, tramite la Eq. 2.3-12 relativa al coefficiente di assorbimento della sostanza gialla al variare della lunghezza d'onda, fissa un valore massimo per  $Y$  pari a:

$$Y = a_{\lambda}(\lambda = 440nm) < 0.043 \quad m^{-1} \quad Eq. 6.2-1$$

La seconda condizione consente di fissare i limiti superiori per i parametri  $C$  ed  $X$ . *Gordon e Morel* in [25], studiando il legame tra lo *scattering* delle particelle e la concentrazione di clorofilla-a+feofitina-a hanno ottenuto la seguente relazione empirica tra  $C$  e la concentrazione di TSM:

$$TSM = 0.30 \cdot C^{0.62}, \quad Eq. 6.2-2$$

dove si è indicato con  $TSM$  la concentrazione di particelle in sospensione espressa in  $g/m^3$ . Imponendo la condizione  $TSM < 0.5 \text{ g/m}^3$  si perviene al limite superiore per il parametro  $C$ :

$$C < 2.26 \text{ mg/m}^3. \quad Eq. 6.2-3$$

Un limite superiore per  $X$  può essere fissato assumendo che la concentrazione massa su volume delle particelle non clorofillacee, che sarà indicata con  $S$ , rispetti il limite considerato per la TSM anche se tale limite risulta meno stringente di quanto non dovrebbe essere nella realtà (l'approssimazione risulterà sempre più precisa al crescere della concentrazione relativa di particelle non clorofillacee rispetto alla concentrazione di particelle di natura clorofillacea). Dai risultati presentati in [26] e [27] e ripresi in [3] riguardanti il coefficiente di attenuazione specifico delle particelle non clorofillacee (somma dei coefficienti specifici di assorbimento e scattering), è possibile ricavare una relazione che lega la concentrazione  $S$  al parametro ottico  $X$ . Gli autori riportano per il coefficiente di attenuazione specifico a 565 nm di lunghezza d'onda un valore pari a  $0.3 \text{ m}^2/g$ . La concentrazione  $S$  risulta in tal modo legata al coefficiente di attenuazione a 565 nm dalla seguente equazione:

$$S = 3.33 \cdot c_{sed}(\lambda = 565 \text{ nm}), \quad Eq. 6.2-4$$

dove con  $c_{sed}$  si è indicato il coefficiente di attenuazione totale delle particelle di natura non clorofillacea;  $S$  è espressa in  $g/m^3$ . Sottraendo da  $c_{sed}(\lambda = 565 \text{ nm})$  il contributo dovuto all'assorbimento, tramite la Eq. 2.3-11, ed assumendo un andamento del coefficiente di *scattering* al variare della lunghezza d'onda del tipo  $\lambda^{-1}$ , la relazione tra  $S$  e  $X$  risulta esprimibile nella seguente forma:

$$S = 3.28 \cdot X, \quad Eq. 6.2-5$$

dove  $S$  è ancora espressa in  $g/m^3$  ed  $X$  in  $m^{-1}$ . La Eq. 6.2-5 insieme alla seconda condizione di Doerferr permette di stabilire il limite per  $X$  per le acque di caso I:

$$X < 0.15 \text{ m}^{-1}. \quad Eq. 6.2-6$$

I limiti introdotti permettono di fissare le statistiche degli OAPs per la generazione del *data set* relativo alle acque di caso I. In particolare in [3] è stata effettuata un'analisi statistica su un *data set* di stime del parametro  $C$  effettuate da dati telerilevati dal sensore CZCS su bacini oceanici. Lo studio ha fornito una stima dei valori di media e varianza del logaritmo di  $C$  che è possibile avere in acque di caso I non troppo oligotrofiche (povere di costituenti). I valori trovati sono -0.86 per il valor medio di  $\text{Log}(C)$  e 0.3 per la varianza. Si assumeranno tali valori come statistiche di  $\text{Log}(C)$  per generare il *data set* relativo alle acque di caso I. La ricerca condotta sempre in [3] ha inoltre confermato che la maggior parte degli oceani in termini di superficie occupata è costituita da acque con tali proprietà statistiche il che le rende di particolare importanza dal punto di vista del telerilevamento. Risulta quindi giustificato l'interesse a testare gli algoritmi di stima degli OAPs in tali condizioni. La deviazione standard per il  $\text{Log}(X)$  ed il  $\text{Log}(Y)$  sarà assunta pari a quella di  $\text{Log}(C)$  evitando in tal modo di privilegiare la variabilità di un

parametro rispetto a quella degli altri. I valori medi di  $\text{Log}(X)$  e  $\text{Log}(Y)$  saranno ricavati ammettendo che il 90% dei valori generati caschi all'interno dei limiti fissati precedentemente per  $X$  ed  $Y$ . Per una distribuzione gaussiana ciò corrisponde a dire che il 90% dei valori sono minori di  $\mu + 1.3 \cdot \sigma$  il che porta a valori media pari a  $\mu_{\log X} = -1.21$  e  $\mu_{\log Y} = -1.75$ .

Le acque di caso I sono caratterizzate da concentrazioni dei costituenti fortemente correlate tra loro. Ciò è confermato sperimentalmente come riscontrabile in letteratura [25]. Risulta quindi opportuno introdurre una certa correlazione nei parametri che saranno generati. In [3] si fa uso di un valore del coefficiente di correlazione tra i logaritmi dei parametri pari a  $\rho_{\log C, \log X} = \rho_{\log C, \log Y} = 0.8$ , valore che nasce da un compromesso tra i coefficienti di correlazione forniti in letteratura [25] e la necessità di non introdurre in ingresso al modello di riflettanza parametri troppo correlati il che porterebbe a sovrastimare gli indici di qualità (in particolare modo per quanto riguarda il coefficiente di correlazione tra OAPs simulati e OAPs stimati) degli algoritmi di inversione. In Figura 6.2-1 sono riassunte le statistiche usate per la generazione del data set di OAPs per acque di caso I e sono mostrati gli scatterogrammi tra i parametri generati da cui si nota la correlazione introdotta.

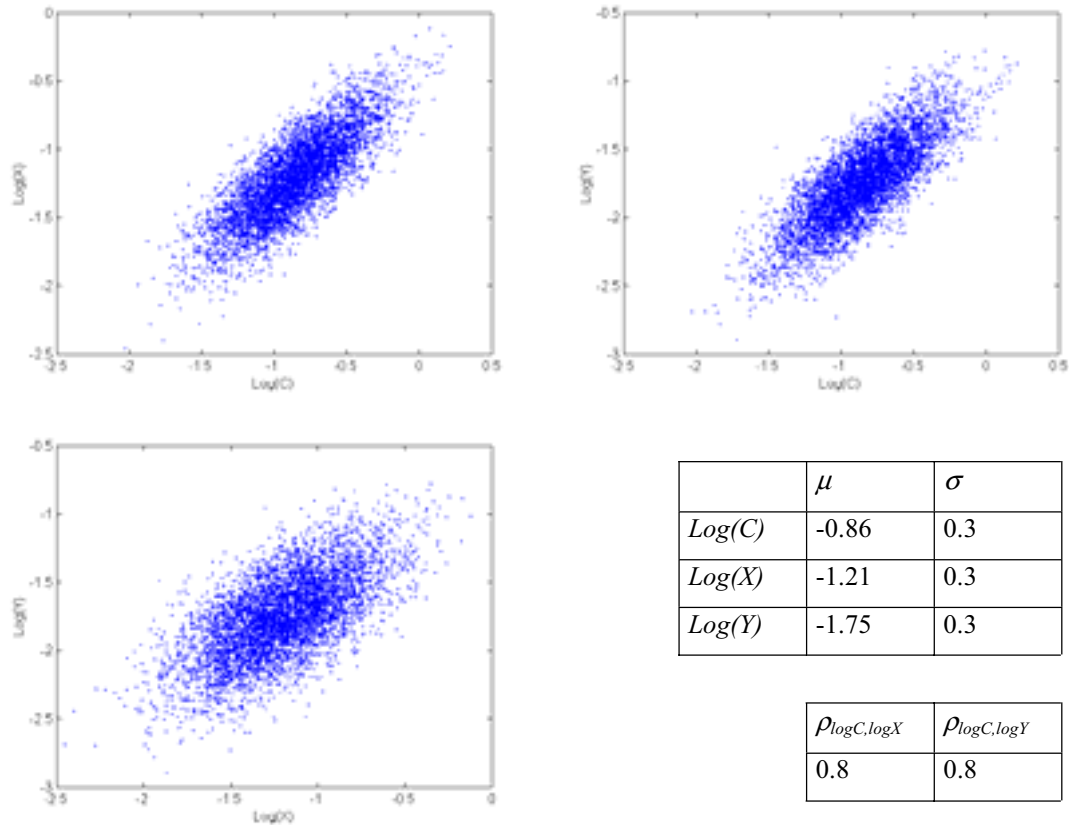


Figura 6.2-1: scatterogrammi dei logaritmi dei parametri otticamenteattivi per il data set acque di caso I.

### 6.2.2 Data set per acque di caso II.

Le acque di caso II sono caratterizzate da una variabilità delle concentrazioni dei costituenti otticamente attivi più ampia rispetto alle acque di caso I precedentemente analizzate. Tali concentrazioni sono inoltre maggiormente incorrelate tra esse.



Le applicazioni di monitoraggio delle zone costiere in cui le acque di caso II sono presenti in larga misura richiedono algoritmi di stima degli OAPs che sia in grado di sfruttare tutta l'informazione spettrale fornita dai sensori. Data l'importanza sempre crescente che tali applicazioni stanno assumendo nel tempo, si ritiene opportuno nell'ambito del presente lavoro di tesi dedicare spazio per testare le metodologie di inversione proposte e comparare le prestazioni dei sensori considerati in condizioni che rispecchiano quelle tipiche delle acque di caso II attraverso un *data set* che simuli abbastanza fedelmente il comportamento di questo tipo di acque.

Data l'alta variabilità dei parametri bio-ottici che è possibile rilevare nelle acque di caso II non sembra esistere in letteratura un *data set* di misurazioni che possa caratterizzare completamente tali acque. In [3] è stata proposta una serie di statistiche per generare un *data set* di OAPs con distribuzione log-normale che in qualche modo potesse avvicinarsi il più possibile ad una situazione reale. Rispetto alla distribuzione usata per generare il *data set* di acque di caso I, per il *data set* di acque di caso II la distribuzione usata risulta essere sempre log-normale, ma con una deviazione standard del logaritmo degli OAPs più alta e pari a  $\sigma_{\log C} = \sigma_{\log X} = \sigma_{\log Y} = 0.5$  ed un coefficiente di correlazione più basso e pari a  $\rho_{\log C, \log X} = \rho_{\log C, \log Y} = -0.5$ . Il valore medio dei logaritmi degli OAPs risulta più alto e pari a  $\mu_{\log C} = 0.0$ ,  $\mu_{\log X} = 0.0$ ,  $\mu_{\log Y} = -0.5$ . Le statistiche per il *data set* di acque di caso II sono riassunte in Figura 6.2-2 in cui sono inoltre riportati gli scatterogrammi dei parametri generati da cui si nota una maggiore dispersione degli stessi rispetto alle acque di caso I.

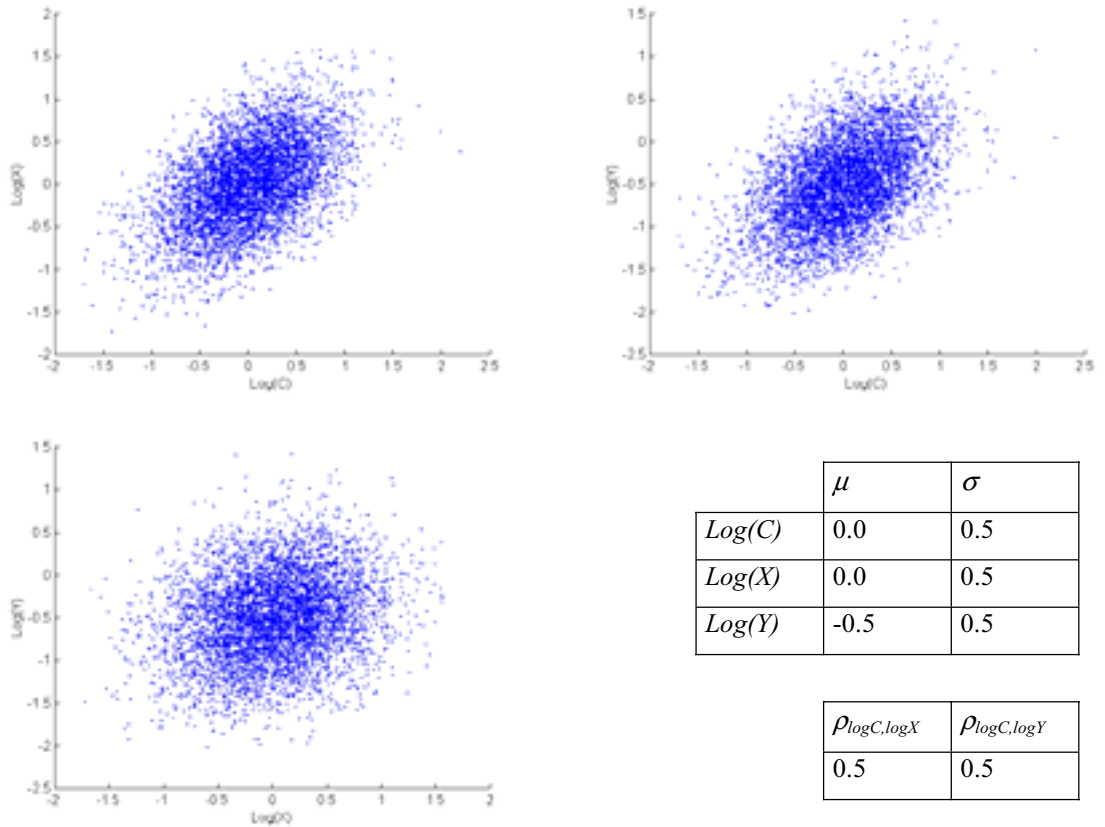


Figura 6.2-2: scatterogrammi dei logaritmi dei parametri otticamenteattivi per il *data set* acque di caso II.

### 6.2.3 Data set per acque di caso I-II.

Gli algoritmi di inversione del modello di riflettanza saranno testati in una situazione più generale prendendo in considerazione un *data set* che simula una condizione più generale di quelle fino ad ora analizzate che comprende sia le acque di caso I che quelle di caso II. Sempre in [3] sono riportate le statistiche per generare una serie di terne  $C, X, Y$  log-normali che simulano la condizione desiderata. I valori delle statistiche sono stati fissati in maniera analoga a quanto fatto per le acque di caso I analizzando il *data set* globale delle stime di clorofilla da dati CZCS e fissando i valori medi di  $\text{Log}(X)$  e  $\text{Log}(Y)$  in base ai valori massimi e minimi riscontrabili in natura (riportati nell'introduzione del presente capitolo). Per quanto riguarda i valori medi Figura 6.2-3 questi sono stati fissati a  $\mu_{\log C} = -0.04$ ,  $\mu_{\log X} = -0.57$  e  $\mu_{\log Y} = -1.05$  mentre le deviazioni standard sono state assunte uguali per i singoli parametri e pari a  $\sigma_{\log C} = \sigma_{\log X} = \sigma_{\log Y} = 0.45$ . I coefficienti di correlazione sono pari a  $\rho_{\log C, \log X} = \rho_{\log C, \log Y} = 0.8$ .

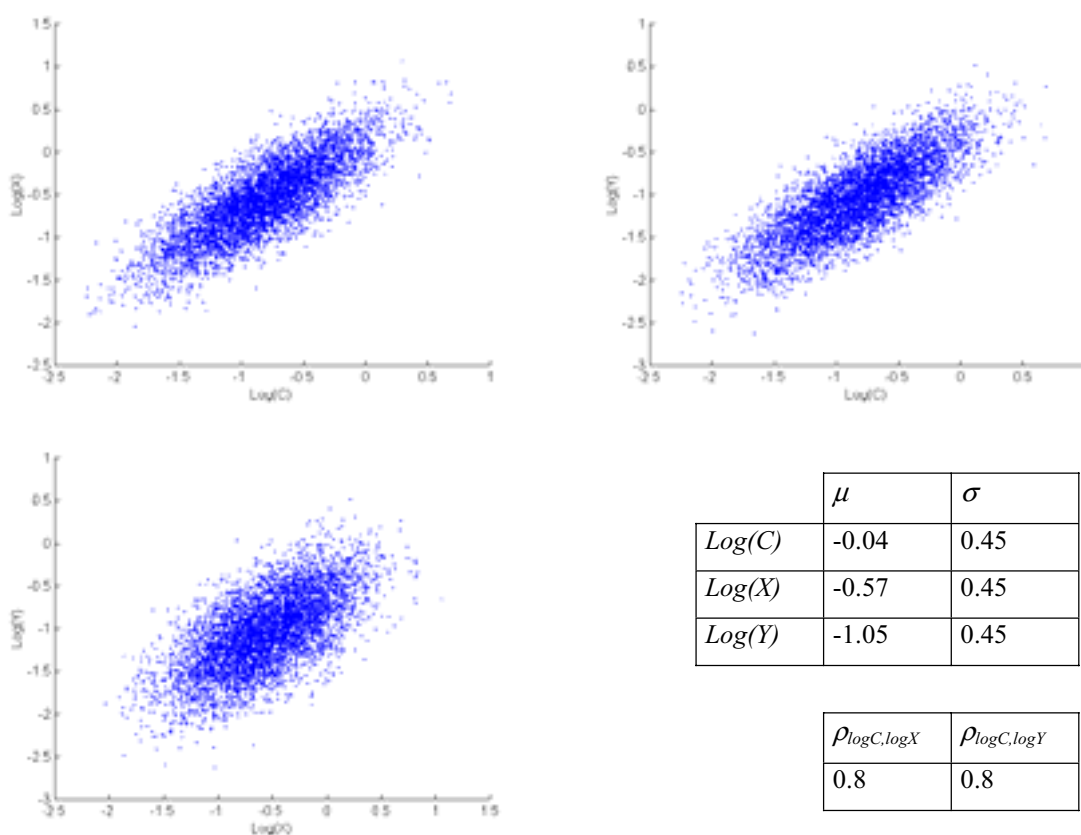


Figura 6.2-3: Scatterogrammi dei logaritmi dei parametri otticamenteattivi per il data set acque di caso I-II.

## 6.3 Generazione delle riflettanze medie sulle bande dei sensori

Le triplette  $C, X, Y$  generate per formare i data set precedentemente descritti sono state fornite in ingresso al modello di riflettanza sub-superficiale per formare altrettanti *data-set* di riflettanze medie sulle bande dei sensori. Ogni *data set* è costituito da 5000 vettori la cui componente  $i$ -esima è la riflettanza media sulla banda del canale  $i$ -esimo del sensore considerato. Le riflettanze sono state generate per i tre sensori SeaWiFS, MOS e MERIS. Nei prossimi sottoparagrafi per ognuno dei *data set* costruiti saranno riportati i grafici delle riflettanze generate (alcuni esempi), dell'intervallo di variazione delle riflettanze (curva di

riflettanza massima minima e media sulle 5000 curve di riflettanza generate), gli scatterogrammi delle riflettanze per una serie di coppie di canali del sensore considerato e la matrice di correlazione tra le riflettanze nei diversi canali per analizzare sia graficamente che quantitativamente il grado ed il tipo di correlazione esistente tra i diversi canali. L'analisi della correlazione tra i dati di riflettanza assume notevole importanza perchè aiuta a riconoscere i casi di singolarità che possono presentarsi quando si ricorre all'uso di metodi di regressione multilineare per (vedere capitolo 3) invertire il modello di colore marino.

### 6.3.1 Data set di riflettanze per il sensore SeaWiFS.

#### 6.3.1.1 SeaWiFS acque di Caso I.

In Figura 6.3-1, sono riportati alcuni esempi di riflettanza generata per i sei canali del sensore SeaWiFS utilizzando il *data set* di OAPs per le acque di caso I. In tali condizioni la riflettanza è maggiore nei canali del blu e del blu-verde caratteristica questa tipica delle acque oligotrofiche ovvero povere di costituenti e quindi con un colore che tende maggiormente al blu (vedere l'andamento del coefficiente di assorbimento e di *backscattering* dell'acqua di mare pura al variare della lunghezza d'onda). Nella stessa Figura 6.3-1 è riportato il grafico delle riflettanze medie, massime e minime in ogni canale del sensore. Per i canali dal primo al quarto la riflettanza è sempre maggiore di 0.01 mentre nel canale 6 essa è praticamente sempre al di sotto di tale limite. La Figura 6.3-2 mostra gli scatterogrammi delle riflettanze di una serie di coppie di canali. Nel primo grafico sono confrontati i valori della riflettanza nei canali 1 e 2 da cui è possibile notare l'esistenza di una forte correlazione lineare. In generale questo comportamento si ha tra una qualsiasi coppia di canali adiacenti come è possibile notare dal terzo e quarto grafico di Figura 6.3-2 relativi agli scatterogrammi tra le riflettanze nei canali 4 e 5 e nei canali 5 e 6. Per canali posizionati ad una certa distanza nello spettro del visibile le riflettanze risultano incorrelate come è possibile notare dal secondo grafico di Figura 6.3-2 in cui sono confrontate le riflettanze tra i canali 1 e 6 e da cui si nota una maggiore dispersione rispetto ai casi precedenti. Il comportamento della correlazione tra i differenti canali descritto precedentemente trova conferma nei valori dei vari elementi della matrice di correlazione riportata in Tabella 6.3-1. Il valore assoluto della correlazione tra canali adiacenti (seconda diagonale della matrice al di sopra della diagonale principale) risulta essere sempre maggiore di 0.74 e raggiunge un valore massimo di circa 0.979. La correlazione in valore assoluto decresce man mano che la distanza tra i canali aumenta come è possibile notare osservando le diagonali successive (si ha un valore minimo di correlazione pari a circa 0.04 tra i canali 2 e 5).

#### 6.3.1.2 SeaWiFS acque di caso II.

In Figura 6.3-3 sono rappresentate un certo numero di curve di riflettanza media generate dal *data set* di OAPs per acque di caso II. Nella stessa figura, sono rappresentate le corrispondenti curve delle riflettanze medie, massime e minime, da cui è possibile notare come l'intervallo di variazione in ogni canale sia molto più ampio rispetto a quello riscontrato per il *data set* che simula il comportamento delle acque di caso I. La riflettanza risulta in media più alta nei canali vicini alla banda del verde (canali 4 e 5), comportamento tipico delle acque di caso II in cui la concentrazione dei costituenti è elevata. Pur essendo i tre parametri otticamente attivi meno correlati, la correlazione tra le riflettanze nei diversi canali è in generale maggiore rispetto alle acque di caso I come è possibile notare dal raffronto fra le matrici di

correlazione riportate in Tabella 6.3-2, relativa alle acque di caso II, e nella Tabella 6.3-1 relativa alle acque di caso I. I canali adiacenti hanno una correlazione che è superiore a 0.99 ad esclusione dei canali 5 e 6 in cui si osserva una correlazione di circa 0.85. La correlazione in generale è elevata anche per canali distanti, si nota ad esempio un valore di correlazione tra i canali 1 e 5 pari a circa 0.9; per le altre coppie di canali la correlazione non scende mai al di sotto di 0.6 ad eccezione del valore relativo ai canali 1 e 6 che è di circa 0.57. Gli scatterogrammi di Figura 6.3-4 danno un'idea della dispersione dei dati di riflettanza. I grafici danno una conferma dell'alta correlazione lineare esistente tra canali adiacenti; una minore dispersione dei dati (sempre rispetto alle acque di caso I) si riscontra anche tra i canali più distanti ovvero il canale 1 ed il canale 6.

#### 6.3.1.3 SeaWiFS acque di caso I-II.

Il *data set* per acque di caso I-II presenta caratteristiche intermedie tra i due *data set* precedenti. In particolare la riflettanza risulta in media più alta nei canali centrali (2, 3 e 4) e presenta una dinamica che è intermedia tra le dinamiche osservate nelle acque di caso I e II (Figura 6.3-5). La correlazione tra i canali adiacenti (Tabella 6.3-3e Figura 6.3-6) è sempre molto alta anche se il valore di correlazione di 0.99, osservato nel *data set* per le acque di caso II nella maggior parte delle coppie di canali adiacenti, viene raggiunto solo tra i canali 1 e 2. Le altre coppie presentano valori tipici del *data set* di acque di caso I. All'aumentare della distanza fra i canali la correlazione scende, ma non alla stessa velocità con cui decresce nelle acque di caso I. Anche per canali distanti osserviamo una leggera correlazione superiore a 0.5.

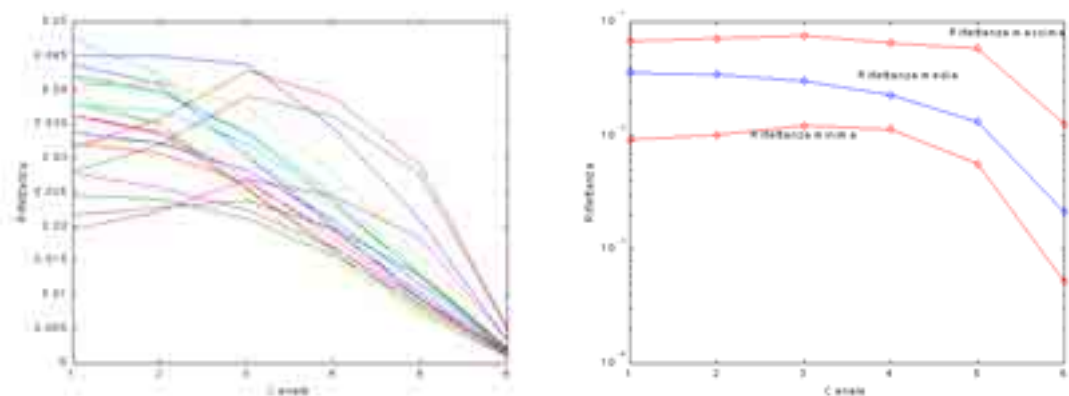


Figura 6.3-1: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso I per il sensore SeaWiFS.

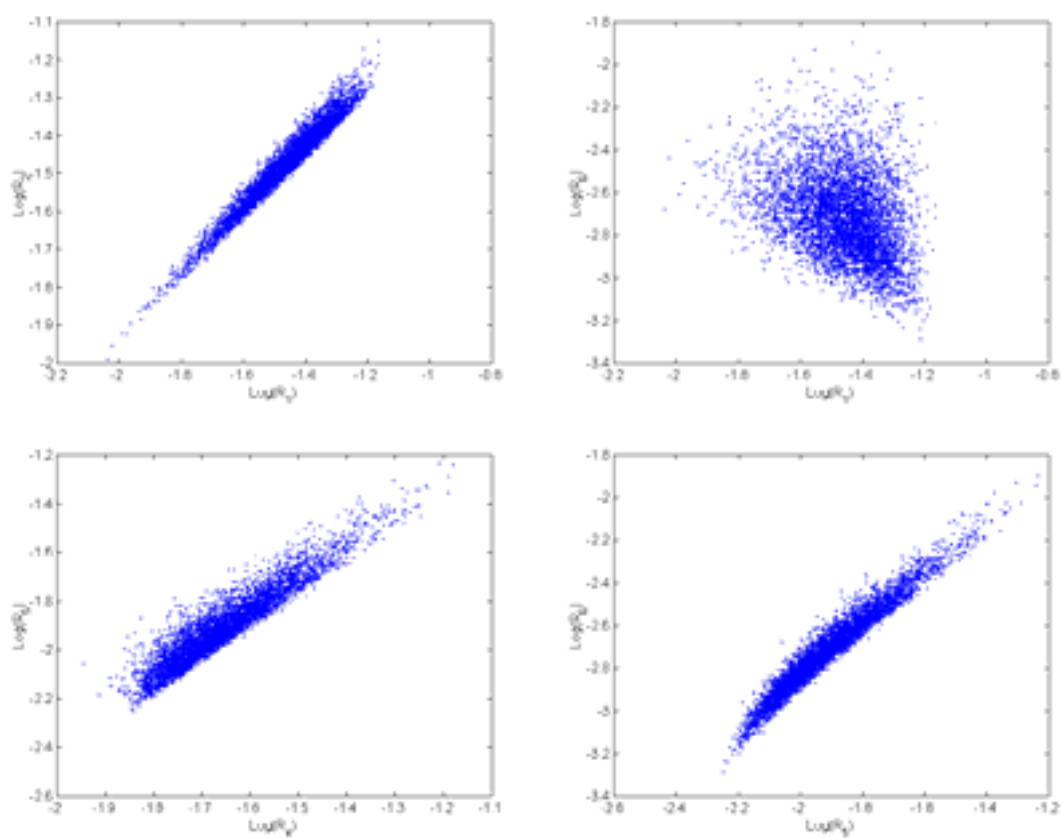


Figura 6.3-2: Scatterogrammi delle riflettanze generate dal data set SeaWiFS-acque di caso I in alcuni canali del sensore.

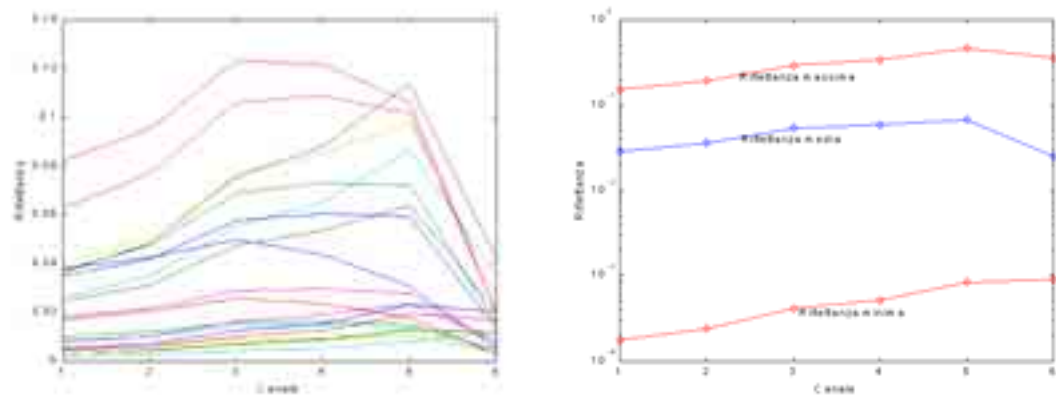


Figura 6.3-3: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima per il data set SeaWiFS acque di caso II.

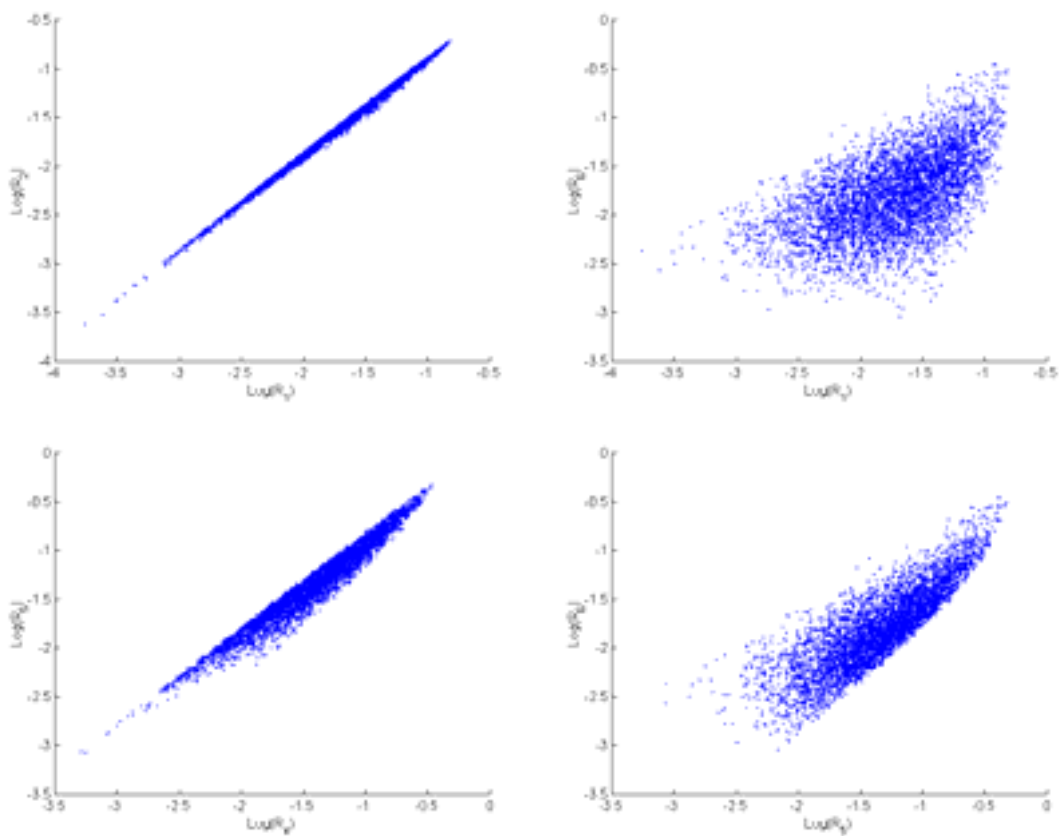


Figura 6.3-4: Scatterogrammi delle riflettanze generate dal data set SeaWiFS-acque di caso II in alcuni canali del sensore.

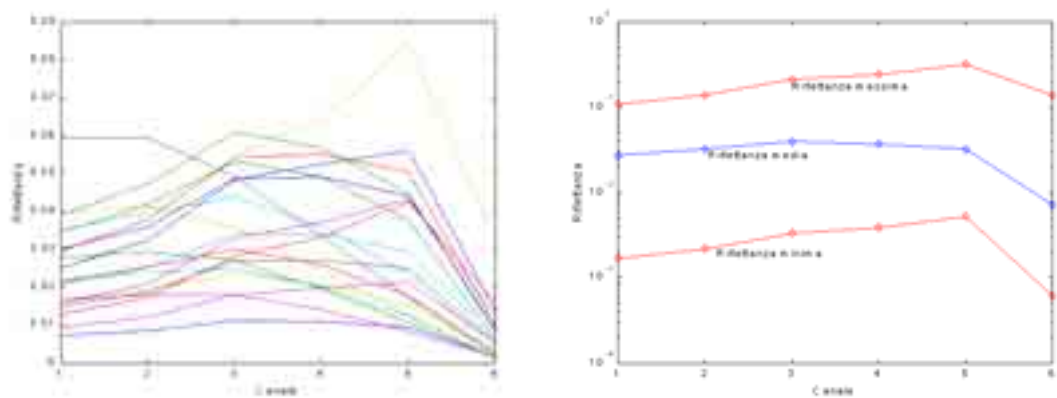


Figura 6.3-5: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima per il data set SeaWiFS acque di caso I-II.

