

# INTERPOLAZIONE MEDIANTE KRIGING DI DATI LAGRANGIANI E CONFRONTO TRA DATI DEL MAR NERO E DEL MAR ADRIATICO

R. BARBANTI, L. URSELLA & P.-M. POULAIN

Approvato da: .....

Dr. Alessandro Crise

Direttore, Dipartimento di Oceanografia

**SOMMARIO:**

1. Introduzione.....Pagina 3

2. Metodi.....Pagina 4

3. Dati .....Pagina 8

4. Risultati.....Pagina 16

5. Conclusioni.....Pagina 25

6. Appendici.....Pagina 26

7. Referenze.....Pagina 34

## 1. Introduzione

Le tecniche di interpolazione possono essere di due tipi: deterministiche o stocastiche. Il *kriging* è di tipo stocastico, ed è un metodo nel quale le proprietà statistiche sono definite a priori da una funzione detta semivariogramma o “structure function”. La caratteristica interessante di questa tecnica sta nel poter disporre, per ogni valore ricostruito, di una stima dell’affidabilità di tale ricostruzione. Ciò consente per esempio, di definire per quest’ultima un intervallo di fiducia o individuare aree (nel caso di interpolazione spaziale) in cui è necessario aumentare o intensificare il campionamento. La semivarianza misura il grado di dipendenza tra i campioni, la sua ampiezza dipende dalla distanza tra i punti e il semivariogramma è il grafico della semivarianza in funzione della distanza. Nel caso della traiettoria di un galleggiante, detto anche drifter, il grafico sarà in funzione di un intervallo di tempo, detto *time lag* ( $t$ ). La grandezza da ricostruire è considerata come una variabile aleatoria di cui sono note stime costituite da un certo numero di valori osservati  $x(t)$ . Sulla base dei dati disponibili vengono poi stimati tutti i valori desiderati, mentre quelli noti non sono ricostruiti esattamente per l’introduzione nel modello di interpolazione della stima dell’errore strumentale.

Tra le tecniche deterministiche possiamo includere: l’interpolazione *lineare*, così detta perché equivale a collegare i punti noti con una funzione lineare, l’interpolazione *polinomiale*, in cui per ogni funzione esiste una classe di polinomi interpolanti che la approssima con precisione sempre maggiore all’aumentare del grado del polinomio, e l’interpolazione *spline*, una procedura alternativa, ampiamente adottata, che consiste nell’approssimare i punti noti utilizzando un polinomio di grado basso fra ogni coppia di punti adiacenti. Il nome di questa tecnica deriva da un particolare strumento (spline) che veniva utilizzato dai disegnatori per tracciare una curva quanto più regolare possibile fra una serie di punti. La procedura più comune utilizza i polinomi di terzo grado; per questo l’interpolazione è anche detta interpolazione spline cubica (Emery e Thomson, 2001).

In questo rapporto sono riportati i risultati riguardanti l’applicazione della tecnica del kriging per l’interpolazione delle traiettorie dei galleggianti e sono stati confrontati i risultati ottenuti per i dati riguardanti il Mar Nero con quelli del Mar Adriatico. Nel capitolo 2 si possono avere informazioni sulla tecnica del kriging e sulla procedura

adottata in pratica , sul calcolo della “structure function” e sull’errore di interpolazione. Nel capitolo 3 mostriamo i dati utilizzati nell’analisi ed alcune statistiche che descrivono i due insiemi di dati. Nel quarto capitolo sono presentati i risultati, tra cui: le “structure function” ed alcuni esempi di traiettorie e serie temporali di temperatura superficiale interpolate nei due mari. Infine nel 5° capitolo sono riportate le conclusioni.

## 2. Metodi

L’interpolazione dei dati dei drifter è stata eseguita utilizzando un programma Matlab che esegue l’interpolazione ogni 2 ore per il Mar Nero e ogni 1/2 ora per il Mar Adriatico delle serie temporali editate di latitudine, di longitudine e di temperatura superficiale (SST). Lo stesso programma esegue un filtraggio con filtro passa basso Hamming a periodo di taglio di 36 ore, in modo da eliminare le componenti della corrente ad alta frequenza, ed infine sotto-campiona le serie ogni 6 ore e calcola la velocità zonale e meridionale del drifter.

Il codice è stato sviluppato tenendo conto delle seguenti assunzioni e vincoli (Hansen and Poulain, 1996):

i valori osservati  $x_i$  (latitudine, longitudine o temperatura, con  $i=1,2,\dots,n$  osservazioni) sono somma di un valore atteso,  $\hat{x}_i$  più un errore,  $e_i$ :

$$x_i = \hat{x}_i + e_i . \quad (1)$$

Una stima della variabile da interpolare al tempo  $t_0$  si ottiene mediante una combinazione lineare di valori noti  $x_i$  in cui  $w_i$  indicano i pesi:

$$\hat{x}_0^* = \sum_{i=1}^n w_i \cdot x_i . \quad (2)$$

I pesi vengono determinati in modo che l’errore, definito come differenza tra la stima e il valore vero non sia sistematico, vale a dire che gli errori di previsione oscillino attorno allo 0:

$$\langle e_0 \rangle = \langle \hat{x}_0 - \hat{x}_0^* \rangle = 0, \quad (3)$$

e in modo tale da ottimizzare la stima  $\hat{x}_0^*$  minimizzando lo scarto quadratico (mediante l'utilizzo dei moltiplicatori di Lagrange):

$$S_k^2 = \langle (\hat{x}_0 - \hat{x}_0^*)^2 \rangle = \min. \quad (4)$$

I pesi  $w_i$  sono calcolati dal semivariogramma empirico (“structure function”) il quale fornisce informazioni molto simili a quelle che può dare la funzione di covarianza, ma rispetto ad essa è indipendente dalle stime del valor medio (Hansen and Poulain 1996), precisamente :

$$S_{ij} = \frac{1}{2} \langle (x_i - x_j)^2 \rangle = \frac{1}{2} \langle [x(t_i) - x(t_j)]^2 \rangle. \quad (5)$$