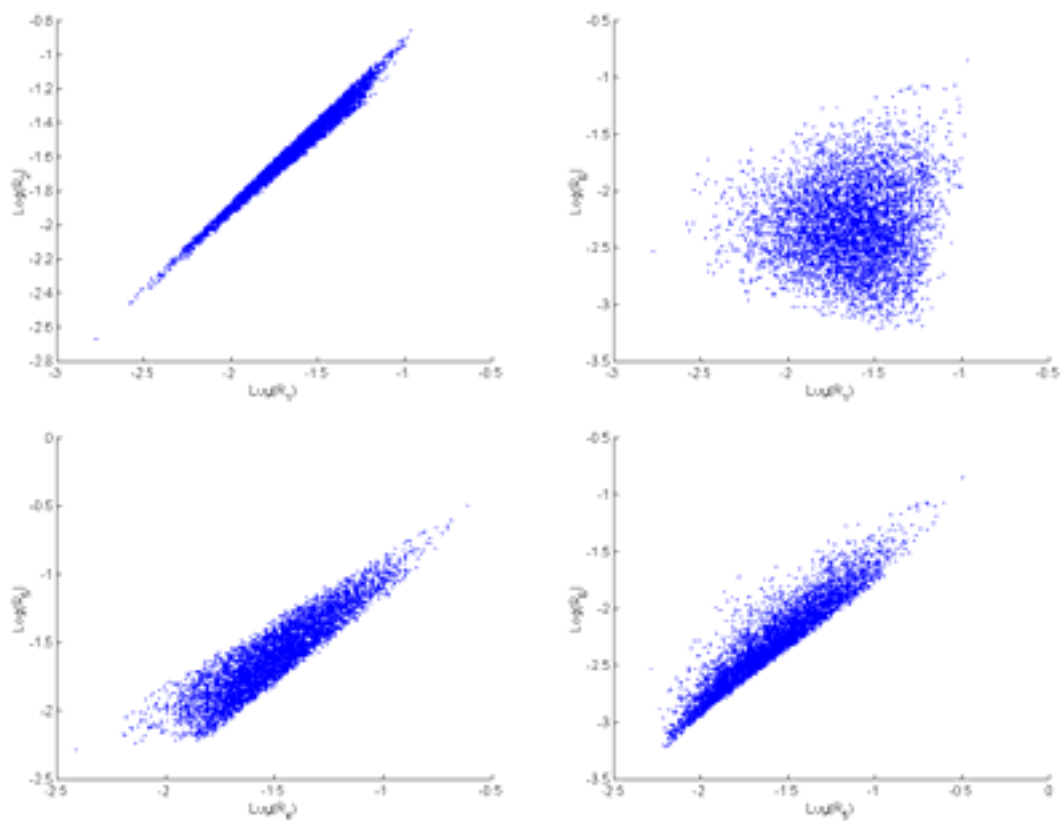


Figura 6.3-6: Scatterogrammi delle riflettanze generate dal data set acque di caso I-II in alcuni canali del sensore SeaWiFS.

*Tabella 6.3-1: Correlazione tra il logaritmo delle riflettanze per i canali del sensore SeaWiFS: data set acque di caso I*

Can./Can.	1	2	3	4	5	6
1	1.0000	0.97870	0.60031	0.17711	-0.14977	-0.34383
2		1.0000	0.74496	0.36187	0.039627	-0.16982
3			1.0000	0.89049	0.69240	0.52044
4				1.0000	0.94433	0.84997
5					1.0000	0.97313
6						1.0000

*Tabella 6.3-2: Correlazione tra il logaritmo delle riflettanze per i canali del sensore SeaWiFS: data set acque di caso II.*

Can./Can.	1	2	3	4	5	6
1	1.0000	0.99859	0.98958	0.97384	0.90395	0.56655
2		1.0000	0.99562	0.98380	0.92317	0.60139
3			1.0000	0.99616	0.95437	0.66605
4				1.0000	0.97666	0.72356
5					1.0000	0.84835
6						1.0000

*Tabella 6.3-3: Correlazione tra il logaritmo delle riflettanze per i canali del sensore SeaWiFS: data set acque di caso I-II*

Can./Can.	1	2	3	4	5	6
1	1.0000	0.99187	0.89240	0.72221	0.38847	0.081234
2		1.0000	0.94216	0.80315	0.49986	0.20085
3			1.0000	0.95601	0.75771	0.50423
4				1.0000	0.91474	0.72842
5					1.0000	0.94008
6						1.0000



### 6.3.2 **Data set di riflettanze per il sensore MOS-B.**

#### 6.3.2.1 MOS-B acque di caso I.

La riflettanza media sulle bande del sensore MOS-B per acque di caso I presenta un comportamento analogo a quello mostrato per le bande del sensore SeaWiFS. La riflettanza in media (Figura 6.3-7) risulta maggiore nelle bande del blu e del blu-verde e inizia a decrescere a partire dal canale 4. Il sensore è dotato di un canale (canale 8) posizionato sul picco di fluorescenza della clorofilla (a circa 685 nm). Dalla Figura 6.3-7 è possibile notare come in generale le diverse curve di riflettanza presentino un picco ben delineato proprio nel canale 8 posizionato sul picco di fluorescenza che risulta quindi rilevabile dal sensore grazie anche all'alta risoluzione spettrale da esso posseduta (pari a 10 nm). Anche nel caso del sensore MOS-B, in maniera analoga al SeaWiFS, la correlazione fra canali adiacenti è elevata (Figura 6.3-8 e Tabella 6.3-4), ma si smorza rapidamente all'aumentare della distanza tra i canali (si ha un minimo assoluto di correlazione pari a circa 0.036 tra la coppia di canali 2 e 7).

#### 6.3.2.2 MOS-B acque di caso II.

Le riflettanze generate dal *data set* per le acque di caso II mostrano i valori più elevati nel canale 5 ovvero alle lunghezze d'onda del verde (Figura 6.3-9). Il comportamento ovviamente rispecchia quello osservato per le bande del sensore SeaWiFS. La correlazione tra le riflettanze nei diversi canali è molto alta anche per canali non adiacenti ed in generale non scende mai al di sotto di 0.5 qualunque sia la coppia di canali considerata (Figura 6.3-10 e Tabella 6.3-5).

#### 6.3.2.3 MOS-B acque di caso I-II.

In Figura 6.3-11 sono riportati alcuni esempi di curve di riflettanza generate tramite il *data set* per acque di caso I-II. In Figura 6.3-12 sono mostrati gli scatterogrammi per tre coppie di canali adiacenti e per la coppia di canali 1 e 6. I valori della riflettanza nelle varie bande mostrano un comportamento analogo a quello riscontrato per le riflettanze medie sulle bande del sensore SeaWiFS; il comportamento sia dal punto di vista della risposta massima che del grado di correlazione tra le varie coppie di canali (Tabella 6.3-6) si pone a livello intermedio tra i due casi precedentemente esaminati.

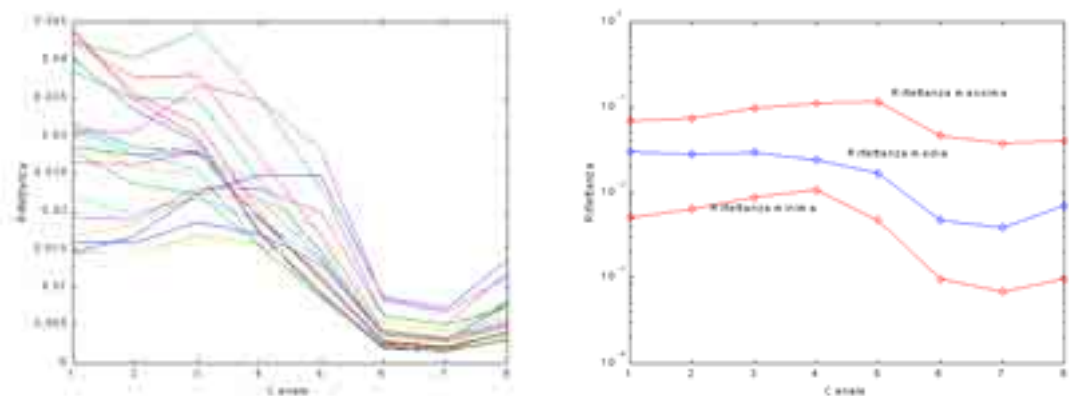


Figura 6.3-7: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso I per il sensore MOS-B.

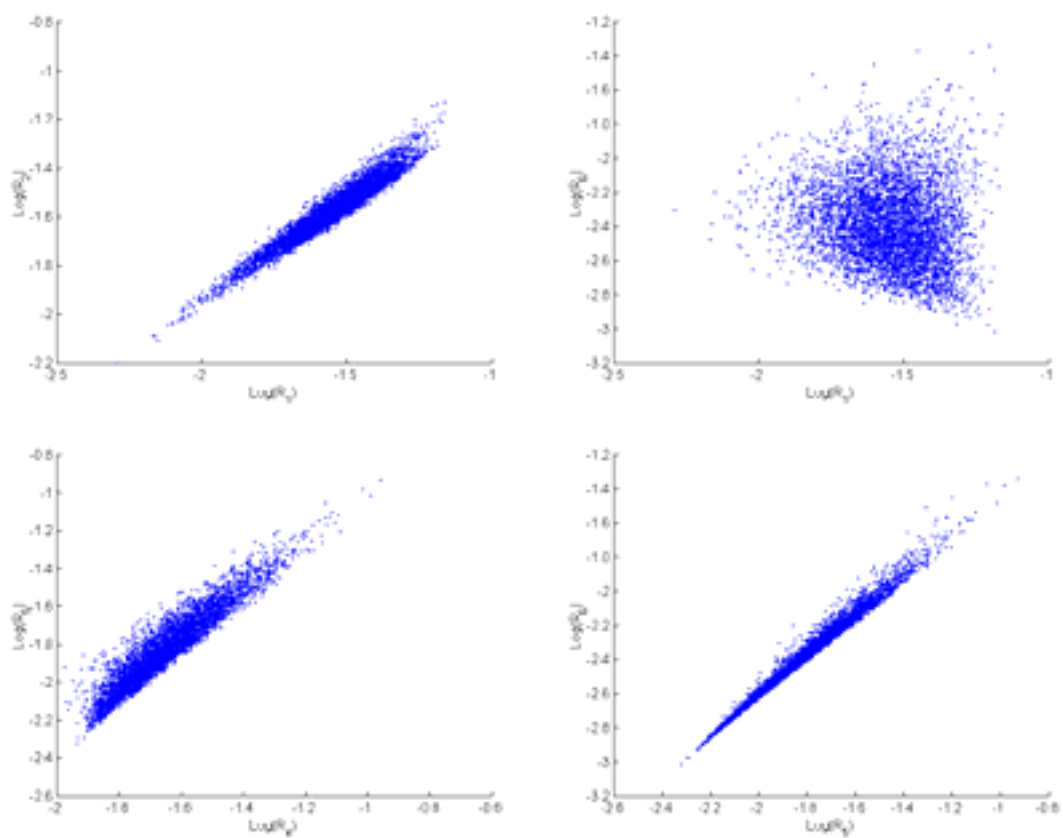


Figura 6.3-8: Scatterogrammi delle riflettanze generate dal data set acque di caso I in alcuni canali del sensore MOS-B.

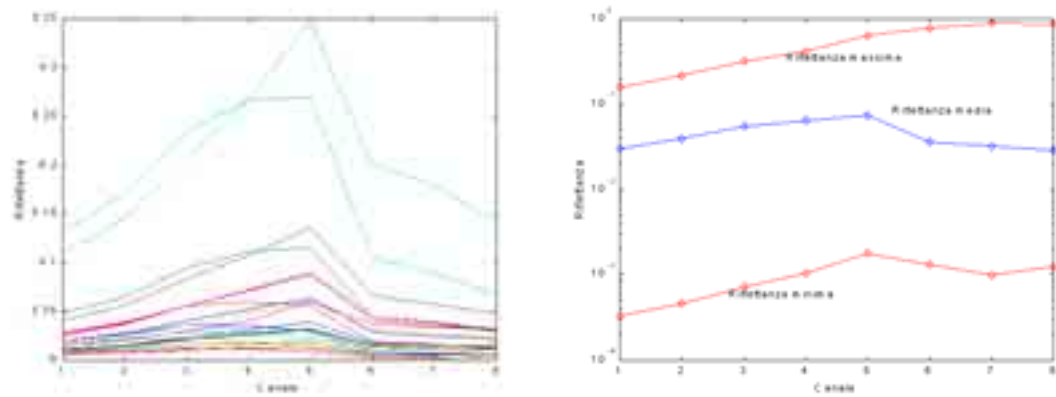


Figura 6.3-9: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso II per il sensore MOS-B.

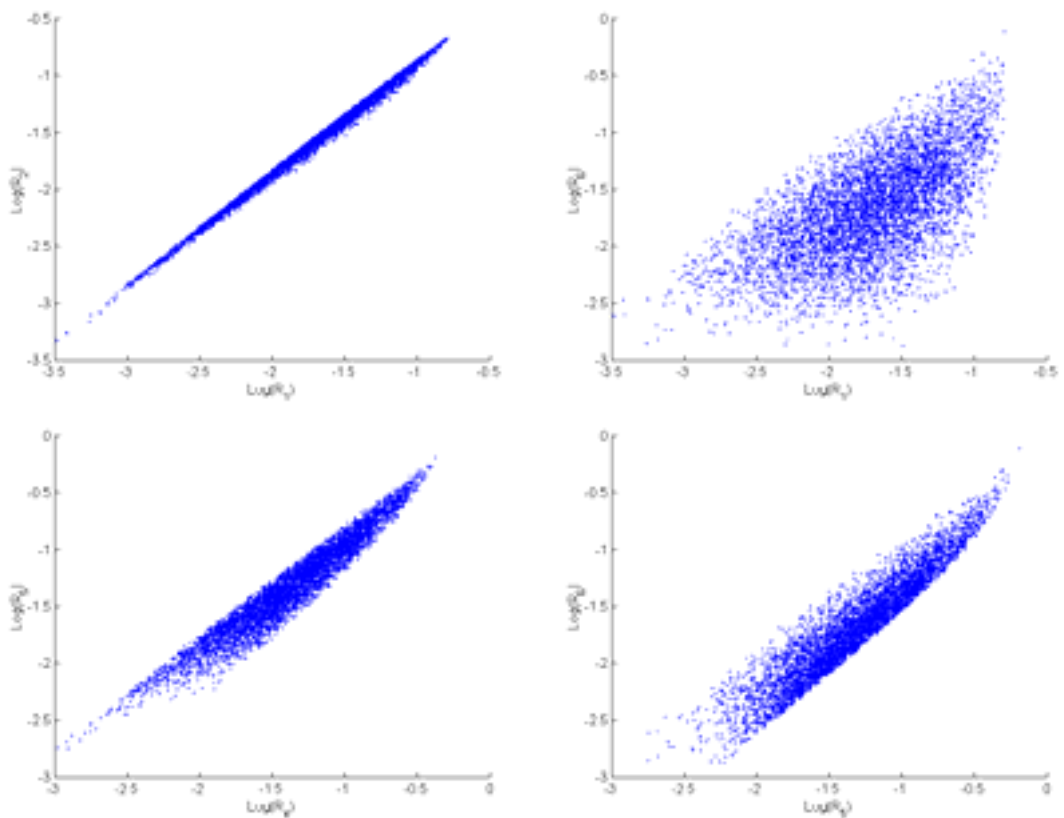


Figura 6.3-10: Scatterogrammi delle riflettanze generate dal data set acque di caso II in alcuni canali del sensore MOS-B.

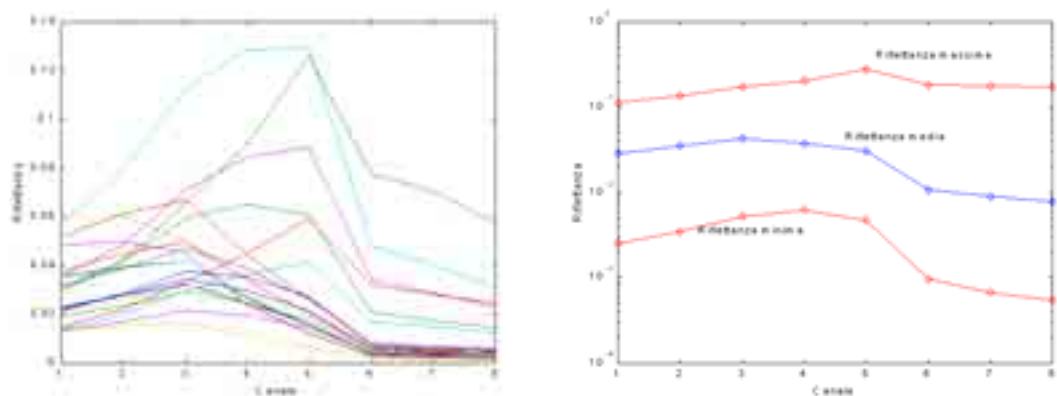


Figura 6.3-11: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso I-II per il sensore MOS-B.

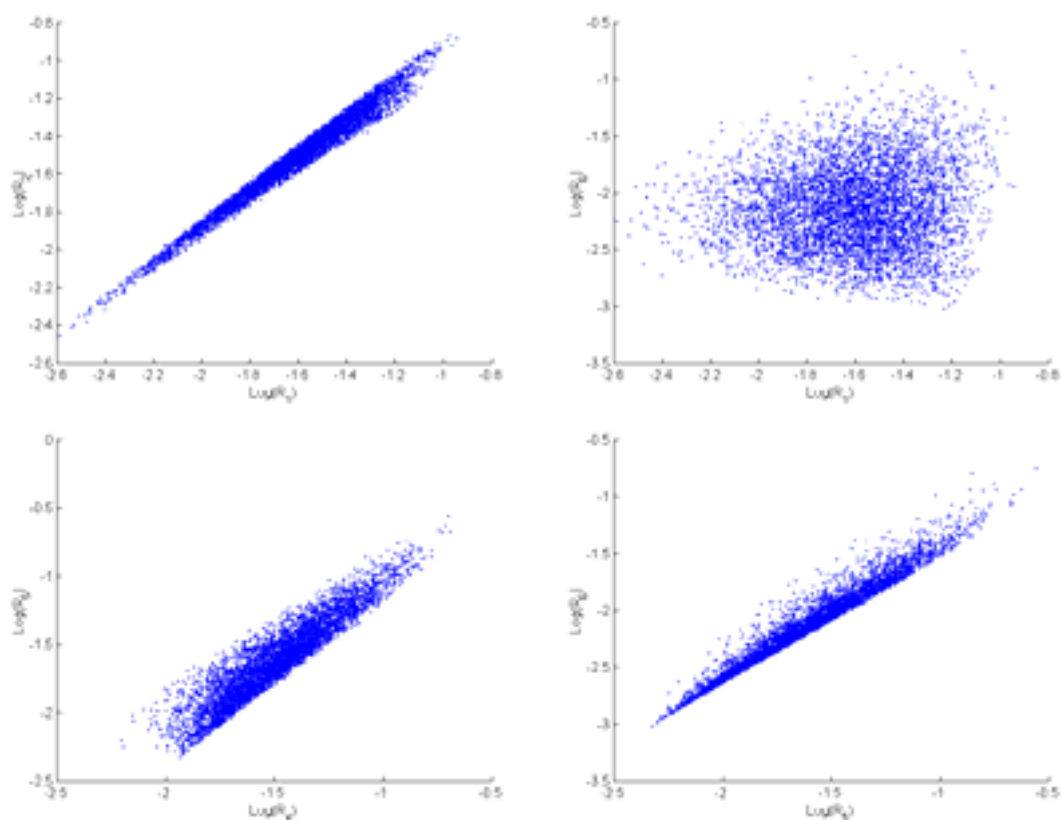


Figura 6.3-12: Scatterogrammi delle riflettanze generate dal data set acque di caso I-II in alcuni canali del sensore MOS-B.

*Tabella 6.3-4: Correlazione tra il logaritmo delle riflettanze per i canali del sensore MOS-B: data set acque di caso I*

Can./Can.	1	2	3	4	5	6	7	8
1	1.0000	0.96753	0.77169	0.26099	-0.059100	-0.16849	-0.18084	-0.40149
2		1.0000	0.89934	0.46817	0.15942	0.048843	0.035866	-0.22620
3			1.0000	0.80633	0.57159	0.47437	0.46310	0.20224
4				1.0000	0.94405	0.89685	0.89164	0.72892
5					1.0000	0.99196	0.99056	0.90586
6						1.0000	0.99988	0.94296
7							1.0000	0.94676
8								1.0000

*Tabella 6.3-5: Correlazione tra il logaritmo delle riflettanze per i canali del sensore MOS-B: data set acque di caso II*

Can./Can.	1	2	3	4	5	6	7	8
1	1.0000	0.99780	0.98965	0.95528	0.85266	0.64547	0.60469	0.50457
2		1.0000	0.99678	0.97126	0.88144	0.68703	0.64787	0.54602
3			1.0000	0.98700	0.91531	0.73873	0.70175	0.60222
4				1.0000	0.96742	0.83146	0.79959	0.71110
5					1.0000	0.94249	0.92176	0.85710
6						1.0000	0.99823	0.97050
7							1.0000	0.97826
8								1.0000

*Tabella 6.3-6: Correlazione tra il logaritmo delle riflettanze per i canali del sensore MOS-B: data set acque di caso I-II*

Can./Can.	1	2	3	4	5	6	7	8
1	1.0000	0.98830	0.90535	0.58282	0.26035	0.084384	0.065831	-0.015766
2		1.0000	0.95915	0.69784	0.40032	0.22826	0.20987	0.12643
3			1.0000	0.87088	0.63774	0.48228	0.46521	0.38705
4				1.0000	0.93134	0.84022	0.82891	0.77809
5					1.0000	0.97852	0.97387	0.94990
6						1.0000	0.99973	0.99044
7							1.0000	0.99225
8								1.0000

### 6.3.3 **Data set di riflettanze per il sensore MERIS.**

Le caratteristiche dei *data set* di riflettanze medie sulle bande del sensore MERIS generati per le acque di caso I, caso II e caso I-II sono riassunte di seguito nelle figure a partire dalla Figura 6.3-13 alla Figura 6.3-18 e nelle tabelle dalla Tabella 6.3-7 alla Tabella 6.3-9. Quanto detto per i casi precedenti, relativi al sensore SeaWiFS e MOS-B, vale anche per il sensore MERIS. Le caratteristiche mostrate dai valori di riflettanza al variare del tipo di acqua preso in considerazione (in particolar modo per quanto riguarda il comportamento della correlazione tra coppie di canali) assumono quindi valore generale per i tre sensori presi in esame. In particolare per acque di caso I i valori di riflettanza sono più alti, come è logico aspettarsi, nei canali del blu e del blu-verde; la correlazione è alta per canali adiacenti, ma decresce rapidamente in valore assoluto (rispetto a quanto riscontrato per le acque di caso II) all'aumentare della distanza tra i canali (distanza in termini di lunghezza d'onda); si osservano inoltre casi di coppie di canali con coefficiente di correlazione negativo. Per quanto riguarda le acque di caso II i canali in cui la riflettanza assume i valori più elevati sono in genere quelli posizionati nella banda del verde ed i valori di correlazione sono sempre positivi e non scendono mai al di sotto di 0.5 fatto questo che rende il problema dell'inversione del modello tramite algoritmi multilineari mal condizionato. Per le acque di caso I-II le riflettanze nei vari canali hanno un comportamento intermedio fra i casi precedenti sia per quanto riguarda l'intensità che per il valore assunto dalla correlazione fra le varie coppie di canali.

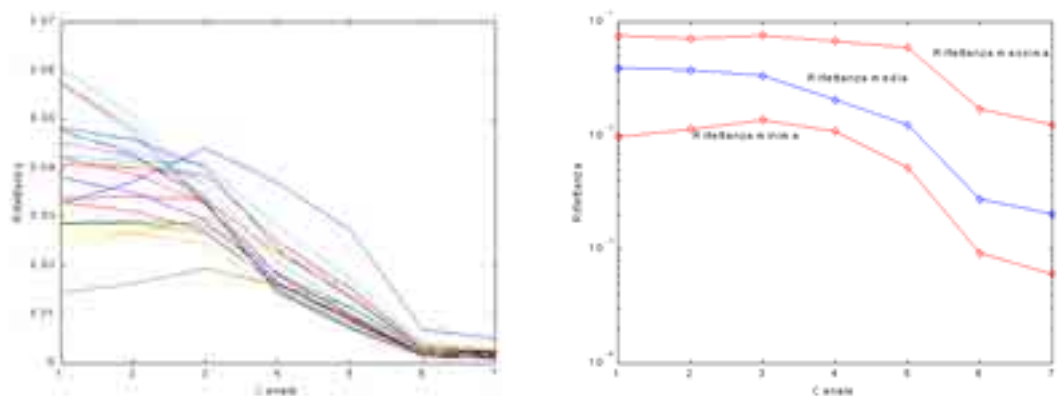


Figura 6.3-13: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso I per il sensore MERIS.

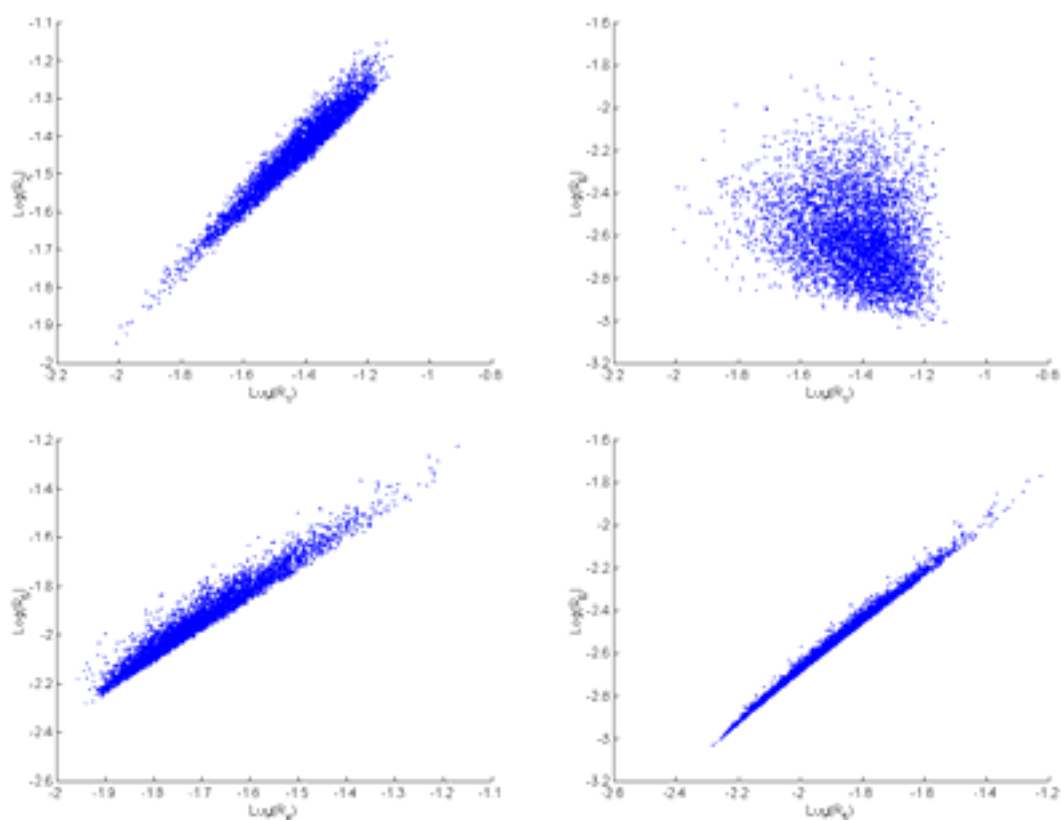


Figura 6.3-14: Scatterogrammi delle riflettanze generate dal data set acque di caso I in alcuni canali del sensore MERIS.

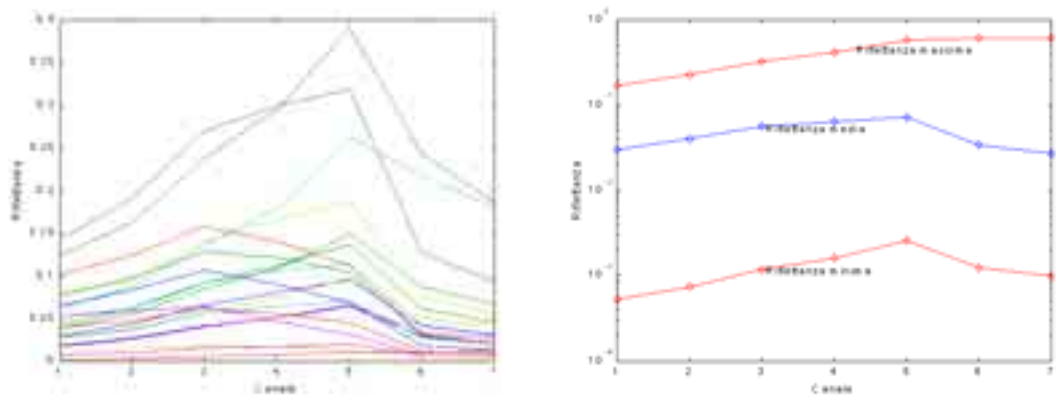


Figura 6.3-15: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso II per il sensore MERIS.

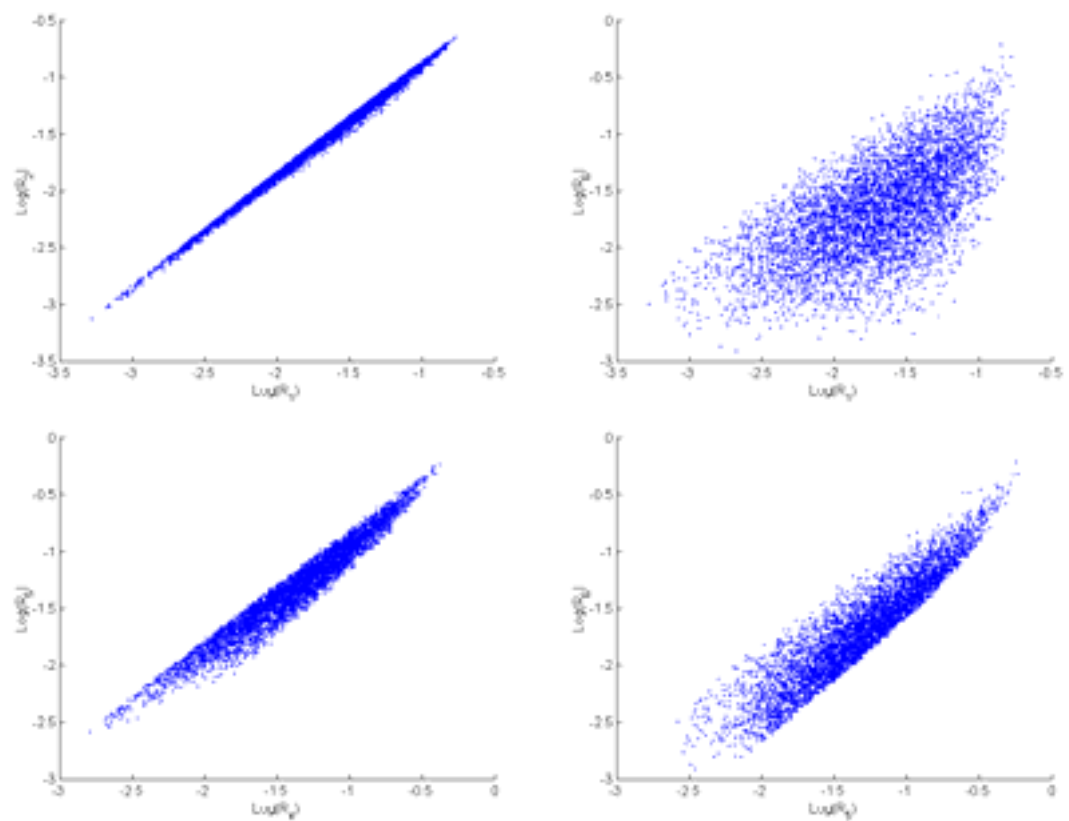


Figura 6.3-16: Scatterogrammi delle riflettanze generate dal data set acque di caso II in alcuni canali del sensore MERIS.



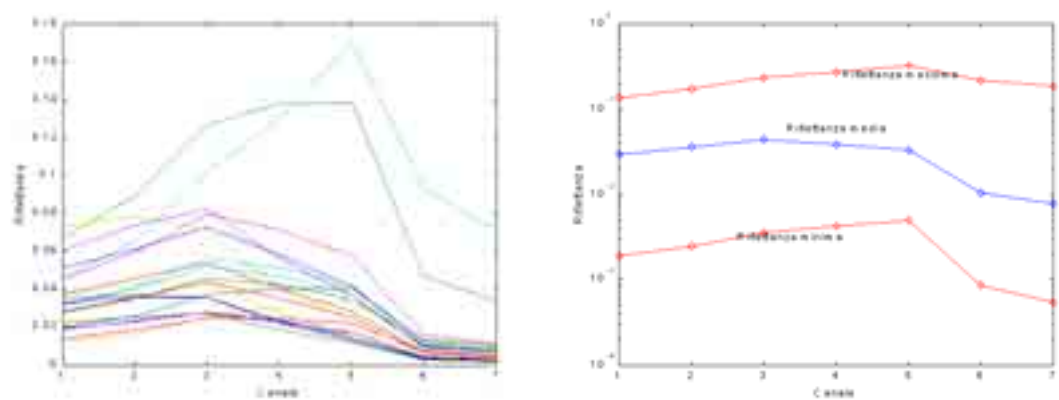


Figura 6.3-17: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set acque di caso I-II per il sensore MERIS.

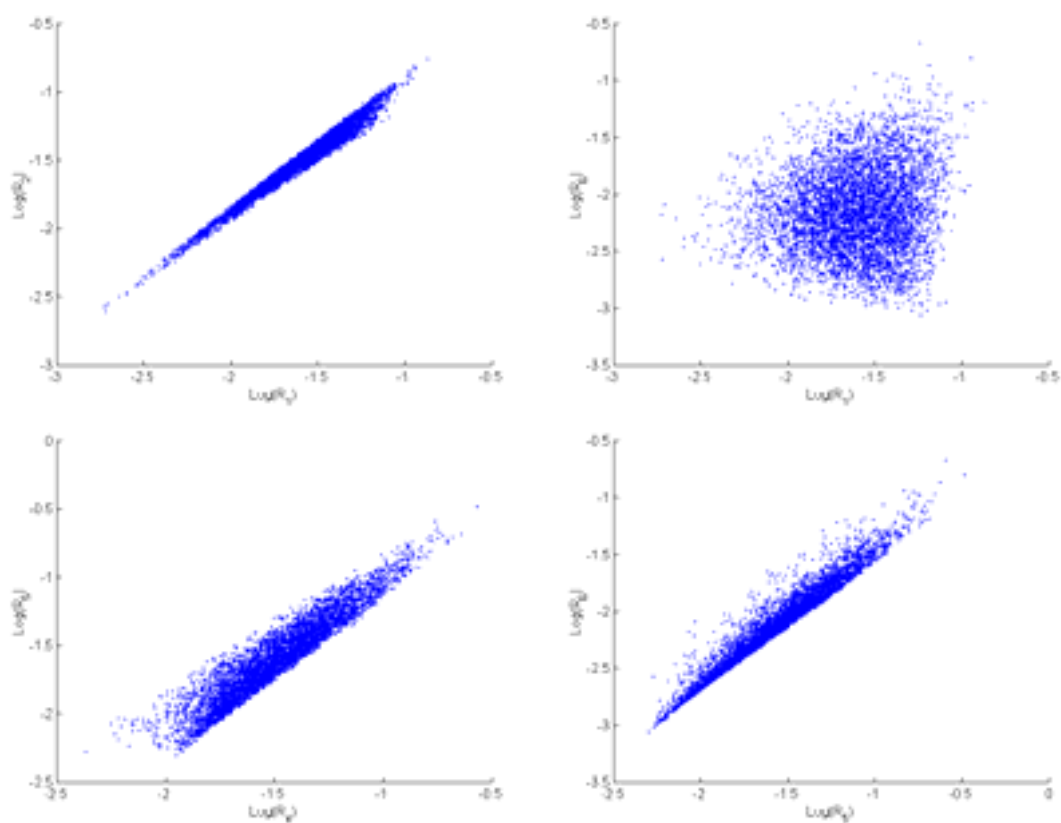


Figura 6.3-18: Scatterogrammi delle riflettanze generate dal data set acque di caso I-II in alcuni canali del sensore MERIS.

Tabella 6.3-7: Correlazione tra il logaritmo delle riflettanze per i canali del sensore MERIS: data set acque di caso I

Can./Can.	1	2	3	4	5	6	7
1	1.0000	0.97086	0.63159	0.020242	-0.19420	-0.26732	-0.31923
2		1.0000	0.79286	0.24309	0.028401	-0.047143	-0.10591
3			1.0000	0.78205	0.62768	0.56624	0.51567
4				1.0000	0.97548	0.95521	0.93612
5					1.0000	0.99685	0.98969
6						1.0000	0.99698
7							1.0000

Tabella 6.3-8: Correlazione tra il logaritmo delle riflettanze per i canali del sensore MERIS: data set acque di caso II.