Questa osservazione è legata al differente modo con cui la tecnica del kriging tiene conto degli errori, rispetto all’analisi oggettiva (altro metodo statistico di interpolazione) che utilizza invece la funzione di covarianza come descrittore del campo statistico. Infatti mentre nell’analisi oggettiva l’errore è presente sia nei dati di partenza che nel covariogramma, il kriging è legato soltanto ai dati (Bergamasco et al., 1993).

La “structure function” empirica è data da:

$$S = S(0) + \hat{S}, \quad (6)$$

dove  $\hat{S}(0)$  indica l’errore di misura a zero time-lag ed  $\hat{S}$  è la structure function teorica priva di errore. Per il kriging dei dati dei drifter con sistema ARGOS,  $S(0)$  è l’errore strumentale o “mean square error” (mse) fornito da tale metodo ed è stimato essere circa 172m per la latitudine e 277m per la longitudine, mentre per l’interpolazione dei dati di drifter equipaggiati di sistema GPS l’errore strumentale della posizione è di 50m.

Per poter essere una funzione continua  $\forall t$ , la structure function sperimentale viene parametrizzata attraverso il calcolo di una “structure function” priva di errore ( $\hat{S}$ ), mediante “best-fitting”:

$$\hat{S} = \mathbf{a} \cdot t^b. \quad (7)$$

I coefficienti  $\mathbf{a}$  e  $\mathbf{b}$  sono determinati dai dati empirici per mezzo di un “best fit”. Per misure lagrangiane il range è  $1 < \mathbf{b} < 2$ , per  $\mathbf{b} = 1$  il moto è completamente caotico mentre se  $\mathbf{b} = 2$  il moto è rettilineo (Hansen and Poulain, 1996). Anche per la SST può venir utilizzata l’equazione (7) con un  $\hat{S}(0)$  pari a 0.01 °C, ma si può anche ricorrere ad una funzione differente dalla precedente quale (Hansen and Poulain 1996):

$$\hat{S} = \mathbf{a} \cdot t^b + \mathbf{g} \cdot (1 - \cos 2pt). \quad (8)$$

L’espressione dello scarto quadratico (Hansen and Poulain 1996) è ottenuta sostituendo l’equazione 2) nella equazione 4) e tenendo conto che la somma dei pesi è uguale ad uno (vedere dettagli in Appendice A),

$$\sum_{i=1}^n w_i = 1, \quad (9)$$

si ha:

$$\mathbf{s}_k^2 = 2 \sum_{i=1}^n w_i \hat{S}_{0i} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij} + \sum_{i=1}^n w_i^2 \langle e^2 \rangle. \quad (10)$$

Dopo aver introdotto nella (10) il moltiplicatore di Lagrange ( $\mathbf{I}$ ) e tenendo conto anche della condizione di normalità (9), l’equazione dello scarto quadratico viene minimizzata

ponendo uguale a zero la sua derivata parziale rispetto a  $w_i$  ( $\frac{\partial \mathbf{s}_k^2}{\partial w_i} = 0$ ). Si ottiene

pertanto:

$$\hat{S}_{0j} = \sum_{i=1}^n w_i \cdot \hat{S}_{ij} - w_i \cdot \langle e^2 \rangle + \mathbf{I}, \quad (11)$$

cioè un sistema lineare di  $n+1$  equazioni per  $n+1$  incognite  $(w_i, \mathbf{I})$ , che di seguito rappresentiamo in forma matriciale (vedere dettagli in Appendice B):

$$\begin{pmatrix} \hat{S}_{11} - \langle e^2 \rangle & \hat{S}_{12} - \langle e^2 \rangle & \dots & \hat{S}_{1n} - \langle e^2 \rangle & 1 \\ \hat{S}_{11} - \langle e^2 \rangle & \hat{S}_{i_j} - \langle e^2 \rangle & & & 1 \\ \hat{S}_{n1} - \langle e^2 \rangle & \hat{S}_{11} - \langle e^2 \rangle & \dots & \hat{S}_{nn} - \langle e^2 \rangle & 1 \\ 1 & 1 & & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_n \\ \mathbf{I} \end{pmatrix} = \begin{pmatrix} \hat{S}_{01} \\ \hat{S}_{02} \\ \dots \\ \hat{S}_{0n} \\ 1 \end{pmatrix}, \quad (12)$$

e moltiplicando ambo i membri della (11) per  $\sum_{i=1}^n w_i$  e aggiungendo ad entrambi i lati dell'equazione  $\sum_{i=1}^n w_i \hat{S}_{0i}$ , si ricava il valore minimo dello scarto quadratico (vedere dettagli in Appendice B):

$$\mathbf{s}_k^2 = \sum_{i=1}^n w_i \hat{S}_{0i} + \mathbf{I}, \quad (13)$$

ed in forma matriciale:

$$\begin{pmatrix} \mathbf{s}_1^2 \\ \mathbf{s}_2^2 \\ \dots \\ \mathbf{s}_n^2 \\ 1 \end{pmatrix} = \begin{pmatrix} \hat{S}_{01} \\ \hat{S}_{02} \\ \dots \\ \hat{S}_{0n} \\ 1 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_n \\ \mathbf{I} \end{pmatrix}. \quad (14)$$

A questo punto poniamo l'attenzione sulla relazione tra la "structure function" e il covariogramma che è (vedere dettagli in Appendice C):

$$S_{ij} = \langle x^{i2} \rangle - cov(x_i, x_j) = \mathbf{s}^2 - cov(x_i, x_j). \quad (15)$$

Il metodo di interpolazione kriging è quindi di poco differente all'analisi oggettiva che come detto in precedenza utilizza invece il covariogramma come descrittore del campo statistico.