Can./Can.	1	2	3	4	5	6	7
1	1.0000	0.99793	0.98911	0.95778	0.87005	0.63797	0.57443
2		1.0000	0.99632	0.97297	0.89668	0.67909	0.61727
3			1.0000	0.98898	0.93034	0.73489	0.67628
4				1.0000	0.97388	0.82131	0.77003
5					1.0000	0.92644	0.89008
6						1.0000	0.99558
7							1.0000

Tabella 6.3-9: Correlazione tra il logaritmo delle riflettanze per i canali del sensore MERIS: data set acque di caso I-II.

Can./Can.	1	3	3	4	5	6	7
1	1.0000	0.98802	0.89443	0.61019	0.32559	0.12966	0.094684
2		1.0000	0.95234	0.72327	0.46392	0.27447	0.23983
3			1.0000	0.89849	0.70732	0.54280	0.51187
4				1.0000	0.94412	0.84782	0.82786
5					1.0000	0.97370	0.96471
6						1.0000	0.99912
7							1.0000

## 6.4 Descrizione di un *data set* di misure in situ di parametri otticamente attivi

Nel presente paragrafo sarà analizzato un *data set* di misure in situ relativo al Mare di Alboran che sarà usato per generare un insieme di algoritmi di stima degli OAPs per un caso reale. Il paragrafo è organizzato in un primo sottoparagrafo in cui si descrive dal punto di vista oceanografico il bacino ed in un secondo sottoparagrafo in cui si analizza dal punto di vista statistico il *data set* di misure in situ di OAPs e le riflettanze tramite il modello.

### 6.4.1 Caratteristiche oceanografiche del Mare di Alboran.

Le strutture oceanografiche di mesoscala sono spesso il risultato dell'instabilità di regioni frontali che marciano i confini tra due o più masse d'acqua di differente origine e densità. Tali processi frontogenetici, la cui struttura è di solito visibile in dati altimetrici e dati telerilevati nell'infrarosso, sono caratterizzati da moti orizzontali e verticali che influenzano i processi biologici. Una serie di immagini satellitari delle zone di interesse fanno da complemento alle misurazioni in situ da un punto di vista della sinotticità e permettono una valutazione della variabilità a lungo termine del sistema oceanografico sotto osservazione. L'analisi di serie temporali di dati di temperatura superficiale (SST: Sea Surface Temperature), stimata da dati telerilevati nell'infrarosso termico, e di dati altimetrici risulta essere efficace per investigare la variabilità spaziale e temporale delle correnti superficiali permettendo una migliore comprensione della fisica dei sistemi di circolazione a mesoscala. D'altra parte, è possibile investigare direttamente la biologia di tali sistemi osservando le variazioni spettrali del mare nella banda di lunghezze d'onda del visibile tramite un radiometro con appropriate caratteristiche radiometriche e spettrali. Il sensore SeaWiFS ad esempio, lanciato in Agosto 1997 ha dato alla comunità scientifica l'opportunità di disporre di uno strumento con tali caratteristiche. Risulta quindi possibile combinare le informazioni derivate da satelliti relative alla fisica con quelle relative alla biologia. Tale approccio sembra essere estremamente promettente per analizzare gli effetti della circolazione sull'attività biologica [28].

Un'area di forte attività di mesoscala dove tale approccio può essere applicato è il Mare di Alboran, nel Mediterraneo occidentale, caratterizzato da un intenso jet associato al flusso di acqua Atlantica che entra nel Mediterraneo attraverso lo Stretto di Gibilterra, il quale di solito forma uno o due vortici anticiclonici (Alboran Gyres) e un fronte con forte gradiente di densità tra la Spagna e l'Africa (Almeria-Oran front).

Le campagne di osservazione del progetto MAST-III OMEGA, tra Ottobre e Dicembre del 1996, hanno fornito un *data set* di misurazioni in situ (di parametri fisici quali la temperatura superficiale e di parametri biologici come la concentrazione di clorofilla-a e di feofitina-a) di considerevole dimensione che ha mostrato la forte relazione tra i moti verticali delle masse d'acqua e l'attività biologica dell'area. La scala, il numero e l'accuratezza delle misurazioni sono cruciali per avere un'idea più fedele possibile della circolazione tridimensionale del Mare di Alboran. Il *data set* rappresenta un buon riferimento da cui partire per definire gli algoritmi di stima dei parametri otticamente attivi dell'acqua marina per la zona del Mare di Alboran. Gli algoritmi definiti permetteranno di integrare le conoscenze sulla circolazione acquisite dall'analisi dei dati in situ con quelle ottenibili dai dati satellitari che hanno la caratteristica di essere sinottici e non limitati nel tempo. Potrà essere possibile inoltre l'elaborazione combinata di dati di SST, ad esempio stimati da dati telerilevati dal radiometro AVHRR, e di colore marino acquisiti durante un ampio

intervallo di tempo in modo da analizzare la variabilità spaziale e temporale dei fenomeni fisici e biologici che caratterizzano il bacino [28].

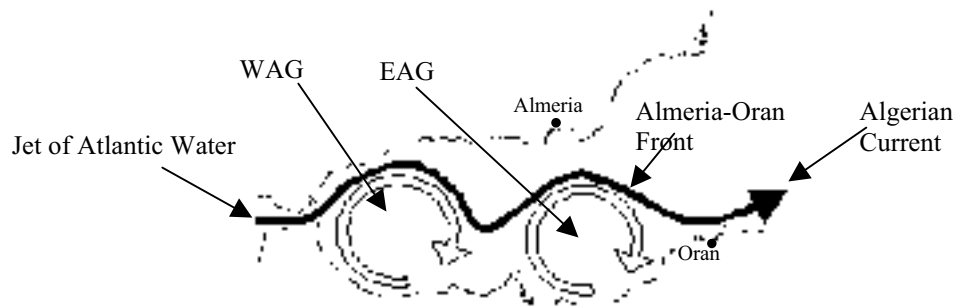


Figura 6.4-1: schema della circolazione superficiale del Mare di Alboran.

Il Mare di Alboran, situato tra la Spagna e l'Algeria Nel Mediterraneo Occidentale, a lungo è stata una regione di particolare interesse per gli oceanografi [29] perchè mostra caratteristiche tipiche della mesoscala. Uno schema della circolazione superficiale del bacino è mostrato in Figura 6.4-1. Il jet di acqua Atlantica (AW: Atlantic Water) che entra nel bacino dallo Stretto di Gibilterra può formare, dipendentemente da un certo numero di fattori i quali non sono stati ancora compresi completamente, uno o due vortici anticiclonici, il Western Alboran Gyre (WAG) e l'Eastern Alboran Gyre (EAG), prima di incanalarsi nella corrente Algerina lungo le coste dell'Africa. Durante questo transito la AW subisce dei mescolamenti con le masse d'acqua circostanti diventando acqua Atlantica Modificata (MAW: *Modified Atlantic Water*). Al confine nord orientale del EAG, la MAW converge con la più densa e salina acqua Mediterranea di superficie (MSW: *Mediterranean Surface Water*), la quale fluisce generalmente verso sud ovest lungo la costa Spagnola. L'incontro delle due masse d'acqua origina un fronte con forte gradiente di densità, il fronte Almeria-Orano (AOF: *Almeria-Oran Front*) il quale è normalmente orientato lungo la direzione nord-ovest sud-est [30]. Un'altra caratteristica del Mare di Alboran è il cosiddetto up-welling, ovvero trasporto in superficie, che avviene lungo le coste della Spagna nel lato nord del WAG, di acqua ricca di nutrienti dalle profondità del bacino; tale fenomeno è spesso visibile nelle mappe di SST e di concentrazione di clorofilla data la maggiore crescita del fitoplancton in tali zone. L'acqua portata in superficie è trascinata dalla circolazione superficiale mescolandosi con la MAW per formare la cosiddetta Atlantic-Mediterranean Interface Water (A-MIW, [31]). I due vortici anticiclonici sono semi-permanenti: in alcuni casi l'EAG ed tal volta anche il WAG, possono collassare per un periodo pari a qualche settimana in maniera tale che la circolazione superficiale diventa principalmente zonale e l'AOF assume approssimativamente un orientamento ovest-est. Nel recente lavoro di Viudez et al. [29], è stato osservato che i vortici possono migrare verso est in maniera tale che l'EAG si mescola e eventualmente scompare dentro il bacino Algerino, il vecchio WAG diventa il nuovo EAG ed il nuovo WAG si forma tramite la circolazione del jet di acqua Atlantica. Durante il periodo di transizione un sistema a tre vortici è stato osservato con una scala temporale di alcuni giorni.

#### 6.4.2 Statistiche del *data set* di misure in situ di OAPs

Le misurazioni dei tre parametri  $C$ ,  $X$  ed  $Y$  sono state acquisite nell'Ottobre 1996 durante una crociera oceanografica effettuata nell'ambito del progetto OMEGA a bordo del vascello Bio-Hespérides. I dati sono stati acquisiti durante tre successive campagne di misura a partire dal 1 Ottobre, 1996 al 11 Ottobre, 1996 lungo le tracce mostrate in Figura 2.3-3 dove sono riportate anche le date di ogni campagna. Come si osserva dalla Figura 2.3-3 la zona interessata dal campionamento è situata nei pressi dello Stretto di Gibilterra tra 6° E e 4° di longitudine.

I valori degli OAPs sono stati acquisiti in continua all'interno dello strato superficiale di acqua tramite una sonda multi-parametrica connessa al sistema di pompaggio del vascello. Le misure relative alla concentrazione di clorofilla e di sostanza gialla sono state effettuate tramite metodi cosiddetti fluorometrici [32] mentre la stima della concentrazione di particolato solido in sospensione è stata ottenuta attraverso il coefficiente di attenuazione del particolato a 660 nm tramite un trasmissiometro avente una cella di propagazione di 25 cm di lunghezza [32]. Tramite opportuni coefficienti ottenuti per regressione [32] i valori acquisiti sono stati trasformati per ottenere un insieme di triplette  $C$ ,  $X$ ,  $Y$ .

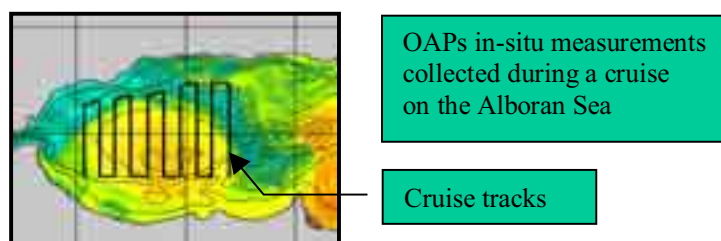


Figura 6.4-2: Omega cruise tracks and surveys date.

Tabella 6.4-1: date delle campagne oceanografiche condotte nel Mare di Alboran.

Survey 1	October 1st 1996	October 4th 1996
Survey 2	October 6th 1996	October 9th 1996
Survey 3	October 9th 1996	October 11th 1996

Le statistiche del logaritmo degli OAPs sono riassunte in Tabella 6.4-2. Durante la crociera sono stati acquisiti circa novemila campioni che formano un *data set* estremamente ampio. I valori dei coefficienti di correlazione tra i parametri risultano pari a  $\rho_{\log C \log X} = 0.6943$ ,  $\rho_{\log C \log Y} = 0.7228$  e  $\rho_{\log G \log Y} = 0.355$ . Globalmente le statistiche in termini di valor medio, deviazione standard e correlazione, sono tipiche di acque oligotrofiche ovvero i valori forniti in Tabella 6.4-2 sono consistenti con ciò che ci aspetteremmo di trovare per le acque di caso I.

La distribuzione degli OAPs, come si nota dalla Figura 6.4-3, Figura 6.4-4 e Figura 6.4-5 in cui sono mostrati gli istogrammi dei valori, non può essere modellata tramite una densità di probabilità log-normale. Le differenti mode degli istogrammi corrispondono ad altrettante sottodistribuzioni (probabilmente di tipo log-normale) che caratterizzano la variabilità statistica dei parametri otticamente attivi in differenti aree della zona interessata al campionamento (ad esempio è plausibile aspettarsi una diversa distribuzione statistica delle concentrazioni in zone in cui è presente acqua atlantica mescolata ad acqua mediterranea rispetto a zone in cui è presente la sola acqua mediterranea). Tipicamente durante

l'arco dell'anno l'interno del WAG è caratterizzato da bassi valori della concentrazione di clorofilla e feofitina (corrispondenti alla moda più bassa dell'istogramma del  $\text{Log}(C)$  in Figura 6.4-3) al contrario ai bordi si trovano i valori più alti dove i moti verticali delle masse d'acqua sono più intensi e quindi è maggiore l'apporto di nutrienti dalle profondità del bacino che favoriscono la crescita del fitoplancton.

Tabella 6.4-2: statistiche del logaritmo degli OAPs.

	$\text{Log}(C)$	$\text{Log}(Y)$	$\text{Log}(X)$
Massimo	-0.96850	-1.5281	-2.3675
Minimo	0.53244	-1.2638	-0.75084
Num. punti	9027.0	9027.0	9027.0
Media	-0.49100	-1.4195	-1.7101
Mediana	-0.56830	-1.4240	-1.7183
RMS	0.57078	1.4207	1.7408
Dev. St.	0.29105	0.057521	0.32529

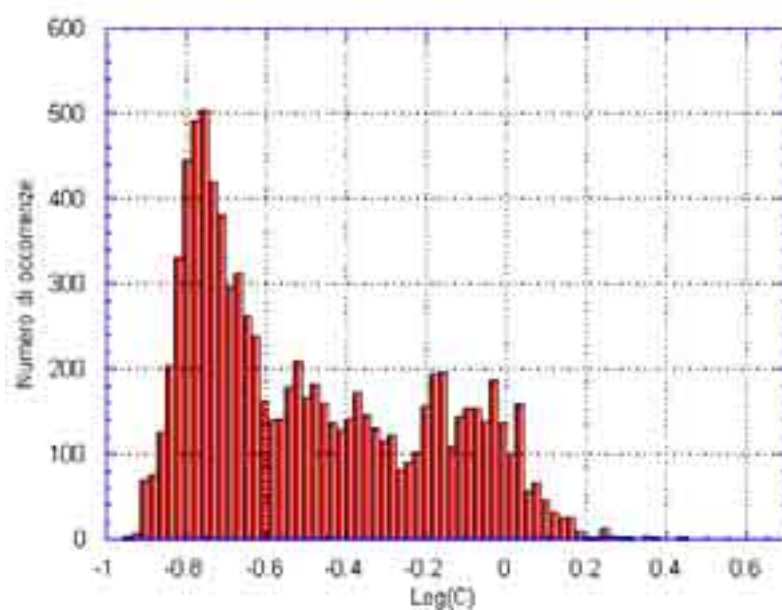


Figura 6.4-3: istogramma del logaritmo della concentrazione  $C$ .

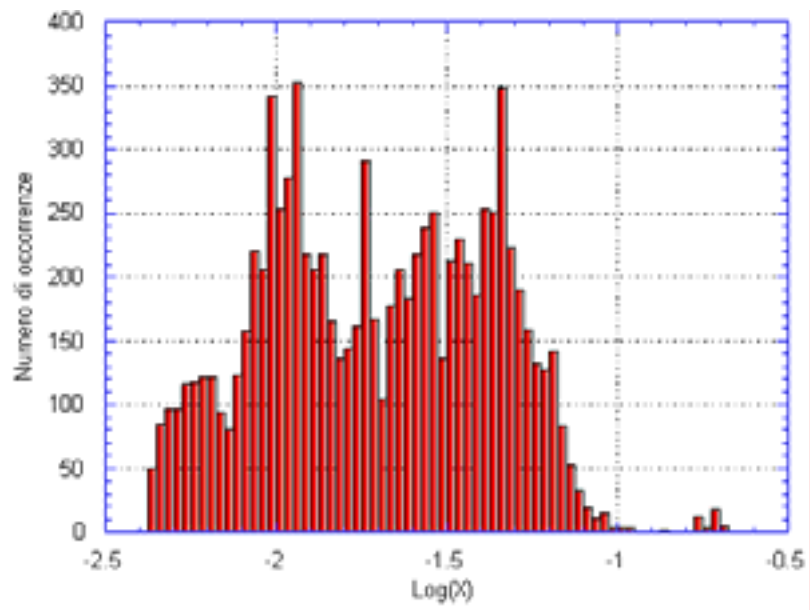


Figura 6.4-4: istogramma del logaritmo di  $X$ .

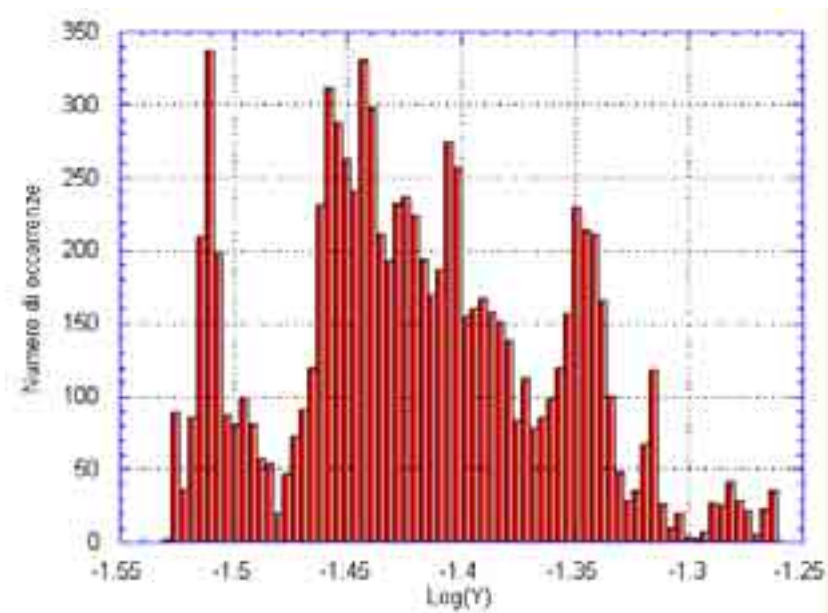


Figura 6.4-5: istogramma del logaritmo di  $Y$ .

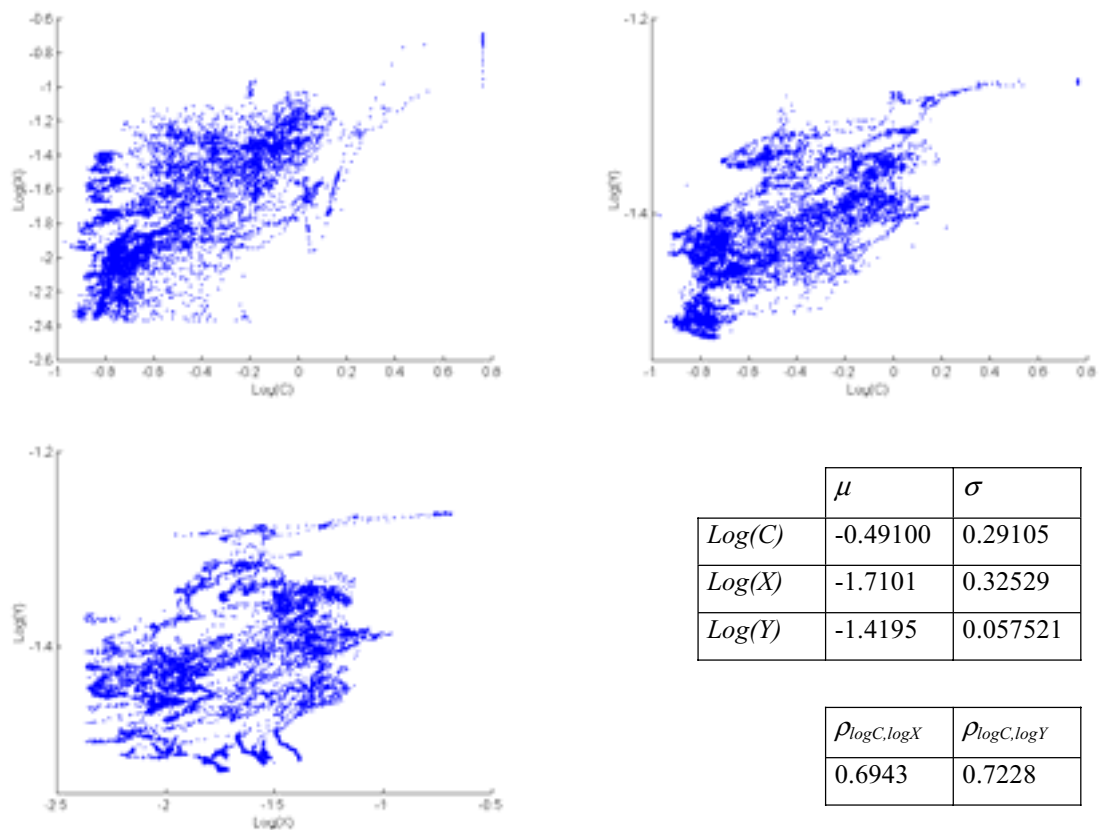


Figura 6.4-6: Scatterogrammi dei logaritmi dei costituenti otticamente attivi per il data set di misure in situ prelevate nel Mare di Alboran.

In Figura 6.4-6 sono riportati gli scatterogrammi dei valori dei logaritmi dei parametri otticamente attivi del *data set* e riassunte le statistiche significative dal punto di vista della generazione delle riflettanze dal modello. Nella successiva Figura 6.4-7 analogamente ai *data set* simulati precedentemente descritti, si trovano graficati alcuni valori della riflettanza media per ciascuna delle bande del SeaWiFS (sulla sinistra) e i valori della riflettanza media, massima e minima su tutte le 5000 curve di riflettanza generate. I grafici mostrano un comportamento analogo a quello riscontrato per i valori di riflettanza calcolati tramite il *data set* degli OAPs simulati per le acque di caso I a conferma di quanto precedentemente affermato. Una ulteriore conferma di ciò è fornita dall'andamento della correlazione tra le coppie di canali (Figura 6.4-8 e Tabella 6.4-3).



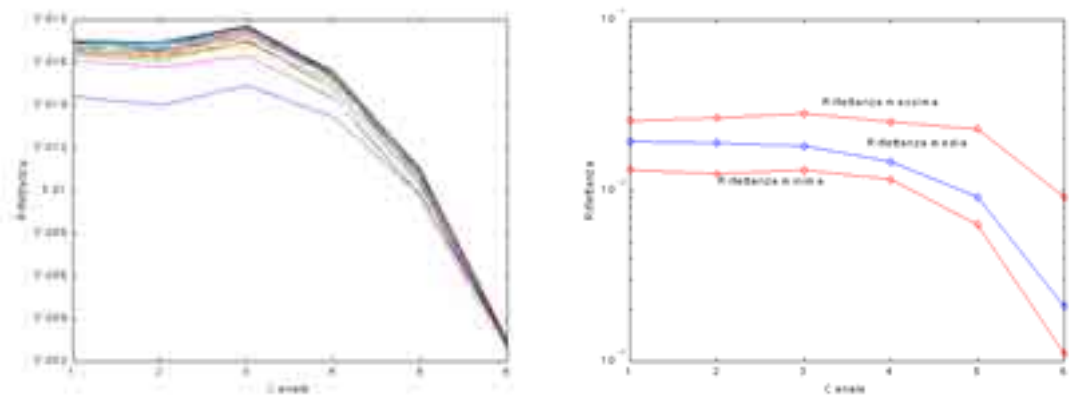


Figura 6.4-7: Esempio di curve di riflettanza e curva di riflettanza media, massima e minima generate dal data set di misure insitu prelevate sul Mare di Alboran per il sensore SeaWiFS.

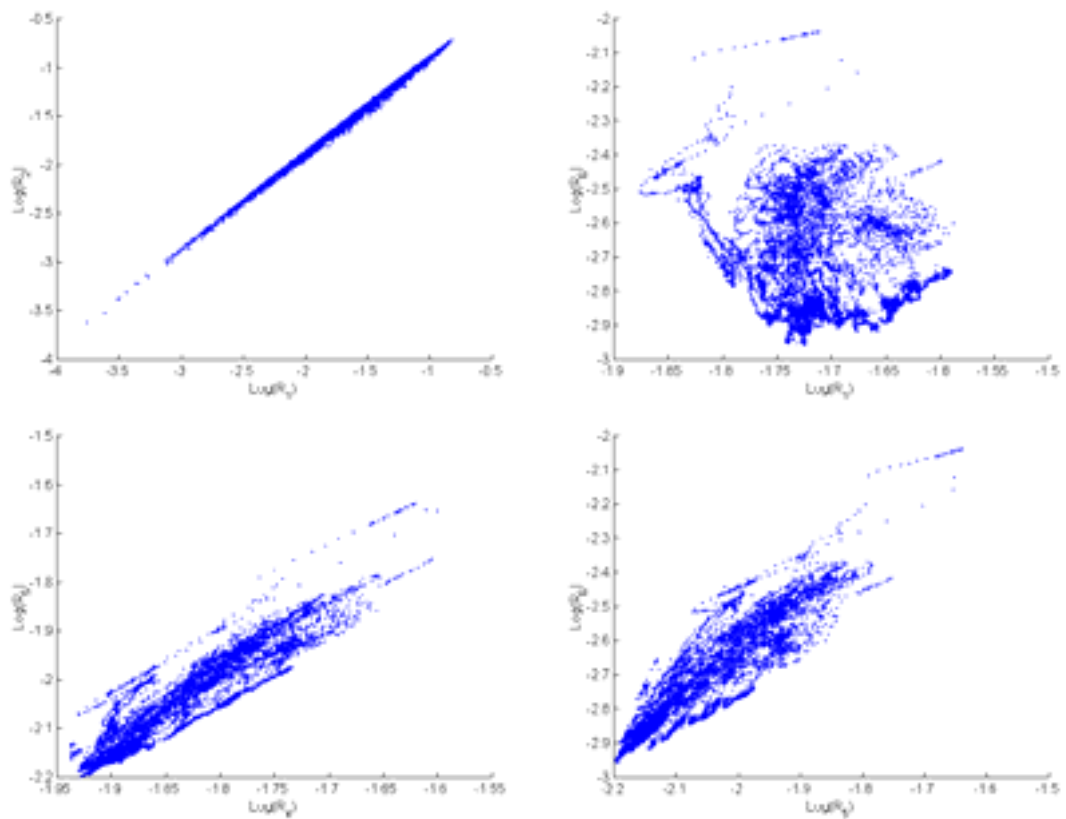


Figura 6.4-8: Scatterogrammi delle riflettanze generate dal data set di misure in situ prelevate sul Mare di Alboran in alcuni canali del sensore SeaWiFS.

*Tabella 6.4-3: Correlazione tra il logaritmo delle riflettanze per i canali del sensore SeaWiFS: data set Mare di Alboran*

Can./Can.	1	2	3	4	5	6
1	1.0000	0.99203	0.81801	0.49806	0.15714	-0.20332
2		1.0000	0.84023	0.51990	0.17457	-0.20337
3			1.0000	0.89902	0.67783	0.34702
4				1.0000	0.93095	0.71972
5					1.0000	0.92127
6						1.0000

---

# 7 Capitolo 5

## 8 Risultati sperimentali

In quest'ultimo capitolo saranno mostrati alcuni dei risultati ottenuti applicando le tecniche di inversione proposte nel capitolo 3 ai *data set* simulati di OAPs e di riflettanze descritti nel capitolo 4. Le stesse tecniche sono state inoltre utilizzate per generare degli algoritmi di stima degli OAPs per il sito di sperimentazione costituito dal bacino del Mare di Alboran, in cui è stato acquisito il *data set* di misurazioni descritto sempre nel precedente capitolo 4. Gli algoritmi generati sono stati utilizzati per stimare gli OAPs da dati reali acquisiti dal sensore SeaWiFS sul sito preso in considerazione. A tale proposito sarà schematizzata la catena di elaborazione completa dei dati telerilevati e sarà fornito un metodo di correzione atmosferica dei dati che si basa sul modello atmosferico descritto nel capitolo 2 e su alcune ipotesi semplificative che permetteranno la stima del carico di aerosol presente nelle immagini considerate.

Il capitolo è organizzato in un primo paragrafo in cui sono riportati i risultati ottenuti dai dati simulati. Per ogni sensore e per i diversi tipi di acque saranno riassunte le prestazioni degli algoritmi di stima in termini di coefficiente di correlazione e di errore quadratico medio privilegiando maggiormente gli algoritmi per stimare il  $\text{Log}(C)$  che risulta il parametro biologicamente più importante. Il secondo paragrafo del capitolo riguarda gli algoritmi di stima degli OAPs per il sensore SeaWiFS ottenuti dai dati rilevati in situ nel Mare di Alboran. Oltre che la serie di algoritmi di tipo PCA e RBF-NN in questa sezione saranno considerati anche algoritmi a rapporto di bande e polinomiali. Alcuni risultati dell'applicazione degli algoritmi proposti a dati reali provenienti dal sensore SeaWiFS, saranno infine riportati e discussi.

### 8.1 Generazione degli algoritmi per la stima degli OAPs da *data set* simulati

Questo paragrafo è dedicato alla presentazione dei risultati ottenuti applicando le tecniche di inversione PCA e RBF-NN ai dati simulati. Lo scopo principale di questo paragrafo è quello di cercare di effettuare un'analisi comparata delle prestazioni delle diverse metodologie di inversione proposte e cercare di individuare la più adatta al tipo di acqua di volta in volta considerato. L'analisi effettuata sui singoli sensori inoltre permetterà di confrontare le prestazioni di questi in maniera tale da individuare le caratteristiche spettrali più adatte a stimare gli OAPs nelle differenti situazioni prese in considerazione. I grafici e le tabelle prestazionali riportate non saranno relative a tutti gli algoritmi che è possibile generare per ogni sensore tramite i *data set* a disposizione. In generale, come già anticipato nell'introduzione del capitolo, sarà privilegiato l'algoritmo per la stima del  $\text{Log}(C)$  o comunque il parametro che a seconda del tipo di acqua preso in considerazione ha la maggiore importanza per le applicazioni. Negli altri casi saranno effettuate considerazioni generali sulle prestazioni raggiunte dai singoli algoritmi.

### 8.1.1 Sensore SeaWiFS: algoritmi multilineari

In questo sottoparagrafo sono riportati alcuni dei risultati ottenuti applicando il metodo di inversione PCA ai *data set* generati per il sensore SeaWiFS. Per ogni *data set* saranno riportati: a) i risultati relativi all'analisi PCA effettuata sulle riflettanze in modo da stabilire la dimensionalità dei dati e quindi il numero di componenti principali sufficienti a stimare i parametri otticamente attivi e b) i risultati delle regressioni multilineari effettuate sui dati proiettati nel sistema di riferimento degli autovettori come spiegato nel capitolo 3.

#### 8.1.1.1 Algoritmi per acque di caso I

In Figura 8.1-9 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del *data set* di riflettanze per le acque di caso I. La somma cumulata è rapportata alla somma totale degli autovalori e rappresenta quindi la percentuale di varianza totale spiegata dalle prime  $n$  componenti principali con  $n=1, 2, \dots, N_c$  dove  $N_c$  è il numero di canali del sensore. Dal grafico è evidente come le prime tre componenti principali siano sufficienti a catturare l'intera variabilità dei dati di riflettanza spiegando queste complessivamente più del 99% della varianza totale. In Figura 8.1-10 nel grafico di sinistra troviamo lo scatterogramma tra il logaritmo di  $C$ , privato della media ed indicato con  $\text{Log}(C)_0$ , e la stessa quantità stimata usando tutti i coefficienti della regressione (Eq. 3.2-33), indicata con  $\text{Log}(C)_0'$  (nel grafico ed in quelli successivi è riportata la retta  $y=x$  intorno alla quale si distribuiscono i punti dello scatterogramma). Nel grafico di sinistra è riportato lo stesso tipo di scatterogramma in cui  $\text{Log}(C)_0'$  è stato stimato utilizzando solo le prime 3 componenti principali del *data set* (Eq. 3.2-40). Dalla Tabella 8.1-4 si può notare come la correlazione tra  $\text{Log}(C)_0$  e  $\text{Log}(C)_0'$  si mantenga alta, maggiore di 0.99, in entrambe i casi (con  $\rho_3$  si è indicato il coefficiente di correlazione relativo alla stima ottenuta con i primi tre coefficienti  $\eta_i$ ); l'errore quadratico medio subisce nel secondo caso un aumento passando da  $\text{MSE}=4.24\text{e-}4$  a  $\text{MSE}_3=0.0021$ . Sempre in Tabella 8.1-4 sono riportati i coefficienti  $\eta_i$  della regressione PCA stimati tramite la Eq. 3.2-36. La stessa tabella contiene inoltre i risultati delle regressioni relative al logaritmo degli altri due OAPs  $\text{Log}(X)_0$  e  $\text{Log}(Y)_0$ . Si osserva anche per questi parametri lo stesso comportamento della correlazione e dell'errore quadratico medio osservato per il  $\text{Log}(C)_0$ . Per quanto riguarda la concentrazione del sedimento solido in sospensione è stato osservato un residuo di non linearità nello scatterogramma  $\text{Log}(X)_0 - \text{Log}(X)_0'$  (non mostrato in questo paragrafo) che mostra come il modello multilineare non sia il più adatto a stimare tale parametro da dati relativi ad acque di caso I.

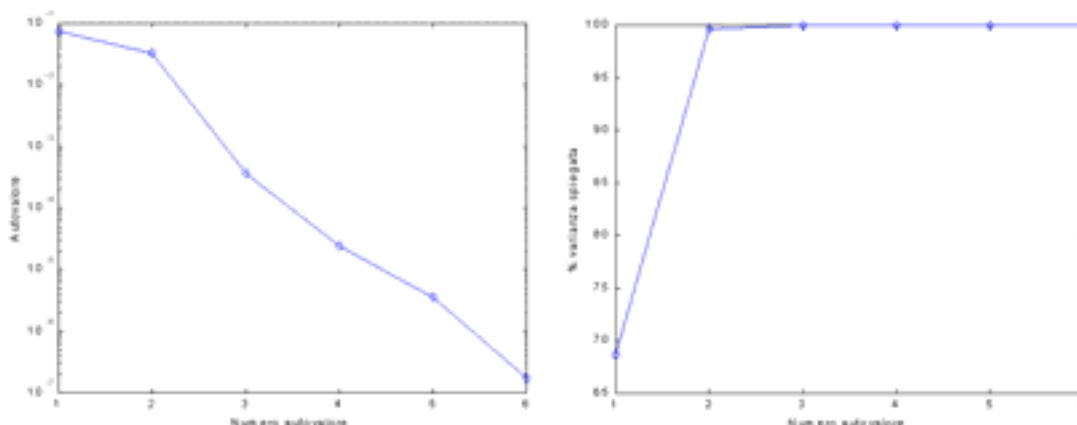


Figura 8.1-9: grafico degli autovalori e della percentuale cumulata della varianza spiegata da ogni componente principale; il data set è relativo alle acque di caso I.

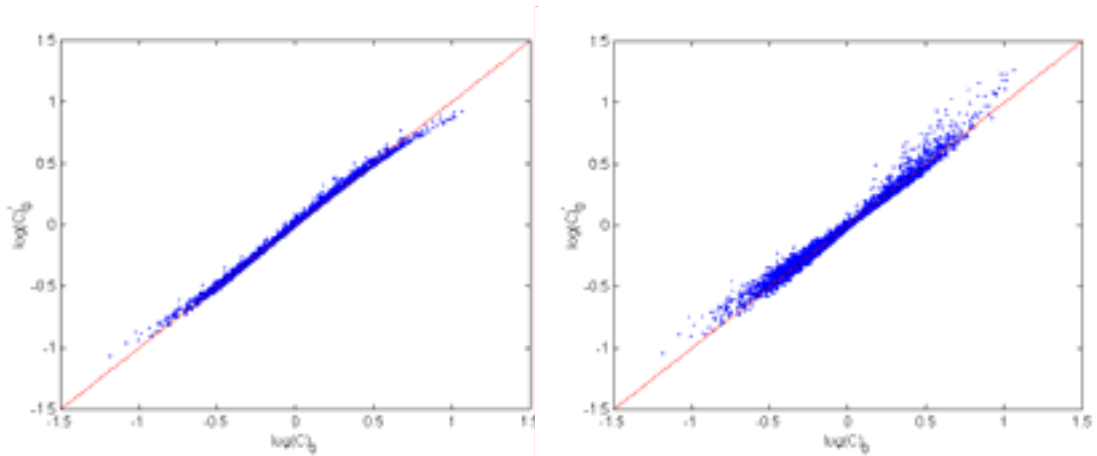


Figura 8.1-10: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I (destra).