### 3. Dati

In questa analisi sono state utilizzate le osservazioni editate, provenienti dai 54 drifter del programma internazionale drifter in Mar Nero dal 1999 al 2003 e quelle fornite da 189 drifter rilasciati in Mar Adriatico fra settembre 2002 e novembre 2003 e facenti parte del progetto DOLCEVITA.

Tra i drifter utilizzati in Mar Nero vi sono: 6 XAN-3 con catena di 10 termistori fino a 50 m della MetOcean (Canada); 38 SVP-B con ancora di tipo "holey-sock" centrata ad una profondità nominale di 15 m, e dotati di sensori per la misurazione della pressione atmosferica e della temperatura dell'aria, di cui 4 americani della Naval Oceanographic Office (NAVO); 10 SVP con una ancora di tipo "holey-sock" centrata ad una profondità nominale di 15 m, di cui 3 ucraini costruiti dalla Marlin (Ucraina) e 7 americani costruiti dalla MetOcean. Questi drifter hanno fornito 55 file per un totale di 36306 osservazioni.

In Mar Adriatico sono stati impiegati principalmente drifter del tipo CODE (Davis, 1985; Poulain, 2001). Tutti erano dotati del sistema di posizionamento ARGOS, alcuni anche equipaggiati di sistema GPS (Global Positioning System) che fornisce una più accurata risoluzione sia temporale che spaziale. I drifter sono stati programmati per campionare posizione e temperatura superficiale a intervalli di 1 ora. Sono stati utilizzati 65 drifter del tipo CODE e 41 CODE dotati anche di GPS. Oltre a questi sono state utilizzati 5 SVP, 2 MINIMET, 2 CODE/Tz con catena di termistori e 2 SVP/OCM con sensori per la misura di radianza "upwelling" e di irradianza "downwelling". Le boe del tipo SVP, SVP/OCM e CODE/Tz avevano sistema GPS e sono stati programmati per campionare ogni mezz'ora.

Tutti i galleggianti adottano il sistema satellitare ARGOS per la telemetria dei dati e la loro localizzazione, e mediante il quale vengono trasmessi i dati relativi a ciascuno di essi. Questo sistema calcola la posizione attraverso la misura dello spostamento Doppler del segnale a frequenza fissa inviato dal drifter.



Le distribuzioni dei due database in esame, costituiti dai soli dati di posizione acquisiti con sistema satellitare ARGOS, sono raffigurati in figura 1. Per il Mar Nero abbiamo avuto più frequentemente 7 osservazioni al giorno, mentre 8 per il Mar Adriatico costituenti rispettivamente il 14% e il 9% del totale, con medie di circa 6 e 7 osservazioni e corrispondenti deviazioni standard pari a 3 e 4 osservazioni al giorno.

In figura 2 rappresentiamo istogrammi che mostrano la frequenza di intervalli temporali tra osservazioni successive nei due database. Per il Mare Adriatico si ha un intervallo massimo di 10 giorni, mentre i dati del Mar Nero mostrano anche intervalli di circa 12 e 13 giorni. Le “timeseries” con intervalli maggiori di 13 giorni sono stati divisi e considerati come drifter differenti.

La figura 3 rappresenta i dati di posizione GPS del Mar Adriatico in cui: in a) sono evidenti distribuzioni dei dati dei CODE/GPS (24 osservazioni al giorno) e degli SVP, SVP-OCM e CODE/Tz (48 osservazioni al giorno), in b) mostriamo gli intervalli temporali tra posizione successive. Il drifter b04020 mostra un buco di 13 giorni ed il b12589 di 19 giorni.

La figura 4 raffigura invece i dati argos di temperatura superficiale. Si nota che sia per il Mar Nero, che per il Mar Adriatico, abbiamo avuto più frequentemente 7 osservazioni al giorno rispettivamente, 15% e 11% del totale, con medie di 6.4 e 7 osservazioni e corrispondenti deviazioni standard pari a 3 e 3.6 osservazioni al giorno.

In figura 5 rappresentiamo gli istogrammi della frequenza di intervalli temporali tra le osservazioni successive di temperatura nei due database. Per il Mare Adriatico si ha un intervallo massimo di 8.5 giorni relativo al drifter 29361, mentre i dati del Mar Nero mostrano un intervallo massimo di 4 giorni relativo al drifter 17484. I dati di temperatura superficiale dei drifter con GPS sono rappresentati in figura 6 Alcune di queste statistiche sono riassunte nella tabella seguente:

		BLACK SEA				ADRIATIC SEA			
		<i>Mode (obs/day)</i>	<i>Mean (obs/day)</i>	<i>Std (obs/day)</i>	<i>#Obs (obs/day)</i>	<i>Mode (obs/day)</i>	<i>Mean (obs/day)</i>	<i>Std (obs/day)</i>	<i>#Obs (obs/day)</i>
<b>Positions</b>	<b>Argos</b>	7	6	3	36306	8	7	4	67145
	<b>GPS</b>	/	/	/	/	24-48	20	13	69863
<b>Temperature</b>	<b>Argos</b>	7	6.5	3	32394	7	7	3.6	34282
	<b>GPS</b>	/	/	/	/	24	14.6	7.7	40751

Tabella 1. Statistiche dei dati di posizione e temperatura calcolata per il Mar Nero e per il Mar Adriatico .

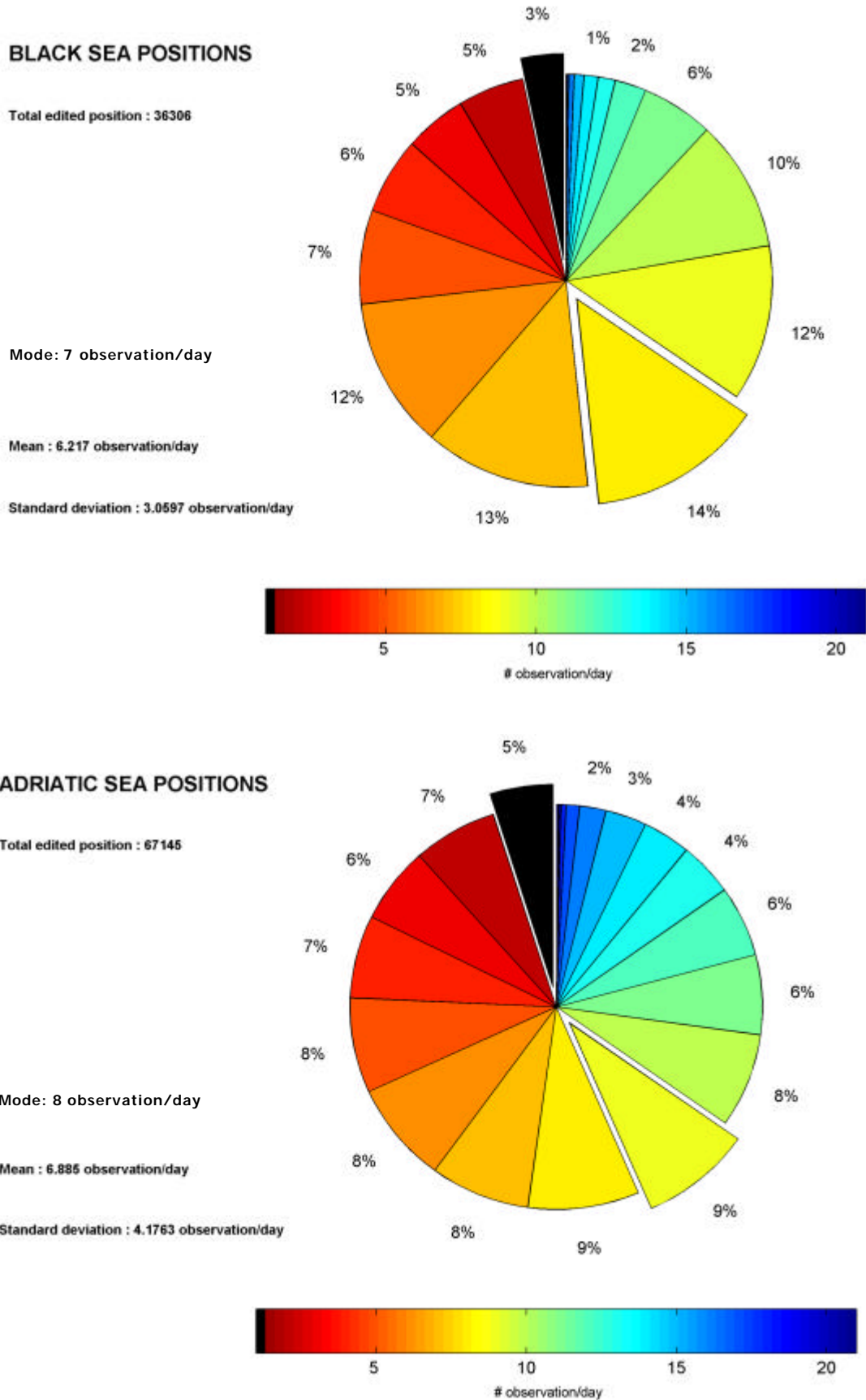


Figura 1. Quantità di osservazioni per giorno nei dati di posizione Argos editati utilizzati nell'interpolazione per ognuno dei due database in esame; in colore nero sono rappresentate 0 osservazioni per giorno corrispondenti a buchi per mancanza di posizioni nei dati.