Tabella 8.1-4: coefficienti degli algoritmi PCA per acque di caso I.

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\log(C)$	-0.9083	0.7442	5.1959	6.3172	12.6065	22.5108	4.24e-4	0.0021	0.997	0.991
$\log(X)$	-1.1004	0.0822	-0.8587	2.1682	-10.239	-80.07	5.95e-4	0.0022	0.980	0.950
$\log(Y)$	-0.6787	1.2382	-2.8558	2.0416	-3.4999	33.6903	8.94e-4	0.0012	0.991	0.96

#### 8.1.1.2 Algoritmi per acque di caso II

In Figura 8.1-12 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del data set di riflettanze per le acque di caso II.

Anche in questo caso le prime tre componenti principali sono sufficienti a descrivere l'intera variabilità dei dati di riflettanza spiegando queste complessivamente più del 99% della varianza totale. La prima componente principale da sola spiega più del 90% della variabilità totale. In Tabella 8.1-5 sono riportati i risultati delle regressioni tra le componenti principali delle riflettanze ed i logaritmi degli OAPs. Il coefficiente di correlazione  $\rho$  relativo al  $\log(C)_0$  è pari a circa 0.9 e scende a circa 0.7 quando si considerano solo i primi tre coefficienti della regressione. Il valore dell'MSE è pari a 0.0413 mentre l'MSE<sub>3</sub> sale a circa 0.1. Tali risultati e gli scatterogrammi di Figura 8.1-14 (notare la notevole dispersione dei dati presente in entrambe i grafici) confermano che per acque di caso II non è possibile la stima del  $\log(C)_0$  con errore ragionevole mediante un algoritmo multilineare, sia che si considerino tutti i coefficienti della regressione che una parte corrispondente al numero di componenti principali che spiegano oltre il 99% della variabilità totale. Il modello multilineare è però adatto alla stima degli altri due parametri otticamente attivi  $\log(X)_0$  e  $\log(Y)_0$  come è possibile notare dall'alto coefficiente di correlazione (sia per quanto riguarda  $\rho$  che  $\rho_3$  superiori a 0.99 per il  $\log(X)_0$  e a 0.97 per il  $\log(Y)_0$ ) e

dalla bassa dispersione degli scatterogrammi di Figura 8.1-16 intorno alla retta  $y=x$  (relativi al solo parametro  $\text{Log}(X)_0$  riportati come esempio).

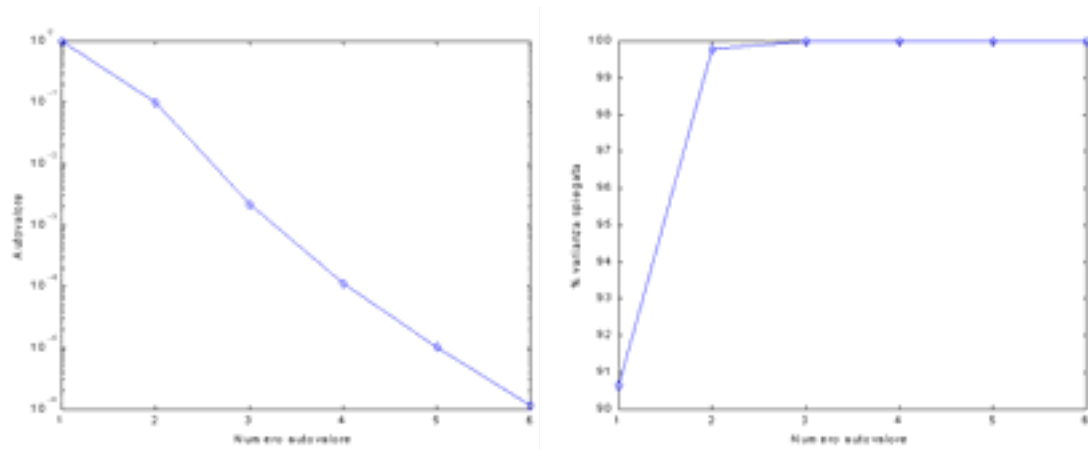


Figura 8.1-12: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso II.

Tabella 8.1-5: coefficienti degli algoritmi PCA per acque di caso II.

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	0.0395	1.0356	2.2240	-18.847	42.4847	-106.49	0.0413	0.1128	0.907	0.7181
$\text{Log}(X)$	0.3631	1.0577	0.3622	3.6534	-0.8985	-22.863	0.0018	0.0040	0.9963	0.9920
$\text{Log}(Y)$	-0.2450	1.3277	-0.7470	5.6848	-4.3481	-11.713	0.0069	0.0108	0.9861	0.9780

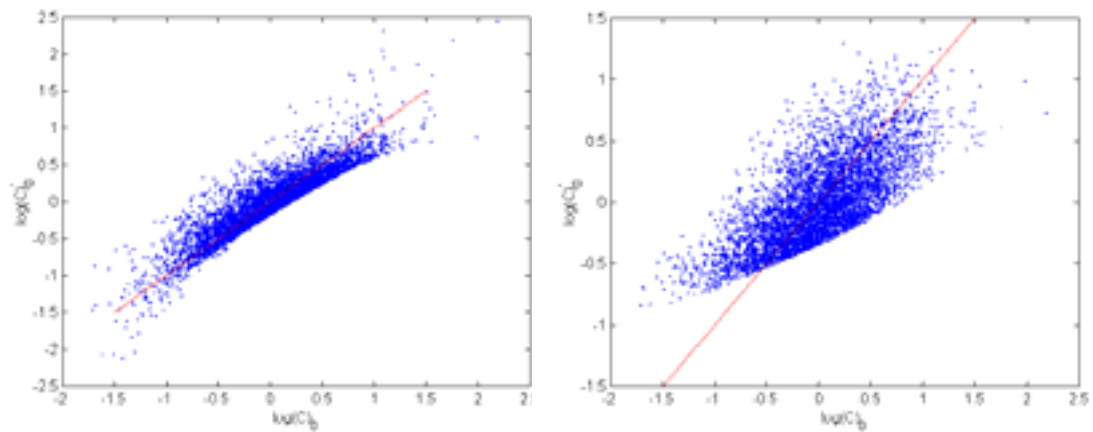


Figura 8.1-14: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso II (destra).

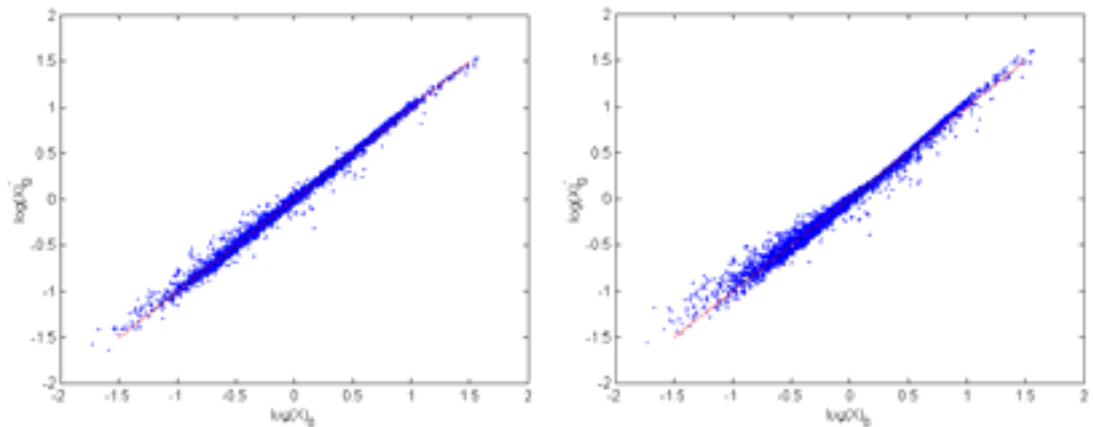


Figura 8.1-16: scatterogrammi del  $\log(X)_0$  (generati da distribuzione log-normale) e del  $\log(X)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I-II (destra).

### 8.1.1.3 Algoritmi per acque di caso I-II

In Figura 8.1-18 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del data set di riflettanze per le acque di caso intermedio I-II.

Le prime tre componenti principali siano sufficienti a descrivere l'intera variabilità dei dati di riflettanza spiegando queste complessivamente più del 99% della varianza totale. La prima componente principale da sola spiega poco più del 70% della variabilità totale. In Tabella 8.1-7 sono riportati i risultati delle regressioni tra le componenti principali delle riflettanze ed i logaritmi degli OAPs. Il coefficiente di correlazione  $\rho$  relativo al  $\text{Log}(C)_0$  è pari a circa 0.9788 e scende a circa 0.91 quando si considerano solo i primi tre coefficienti della regressione quindi, rispetto alle acque di caso II, la correlazione si mantiene più alta. Il valore dell'MSE si mantiene più basso di circa un ordine di grandezza (sia per quanto riguarda il valore di MSE che di  $\text{MSE}_3$ ). Gli scatterogrammi di Figura 8.1-20, relativi alla stima del  $\text{Log}(C)_0$ , mostrano una certa dispersione anche se minore di quella presente nei corrispondenti grafici relativi alle acque di caso II. Per quanto riguarda i parametri  $\text{Log}(X)_0$  e  $\text{Log}(Y)_0$  è stato riscontrato un comportamento analogo a quello riscontrato per le acque di caso II.

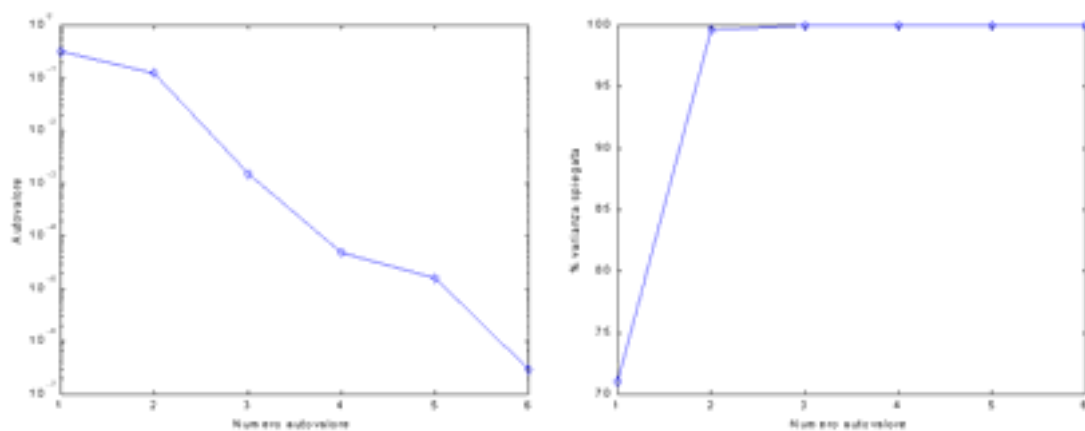


Figura 8.1-18: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso I-II.

Tabella 8.1-7: coefficienti degli algoritmi PCA per acque di caso I-II.

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	0.3490	0.9945	1.3736	-6.3627	37.7189	-51.992	0.0084	0.0335	0.9788	0.9122
$\text{Log}(X)$	0.6362	0.7293	0.3111	2.9936	-4.0970	-42.414	4.01e-4	0.0016	0.9990	0.9959
$\text{Log}(Y)$	0.0638	1.2280	-0.2197	1.5902	-5.1546	52.0342	0.0021	0.0034	0.9947	0.9913

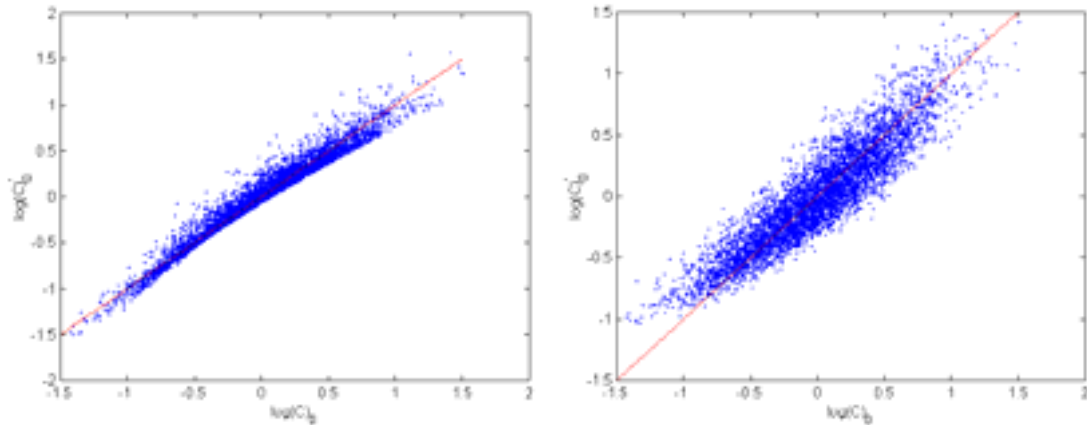


Figura 8.1-20: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I-II (destra).

## 8.1.2 Sensore SeaWiFS: algoritmi RBF-NN

Questo sottoparagrafo è dedicato alla breve descrizione di alcuni risultati ottenuti applicando il metodo di inversione RBF-NN al modello di riflettanza. Per brevità i risultati riportati sono relativi alla sola concentrazione cumulata di clorofilla-a e feofitina-a. In generale i risultati ottenuti con tale metodo di inversione hanno mostrato una buona capacità delle RBF-NN di risolvere i problemi di stima dei parametri otticamente attivi laddove gli algoritmi multilineari non sono in grado di ottenere risultati soddisfacenti (si pensi ad esempio al residuo di non linearità che presenta l'algoritmo multilineare per la stima del  $\text{Log}(X)_0$  nelle acque di caso I o alla stima del  $\text{Log}(C)_0$  per le acque di caso II).

Le reti per la stima dei parametri sono state addestrate utilizzando un *data set* di *training* ottenuto scegliendo casualmente 500 punti dai *data set* completi (costituiti da 5000 punti) generati per le acque di caso I, caso II e caso I-II. Le prestazioni delle reti addestrate secondo la procedura di *forward selection* descritta nel capitolo 3, sono state successivamente calcolate (sempre in termini di MSE e di coefficiente di correlazione) sull'intero *data set* di 5000 punti per testare le proprietà di generalizzazione.

### 8.1.2.1 Algoritmi per acque di caso I

In Figura 8.1-22 sulla sinistra è mostrato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* per acque di caso I (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set* generato per le acque di caso I. La rete addestrata presenta buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al *data set* completo. La Figura 8.1-24 mostra la sequenza dei pesi della rete mentre nella Tabella 8.1-9 sono riportati alcuni parametri



della rete e gli indici prestazionali. Con  $MSE_{ck}$  si è indicato l'errore quadratico medio calcolato sul *data set* completo mentre con  $MSE_{tr}$  lo stesso parametro, ma calcolato sul *data set* di addestramento. La stessa notazione è stata usata per i coefficienti di correlazione indicati con  $\rho_{tr}$  e  $\rho_{ck}$ . Con  $r$  si è indicato il valore del parametro di dispersione della funzione radiale di base. La procedura di *forward selection* ha restituito in uscita una rete contenente 65 neuroni nello strato nascosto (o equivalentemente numero di centri). I valori dei coefficienti di correlazione per la rete RBF e per il corrispondente algoritmo PCA definito per lo stesso *data set* sono tra loro confrontabili; nel caso di rete RBF i valori dell'MSE sono un ordine di grandezza sotto rispetto all'MSE ottenuto con l'algoritmo PCA. La rete RBF inoltre risolve i problemi relativi alla non linearità residua che si ha quando si cerca di stimare il  $\text{Log}(X)$  con algoritmo PCA. Il *fitting* ottenuto è ottimo come è evidente dagli scatterogrammi di Figura 8.1-22. La rete RBF rappresenta quindi un modello più flessibile che si adatta ad una classe di modelli inversi più ampia (linari e non lineari).

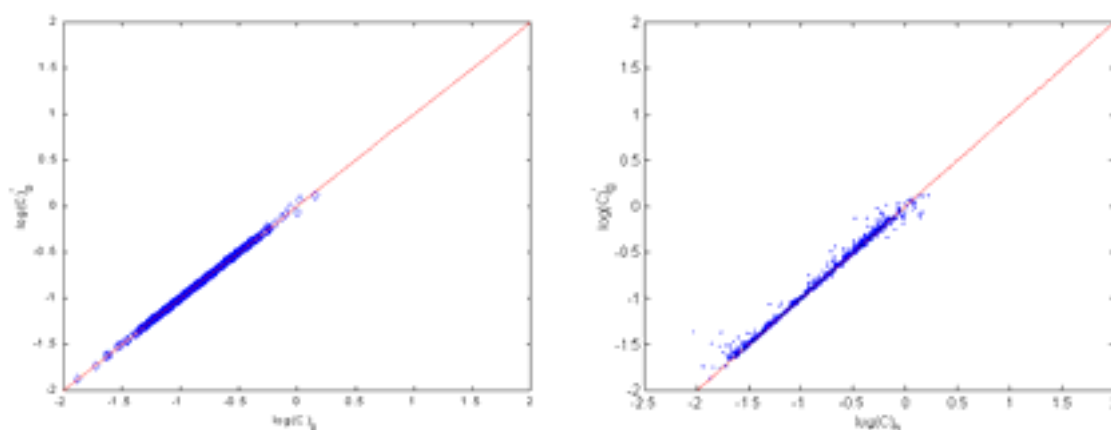


Figura 8.1-22: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo alle acque di caso I (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.1-9: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso I

	$r$	Num Centri	$MSE_{tr}$	$MSE_{ck}$	$\rho_{tr}$	$\rho_{ck}$
$\text{Log}(C)$	0.3	65	4.5224e-005	5.5137e-004	0.9997	0.9970

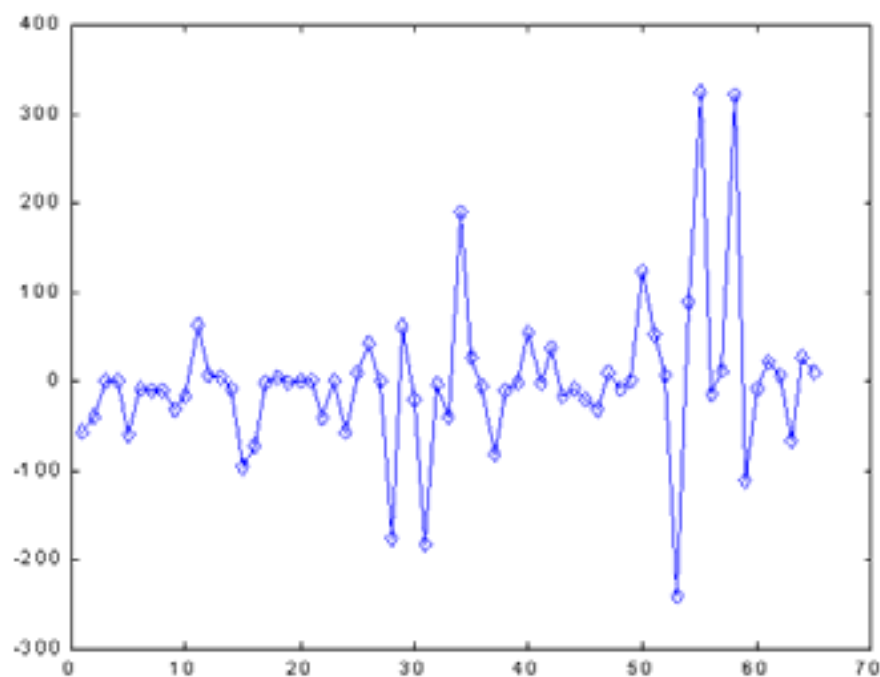


Figura 8.1-24: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso I

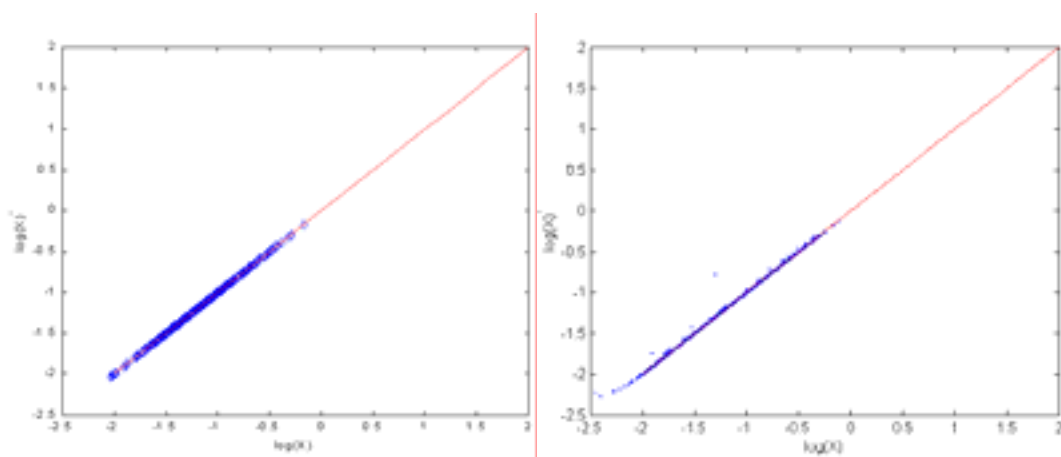


Figura 8.1-26: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(X)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(X)$  del training set relativo alle acque di caso I (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

#### 8.1.2.2 Algoritmi per acque di caso II

In Figura 8.1-27 sulla sinistra è mostrato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set (in ascisse) relativi al data set per acque di caso II; sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set generato per le acque di caso II. Rispetto al corrispondente algoritmo PCA, tramite la rete RBF si ottiene un buon *fitting* dei dati. Anche in questo caso la rete addestrata presenta buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al data set completo. La Figura 8.1-28 mostra la sequenza dei pesi della rete mentre nella Tabella

8.1-11 sono riportati i parametri della rete e gli indici prestazionali. I valori dei coefficienti di correlazione per la rete RBF sono più alti di quelli ottenuti con il corrispondente algoritmo PCA, definito per lo stesso *data set*; l'errore quadratico medio nel caso di rete RBF è di un ordine di grandezza sotto rispetto all'MSE ottenuto con l'algoritmo PCA. Ciò dimostra come un modello inverso non lineare sia più adatto a stimare la concentrazione di clorofilla più feofitina nelle acque di caso II. La complessità della rete rispetto a quella relativa ad acque di caso I è maggiore essendo il numero di neuroni dello strato nascosto pari a 181.

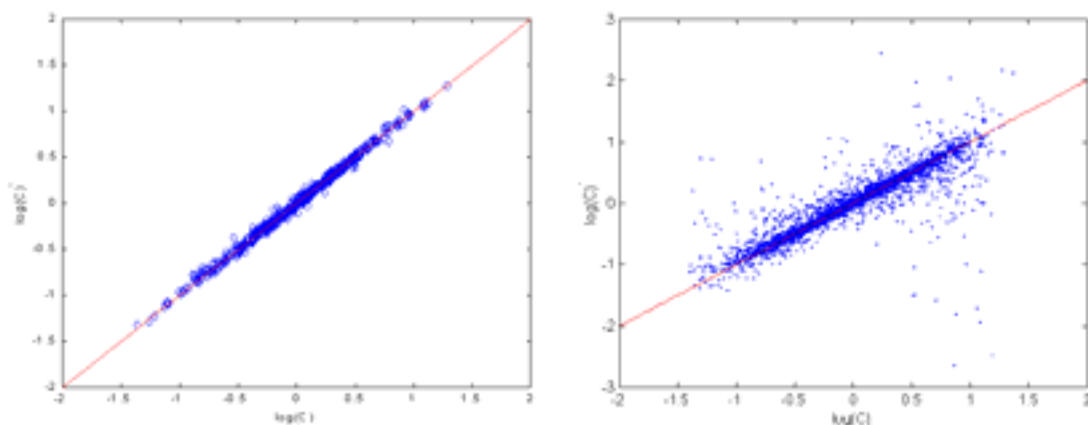


Figura 8.1-27: sulla sinistra è riportato lo scattergramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo alle acque di caso II (in ascisse); sulla destra si trova graficato lo stesso tipo di scattergramma, ma relativo all'intero data set.

Tabella 8.1-11: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso II

	r	Num Centri	MSE <sub>tr</sub>	MSE <sub>ck</sub>	$\rho_{tr}$	$\rho_{ck}$
$\text{Log}(C)$	0.4	181	0.0012	0.0278	0.9972	0.9381

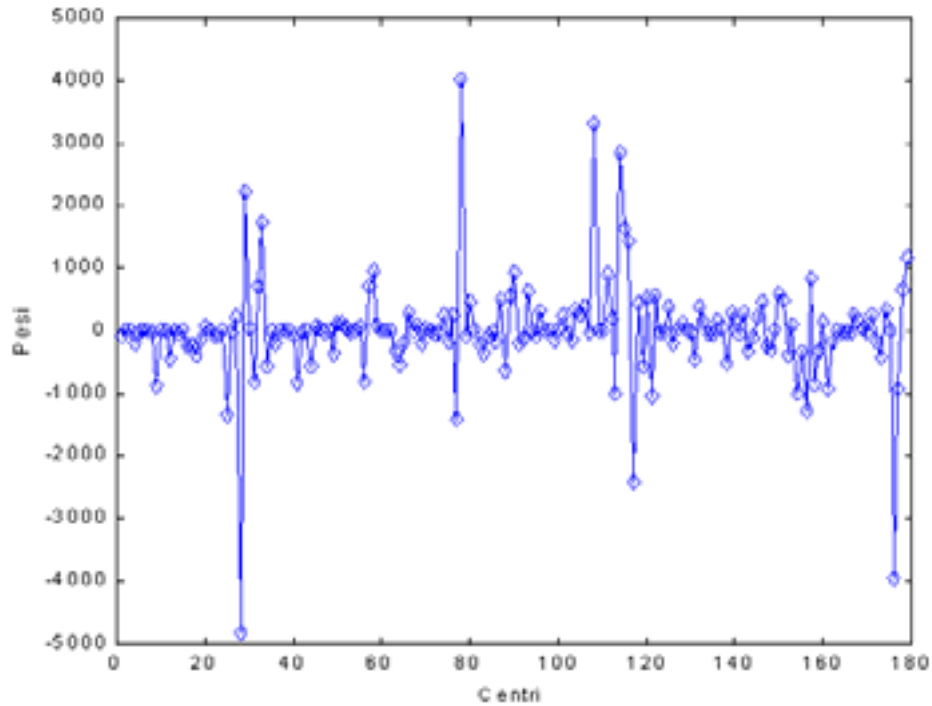


Figura 8.1-28: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso II

### 8.1.2.3 Algoritmi per acque di caso I-II

I risultati relativi all'addestramento della rete RBF per le acque di caso I-II sono mostrati in Figura 8.1-30 ed in Tabella 8.1-13. Analogamente ai precedenti casi la Figura 8.1-30 presenta sulla sinistra lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set* per le acque di caso I-II. Anche in questo caso la rete addestrata presenta buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al *data set* completo. Come risulta evidente dalla Tabella 8.1-13 la rete alla fine dell'addestramento contiene 100 neuroni nello strato nascosto. La complessità della struttura della rete risulta quindi di tipo intermedio fra i due casi precedentemente esaminati. I valori dei coefficienti di correlazione per la rete RBF sono maggiori rispetto al corrispondente algoritmo PCA definito per lo stesso *data set*; ancora una volta la rete RBF presenta valori dell'MSE che sono un ordine di grandezza sotto all'MSE ottenuto con l'algoritmo PCA.

I risultati mostrati per i vari casi di acque mettono in evidenza come in generale una rete RBF sia in grado di stimare con maggiore accuratezza il  $\text{Log}(C)$  nelle condizioni tipiche delle acque di caso I e di caso I-II. Nelle acque di caso I il modello multilineare è sufficiente ad ottenere una buona stima del  $\text{Log}(C)$  mentre non lo è per il  $\text{Log}(X)$  il quale richiede un modello inverso non lineare. In questo caso la rete RBF addestrata per la stima del  $\text{Log}(X)$  ha dimostrato di possedere prestazioni adeguate allo scopo.

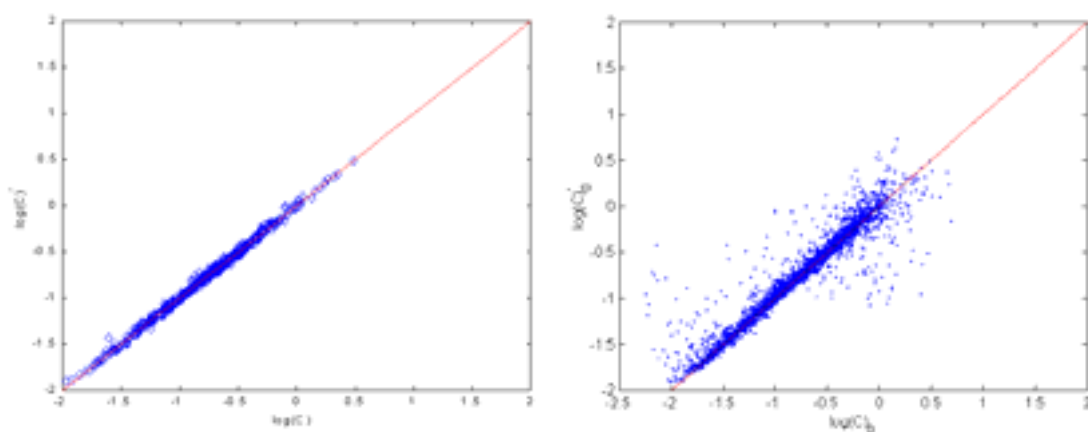


Figura 8.1-30: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo alle acque di caso I-II (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.1-13: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso I-II

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.5	100	3.4537e-004	0.0027	0.9990	0.9935

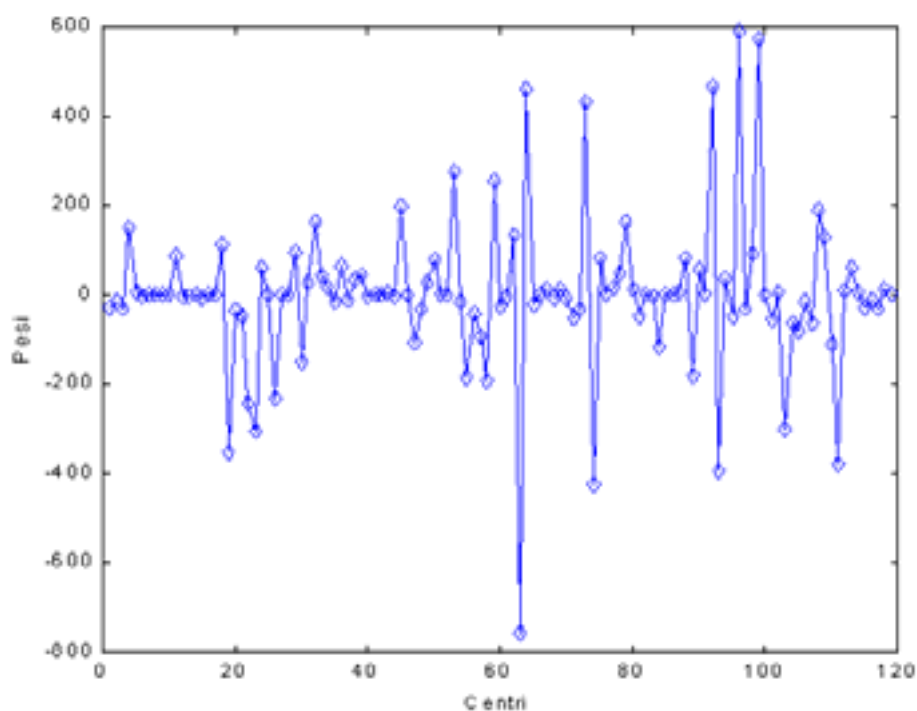


Figura 8.1-32: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ ; data set acque di caso I-II

### 8.1.3 Sensore MOS-B: algoritmi multilineari

In questo sottoparagrafo sono riportati alcuni dei risultati ottenuti applicando il metodo di inversione PCA ai *data set* generati per il sensore MOS-B. Per ogni *data set* in maniera analoga a quanto fatto per il sensore SeaWiFS saranno riportati: a) i risultati relativi all'analisi PCA effettuata sulle riflettanze in modo da stabilire la dimensionalità dei dati e quindi il numero di componenti principali sufficienti a stimare i parametri otticamente attivi e b) i risultati delle regressioni multilineari effettuate sui dati proiettati nel sistema di riferimento degli autovettori come spiegato nel capitolo 3. Saranno effettuate inoltre alcune brevi considerazioni riguardanti il raffronto tra le prestazioni raggiunte con il sensore SeaWiFS.

#### 8.1.3.1 Algoritmi per acque di caso I

In Figura 8.1-33 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del *data set* di riflettanze per le acque di caso I. In maniera analoga al caso del sensore SeaWiFS la somma cumulata riportata è rapportata alla somma totale degli autovalori e rappresenta quindi la percentuale di varianza totale spiegata dalle prime  $n$  componenti principali con  $n=1, 2, \dots, N_c$  dove  $N_c$  è il numero di canali del sensore. Dal grafico è evidente come le prime tre componenti principali siano sufficienti a catturare l'intera variabilità dei dati di riflettanza spiegando queste complessivamente più del 99% della varianza totale. Per quanto riguarda la stima del  $\text{Log}(C)_0$  come è possibile notare dallo scatterogramma di Figura 8.1-34 e dai risultati riportati in Tabella 8.1-15 le prestazioni del relativo algoritmo sono superiori a quelle del corrispondente algoritmo SeaWiFS. Il valore di MSE è pari a  $2e-4$ , circa la metà rispetto al valore ottenuto per il SeaWiFS; mentre  $\text{MSE}_3$  è pari a  $3.44e-4$  ovvero un ordine di grandezza inferiore al corrispondente valore relativo sempre al sensore SeaWiFS. Per il sensore MOS-B quindi, l'aumento che il valore di MSE subisce quando si stima il  $\text{Log}(C)_0$  tramite i primi tre coefficienti  $\eta_i$  è decisamente inferiore dell'incremento osservato per il sensore SeaWiFS. La Tabella 8.1-15 contiene i risultati delle regressioni relative al logaritmo degli altri due OAPs. Il comportamento del valore di MSE e  $\text{MSE}_3$  e dei coefficienti di correlazione è analogo a quanto riscontrato per il sensore SeaWiFS. Le prestazioni raggiunte con la configurazione spettrale del MOS-B sono, come appare evidente dal riscontro con i dati di Tabella 8.1-4, dello stesso ordine di grandezza di quelle ottenute per il SeaWiFS con la differenza che il problema della non linearità, relativa alla stima del  $\text{Log}(X)_0$ , nel caso del MOS-B, risulta meno importante.

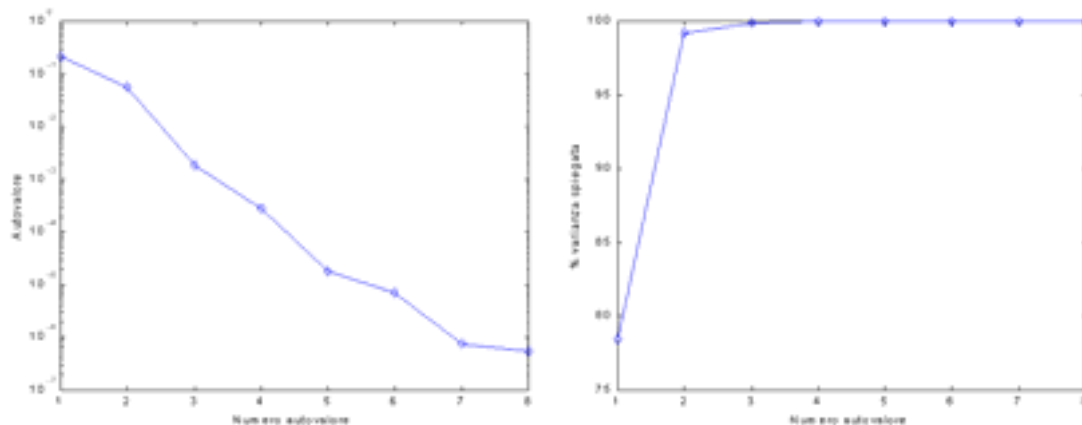


Figura 8.1-33: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso I.

Tabella 8.1-15: coefficienti degli algoritmi PCA per acque di caso I.

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	$\eta_7$	$\eta_8$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	-0.6340	0.6725	2.2688	0.2338	0.5615	-0.3031	-2.7044	14.4947	2.0e-4	3.44e-4	0.9992	0.9986
$\text{Log}(X)$	-0.7497	-0.0258	-0.7367	0.0482	3.5199	-11.191	-21.464	-6.5390	2.60e-4	0.0017	0.9989	0.9930
$\text{Log}(Y)$	-0.4767	1.0121	-2.0337	0.0333	-10.515	5.6219	12.3018	12.4892	5.34e-4	0.0029	0.9977	0.9875

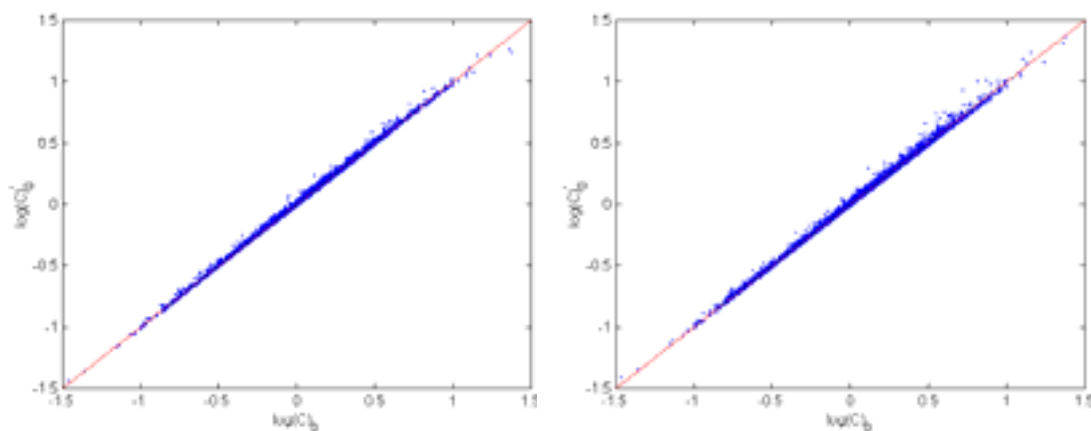


Figura 8.1-34: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I (destra).

### 8.1.3.2 Algoritmi per acque di caso II

In Figura 8.1-36 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del data set di riflettanze per le acque di caso II generato per il sensore MOS-B.

Le prime tre componenti principali sono sufficienti a descrivere l'intera variabilità dei dati di riflettanza spiegando queste complessivamente più del 99% della varianza totale. La prima componente principale spiega da sola circa l'85% della variabilità totale. In Tabella 8.1-17 sono riportati i risultati delle regressioni tra le componenti principali delle riflettanze ed i logaritmi degli OAPs. Il coefficiente di correlazione  $\rho$  relativo al  $\text{Log}(C)_0$  è pari a circa 0.94 e scende a circa 0.83 quando si considerano solo i

primi tre coefficienti della regressione. Il valore dell'MSE è pari a 0.0286 mentre l'MSE<sub>3</sub> sale a circa 0.08. Pur essendo i valori degli errori ancora una volta più bassi rispetto al caso del SeaWiFS, gli scatterogrammi di Figura 8.1-38 mostrano che anche per il sensore MOS-B (pur dotato di una migliore risoluzione spettrale e di un maggiore numero di canali) non è possibile la stima del  $\text{Log}(C)_0$  con errore accettabile mediante un algoritmo multilineare. Il modello multilineare ha ancora prestazioni buone quando è usato per la stima degli altri due parametri otticamente attivi  $\text{Log}(X)_0$  e  $\text{Log}(Y)_0$  come è possibile notare dall'alto coefficiente di correlazione (sia per quanto riguarda  $\rho$  che  $\rho_3$  superiori a 0.99 per il  $\text{Log}(X)_0$  e a 0.98 per il  $\text{Log}(Y)_0$ ) e dalla bassa dispersione degli scatterogrammi di Figura 8.1-39 intorno alla retta  $y=x$  (relativi al solo parametro  $\text{Log}(X)_0$ ).

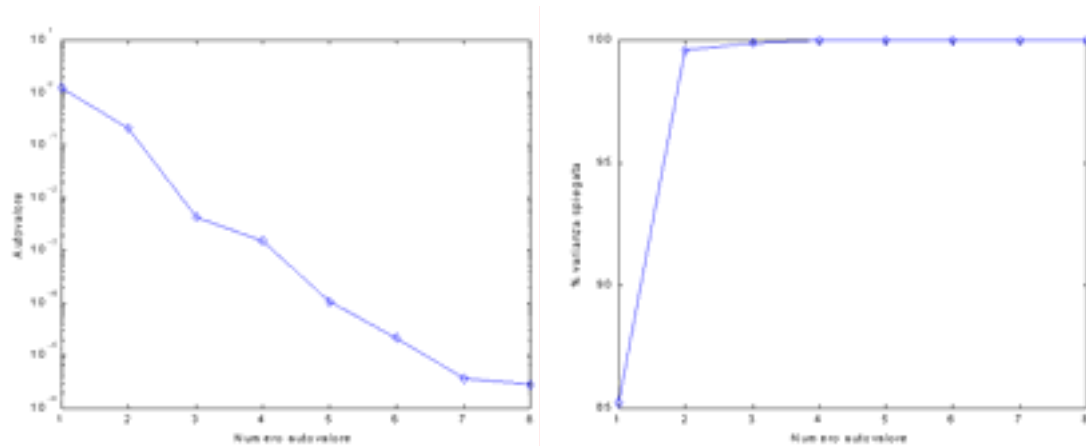


Figura 8.1-36: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso II.

Tabella 8.1-17: coefficienti degli algoritmi PCA per acque di caso II

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	$\eta_7$	$\eta_8$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	0.1061	0.7224	3.5424	4.0356	6.8650	-3.3751	56.4415	-54.818	0.0286	0.0782	0.9424	0.8330
$\text{Log}(X)$	0.3928	0.5641	-0.2547	-1.2468	-1.1596	-3.5193	3.9384	-10.443	2.69e-4	0.0033	0.9995	0.9936
$\text{Log}(Y)$	-0.1209	1.0147	-0.8989	-0.5435	-3.6879	2.9035	9.8030	3.9948	0.0072	0.0096	0.9855	0.9804



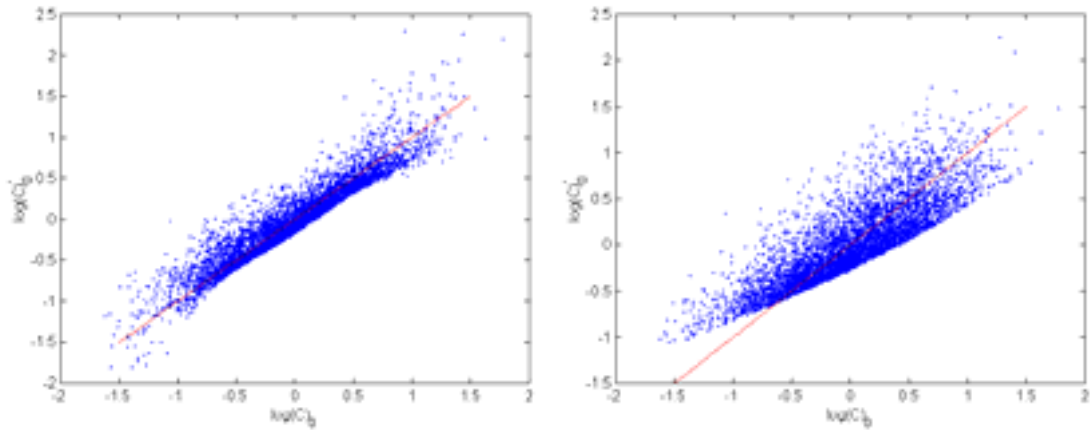


Figura 8.1-38: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso II (destra).

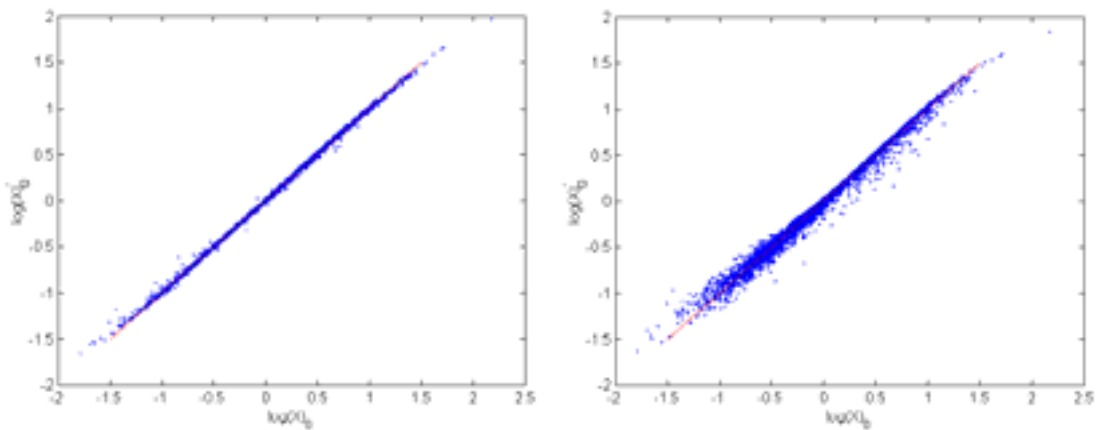


Figura 8.1-39: scatterogrammi del  $\log(X)_0$  (generati da distribuzione log-normale) e del  $\log(X)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso II (destra).

### 8.1.3.3 Algoritmi per acque di caso I-II

In Figura 8.1-40 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del *data set* di riflettanze per le acque di caso I-II.

Al solito le prime tre componenti principali spiegano complessivamente più del 99% della varianza totale. La prima componente principale da sola spiega poco più del 75% della variabilità totale. In Tabella 8.1-19 sono riportati i risultati delle regressioni tra le componenti principali delle riflettanze ed i logaritmi degli OAPs. Il coefficiente di correlazione  $\rho$  relativo al  $\log(C)_0$  è pari a circa 0.987 e scende a circa 0.91 quando si considerano solo i primi tre coefficienti della regressione. I valori dell'MSE sono confrontabili con i valori ottenuti per il sensore SeaWiFS. Gli scatterogrammi di Figura 8.1-42, relativi alla stima del  $\log(C)_0$ , mostrano una certa dispersione anche se minore di quella presente nei corrispondenti grafici relativi alle acque di caso II (analogamente a quanto ottenuto per il SeaWiFS). Per quanto riguarda i parametri  $\log(X)_0$  e  $\log(Y)_0$  è stato riscontrato un comportamento analogo a quello riscontrato per le acque di caso II con prestazioni confrontabili con quelle del sensore SeaWiFS.

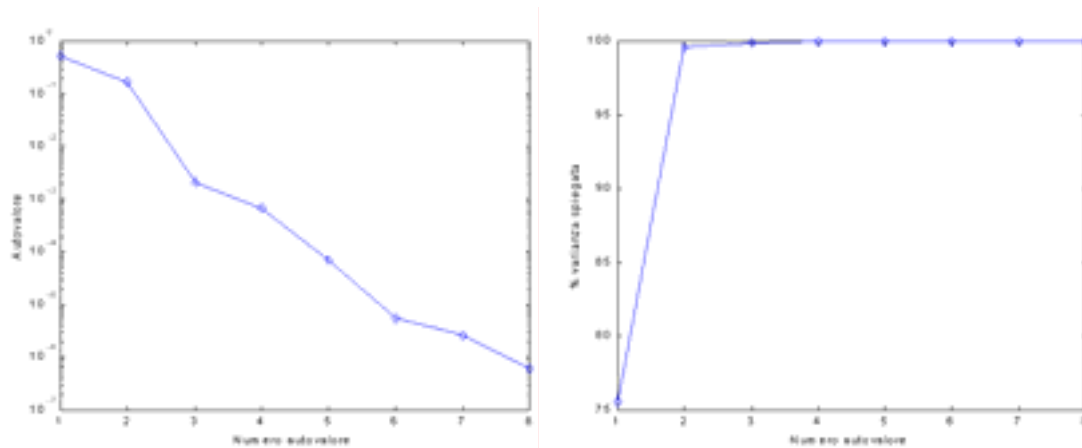


Figura 8.1-40: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso I-II.

Tabella 8.1-19: coefficienti degli algoritmi PCA per acque di caso I-I.I

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	$\eta_7$	$\eta_8$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	0.4334	0.6194	1.1581	6.3643	2.1168	4.6984	13.1177	34.1766	0.0053	0.0342	0.9869	0.9121
$\text{Log}(X)$	0.5857	0.2400	0.0256	0.6786	2.4147	1.7603	16.8970	4.0037	9.84e-5	0.0016	0.9997	0.9959
$\text{Log}(Y)$	0.2945	0.9607	0.4633	0.5165	1.8519	2.3625	28.2164	1.9726	0.0011	0.0036	0.9975	0.9913

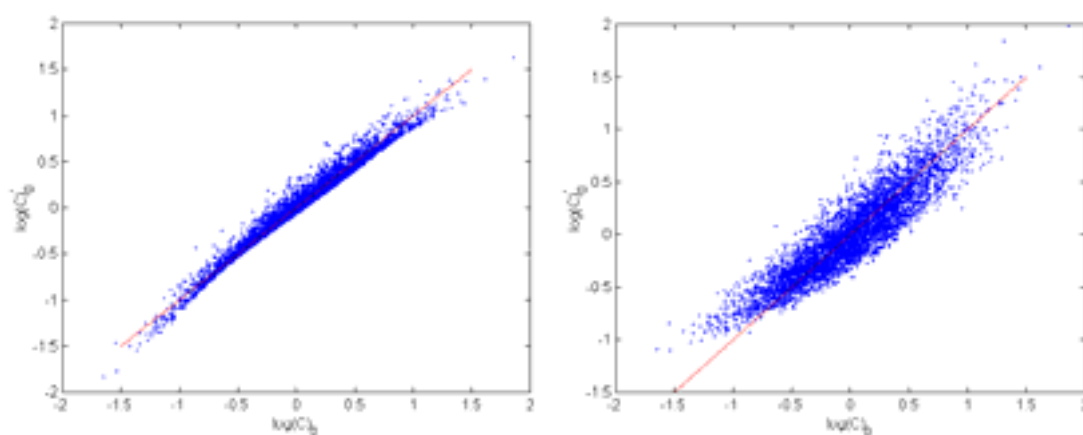


Figura 8.1-42: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I-II (destra).

#### 8.1.4 Sensore MOS-B: algoritmi RBF-NN

In maniera analoga a quanto fatto per il sensore SeaWiFS in questo sottoparagrafo saranno riportati alcuni dei risultati ottenuti applicando il metodo di inversione RBF-NN al problema della stima degli OAPs dai dati di riflettanza subsuperficiale. Valgono le stesse considerazioni fatte per il sensore SeaWiFS sia per quanto riguarda le metodiche di addestramento delle reti RBF, sia per quanto riguarda il tipo di risultati che verranno mostrati nel seguito. In generale i risultati ottenuti confermano anche per il sensore MOS-B

la buona capacità delle RBF-NN di risolvere i problemi di stima dei parametri otticamente attivi nei casi di spiccata non linearità.

#### 8.1.4.1 Algoritmi per acque di caso I

In Figura 8.1-43 sulla sinistra è mostrato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* per acque di caso I (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set*. La rete addestrata presenta buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al *data set* completo anche se rispetto al caso del sensore SeaWiFS il valore di  $\text{MSE}_{\text{ck}}$  risulta più alto a causa di alcuni *outliers* che contribuiscono ad aumentarne in maniera artificiosa il valore (notare che per quanto riguarda il valore di  $\text{MSE}_{\text{tr}}$  questo è un ordine di grandezza più basso nel MOS-B rispetto al SeaWiFS). La Figura 8.1-44 mostra la sequenza dei pesi della rete mentre nella Tabella 8.1-21 sono riportati i parametri della rete e le prestazioni raggiunte. La procedura di *forward selection* ha scelto 113 centri (neuroni) dal data set di *training* da inserire nello strato nascosto (la rete è più complessa rispetto a quanto ottenuto con il sensore SeaWiFS). Il parametro di dispersione è stato scelto pari a  $r=0.3$ .

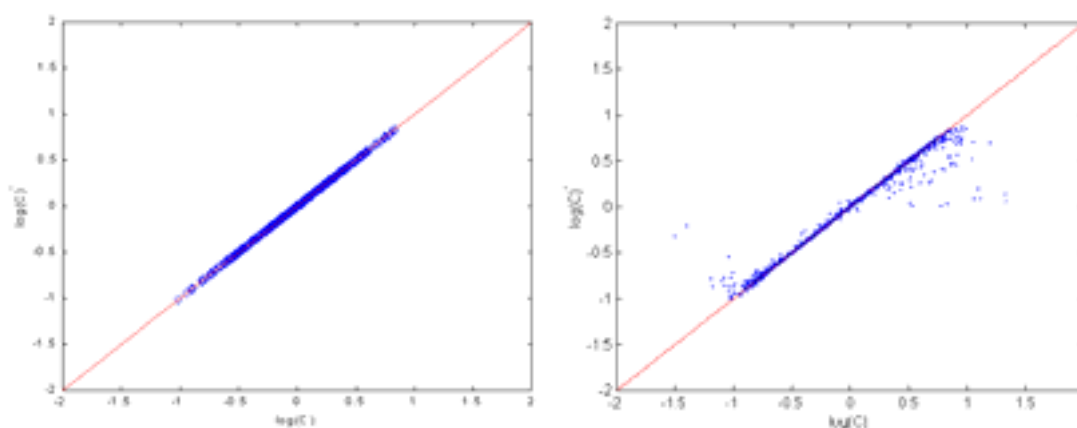


Figura 8.1-43: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* relativo alle acque di caso I (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set*.

Tabella 8.1-21: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso I

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.3	113	9.6631e-006	0.0031	0.9999	0.9875

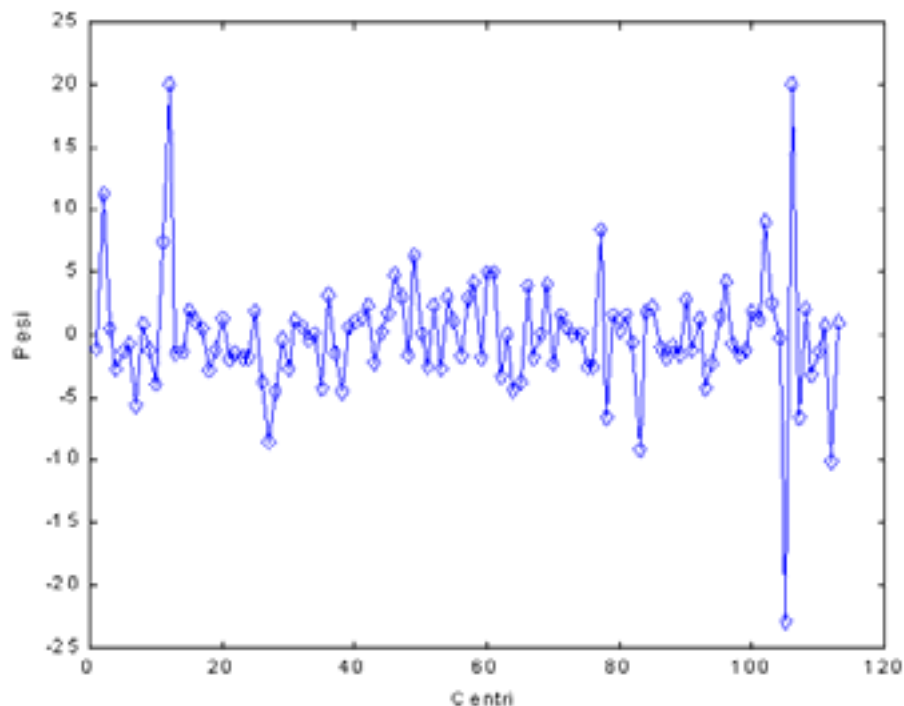


Figura 8.1-44: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso I

#### 8.1.4.2 Algoritmi per acque di caso II

In Figura 8.1-45 sulla sinistra è mostrato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* (in ascisse) relativo al *data set* per acque di caso II; sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set*. In maniera analoga a quanto ottenuto per il SeaWiFS, rispetto al corrispondente algoritmo PCA, la rete RBF produce un buon *fitting* dei dati presentando buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al *data set* completo. La sequenza dei pesi della rete è mostrata in Figura 8.1-46. La Tabella 8.1-22 contiene i parametri della rete e gli indici prestazionali. Dal confronto con i dati riportati in Tabella 8.1-17 notiamo come le prestazioni della rete RBF siano superiori rispetto all'algoritmo PCA. Rispetto alla rete addestrata con i dati relativi al sensore SeaWiFS, la rete per il sensore MOS-B possiede prestazioni migliori sia dal punto di vista del valore dell'MSE che del coefficiente di correlazione e ciò a conferma del fatto che la configurazione spettrale del MOS-B rende il sensore più adatto del SeaWiFS allo studio delle acque di caso II. La rete inoltre presenta una minore complessità essendo il numero di neuroni nello strato nascosto pari a 139 contro i 181 necessari per il SeaWiFS.

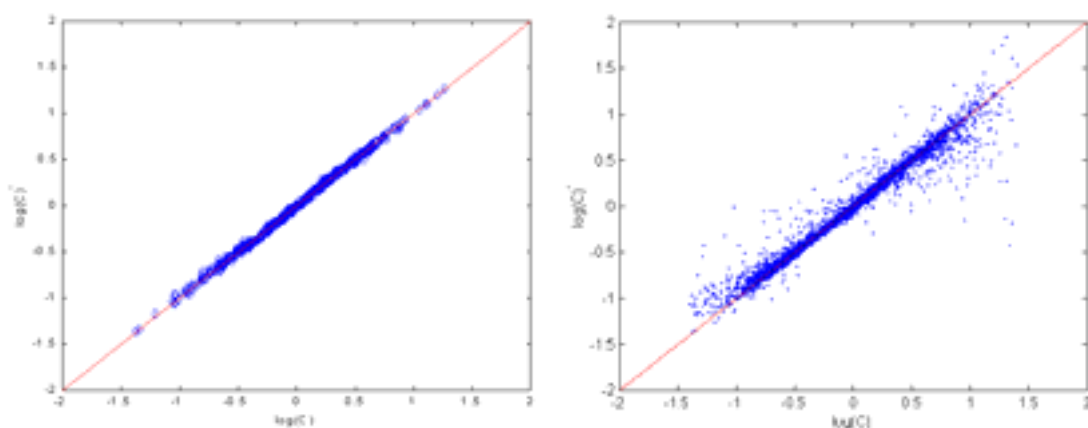


Figura 8.1-45: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo alle acque di caso II (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.1-22: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso II

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.6	139	5.2033e-004	0.0109	0.9989	0.9780

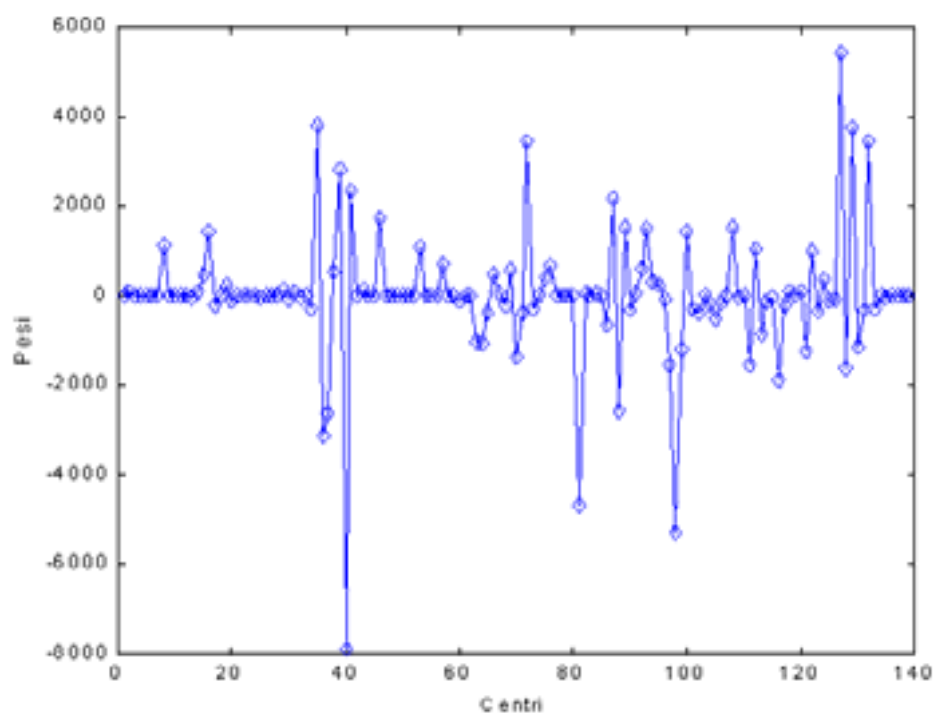


Figura 8.1-46: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso II

#### 8.1.4.3 Algoritmi per acque di caso I-II

I risultati relativi all'addestramento della rete RBF per le acque di caso I-II sono mostrati in Figura 8.1-48 ed in Tabella 8.1-24. Analogamente ai precedenti casi la Figura 8.1-48 presenta sulla sinistra lo

scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma relativo all'intero *data set* per le acque di caso I-II. La rete addestrata presenta buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al *data set* completo. La rete alla fine dell'addestramento contiene 177 neuroni nello strato nascosto (i pesi sono mostrati in Figura 8.1-49). Ancora una volta le prestazioni della rete RBF rispetto al corrispondente algoritmo PCA si sono dimostrate superiori. Il confronto con l'algoritmo RBF per il sensore SeaWiFS non mette in evidenza miglioramenti apprezzabili.

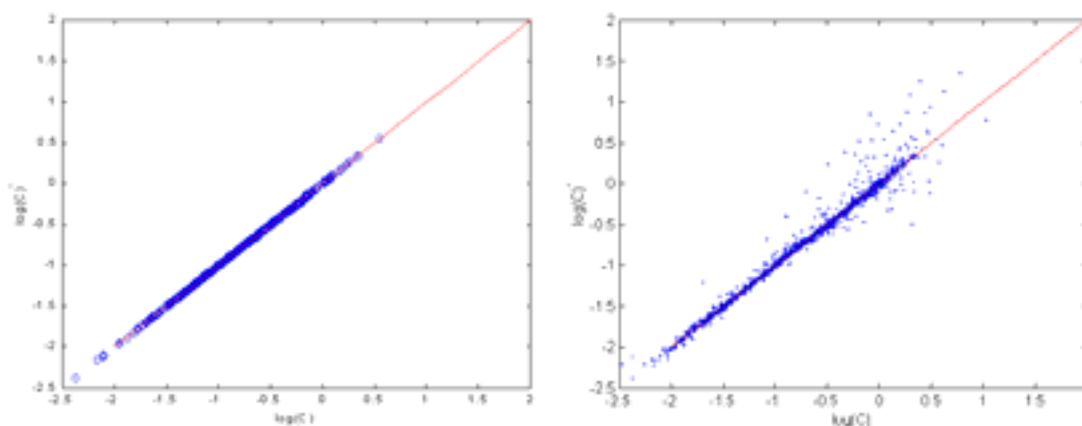
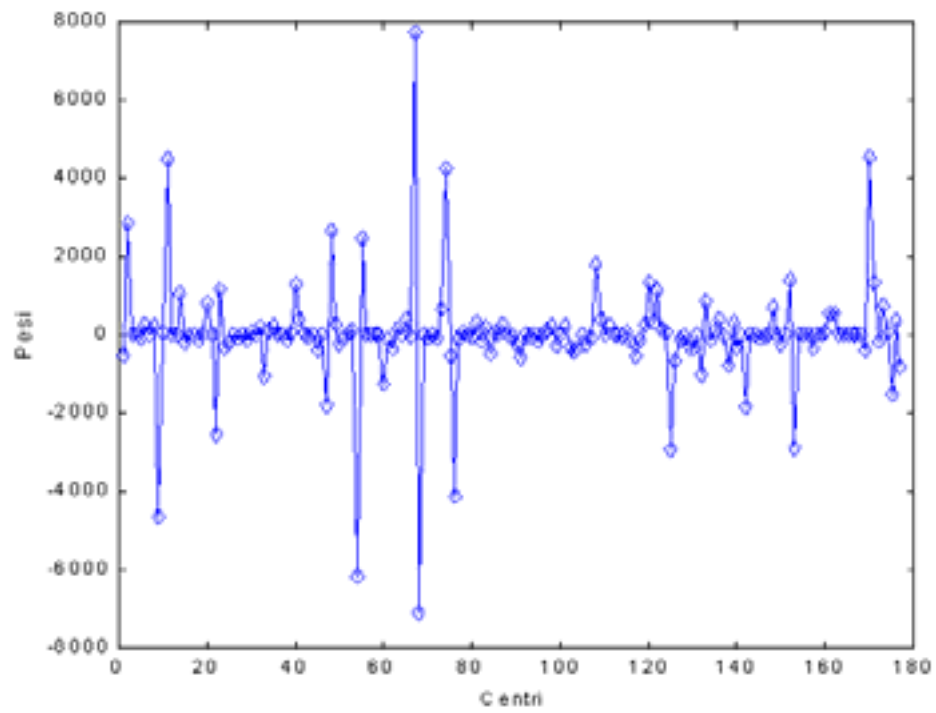


Figura 8.1-48: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo alle acque di caso I-II (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.1-24: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso I-II

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.5	177	3.8227e-005	0.0031	0.9999	0.9927



*Figura 8.1-49: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso I-II*

### 8.1.5 Sensore MERIS: algoritmi multilineari

Nei seguenti sottoparagrafi sono riportati alcuni dei risultati ottenuti applicando il metodo di inversione PCA ai *data set* generati per il sensore MERIS. Le modalità di presentazione rimangono invariate rispetto ai casi precedenti.

#### 8.1.5.1 Algoritmi per acque di caso I

In Figura 8.1-50 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del *data set* di riflettanze per le acque di caso I. Le prime tre componenti principali catturano più del 99% della varianza totale del *data set* di riflettanze. Per quanto riguarda la stima del  $\text{Log}(C)_0$ , dal confronto con i risultati della regressione ottenuti per il sensore SeaWiFS (Figura 8.1-10 e Tabella 8.1-4) e per il sensore MOS-B (Figura 8.1-34 e Tabella 8.1-15) è possibile notare come quest'ultimo sensore presenti le prestazioni migliori tra i tre presi in considerazione. Pur possedendo il MERIS una migliore risoluzione spettrale ed una banda in più rispetto al SeaWiFS le prestazioni risultano confrontabili. Per quanto riguarda la stima del  $\text{Log}(X)_0$  e  $\text{Log}(Y)_0$  le prestazioni che si ottengono per i tre sensori sono dello stesso ordine di grandezza.

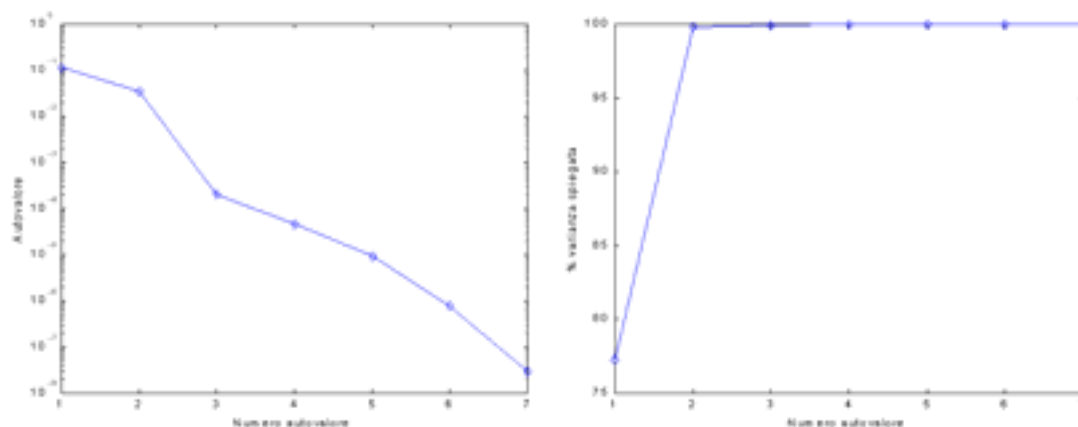


Figura 8.1-50: grafico degli autovalori e della varianza spiegata da ogni componente principale; *data set* di riflettanze per acque di caso I.

Tabella 8.1-25: coefficienti degli algoritmi PCA per acque di caso I

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	$\eta_7$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	-0.7083	0.7186	6.5773	8.5088	5.8510	0.8818	27.2749	4.56e-4	0.0042	0.9974	0.9765
$\text{Log}(X)$	-0.8712	0.0366	0.5283	0.3372	10.9963	20.6455	3.1724	4.92e-4	0.0019	0.9973	0.9894
$\text{Log}(Y)$	-0.5549	1.2295	3.4221	1.4255	0.4243	10.7094	33.2492	8.66e-4	0.0011	0.9952	0.9940



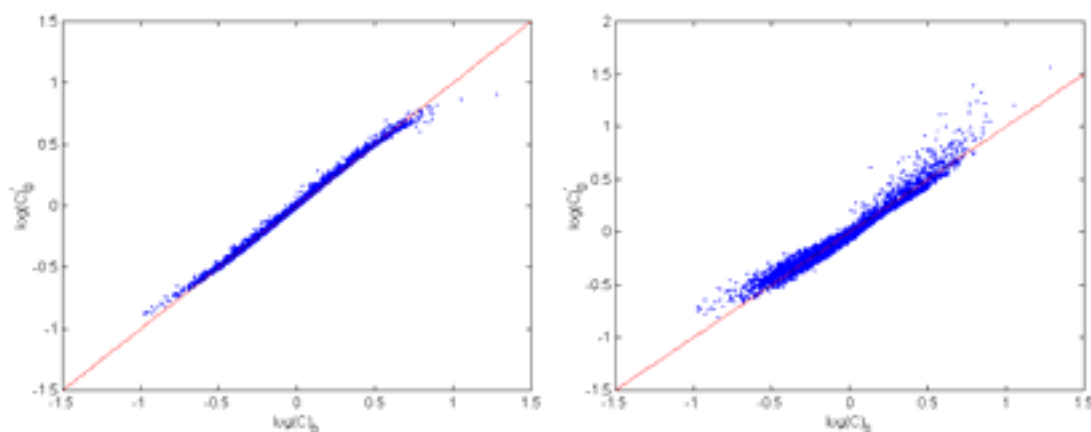


Figura 8.1-52: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I (destra).

#### 8.1.5.2 Algoritmi per acque di caso II

Come risulta evidente dalla Figura 8.1-53, il data set per acque di caso II, relativo al sensore MERIS, è caratterizzabile interamente dalle prime tre componenti principali che da sole spiegano più del 99% della variabilità totale del data set stesso. La prima componente principale da sola spiega poco meno del 88% della variabilità totale. I risultati riportati in Tabella 8.1-26 e i grafici di Figura 8.1-53 mostrano come anche nel caso del sensore MERIS non si possa essere in grado di stimare con adeguata precisione il  $\text{Log}(C)_0$  tramite un algoritmo multilineare. Il modello multilineare risulta ancora adatto alla stima del  $\text{Log}(X)_0$  e del  $\text{Log}(Y)_0$  come è evidente dallo scatterogramma di Figura 8.1-55 relativo alla regressione per la stima del  $\text{Log}(X)_0$ . Le prestazioni in questo caso sono confrontabili con quelle ottenute con i sensori SeaWiFS e MOS-B.