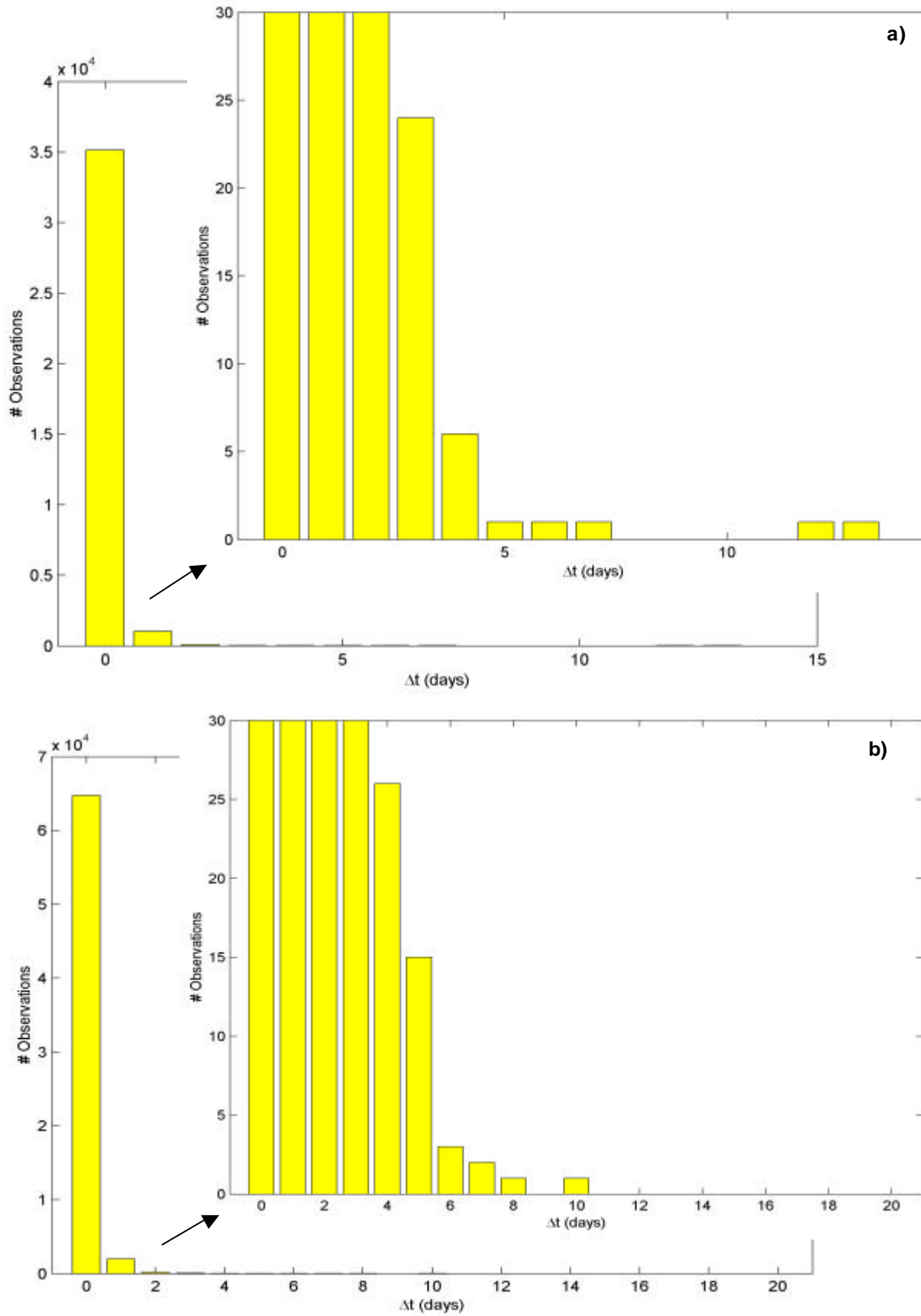


Figura 2. Istogrammi degli intervalli in giorni tra posizioni successive, a) Mar Nero b) Mar Adriatico.