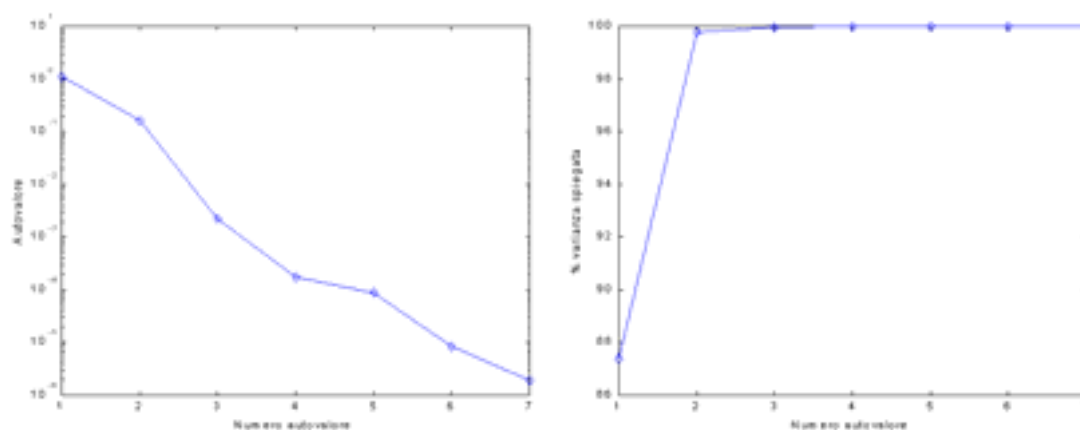


Figura 8.1-53: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso II.

Tabella 8.1-26: coefficienti degli algoritmi PCA per acque di caso II

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	$\eta_7$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	0.0754	0.8125	1.4731	18.5369	0.4364	24.7409	110.9047	0.0435	0.1312	0.9077	0.6848
$\text{Log}(X)$	-0.8712	0.0366	0.5283	0.3372	10.9963	20.6455	3.1724	4.92e-4	0.0019	0.9973	0.9894
$\text{Log}(Y)$	-0.1771	1.1393	0.4271	3.7407	0.5233	2.5781	0.7170	0.0067	0.0093	0.9864	0.9813

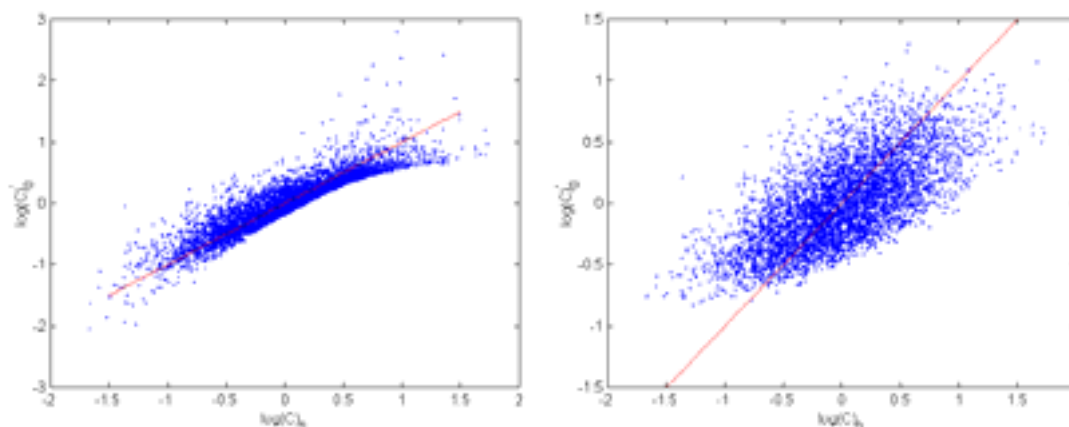


Figura 8.1-39: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso II (destra).

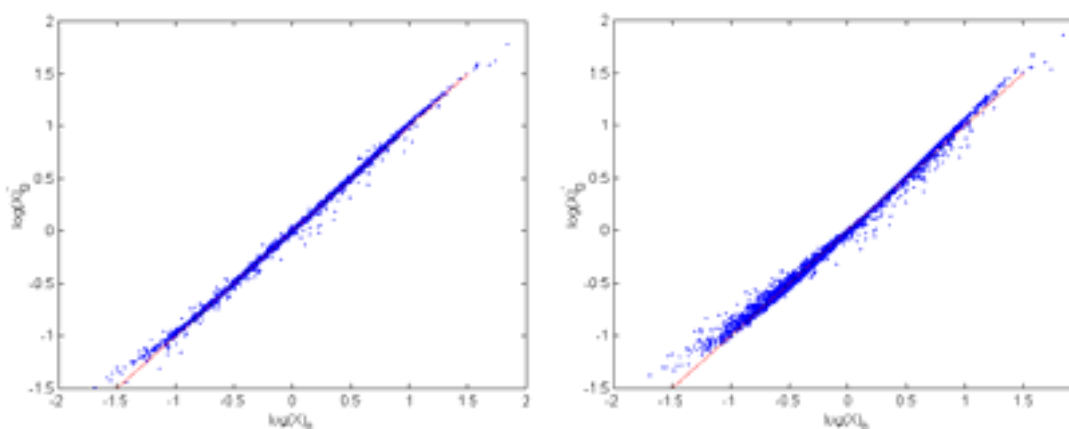


Figura 8.1-55: scatterogrammi del  $\log(X)_0$  (generati da distribuzione log-normale) e del  $\log(X)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso II (destra).

### 8.1.5.3 Algoritmi per acque di caso I-II

Dal confronto dei risultati riportati nella Tabella 8.1-7, Tabella 8.1-19 e Tabella 8.1-27 si nota come gli algoritmi per acque di caso I-II relativi al sensore MERIS hanno praticamente le stesse prestazioni, in termini di MSE e coefficiente di correlazione, degli analoghi ottenuti per i sensori MOS-B e SeaWiFS. L'algoritmo multilineare ancora una volta non è capace di stimare il  $\text{Log}(C)_0$  con precisione adeguata. Lo scatterogramma riportato in Figura 8.1-42 mostra infatti una certa dispersione intorno alla retta  $y=x$ . Per quanto riguarda i parametri  $\text{Log}(X)_0$  e  $\text{Log}(Y)_0$  è stato riscontrato un comportamento analogo a quello riscontrato per le acque di caso II e per gli altri sensori presi in considerazione.

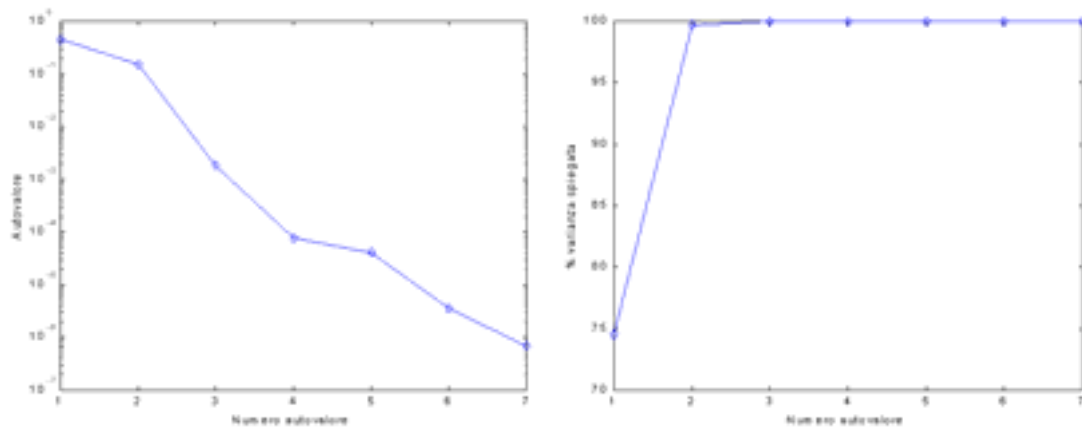


Figura 8.1-41: grafico degli autovalori e della varianza spiegata da ogni componente principale; data set di riflettanze per acque di caso I-II.

Tabella 8.1-27: coefficienti degli algoritmi PCA per acque di caso I-II

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	$\eta_7$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	0.4079	0.7378	0.4436	11.8761	21.8321	2.5797	36.3692	0.0088	0.0402	0.9777	0.8943
$\text{Log}(X)$	0.6237	0.4018	0.1826	2.6452	1.0229	7.6105	20.3796	5.06e-4	0.0016	0.9988	0.9961
$\text{Log}(Y)$	0.2260	1.0660	0.2769	1.9059	2.1686	14.0021	18.9528	0.0020	0.0034	0.9951	0.9915

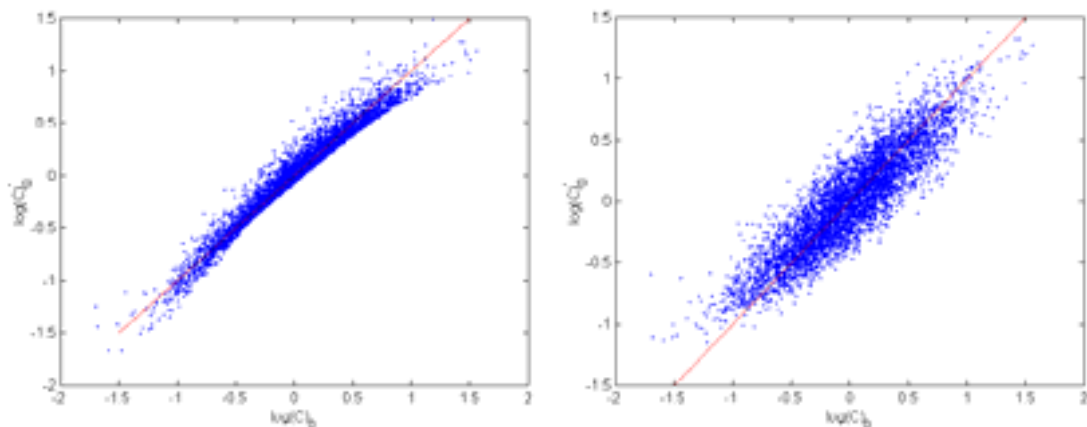


Figura 8.1-42: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze per le acque di caso I-II (destra).

### 8.1.6 Sensore MERIS: algoritmi RBF-NN

Questo sottoparagrafo è dedicato alla breve descrizione di alcuni risultati ottenuti applicando il metodo di inversione RBF-NN al modello di riflettanza. Per brevità i risultati riportati sono relativi alla sola concentrazione cumulata di clorofilla-a e feofitina-a. In generale i risultati ottenuti con tale metodo di inversione hanno mostrato una buona capacità delle RBF-NN di risolvere i problemi di stima dei parametri otticamente attivi laddove gli algoritmi multilineari non sono in grado di ottenere risultati soddisfacenti (si pensi ad esempio al residuo di non linearità che presenta l'algoritmo multilineare per la stima del  $\text{Log}(X)_0$  nelle acque di caso I o alla stima del  $\text{Log}(C)_0$  per le acque di caso II).

Le reti per la stima dei parametri sono state addestrate utilizzando un *data set* di *training* ottenuto scegliendo casualmente 500 punti dai *data set* completi (costituiti da 5000 punti) generati per le acque di caso I, caso II e caso I-II. Le prestazioni delle reti addestrate secondo la procedura di *forward selection* descritta nel capitolo 3, sono state successivamente calcolate (sempre in termini di MSE e di coefficiente di correlazione) sull'intero *data set* di 5000 punti per testare le proprietà di generalizzazione.

#### 8.1.6.1 Algoritmi per acque di caso I

In Figura 8.1-43 sulla sinistra è mostrato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma e relativo all'intero *data set* per le acque di caso I. La rete presenta buone caratteristiche di generalizzazione come è evidente dalla distribuzione dei punti nello scatterogramma relativo al *data set* completo. La Figura 8.1-44 mostra la sequenza dei pesi della rete mentre nella Tabella 8.1-28 sono riportati i parametri della rete e gli indici prestazionali. Il numero di centri della rete è pari a 88 un valore intermedio tra quelli ottenuti per i sensori SeaWiFS e MOS-B. I valori dei coefficienti di correlazione e di MSE sono paragonabili a quelli ottenuti per il sensore SeaWiFS. Le prestazioni sembrano superiori a quelle mostrate dal MOS-B anche se per quest'ultimo vale la considerazione fatta nel paragrafo 8.1.4.1.

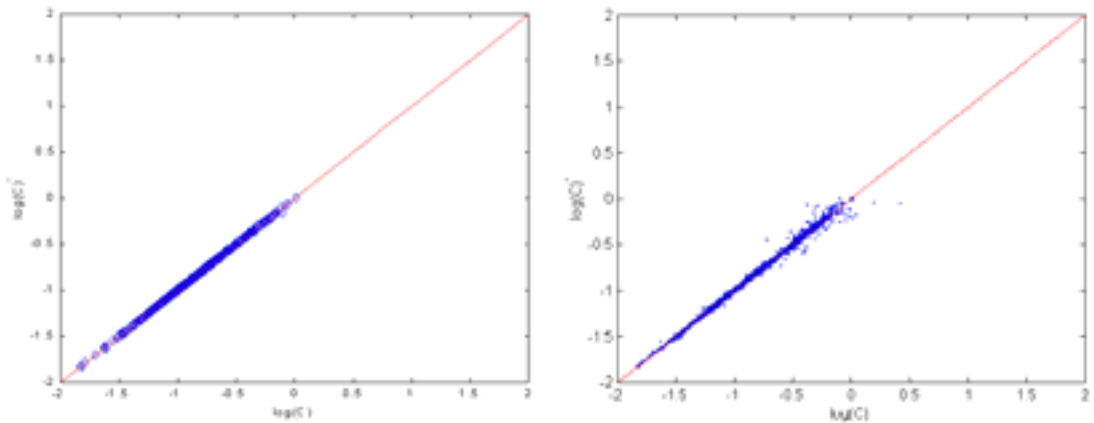


Figura 8.1-43: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del *training set* relativo alle acque di caso I (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set*.

Tabella 8.1-28: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso I

	r	Num Centri	MSE <sub>tr</sub>	MSE <sub>ck</sub>	$\rho_{tr}$	$\rho_{ck}$
$\text{Log}(C)$	0.3	88	3.4777e-005	3.4172e-004	0.9998	0.9981

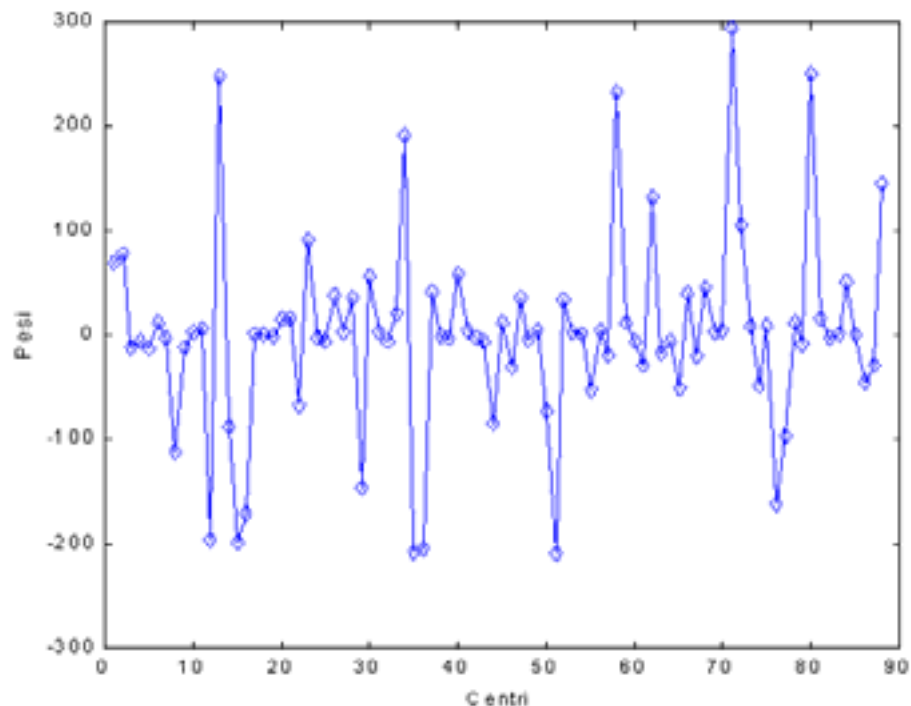


Figura 8.1-44: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso I

#### 8.1.6.2 Algoritmi per acque di caso II

Per le acque di caso II e relativamente al sensore MERIS, l'algoritmo RBF per la stima del  $\text{Log}(C)$  si è dimostrato ancora una volta migliore del corrispondente algoritmo PCA. I risultati dell'addestramento della rete mostrati in Figura 8.1-45 ed in Tabella 8.1-29, confermano tale affermazione. La Figura 8.1-46 mostra la sequenza dei pesi della rete. Le prestazioni della rete sono dello stesso ordine di grandezza di quelle ottenute per il SeaWiFS. Il sensore MOS-B resta il più adatto a stimare la concentrazione cumulata di clorofilla e feofitina.

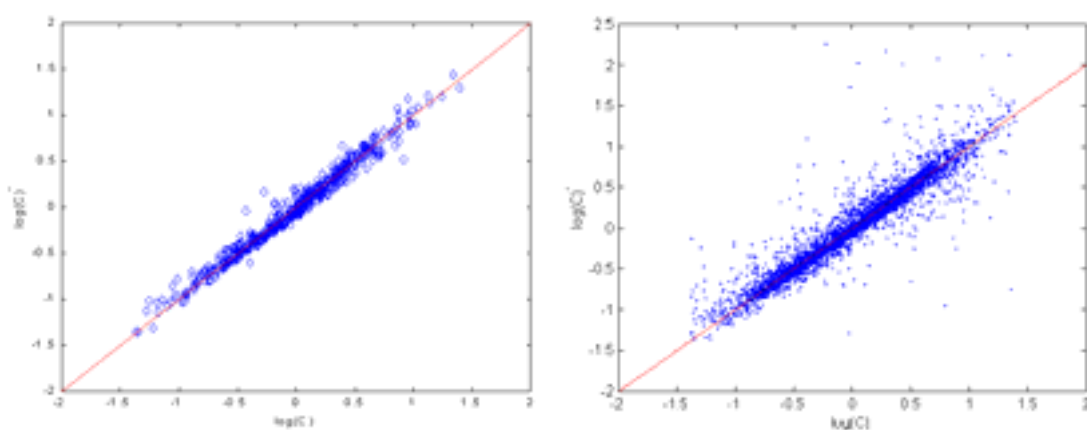


Figura 8.1-45: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo alle acque di caso II (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.1-29: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso II

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.6	91	0.0063	0.0273	0.9875	0.9435

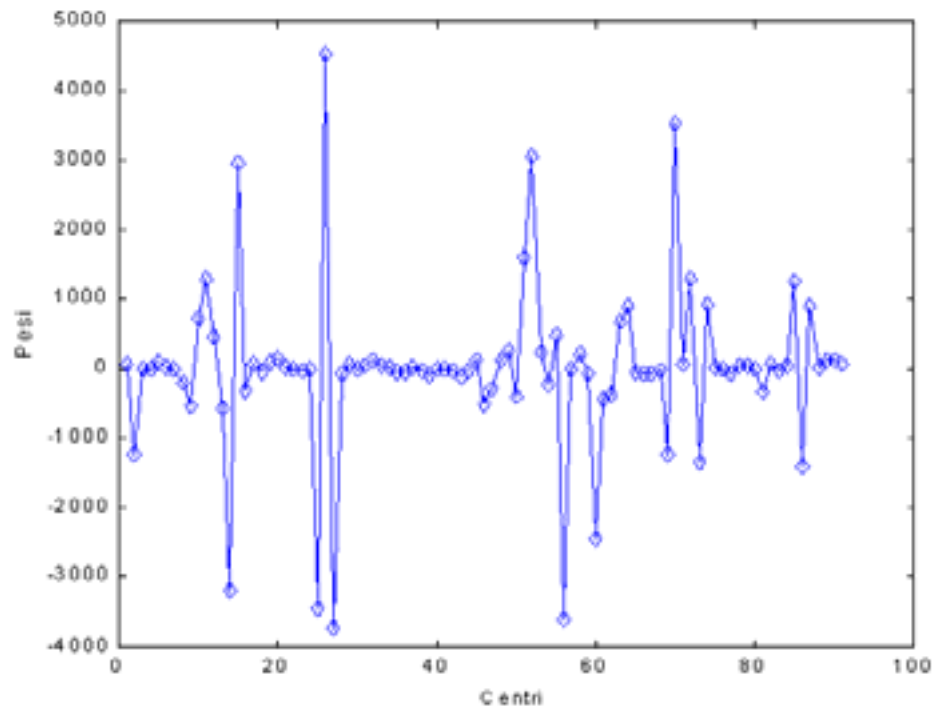


Figura 8.1-46: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso II

### 8.1.6.3 Algoritmi per acque di caso I-II

Per le acque di caso I-II l'algoritmo RBF per la stima del  $\text{Log}(C)$  da dati MERIS è migliore rispetto al corrispondente algoritmo PCA, ma sembra possedere prestazioni inferiori alle reti RBF generate per i sensori SeaWiFS e MOS-B per lo stesso *data set*. Il valore di  $\text{MSE}_{\text{ck}}$  riportato in Tabella 8.1-30 è pari a 0.014 ovvero un ordine di grandezza superiore rispetto ai valori di  $\text{MSE}_{\text{ck}}$  ottenuti per i sensori SeaWiFS e MOS-B. La rete quindi presenta caratteristiche di generalizzazione inferiori.

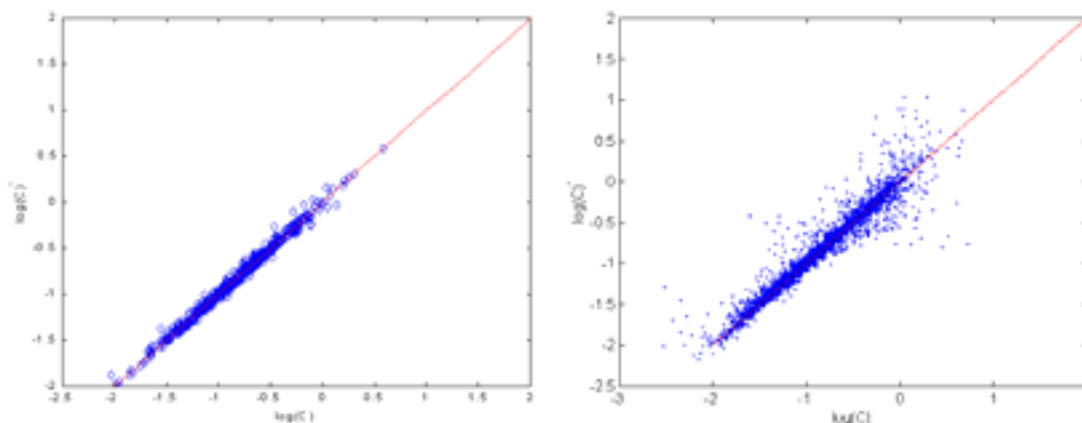


Figura 8.1-47: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relative alle acque di caso I-II (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.1-30: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ ; acque di caso I-II

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.4	94	0.0018	0.0140	0.9950	0.9651

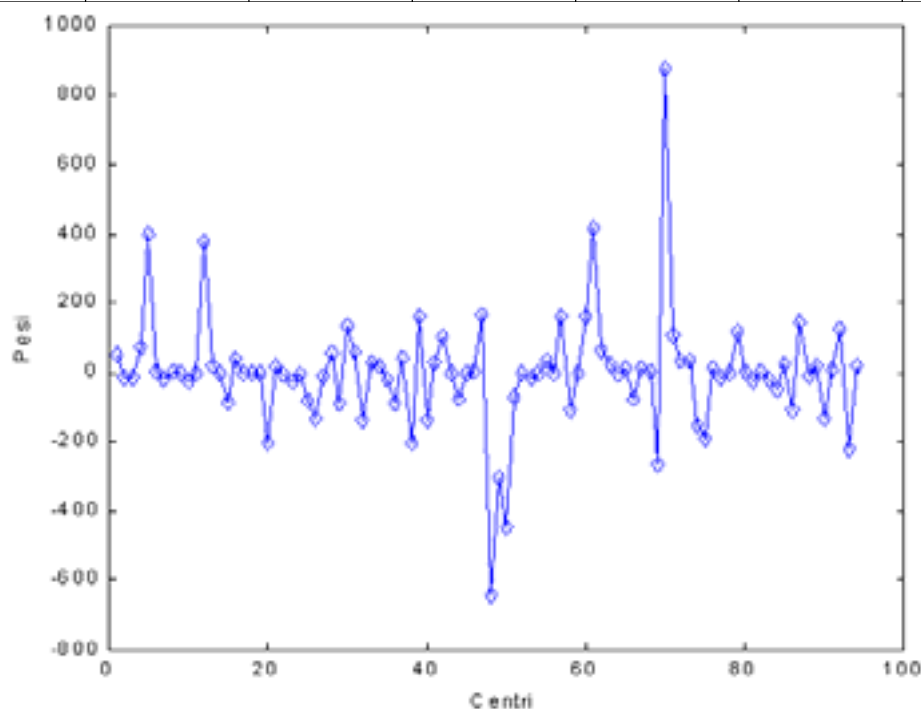


Figura 8.1-48: sequenza dei pesi della rete RBF per la stima del  $\text{Log}(C)$ : data set acque di caso I-II

## 8.2 La stima dei parametri otticamente attivi da dati SeaWiFS: applicazione al sito di sperimentazione del Mare di Alboran

Nel presente paragrafo saranno riportati i risultati dell'applicazione delle tecniche di inversione al *data set* di misure in situ relativo al Mare di Alboran analizzato nel capitolo 4. Il sensore che sarà preso in considerazione è il SeaWiFS dato che è il più adatto tra quelli esaminati ed attualmente operativi, al monitoraggio del bacino di interesse sia dal punto di vista delle caratteristiche geometriche che della frequenza di passaggio. Nei prossimi sottoparagrafi quindi saranno presentati i risultati relativi all'inversione con algoritmi PCA e RBF-NN e una serie di algoritmi a rapporto di bande e polinomiali. Saranno presentati infine alcuni risultati dell'applicazione degli algoritmi definiti a dati reali SeaWiFS.

### 8.2.1 Algoritmi PCA

In Figura 8.2-1 sono mostrati i grafici degli autovalori e della somma cumulata degli autovalori relativi all'analisi PCA del *data set* di riflettanze generato tramite il modello di riflettanza, per il sensore SeaWiFS, dai dati *in situ* acquisiti nel Mare di Alboran. Il grafico mostra come le prime tre componenti principali siano sufficienti a catturare l'intera variabilità dei dati di riflettanza spiegando queste complessivamente più del 99% della varianza totale. Dal grafico di Figura 8.2-2 e di Figura 8.2-4 è possibile notare come il modello inverso multilineare sia adeguato alla stima del  $\text{Log}(C)_0$  e  $\text{Log}(Y)_0$ . Il valore di MSE, riportato in Tabella 8.2-1, è pari a  $1.54\text{e-}4$  per il  $\text{Log}(C)_0$  e  $1.51\text{e-}5$  per il  $\text{Log}(Y)_0$ . Tali valori non subiscono sostanziali incrementi quando si considerano solo i primi tre coefficienti  $\eta_i$ . Come già anticipato nel paragrafo 8.1.1.1 anche per questo *data set*, che può essere preso come esempio di acque di caso I, è stato osservato un residuo di non linearità nello scatterogramma  $\text{Log}(X)_0 - \text{Log}(X)'_0$  (Figura 8.2-3) che mostra come il modello multilineare non sia il più adatto a stimare tale parametro da dati relativi a questo tipo di acque.

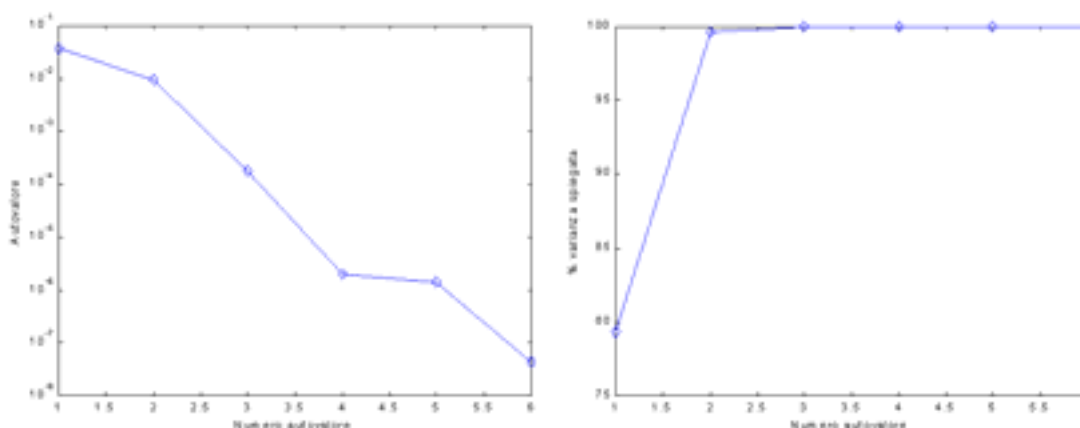


Figura 8.2-1: grafico degli autovalori e della varianza spiegata da ogni componente principale.



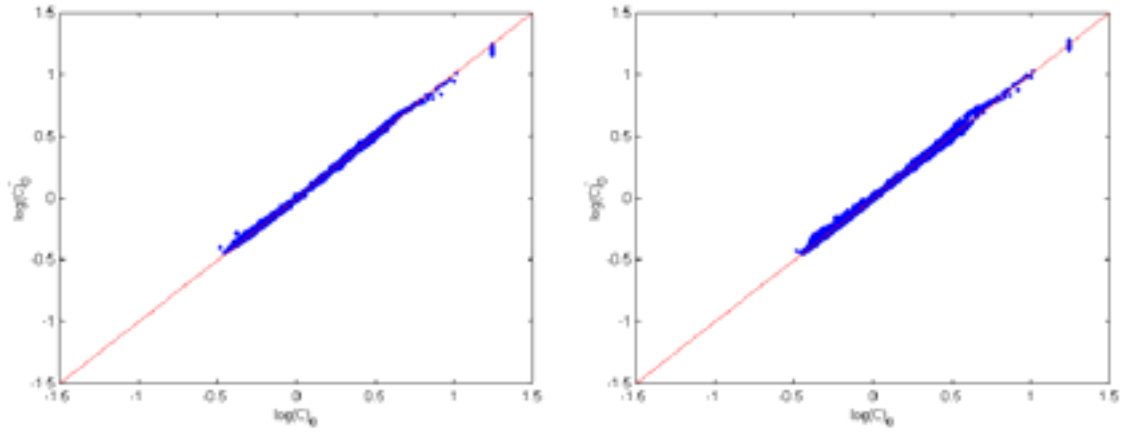


Figura 8.2-2: scatterogrammi del  $\log(C)_0$  (generati da distribuzione log-normale) e del  $\log(C)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze (destra).

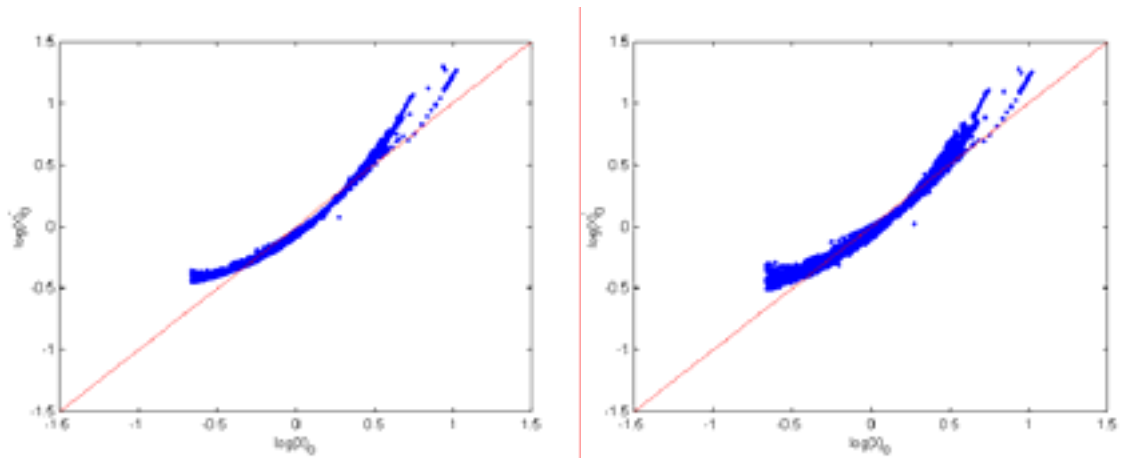


Figura 8.2-3: scatterogrammi del  $\log(X)_0$  (generati da distribuzione log-normale) e del  $\log(X)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze (destra).

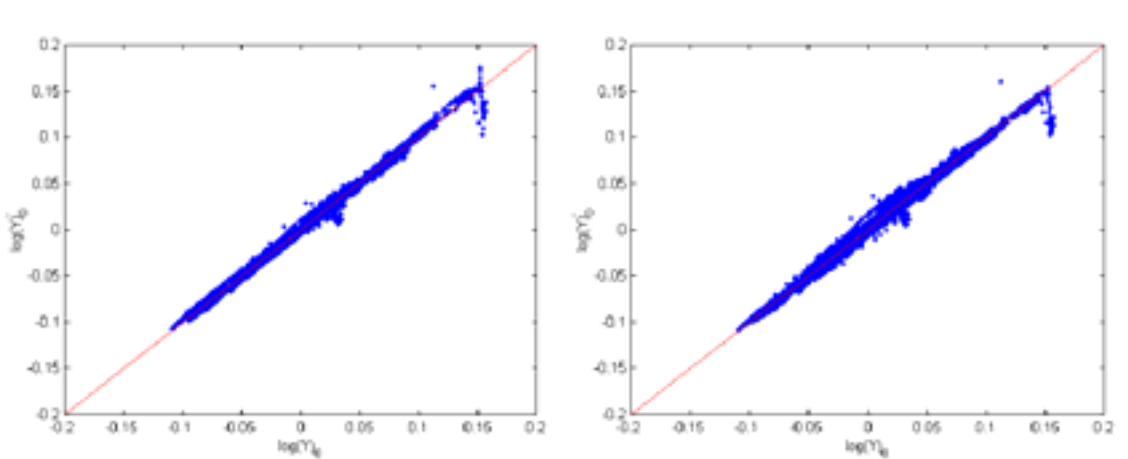


Figura 8.2-4: scatterogrammi del  $\log(Y)_0$  (generati da distribuzione log-normale) e del  $\log(Y)_0$  stimato tramite algoritmo multilineare (sinistra) e tramite le prime tre componenti principali del data set di riflettanze (destra).

Tabella 8.2-1: coefficienti degli algoritmi PCA per il data set del Mare di Alboran

	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$	$\eta_6$	MSE	MSE <sub>3</sub>	$\rho$	$\rho_3$
$\text{Log}(C)$	-1.4097	1.3205	2.6277	-5.9774	6.7540	-21.878	1.54e-4	3.14e-4	0.9992	0.9983
$\text{Log}(X)$	-1.5961	-0.9423	-3.1471	13.9142	-2.3224	25.4065	0.0045	0.0049	0.9793	0.9772
$\text{Log}(Y)$	-0.1620	0.4200	-2.0223	1.6087	-1.5678	0.8149	1.51e-5	2.4e-5	0.9978	0.9965

### 8.2.2 Algoritmi RBF-NN

In Figura 8.2-5 sono mostrati i risultati ottenuti applicando il metodo di inversione RBF al *data set* relativo al Mare di Alboran. La figura mostra quanto ottenuto addestrando una rete RBF per la stima del  $\text{Log}(C)$ .

L'addestramento è stato effettuato con modalità analoghe a quelle usate per i *data set* simulati (il numero di punti del *data set* di *training* è stato fissato a 1000). Il valore di MSE ottenuto sul *data set* di *training* (Tabella 8.2-2) è pari a  $\text{MSE}_{\text{tr}}=1.0157\text{e-}5$  mentre quello ottenuto sull'intero *data set* è pari a  $\text{MSE}_{\text{ck}}=2.3820\text{e-}4$  prestazioni queste che sono comparabili a quelle ottenute con il *data set* simulato per acque di caso I. La rete addestrata ha un numero di neuroni nello strato nascosto pari a 127 ed un parametro di *spreading* pari a  $r=0.1$ .