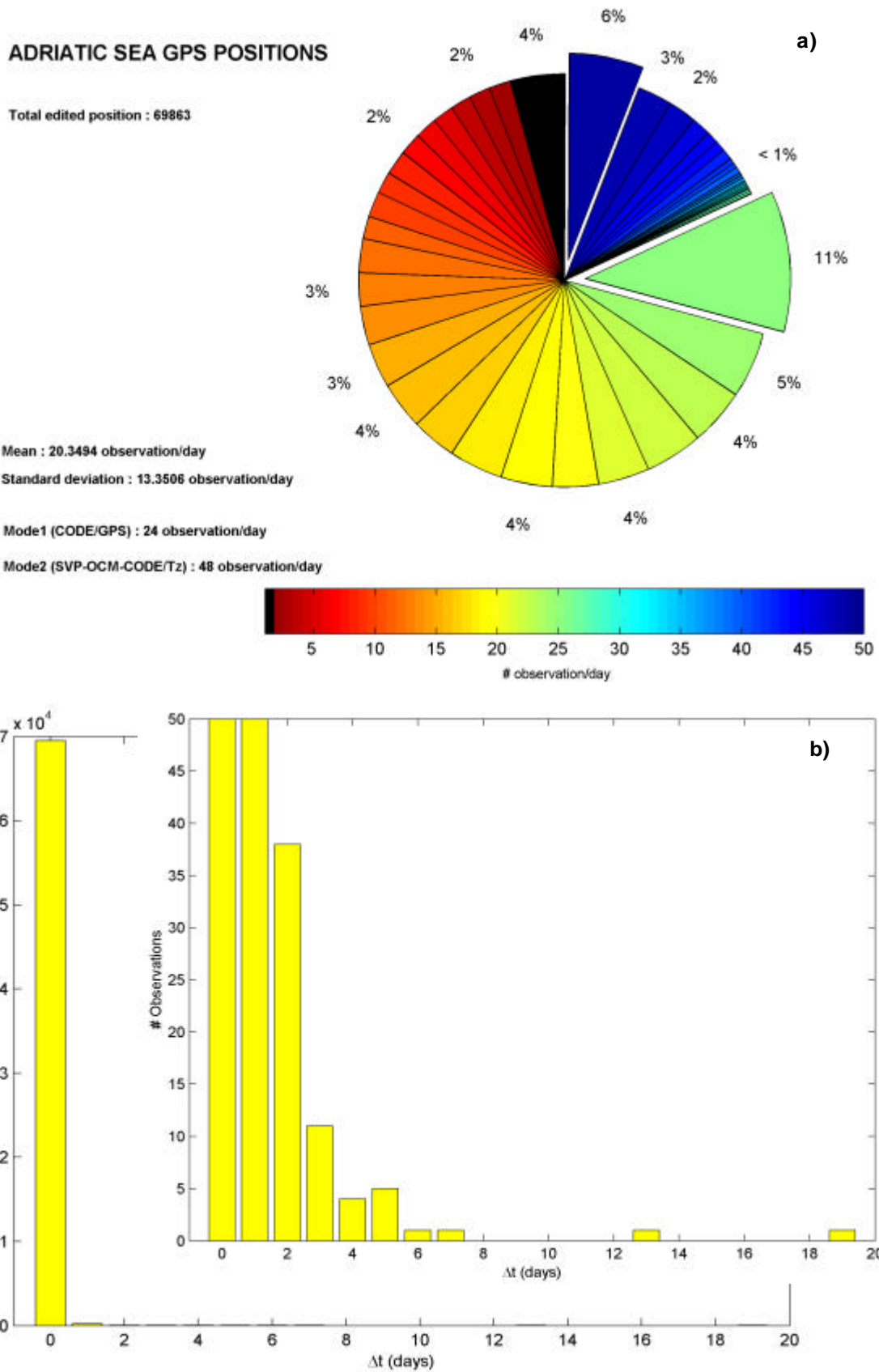


Figura 3. Come figura 1e figura 2, ma utilizzando i dati di posizione GPS del Mar Adriatico.

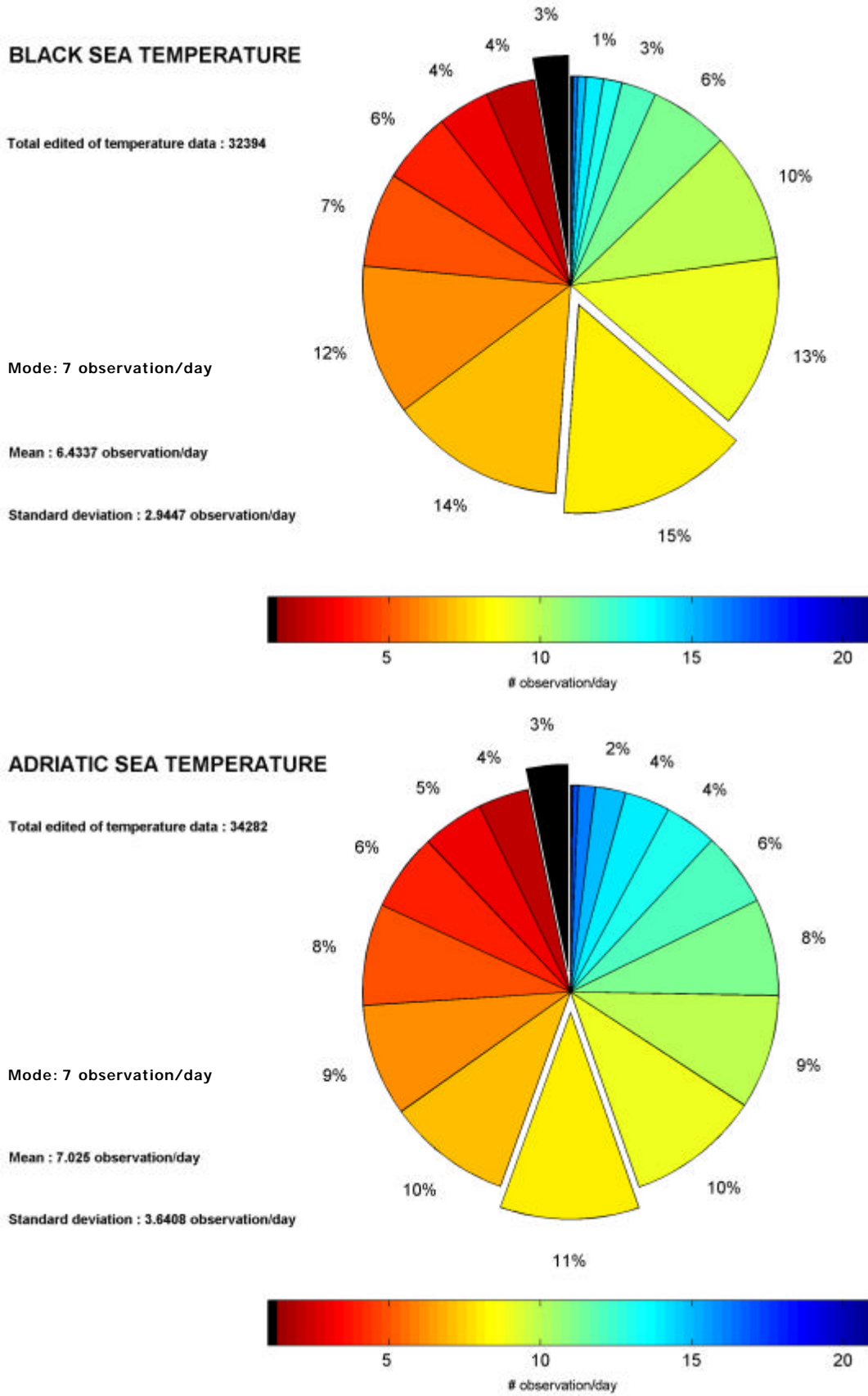


Figura 4. Come figura 1 ma calcolati utilizzando i dati di temperatura (“Argos”) superficiale.

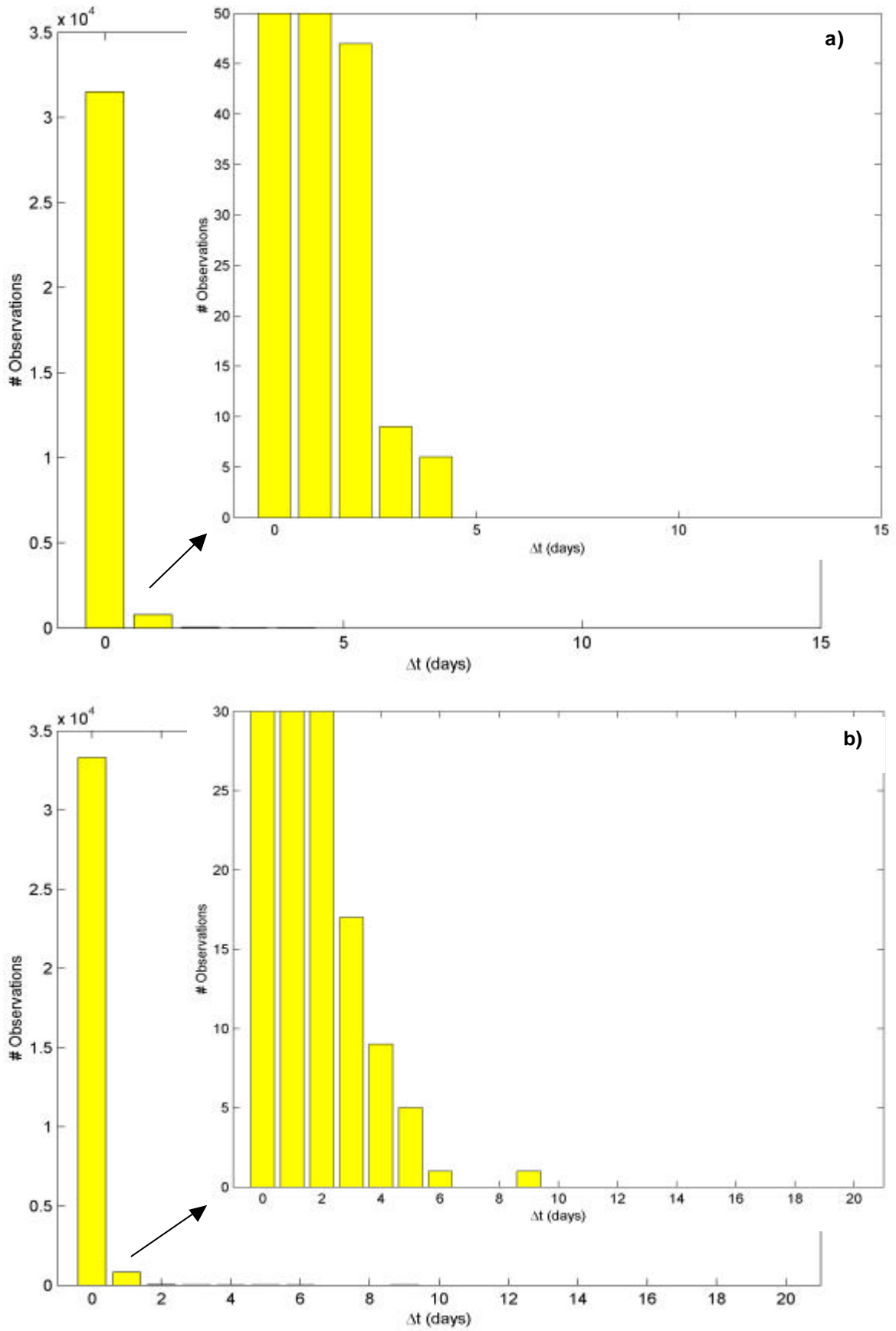


Figura 5. Come figura 2 ma calcolati utilizzando i dati di temperatura (“Argos”) superficiale.