La Figura 8.2-6 mostra gli analoghi risultati per il  $\text{Log}(X)$ . La rete addestrata presenta 92 neuroni nello strato nascosto ed un raggio  $r=0.25$ . Anche in questo caso il *fitting* prodotto presenta caratteristiche molto buone essendo i valori di MSE dell'ordine di grandezza di  $\text{e-}6$  ed  $\text{e-}5$ .

La Figura 8.2-7 mostra infine i risultati ottenuti per l'algoritmo RBF relativo al  $\text{Log}(Y)$ . Le prestazioni in maniera analoga alle altre due reti RBF l'algoritmo presente

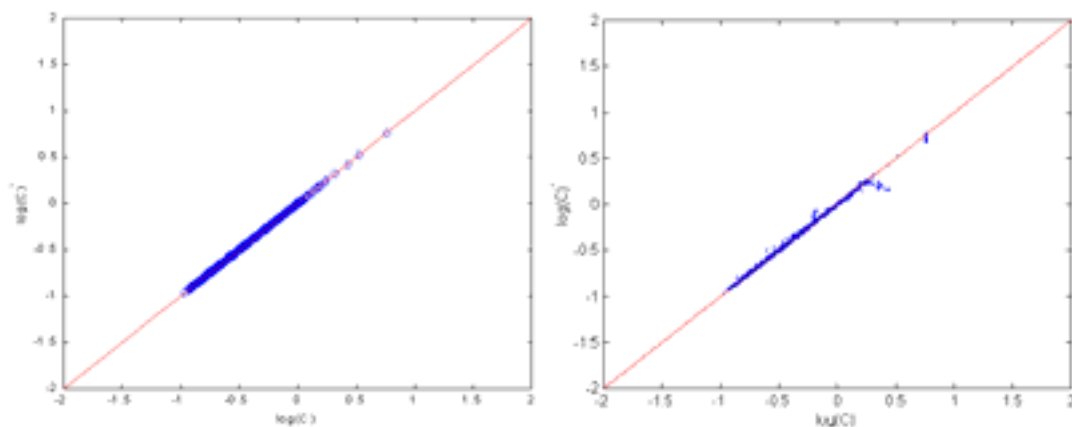


Figura 8.2-5: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(C)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(C)$  del training set relativo al *data set* di misure in situ acquisite nel Mare di Alboran (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero *data set*.

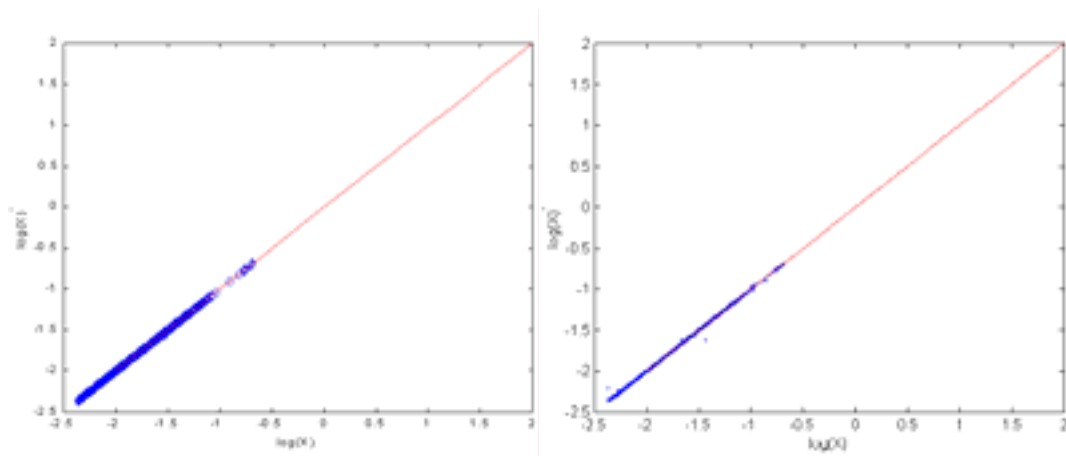


Figura 8.2-6: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(X)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(X)$  del training set relativo al data set di misure in situ acquisite nel Mare di Alboran (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

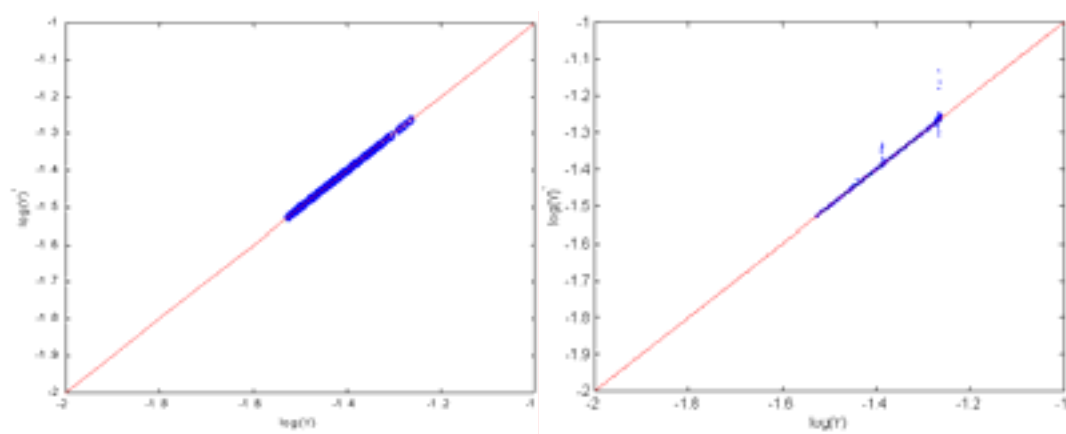


Figura 8.2-7: sulla sinistra è riportato lo scatterogramma del  $\text{Log}(Y)$  in uscita alla rete RBF alla fine dell'addestramento (in ordinate) e dei corrispondenti valori del  $\text{Log}(Y)$  del training set relativo al data set di misure in situ acquisite nel Mare di Alboran (in ascisse); sulla destra si trova graficato lo stesso tipo di scatterogramma, ma relativo all'intero data set.

Tabella 8.2-2: parametri della rete RBF generata per la stima del  $\text{Log}(C)$ .

	r	Num Centri	$\text{MSE}_{\text{tr}}$	$\text{MSE}_{\text{ck}}$	$\rho_{\text{tr}}$	$\rho_{\text{ck}}$
$\text{Log}(C)$	0.1	127	1.0157e-5	2.3820e-4	0.9999	0.9987
$\text{Log}(X)$	0.25	92	3.9851e-6	1.2782e-5	0.9999	0.9985
$\text{Log}(Y)$	0.2	129	6.4831e-8	5.6525e-6	0.9999	0.9992

### 8.2.3 Algoritmi polinomiali ed a rapporto di bande

Tramite il data set di misure in situ relativo al mare di Alboran, sono stati definiti una serie di algoritmi classici a rapporto di bande e polinomiali per il sensore SeaWiFS ([33], [34] e [35]). In Figura 8.2-8, in alto a sinistra, è mostrato lo scatterogramma del  $\text{Log}(C)$  in funzione del logaritmo del rapporto tra le

riflettanze nelle bande 3 e 5 in cui è possibile intravedere una caratteristica lineare (più una costante) esistente tra i due parametri. Lo stesso comportamento può essere notato per quanto riguarda il  $\text{Log}(Y)$  in funzione del logaritmo del rapporto tra le riflettanze nelle bande 2 e 5 (lo scatterogramma relativo è mostrato in Figura 8.2-8 in basso a sinistra).

Lo scatterogramma in alto a destra è relativo al  $\text{Log}(X)$  in funzione del logaritmo della riflettanza in banda 5 e mostra chiaramente un comportamento non lineare tra questi due parametri.

Da tali considerazioni è possibile definire due algoritmi a rapporto di bande per la stima del  $\text{Log}(C)$  e del  $\text{Log}(Y)$  ed un algoritmo polinomiale per la stima del  $\text{Log}(X)$ . Gli algoritmi proposti hanno la seguente forma:

$$\text{Log}(C) = a_0 + a_1 \cdot \text{Log}\left(\frac{R_3}{R_5}\right), \quad \text{Eq. 8.2-1}$$

$$\text{Log}(X) = a_0 + \sum_{n=1}^5 a_n \cdot \text{Log}^n(R_5), \quad \text{Eq. 8.2-2}$$

$$\text{Log}(Y) = a_0 + a_1 \cdot \text{Log}\left(\frac{R_2}{R_5}\right). \quad \text{Eq. 8.2-3}$$

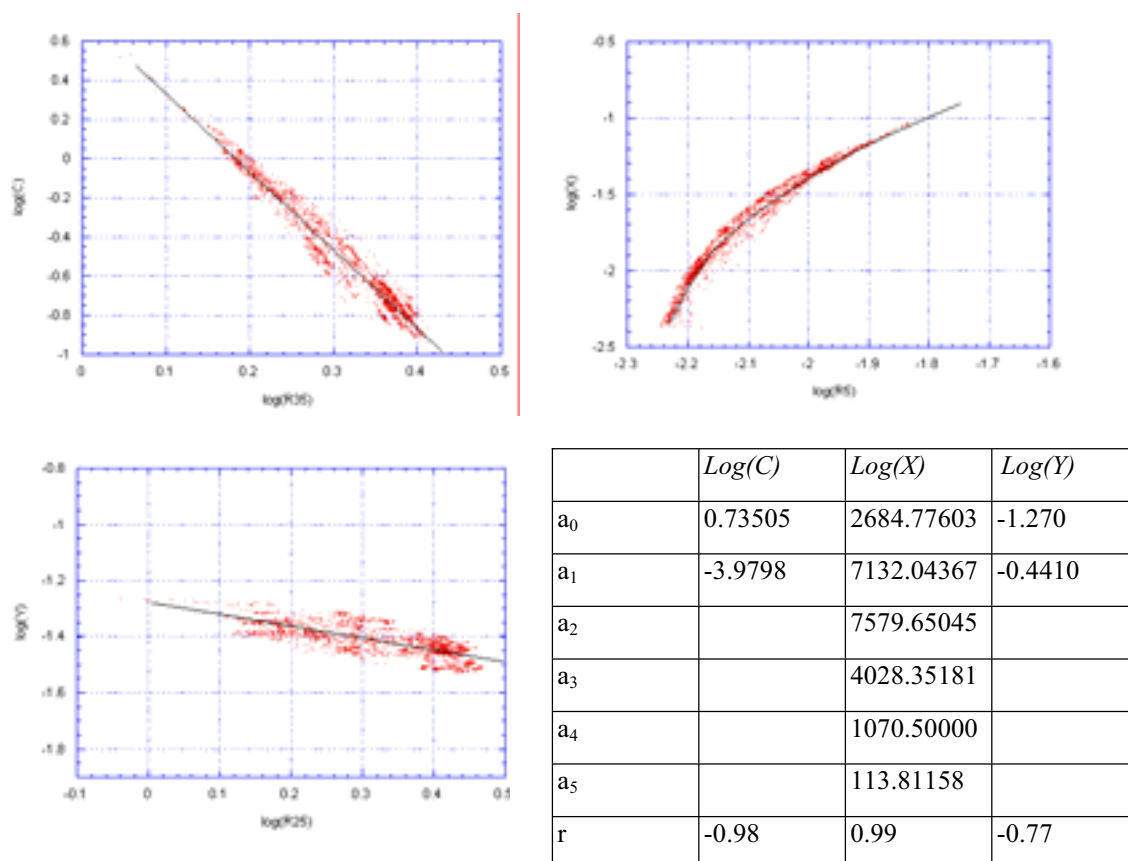


Figura 8.2-8: algoritmi polinomiali ed a rapporto di bande per il sensore SeaWiFS ricavati dal data set di misure in situ acquisite nel Mare di Alboran

I coefficienti delle regressioni sono riportati nella tabella presente nella Figura 8.2-8 in basso a destra. La precisione raggiunta non è paragonabile a quella ottenuta con algoritmi multilineari e a rete RBF. In particolare per quanto riguarda la stima del  $\text{Log}(Y)$  il coefficiente di correlazione ottenuto dalla regressione risulta pari a  $-0.77$  contro lo  $0.99$  osservato per gli algoritmi multilineari ed RBF.

Tali algoritmi sono stati usati per elaborare una sequenza temporale di un anno di immagini relative al Mare di Alboran per studiare la variabilità bio-ottica della zona. Alla fine del capitolo saranno riportati alcuni dei risultati ottenuti da tale elaborazione.

#### 8.2.4 La catena di elaborazioni dei dati telerilevati

In questo sottoparagrafo sono descritte brevemente le operazioni di elaborazione dei dati telerilevati dal sensore SeaWiFS che permettono di estrarre le mappe finali dei parametri otticamente attivi. La Stazione di ricezione da cui sono stati acquisiti i dati è quella del NERC di Dundee in Scozia. In Figura 8.2-9 è rappresentata la zona di copertura a terra di tale stazione. I dati sono organizzati secondo il formato HDF (Hierarchical Data Format) ed hanno livello di elaborazione 1A (Level 1A).

L'HDF è il formato dati sviluppato dal National Center for Supercomputing Applications (NCSA). L'oggetto principale della struttura HDF dei dati SeaWiFS Level 1A, è un array tridimensionale che prende il nome di *Scientific Data Set*. Questo insieme è costituito dai valori delle radianze digitalizzate, per ogni pixel, per ogni linea di scansione e per ogni banda. Tali informazioni vengono accompagnate, a livello 1A, dai dati riguardanti la calibrazione del sensore e la navigazione del satellite.

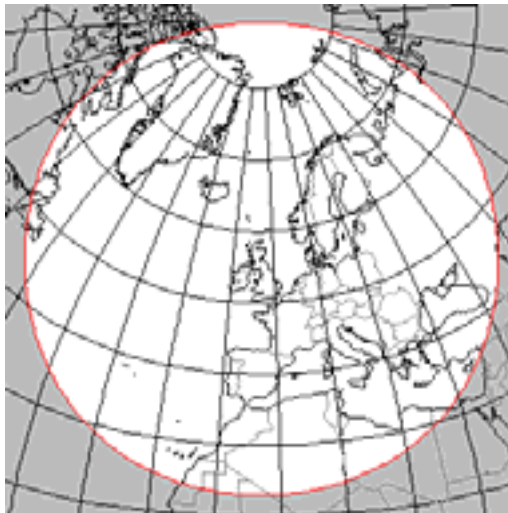


Figura 8.2-9: Zona di copertura a terra della Stazione di ricezione di Dundee

La situazione di partenza prevede otto immagini, pari al numero di bande spettrali del SeaWiFS. Si parla, allo stato iniziale dell'elaborazione, di *immagini grezze o raw*. Ciascuna di queste immagini è costituita da numeri naturali compresi tra 0 e 1023, dal momento che il sensore in esame digitalizza il dato di radianza su 10 bit; ogni livello su cui è quantizzato il segnale viene chiamato *conteggio o digital number*. Le dimensioni delle immagini lette da file sono 1000x1285, dove 1285 sono i pixel che costituiscono ogni linea di scansione del SeaWiFS. Nella Figura 8.2-10 è mostrato, a titolo di esempio, il canale 2, corrispondente a 443 nm, estratto dal file del giorno 10 giugno 1998.

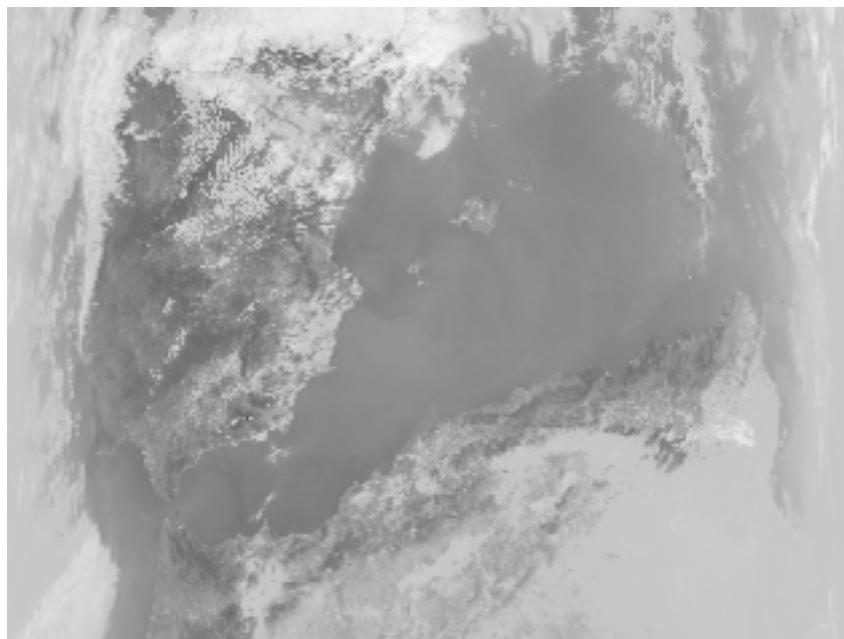


Figura 8.2-10: esempio di immagine SeaWiFS a livello 1A. Canale: 2; data di acquisizione 10-6-1998.

Le operazioni che compongono la catena di elaborazione sono:

- correzione radiometrica;
- classificazione di terra e nubi;
- correzione atmosferica;
- correzione geometrica.

Le prime tre operazioni saranno descritte nei paragrafi seguenti.

#### 8.2.4.1 La correzione radiometrica

Con l'espressione *calibrazione radiometrica*, o *correzione radiometrica*, si identifica quella fase di elaborazione dell'immagine che consente di ottenere i valori di radianza dai *digital number* per ogni banda.

La prima operazione che viene eseguita è il passaggio dal *digital number* misurato a quello corretto. Infatti l'acquisizione dei dati è accompagnata da inevitabili errori di offset introdotti dagli strumenti elettronici a bordo del satellite. La relazione, relativa ad un generico canale, tra il *digital number* corretto ed il *digital number* misurato, è

$$DN_{\text{misurato}} = DN_{\text{corretto}} + \text{offset}_{\text{digitale}} \quad \text{Eq. 8.2-4}$$

Si può notare che quando  $DN_{\text{corretto}}$  vale zero, ad esempio in assenza di segnale, la misura non è nulla ma pari proprio all'offset digitale. Questa considerazione offre lo spunto per correggere i dati; infatti la procedura che viene seguita per stimare l'errore di offset consiste nel compiere la misura nell'intervallo di tempo in cui il sensore non sta puntando terra ed il telescopio ha l'apertura bloccata in modo tale da non poter ricevere segnali. Questo viene fatto per ogni canale e per ogni linea di scansione ed i valori registrati sono definiti *dark pixel*. Tali pixel rappresentano dunque una stima di quello che abbiamo chiamato offset digitale.

L'operazione che viene dunque compiuta è la seguente:

$$\hat{DN}_{corr}[i, j, l] = DN_{mis}[i, j, l] - DARK[i, l] \quad Eq. 8.2-5$$

in cui  $DN_{mis}[i, j, l]$  è il digital number misurato corrispondente alla linea  $i$ , al pixel  $j$ , ed al canale  $l$ ;

$\hat{DN}_{corr}[i, j, l]$  rappresenta il digital number corretto stimato corrispondente alla linea  $i$ , al pixel  $j$ , ed al canale  $l$ ;  $DARK[i, l]$  rappresenta il dark pixel corrispondente alla linea  $i$ , ed al canale  $l$ ; con  $j=1, \dots, 1285$   $l=1, \dots, 8$ . Per semplicità di notazione sarà identificato da adesso il *digital number* corretto stimato con il digital number corretto.

Una volta applicata questa procedura ad ogni pixel degli 8 canali bisogna passare dai conteggi alle radianze. La radianza è la grandezza direttamente misurata dal sensore. A bordo del satellite avviene una quantizzazione su 10 bit del segnale misurato ad opera di un convertitore analogico-digitale: le radianze sono quindi digitalizzate su 1024 livelli. Radianza e conteggio sono legate tra loro da una grandezza chiamata guadagno. In formule si ha dunque:

$$L[i, j, l] = G[l] \times DN_{corr}[i, j, l] = G[l] \times [DN_{mis}[i, j, l] - DARK[i, l]] \quad Eq. 8.2-6$$

$L[i, j, l]$  rappresenta la radianza corrispondente alla linea  $i$ , al pixel  $j$ , ed al canale  $l$ ;  $G[l]$  rappresenta il guadagno corrispondente al canale  $l$ ; con gli indici che variano tra gli estremi prima indicati.

Operativamente il guadagno non viene direttamente utilizzato; piuttosto si lavora su una Tabella di corrispondenza costituita da 1024 righe (pari ai livelli di quantizzazione) ed 8 colonne (pari al numero delle bande). Se viene pensata come una matrice LUT di dimensioni 1024x8, il generico elemento  $LUT(c, l)$  rappresenta il valore di radianza in corrispondenza del conteggio  $c$  e della banda  $l$ , ovvero:

$$L[i, j, l] = LUT[DN_{corr}[i, j], l] \quad Eq. 8.2-7$$

La Tabella in questione è aggiornata periodicamente perché gli strumenti di bordo vengono continuamente calibrati.

Nella Figura 2.4 è riportato un diagramma di flusso che riassume e chiarisce le correzioni fatte sino ad ora per ogni pixel dell'immagine.

#### 8.2.4.2 La classificazione delle nubi e della terra

L'obiettivo di questo lavoro è lo studio del colore marino e perciò si rende necessaria la discriminazione tra mare e terra. Bisogna inoltre evidenziare le zone nuvolose e mascherarle poiché queste non portano alcuna informazione utile.

In letteratura si trovano molti metodi di classificazione basati su algoritmi di una certa complessità. In questo lavoro è stata seguita una procedura molto semplice che ha fornito comunque buoni risultati.

La prima operazione compiuta riguarda appunto la discriminazione dei pixel di mare dai pixel di terra e nuvolosi. Nelle bande del visibile e dell'infrarosso vicino, il mare assume valori più bassi rispetto a quelli relativi a terra e nubi. Si è osservato che questo fenomeno è più accentuato in banda 7, vale a dire nel primo dei due canali del NIR; la mascheratura delle zone non interessanti per l'elaborazione è stata allora fatta analizzando l'istogramma del canale 7 a livello di digital number corretti. In Figura 8.2-11 è raffigurato l'istogramma relativo al giorno 10 giugno 1998.

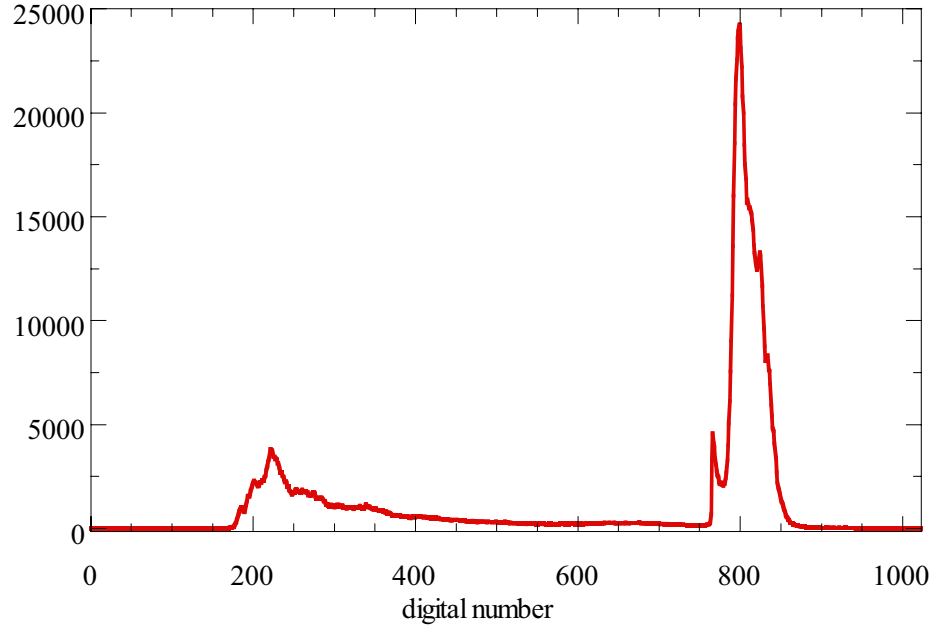


Figura 8.2-11: Canale 7: istogramma dei digital number

Da un esame della Figura si evince che conviene applicare una legge di classificazione a soglia dal momento che l'istogramma risulta costituito da due zone ben separate tra loro. Occorre però evidenziare come uno degli aspetti negativi di tale approccio sia l'intervento di un supervisore che stabilisca la soglia seguendo però criteri soggettivi. Gli inconvenienti legati a tale metodologia sono dunque l'inaccuratezza con cui viene scelto il valore di tale soglia, ma soprattutto l'impossibilità di automatizzare completamente la catena di elaborazione; quest'ultimo fatto è particolarmente pesante se si pensa che bisogna ripetere l'operazione per tutto il data-set. Il valore della soglia è dunque variabile da immagine a immagine ma, come si nota anche dalla Figura 2.5, è sempre compreso tra i 680 ed i 760 digital number. Il test a cui viene sottoposto ogni pixel della scena, una volta stabilita la soglia, è quindi molto semplice:

$$\begin{aligned} \text{DN}_{\text{corr}}[i, j, 7] > \text{soglia} &\Rightarrow M[i, j] = 0 \quad \Leftrightarrow [i, j] \text{ pixel di terra o nuvoloso} \\ \text{DN}_{\text{corr}}[i, j, 7] < \text{soglia} &\Rightarrow M[i, j] = 255 \Leftrightarrow [i, j] \text{ pixel di mare} \end{aligned} \quad \text{Eq. 8.2-8}$$

con gli indici che variano tra i soliti estremi indicati al Paragrafo 2.2. Si genera allora una nuova matrice di dimensioni 1000x1285, che viene indicata con M ed è chiamata genericamente maschera, in cui viene memorizzata l'informazione riguardante la natura dei pixel della scena. Sono stati scelti, come valori caratteristici, 0 per indicare i pixel non significativi e 255 per indicare quelli di mare. Le successive elaborazioni faranno uso di tale maschera in modo tale che solo i pixel utili saranno processati.



#### 8.2.4.3 La correzione atmosferica dei dati

La radianza legata ai costituenti del corpo idrico come messo in evidenza nel capitolo 2, all'interno del paragrafo dedicato al modello atmosferico, è detta *water-leaving*. A questa si aggiungono tutti i contributi dovuti ai fenomeni di diffusione ed assorbimento della radiazione solare presenti in atmosfera. La radianza misurata dal sensore durante la scansione quindi è un segnale corrotto da disturbi di natura atmosferica. Nel capitolo 2 è stato introdotto un modello atmosferico diretto che permette di calcolare la radianza al sensore fissati i valori di un certo numero di parametri oltre che la geometria di ripresa. In questa sezione verranno forniti gli strumenti per risolvere il problema della determinazione della radianza utile. Tale procedura prende il nome di *correzione atmosferica*.

La prima ipotesi di lavoro è quella di considerare trascurabile il termine dovuto alla radianza glitter  $L_g$ . Questo è giustificato dal fatto che il SeaWiFS ha la possibilità di effettuare il tilt, ovvero di ruotare il piano di scansione proprio con lo scopo di "sfuggire" al fascio solare diretto riflesso dalla superficie marina.

La seconda ipotesi semplificativa consiste nel trascurare gli effetti di quel fenomeno che va sotto il nome di *scattering multiplo*. Questo tiene conto, come già messo in luce nei precedenti capitoli, di quei fotoni che vengono scatterati più di una volta prima di entrare nel campo di vista del sensore. Possono avvenire diffusioni multiple ad opera delle molecole dei gas permanenti e delle particelle di aerosol. Nonostante il SeaWiFS, in virtù di una sensibilità superiore rispetto a quella dei precedenti sensori *ocean-color* come il CZCS, necessiti di un'analisi più rigorosa per potere sfruttare al meglio le sue potenzialità, è stata adottata questa semplificazione. La radianza totale al sensore quindi è esprimibile tramite la Eq. 2.2-2.

La seconda ipotesi di lavoro consente di esprimere la radianza dovuta alla diffusione ad opera delle particelle di aerosol e delle particelle di gas permanenti secondo la Eq. 2.2-3.

Per ricavare la radianza *water leaving* dai dati calibrati è necessario sottrarre dalla Eq. 2.2-2 il contributo della radianza di Rayleigh e della radianza dell'aerosol. Il primo contributo è calcolabile in forma chiusa tramite la Eq. 2.2-3 una volta conosciuta la pressione atmosferica, che permette di calcolare lo spessore ottico dei gas permanenti (Eq. 2.2-9), la concentrazione di ozono (fornita come dato ancillare nei files che contengono i dati acquisiti dal sensore), la funzione di fase (Eq. 2.2-8) e gli angoli di zenit e azimut del sole e del sensore. In linea di principio ciò sarebbe possibile anche per il contributo dell'aerosol. Purtroppo non è possibile calcolare queste grandezze; infatti, sia la funzione di fase, sia lo spessore ottico dell'aerosol, dipendono dalle dimensioni delle particelle e dalla loro composizione chimica, quantità che, per l'aerosol, sono altamente variabili nel tempo e nello spazio. Di conseguenza questi termini non sono calcolabili a priori, come invece accade nel caso dello *scattering* di Rayleigh, e quindi il contributo della radianza  $L_a$  non è direttamente determinabile.

La strada che generalmente viene intrapresa per la determinazione del contributo dovuto allo *scattering* dell'aerosol, è quella di utilizzare bande spettrali in cui la radianza *water-leaving* possa essere considerata nota. A questo proposito viene fatta una terza ipotesi di lavoro: si suppone trascurabile la radianza *water-leaving* nelle bande dell'infrarosso vicino ipotesi questa che è verificata sicuramente in acque di caso I [4]. Per quanto detto vengono assunte allora  $L_w(765)$  e  $L_w(865)$  pari a zero per ogni pixel dell'immagine. L'informazione relativa alle radianze *water-leaving* nel NIR, viene integrata dalla determinazione di un coefficiente spettrale che permette di risalire alla radianza  $L_a$  per ogni banda del visibile e per ogni pixel dell'immagine. Si definisce [4] il coefficiente  $\epsilon$  relativo alle bande  $\lambda_i$ ,  $\lambda_j$ , come

$$\varepsilon(\lambda_i, \lambda_j) = \frac{L_a(\lambda_i)}{L_a(\lambda_j)} \times \frac{F_0'(\lambda_j)}{F_0'(\lambda_i)} = \frac{\omega_a(\lambda_i) \tau_a(\lambda_i) p_a(\theta, \theta_0, \lambda_i)}{\omega_a(\lambda_j) \tau_a(\lambda_j) p_a(\theta, \theta_0, \lambda_j)} \quad \text{Eq. 8.2-9}$$

Per descrivere la variazione di  $\varepsilon$  sull'intero dominio delle lunghezze d'onda del SeaWiFS viene considerata la relazione seguente [4]:

$$\varepsilon(\lambda_i, \lambda_j) = \exp\{c(\lambda_j - \lambda_i)\}, \quad \text{Eq. 8.2-10}$$

dove  $c$  dipende dalla geometria (ovvero dagli angoli di zenith ed azimuth di sole e sensore) e dal modello di aerosol, ma non dalla lunghezza d'onda.

Il calcolo di  $\varepsilon(\lambda_7, \lambda_8)$  può essere effettuato tramite la Eq. 8.2-9; infatti, per differenza, vengono determinate le grandezze relative al contributo dell'aerosol nelle bande dell'infrarosso vicino:

$$L_a(\lambda_7) = L_t(\lambda_7) - L_r(\lambda_7), \quad \text{Eq. 8.2-11}$$

$$L_a(\lambda_8) = L_t(\lambda_8) - L_r(\lambda_8). \quad \text{Eq. 8.2-12}$$

Una volta in possesso di  $\varepsilon(\lambda_7, \lambda_8)$ , il coefficiente  $c$  è ricavabile invertendo la (2.23):

$$c = \frac{\ln[\varepsilon(\lambda_7, \lambda_8)]}{\lambda_8 - \lambda_7} \quad \text{Eq. 8.2-13}$$

Si può adesso procedere alla determinazione di  $L_a$  per ogni canale nel visibile invertendo la relazione Eq. 8.2-9:

$$L_a(\lambda_i) = \varepsilon(\lambda_i, \lambda_8) \times L_a(\lambda_8) \times \frac{F_0'(\lambda_i)}{F_0'(\lambda_8)} \quad \text{Eq. 8.2-14}$$

in cui  $\varepsilon(\lambda_i, \lambda_8)$  è calcolato utilizzando la Eq. 8.2-10 con  $c$  stimato tramite la Eq. 8.2-13; l'indice  $i$  varia tra i valori 1 e 6. Questa procedura deve essere ripetuta per ogni pixel dell'immagine. Si osservi come i canali sette ed otto svolgano un ruolo fondamentale nella correzione atmosferica. Le bande spettrali dell'infrarosso vicino esauriscono qui la loro funzione.

Una volta eliminato il contributo della radianza di Rayleigh e dell'aerosol dalla radianza totale al sensore, è possibile stimare la radianza *water leaving* rimuovendo la trasmittanza atmosferica in salita tramite la Eq. 2.2-12, in cui si trascura il termine relativo allo spessore ottico dell'aerosol che risulta di difficile determinazione [5]. Tramite la Eq. 2.3-23 infine è possibile stimare la riflettanza sub-superficiale dalla radianza *water leaving* per ogni canale del sensore.

Rimane da chiarire come calcolare gli angoli di zenith ed azimuth del sole e del sensore che compaiono nelle relazioni riguardanti la correzione atmosferica. Il file da cui vengono letti i dati immagine contiene anche informazioni riguardo alcune caratteristiche del satellite come posizione, velocità ed assetto. Per assetto si intende un vettore di tre componenti che rappresentano i valori di yaw, pitch e roll, ovvero i tre angoli che caratterizzano la posizione del sensore rispetto ad un sistema di riferimento solidale con la piattaforma stessa. È stato implementato un algoritmo, [36], che, accettando queste informazioni come dati di ingresso, restituisce, per ogni pixel dell'immagine, i valori degli angoli  $\theta$ ,  $\theta_0$ ,  $\phi$ ,  $\phi_0$  e le coordinate geografiche latitudine e longitudine che permetteranno di geolocalizzare l'immagine.

### 8.2.5 Alcuni risultati ottenuti da dati reali

La procedura di elaborazione dei dati telerilevati è stata applicata a dati reali relativi al Mare di Alboran. Di seguito sono riportate alcune mappe di parametri otticamente attivi ricavate applicando gli algoritmi RBF-NN ed a rapporto di bande precedentemente definiti. Data la mancanza di dati reali SeaWiFS contemporanei alla campagna oceanografica in cui è stato acquisito il *data set* di misurazioni in situ di parametri otticamente attivi, non è possibile effettuare una validazione degli algoritmi proposti tramite il confronto tra le misure e le stime effettuate dai dati satellitari (si ricordi che il SeaWiFS è operativo dal Settembre 1997 mentre la campagna di misura ha avuto luogo nell'anno 1996). Il problema della validazione resta quindi in sospeso e costituisce la naturale prosecuzione del presente lavoro di ricerca.

Gli algoritmi RBF definiti per la stima della concentrazione cumulata di clorofilla-a e feofitina-a e per la stima del parametro  $X$  sono stati applicati a dati relativi al Mare di Alboran acquisiti il 13 Settembre 1998. In Figura 8.2-13 è mostrata la mappa del parametro  $C$  relativa ad una porzione di immagine situata nei pressi dell'WAG (Western Alboran Gyre). Il *range* di valori riscontrato va da circa  $0.85 \text{ mg/m}^3$  e  $1.2 \text{ mg/m}^3$  che rappresenta un intervallo consistente con quanto riscontrato nel *data set* di misure in situ.

La Figura 8.2-13 mostra la mappa del parametro  $X$  relativa alla medesima zona. L'intervallo di valori registrato va da circa  $0.35 \text{ m}^{-1}$  a  $1 \text{ m}^{-1}$ ; il che non sarebbe in linea con quanto osservato dal *data set* di misure situ.

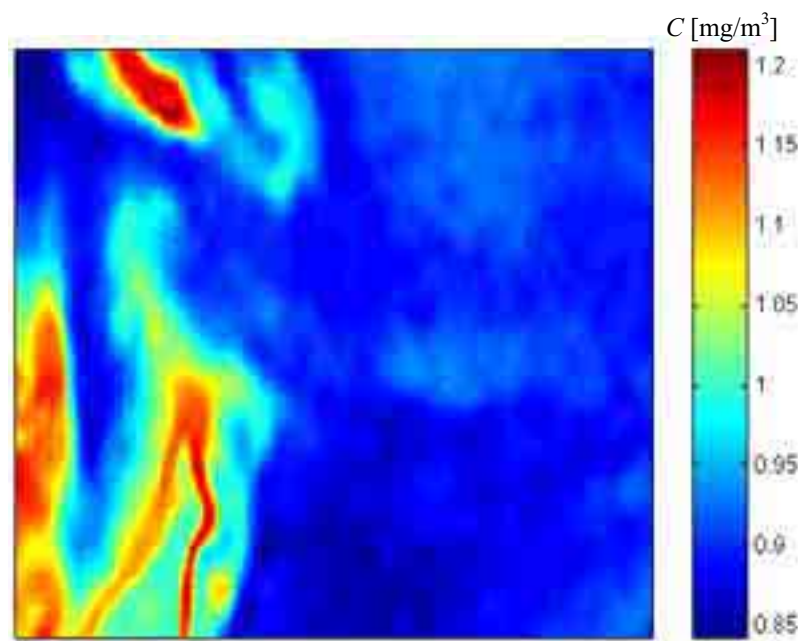


Figura 8.2-12: mappa di concentrazione di clorofilla stimata tramite algoritmo RBF-NN da dati SeaWiFS relativi al Mare di Alboran ed acquisiti il 13 Settembre 1998.

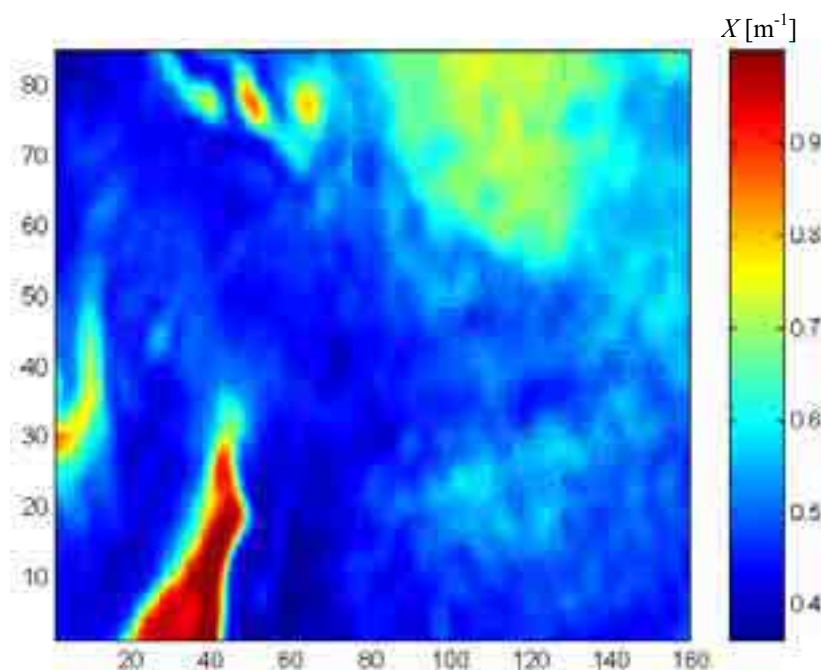


Figura 8.2-13: : mappa di concentrazione di sedimento solido in sospensione espressa in termini del parametro ottico  $X$  stimata tramite algoritmo RBF-NN da dati SeaWiFS relativi al Mare di Alboran ed acquisiti il 13 Settembre 1998.

Nell'ambito del progetto OMEGA, è stato effettuato uno studio relativo alla variabilità bio-ottica, nel tempo e nello spazio, del Mare di Alboran tramite l'analisi EOF (Empirical Orthogonal Function) di una serie temporale di parametri otticamente attivi, stimati da dati SeaWiFS che coprono un periodo di tempo di circa un anno tra il '97 ed il '98 [28]. Lo studio ha messo in luce le connessioni tra la fisica della circolazione e l'attività biologica che caratterizza il sito di sperimentazione. Tale lavoro ha richiesto l'elaborazione di una notevole mole di dati satellitari per arrivare a stimare la serie di mappe di parametri otticamente attivi. La stima di questi dalle riflettanze spettrali è stata effettuata tramite gli algoritmi a rapporto di bande e polinomiali che pur essendo meno precisi dei corrispondenti algoritmi RBF, si sono dimostrati molto robusti nei confronti del tipico rumore di *striping* ([37], [38]) riscontrato nei sensori SeaWiFS e MOS-B. Tali algoritmi quindi non hanno richiesto elaborazioni preliminari di filtraggio cosa che sarebbe stata invece assolutamente necessaria nel caso si fossero usati gli algoritmi RBF. In Figura 2.3-3 è mostrato un esempio di come la conoscenza della distribuzione spaziale della concentrazione dei costituenti otticamente attivi possa essere usata per tracciare le principali strutture oceanografiche presenti nel bacino di interesse. Nelle tre mappe sono chiaramente evidenziati i bordi dei due vortici WAG ed EAG marcati dai valori più alti di concentrazione. La frequenza di passaggio del sensore SeaWiFS permette di monitorare nel tempo l'evoluzione delle strutture oceanografiche di interesse. La Figura 8.2-15 mostra una serie temporale di quattro mappe di concentrazione di clorofilla tra la fine di Aprile e l'inizio di Maggio 1998. L'evoluzione del sistema culmina in un'intensa fioritura di fitoplancton durante il giorno 27 Aprile. Tale fioritura è rilevabile inoltre nel grafico dell'andamento temporale della media spaziale della concentrazione di clorofilla relativo all'intera sequenza (Figura 8.2-16). Intorno al giorno dell'anno 180 (l'origine dell'asse dei tempi è posizionata intorno al 22 Novembre 1997), ovvero intorno al periodo considerato nella sequenza di Figura 8.2-15, è posizionato un picco di concentrazione pari a circa  $0.6 \text{ mg/m}^3$ . La Figura 8.2-17 mostra inoltre l'andamento temporale della deviazione standard spaziale della sequenza. La deviazione standard ha valori nel tempo che sono paragonabili a quelli della corrispondente

media a dimostrazione dell'alta variabilità che la concentrazione di clorofilla presenta nel bacino. I grafici mettono in luce un secondo periodo di fioritura in inverno con un picco di concentrazione media pari a circa  $0.8 \text{ mg/m}^3$ .

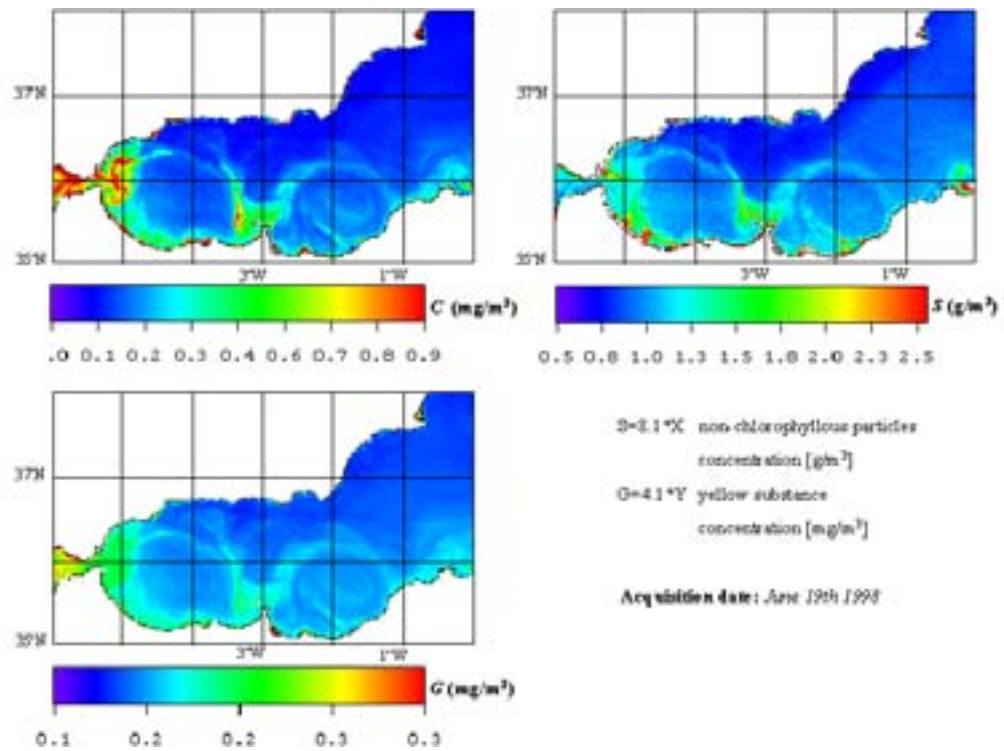


Figura 8.2-14: mappe dei tre parametri otticamente attivi stimati tramite algoritmi a rapporto di bande e polinomiali; l'esempio mostra come la conoscenza delle concentrazioni dei parametri otticamente attivi sull'area in studio possa rivelarsi utile al tracciamento delle principali strutture oceanografiche presenti.

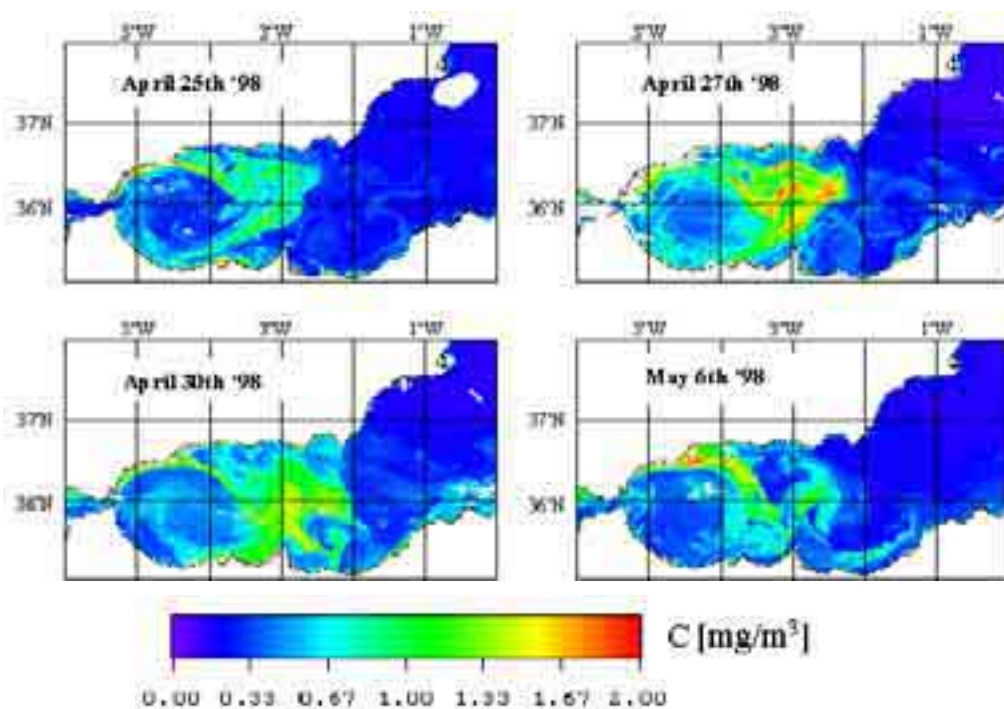


Figura 8.2-15: serie temporale di mappe di concentrazione di clorofilla stimata tramite algoritmo a rapporto di bande.

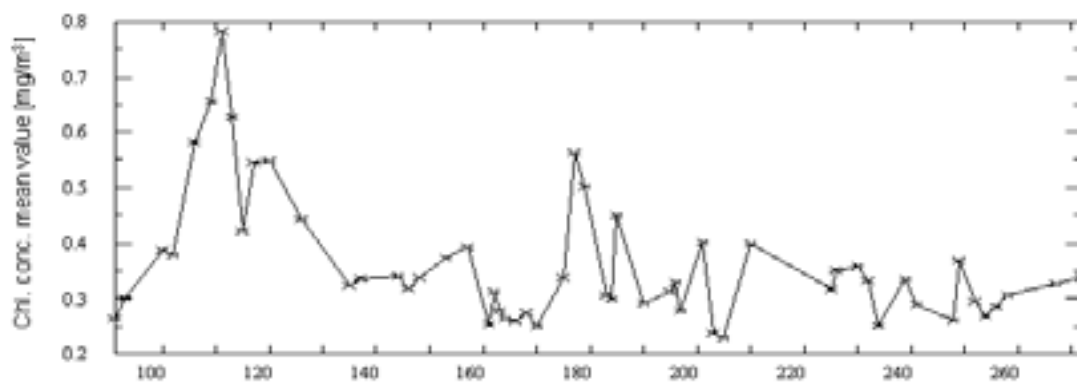


Figura 8.2-16: andamento temporale della media spaziale della concentrazione di clorofilla sul Mare di Alboran. Le medie spaziali sono state effettuate sulle mappe stimate da una serie temporale di dati SeaWiFS che copre il periodo Novembre 1997-Ottobre 1998.

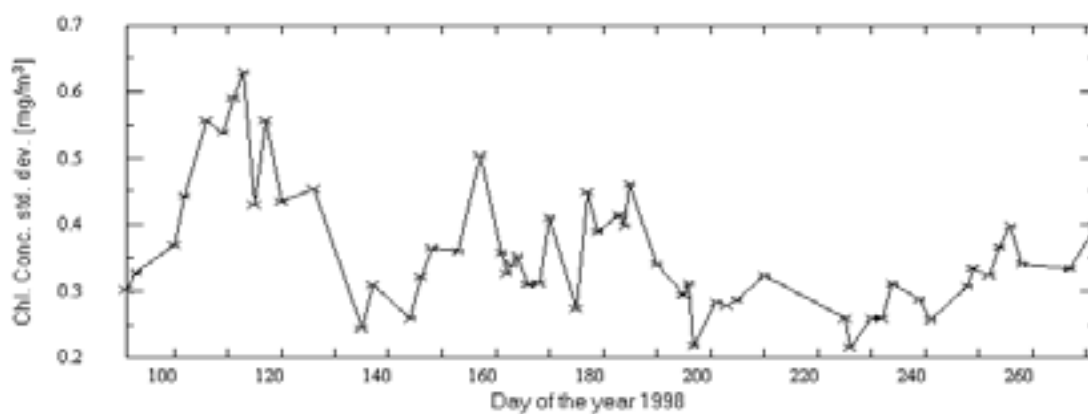


Figura 8.2-17: andamento temporale della deviazione standard spaziale della concentrazione di clorofilla sul Mare di Alboran. La deviazioni standard spaziali sono state ricavate dalle mappe stimate da una serie temporale di dati SeaWiFS che copre il periodo Novembre 1997-Ottobre 1998.

---

# Conclusioni

Nel presente lavoro di tesi è stato affrontato il problema della stima dei parametri otticamente attivi dell'acqua marina da dati passivi multispettrali acquisiti da sensori montati su piattaforma satellitare. L'attività di ricerca ha avuto come scopo quello di definire uno schema d'inversione che permetta di risalire alle concentrazioni dei costituenti otticamente attivi dell'acqua marina dalla riflettanza sub-superficiale stimata dalle misure satellitari. L'approccio usato per la soluzione del problema è stato di tipo *model-based*: attraverso un esame approfondito dei risultati presenti in letteratura è stato definito un modello diretto che, dai valori dei tre parametri otticamente attivi (OAPs)  $C$ ,  $X$ ,  $Y$  (rispettivamente concentrazione cumulata massa su volume di clorofilla-a e feofitina-a, coefficiente di backscattering misurato a 550 nm di lunghezza d'onda del sedimento solido in sospensione e coefficiente di assorbimento della sostanza gialla a 440 nm di lunghezza d'onda) e dai parametri che regolano la propagazione della radiazione luminosa attraverso l'atmosfera, permette di calcolare i contributi alla radianza totale che giunge sulla superficie del sensore. Questa risulta costituita dalla somma di due contributi quello costituito dalla radianza atmosferica ed il contributo di radianza *water-leaving* che rappresenta il segnale utile per la stima degli OAPs. Il modello diretto è costituito dal modello atmosferico e dal modello di riflettanza spettrale sub-suoperficiale, il cosiddetto modello di colore marino legato direttamente ai parametri otticamente attivi. L'inversione del modello atmosferico, che costituisce la cosiddetta fase di correzione atmosferica dei dati telerilevati, permette la stima della riflettanza sub-superficiale in ciascuno dei canali spettrali del sensore. L'inversione del modello di colore marino rende possibile infine la stima dei tre OAPs dai dati telerilevati. Il lavoro di ricerca è stato focalizzato in maniera particolare su quest'ultima tematica ovvero, sullo studio di tecniche di inversione del modello di colore marino che permettano la stima degli OAPs sfruttando tutta l'informazione spettrale messa a disposizione dai dati telerilevati nei differenti canali del sensore. Tali algoritmi multicanale permettono di ottenere una maggiore precisione di stima rispetto ai tradizionali algoritmi che fanno uso di uno o due canali messi a disposizione dal sensore e che non posseggono le prestazioni adeguate per le applicazioni rivolte al monitoraggio delle acque costiere (in cui sono maggiormente presenti le acque di caso II). L'inversione del modello di colore marino è stata effettuata tramite la stima dei coefficienti di un modello inverso, scelto a priori, che accetta in ingresso il vettore delle riflettanze spettrali stimate dai dati telerilevati e restituisce in uscita la stima del valore degli OAPs. La stima dei coefficienti è stata effettuata presentando in ingresso al modello inverso una serie di vettori di riflettanze spettrali, appartenenti ad un *training set* opportunamente generato, e confrontando l'uscita del modello con l'uscita desiderata. I coefficienti vengono quindi stimati minimizzando l'errore quadratico medio sull'intero *training set* tra l'uscita desiderata e l'uscita attuale.

Le seguenti tecniche di inversione sono state prese in considerazione:

- modelli inversi multilineari e regressione PCA

- modelli inversi a rete neurale RBF (*Radial Basis Function*).

L'uso di reti neurali RBF per la stima dei parametri otticamente attivi dell'acqua marina costituisce l'aspetto più innovativo del lavoro di ricerca. I risultati presentati dimostrano che l'uso di reti RBF consente di stimare gli OAPs con prestazioni in generale superiori rispetto ai modelli inversi multilineari (e quindi anche rispetto ai semplici algoritmi a rapporto di bande che costituiscono un caso particolare dei modelli multilineari).

La sperimentazione delle tecniche prese in considerazione ha richiesto la generazione di *training sets* che permettessero la stima dei coefficienti dei modelli inversi. Tali *data sets* sono stati generati presentando in ingresso al modello di colore marino una serie di triplette  $C, X, Y$  e calcolando la riflettanza media sulle bande del sensore preso di volta in volta in considerazione. La generazione dell'insieme di triplette ha richiesto lo studio della variabilità statistica dei valori degli OAPs nelle varie condizioni che possono verificarsi in natura. A tale scopo è stato considerato un modello di densità di probabilità degli OAPs di tipo log-normale. Su tale modello sono state effettuate una serie di verifiche sperimentali tramite lo studio statistico di un certo numero di misurazioni *in situ* prelevate in vari punti del Mar Mediterraneo. Tali verifiche costituiscono una validazione sia pur di tipo locale del modello statistico preso in esame, cosa che di per se costituisce un importante risultato del lavoro di ricerca svolto. Il modello è stato successivamente utilizzato per generare le triplette di OAPs tramite un generatore di variabili casuali per diverse condizioni che è possibile riscontrare in natura (acque di caso I, caso II, caso I-II).

Le metodiche proposte sono state testate su tre sensori di ultima generazione: il SeaWiFS ed il MOS-B attualmente operativi, ed il MERIS di prossima messa in orbita.

I risultati ottenuti per le acque di caso I hanno messo in evidenza le buone prestazioni degli algoritmi multilineari quando sono utilizzati per la stima del parametro  $C$ . Il parametro  $X$ , sempre nelle condizioni operative di acque di caso I, non è stimabile in maniera adeguata tramite un algoritmo multilineare essendo stato osservato un residuo di non linearità negli scatterogrammi del  $\text{Log}(X)$  stimato in funzione dei valori del  $\text{Log}(X)$  desiderato. Tale comportamento risulta più marcato nel caso del sensore SeaWiFS; nei sensori MOS-B e MERIS, data la maggiore risoluzione spettrale ed il maggior numero di canali, gli effetti della non linearità sono meno evidenti. Nel caso in cui la non linearità è presente, la rete neurale RBF opportunamente addestrata costituisce un buon modello inverso.

Per quanto riguarda le acque di caso II è stato osservato che non è possibile stimare il  $\text{Log}(C)$  con un semplice modello multilineare. Per acque di caso II è necessario quindi l'utilizzo di modelli non lineari. Le reti RBF data la loro flessibilità possono adattarsi ad una ampia gamma di modelli inversi. Per acque di caso II e per la stima del  $\text{Log}(C)$  le reti RBF hanno mostrato di possedere prestazioni superiori ai modelli multilineari. Al contrario i modelli multilineari hanno mostrato di essere sufficientemente precisi quando sono utilizzati per stimare il  $\text{Log}(X)$  ed il  $\text{Log}(Y)$ .

In generale gli algoritmi a rete neurale RBF hanno mostrato di possedere prestazioni (in termini di errore quadratico medio e di coefficiente di correlazione) superiore ai modelli multilineari. La configurazione spettrale che in generale risulta avere le migliori prestazioni in tutte le condizioni operative esaminate è quella del sensore MOS-B.

Le tecniche analizzate sono state successivamente applicate ad una *data set* di misurazioni *in situ* prelevate durante una serie di campagne oceanografiche condotte nel Mare di Alboran nell'ambito del progetto europeo OMEGA. Tale *data set* ha permesso di definire una serie di algoritmi per la stima dei parametri



otticamente attivi in tale bacino. Il sensore preso in considerazione in questa ultima fase è stato il SeaWiFS che presenta le caratteristiche più adatte al monitoraggio di un bacino quale quello del Mare di Alboran. Per tale sensore sono stati definiti una serie di algoritmi multilineri, a rete neurale RBF e a rapporto di bande ed a singola banda. Nell'ambito del progetto OMEGA è stato condotto uno studio sulla variabilità bio-ottica sia spaziale che temporale del Mare di Alboran. Gli algoritmi a rapporto di bande sono stati utilizzati per elaborare una serie temporale di dati SeaWiFS che copre un periodo di circa un anno tra il 1997 ed il 1998. In tal modo è stato possibile individuare i periodi di massima produttività del fitoplancton e analizzare l'evoluzione temporale delle tipiche strutture oceanografiche che caratterizzano il bacino.

Gli sviluppi futuri del presente lavoro dovranno prevedere un'intensa attività di validazione degli algoritmi definiti da dati misurati *in situ* ed un confronto con gli algoritmi attualmente in uso per gli studi di colore marino. Un ulteriore aspetto che merita una più approfondita analisi nel proseguo dell'attività di ricerca è quello relativo alle prestazioni degli algoritmi multicanale in presenza di rumore nei dati di riflettanza.

---

# Bibliografia

- [1] Sathyendranath, S., L. Prieur, A. Morel, "A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters," *Int. J. Remote Sensing*, vol. 10, no. 8, pp. 1373-1394, 1989.
- [2] Watson W. Gregg, Nicolas Hoepffner, John Parslow, Trevor Platt, Michael Rast, Shubha Sathyendranath, Tasuku Tanaka, James A. Yoder, "Status and plans for Satellite Ocean-Colour Missions: Considerations for Complementary Missions", *Report of the International Ocean-Colour Coordinating Group*, no. 2, 1999.
- [3] P. Cipollini, "Valutazione delle prestazioni di sensori per il telerilevamento del mare mediante un modello di riflettanza", *tesi di dottorato*, Pisa, 1996.
- [4] Gordon, H. R., M. Wang, "Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm", *Applied Optics*, vol. 33, no. 3, pp. 443-452, 1994.
- [5] Gordon, H. R., D. K. Clark, J. W. Brown, O. B. Brown, R. H. Evans, W. W. Broenkow, "Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates," *Applied Optics*, vol. 22, no. 1, pp. 20-36, 1983.
- [6] Vigroux, "Contribution a l'etude experimentale de l'absorption de l'ozone", *Ann. Physic.*, 8, pp. 709-762, 1953.
- [7] Zibordi G., "Correzione degli effetti atmosferici in dati telerilevati sul mare nel visibile e vicino infrarosso", IMGA-CNR, *rapporto tecnico interno*.
- [8] Shubha Sathyendranath, Trevor Platt, "Analytic model of ocean color", *Applied Optics*, vol. 36, no. 12, pp. 2620-2629, 1997.
- [9] Gordon, H. R., O. B. Brown, M. M. Jacobs, "Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean," *Applied Optics*, vol.14, pp. 417-427, 1975.
- [10] Morel, A., L. Prieur, "Analysis of variations in ocean colour," *Limnology and Oceanography*, vol. 22, no. 4, pp. 709-722, 1977.
- [11] Smith, R. C., K. S. Baker, "Optical properties of the clearest natural waters," *Applied Optics*, vol. 20, no. 2, pp. 177-184, 1981.
- [12] Bricaud, A., M. Babin, A. Morel, H. Claustre, "Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization," *Journal of Geophysical Research*, vol. 100, no. C7, pp. 13321-13332, 1995.

- [13] Prieur, L., S. Sathyendranath, "An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials," *Limnology and Oceanography*, vol. 26, no. 4, pp. 671-689, 1981.
- [14] Bricaud, A., A. Morel, M. Babin, K. Allali, H. Clauser, "Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models", *Journal of Geophysical Research*, vol. 103, no. C13, pp. 31033-31044, 1998.
- [15] Morel, A., L. Prieur, "Analyse spectrale de l'absorption par les substances dissoutes (substances jaunes)," *Publ. CNEXO 10*, Sect. 1.1.11, pp. 1-9, 1976.
- [16] Aiken, J., G. F. Moore, C. C. Trees, S. B. Hooker, D. K. Clark, "The SeaWiFS CZCS-type pigment algorithm," *NASA Tech. Memorandum 104566*, SeaWiFS Technical Report Series, vol. 29, June 1995.
- [17] Morel, A., "Optical properties of pure water and pure sea water," in Jerlov, N. G., Steemann Nielsen, E., *Optical aspects of oceanography*, pp 1-24, Academic Press, 1974.
- [18] Morel, A., "Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters)," *Journal of Geophysical Research*, vol. 93, no. C9, pp. 10749-10768, 1988.
- [19] Ashish Sen, Muni Srivastava, "Regression analysis: Theory, methods, and applications", Springer-Verlag New York Inc., 1990
- [20] Hayken, S., "Neural networks: a comprehensive foundation", Prentice Hall, 1994.
- [21] M. J. L. Orr, "Introduction to Radial Basis Function Networks", Centre for Cognitive Science, University of Edinburgh, *rapporto tecnico interno*, Edinburgh, 1996.
- [22] Campbell, J. W., "The lognormal distribution as a model for bio-optical variability in the sea," *Journal of Geophysical Research*, vol. 100, no. C7, pp. 13237-13254, 1995.
- [23] Doerffer, R., "Untersuchungen über die Verteilung oberflächennaher Substanzen im Elbe-Ästuar mit Hilfe von Fernmessverfahren (the distribution of substances in the Elbe estuary determined by remote sensing)," *Archiv für Hydrobiologie Suppl.*, vol. 43, 119, 1979.
- [24] Mueller, J. L., R. W. Austin, "Ocean optics protocols for SeaWiFS validation, revision 1," *NASA Tech. Memorandum 104566*, SeaWiFS Technical Report Series, vol. 25, February 1995.
- [25] Gordon, H. R., A. Morel, "Remote assessment of ocean color for interpretation of satellite visible imagery, a review," in *Lecture notes on Coastal and Estuarine Studies*, 114 p., Springer-Verlag, New York, 1983.

- [26] Schmitz-Peiffer, A., H. Grassl, "Remote sensing of coastal waters by airborne lidar and satellite radiometer. Part 1: a model study," *Int. J. Remote Sensing*, vol. 11, no. 12, pp. 2163-2184, 1990.
- [27] Schmitz-Peiffer, A., T. Viehoff, H. Grassl, "Remote sensing of coastal waters by airborne lidar and satellite radiometer. Part 2: measurements," *Int. J. Remote Sensing*, vol. 11, no. 12, pp. 2185-2204, 1990.
- [28] A. Baldacci, G. Corsini, R. Grasso, G. Manzella, J. T. Allen, P. Cipollini, T. H. Guymmer, H. M. Snaith, "A study of the Alboran sea mesoscale system by means of EOF (Empirical Orthogonal Function) decomposition of satellite data", *sottomesso al Journal of Marine System*, 1999
- [29] Viudez, A., J. M. Pinot, R. L. Haney, 1998. On the upper layer circulation in the Alboran sea, *J. Geophys. Res.*, 103, 21653-21666.
- [30] Tintorè J., P. E. La Violette, I. Blade and A. Cruzado, 1988. A study of an intense density front in the eastern Alboran Sea: the Almeria-Oran front, *J. Phys. Oceanogr.*, 18, 1384-1397.
- [31] Gascard, J. C. and C. Richez, 1985. Water masses and circulation in the western Alboran Sea and in the Strait of Gibraltar, *Prog. Oceanogr.*, 15, 157-216.
- [32] Ruiz J. *et al.*, 1999. "Surface distribution of chlorophyll particles and gelbstoff in the Atlantic jet of the Alboran Sea: from submesoscale to subinertial scales of variability". *Submitted to the special issue of the J. Mar. Syst. on 31<sup>st</sup> International Liege Colloquium*.
- [33] G. Corsini, P. Cipollini, M. Diani, R. Grasso, "Semi-Analytical algorithms for ocean colour analysis from SeaWiFS data in the Western Mediterranean Sea", *Proc. 3rd International Workshop on MOS-IRS and Ocean Colour*, Institute of Space Sensor Technology, Berlin, 1999.
- [34] G. Corsini, R. Grasso, J. Ruiz, P. Cipollini, "Retrieval of the concentration of optically active constituents in the Alboran Sea: a semi-analytic approach", *Proc. IEEE 1999 International Geoscience and Remote Sensing Symposium*, Amburgo, 28 Giugno-2 Luglio, 1999.
- [35] G. Corsini, M. Diani, R. Grasso, C. Santella, "The potential of SeaWiFS and MOS for the investigation of the bio-optical variability in mesoscale systems", *Proc. 4th International Congress on Energy, Environment and Technological Innovation*, Roma, Settembre 19-24, 1999.
- [36] Patt, F. S., W. W. Greg, "Exact closet-form geolocation algorithm for Earth survey sensors", *International Journal of Remote Sensing*, vol. 15, no. 18, pp. 3719-3734, 1994.
- [37] Giovanni Corsini, Marco Diani, Raffaele Grasso, Rinaldo Rinaldi, Thomas Walzel, "A method for removing striping in MOS-B data", *Proc. 2nd International Workshop on MOS-IRS and Ocean Colour*, Institute of Space Sensor Technology, Berlin, June 10-12 1998, ISBN 3-89685-559-X, pp. 53-63.

- [38] G. Corsini, M. Diani, R. Grasso, R. Rinaldi, T. Walzer, "MOS-B sensor calibration: a signal processing approach", *Proc. IEEE 1999 International Geoscience and Remote Sensing Symposium*, Amburgo, 28 Giugno-2 Luglio, 1999.

### **Lavori prodotti durante l'attività di dottorato**

#### **Lavori a congressi internazionali**

- [1] Giovanni Corsini, Marco Diani, Raffaele Grasso, Rinaldo Rinaldi, Thomas Walzel, "A method for removing striping in MOS-B data", *Proc. 2nd International Workshop on MOS-IRS and Ocean Colour*, Institute of Space Sensor Technology, Berlin, June 10-12 1998, ISBN 3-89685-559-X, pp. 53-63.
- [2] G. Corsini, P. Cipollini, M. Diani, R. Grasso, "Semi-Analytical algorithms for ocean colour analysis from SeaWiFS data in the Western Mediterranean Sea", *Proc. 3rd International Workshop on MOS-IRS and Ocean Colour*, Institute of Space Sensor Technology, Berlin, 1999.
- [3] G. Corsini, R. Grasso, J. Ruiz, P. Cipollini, "Retrieval of the concentration of optically active constituents in the Alboran Sea: a semi-analytic approach", *Proc. IEEE 1999 International Geoscience and Remote Sensing Symposium*, Amburgo, 28 Giugno-2 Luglio, 1999.
- [4] G. Corsini, M. Diani, R. Grasso, R. Rinaldi, T. Walzer, "MOS-B sensor calibration: a signal processing approach", *Proc. IEEE 1999 International Geoscience and Remote Sensing Symposium*, Amburgo, 28 Giugno-2 Luglio, 1999.
- [5] G. Corsini, M. Diani, R. Grasso, C. Santella, "The potential of SeaWiFS and MOS for the investigation of the bio-optical variability in mesoscale systems", *Proc. 4th International Congress on Energy, Environment and Technological Innovation*, Roma, Settembre 19-24, 1999.

#### **Lavori su rivista internazionale**

- [6] Baldacci, G. Corsini, R. Grasso, G. Manzella, J. T. Allen, P. Cipollini, T. H. Guymmer, H. M. Snaith, "A study of the Alboran sea mesoscale system by means of EOF (Empirical Orthogonal Function) decomposition of satellite data", *sottomesso al Journal of Marine System*, 1999

## Lavori in fase di stesura

- [7] G. Corsini, R. Grasso, J. Ruiz, P. Cipollini, "Estimation of the concentration of optically active constituents in the Alboran Sea.", *da sottomettere all'Int. Jour. Of Remote Sensing*.
- [8] A. Baldacci, G. Corsini, R. Grasso, F. Pardini, C. Santella, "EOF (Empirical Orthogonal Function) Analysis of Sequences of AVHRR and SeaWiFS Images to gain insight on the Circulation on the Alboran Sea", *Dipartimento di Ingegneria dell'Informazione-Università di Pisa, rapporto finale progetto OMEGA (MAS3-CT95-001)*.

Durante alcuni mesi precedenti l'inizio dell'attività di dottorato di ricerca e nella fase preliminare di questa, nell'ambito delle tematiche che riguardano il telerilevamento dell'ambiente marino, è stato affrontato lo studio dei metodi di stima dello spettro direzionale della superficie marina da dati SAR. Tale attività ha portato alla stesura dei seguenti lavori:

- [9] Giovanni Corsini, Raffaele Grasso, Giuliano Manara, Agostino Monorchio, "Methods for the estimation of the directional sea-spectrum from SAR image: an application to Northern Tyrrhenian Sea", ESA experiment ERS-1/2 n° A02.I111: Monitoring and detection of sea surface phenomena by SAR data analyses: application to a Tyrrhenian area, PI Prof. Giovanni Corsini, *Progress report n.1*, June 1996.
- [10] Giovanni Corsini, Raffaele Grasso, Giuliano Manara, Agostino Monorchio, "Estimation of directional sea spectra from ERS/SAR images of Mediterranean areas: a case study", *Proc. 3rd Symp. on Space at the service of our Environment*, Florence, Italy, 17-21 March 1997, (ESA SP-414, 3 Vols., May 1997), pp. 1423-1426.
- [11] G. Corsini, R. Grasso, G. Manara, A. Monorchio, "Estimation of directional sea wave spectra from radar images: a Mediterranean Sea case study", *sottomesso al Nuovo Cimento C*, 1999.

---

# Ringraziamenti

Desidero ringraziare il Dott. Bucci dell'Agenzia Regionale per la Protezione dell'Ambiente della Toscana (ARPAT) ed il Dott. Javier Ruiz dell'Università di Cadice per aver messo a disposizione la serie di misurazioni *in situ* utilizzata nel presente studio.

Raffaele Grasso, Dicembre 1999