ADRIATIC SEA GPS TEMPERATURE

Total edited of temperature data : 40751

Mode : 24 observation/day

Mean : 14.6482 observation/day

Standard deviation : 7.7177 observation/day

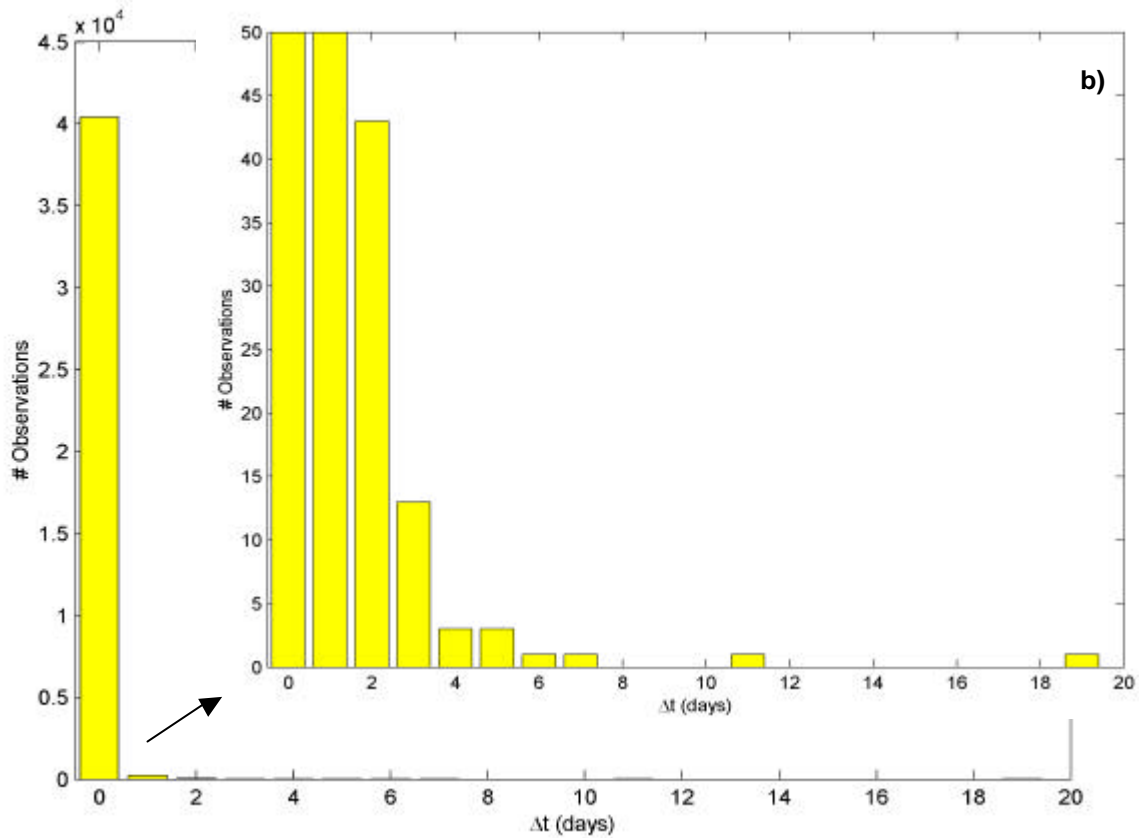
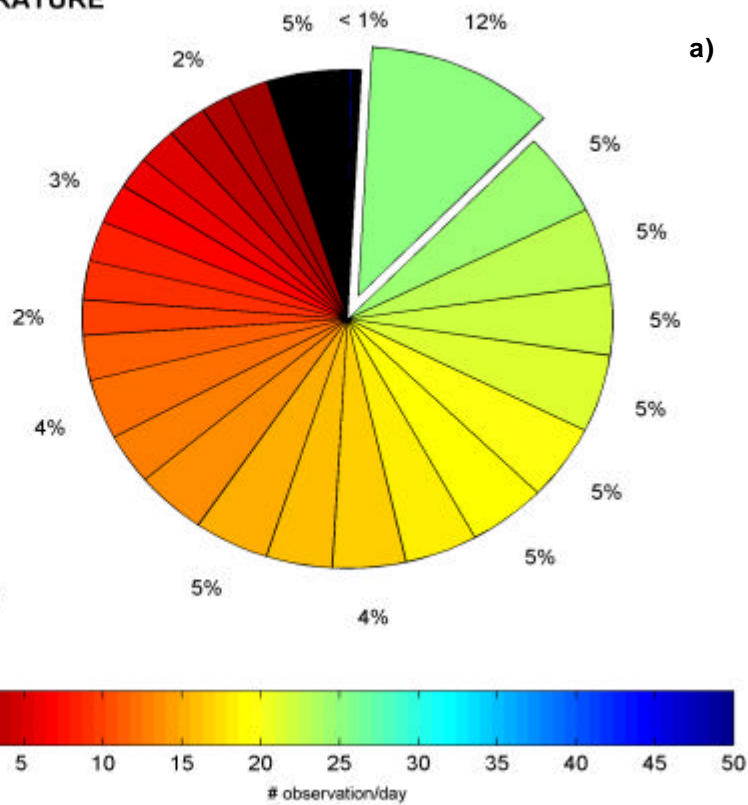


Figura 6. Come figura 4 e figura 5 ma calcolati utilizzando i dati di temperatura (“GPS”) superficiale del Mar Adriatico.

## 4. Risultati

Sono state calcolate le “structure function” sperimentali e, mediante “fitting” dei dati, abbiamo ricavato i parametri del modello descritto nel capitolo 2 (equazione 7). Essi saranno utili, nel processo di interpolazione, per il calcolo della “structure function” teorica. I valori dei parametri **a** e **b** per la longitudine, la latitudine e la temperatura superficiale sono mostrati in tabella 2. Le structure function sono state calcolate con i dati ARGOS e GPS. Si nota che per il Mar Nero si ha maggior differenza tra latitudine e longitudine per quanto riguarda il valore del coefficiente **b** :

		<i>Black Sea</i>			<i>Adriatic Sea</i>		
		<b>a</b>	<b>b</b>	$S(0)$	<b>a</b>	<b>b</b>	$S(0)$
<i>Longitudine</i>	<i>Argos</i>	0.024	1.72	2.40E-06	0.006	1.48	2.40E-06
	<i>GPS</i>	/	/	/	0.006	1.56	2.02E-07
<i>Latitudine</i>	<i>Argos</i>	0.007	1.47	6.20E-06	0.003	1.5	6.20E-06
	<i>GPS</i>	/	/	/	0.005	1.47	3.60E-07
<i>SST</i>	<i>Argos</i>	0.191	0.93	0.01	0.126	1.23	0.01
	<i>GPS</i>	/	/	/	0.118	1.08	

Tabella 2 Parametri ottenuti mediante fitting delle “structure function” sperimentali per due database e valori di  $S(0)$  per i sistemi ARGOS, GPS e per la temperatura.

In Figura 7,8 e 9 sono rappresentate, per i due mari, le “structure function” sperimentali ed il fitting , limitatamente ad un intervallo di 10 giorni. Si vede una maggiore differenza tra le “structure function” di longitudine e di latitudine in Mar Nero che per l’Adriatico.

Nel grafico della “structure function” di temperatura oltre al modello utilizzato (7), mostriamo anche il “fitting” effettuato mediante la (8) e vediamo come questo ultimo segue meglio l’oscillazione diurna nei dati.



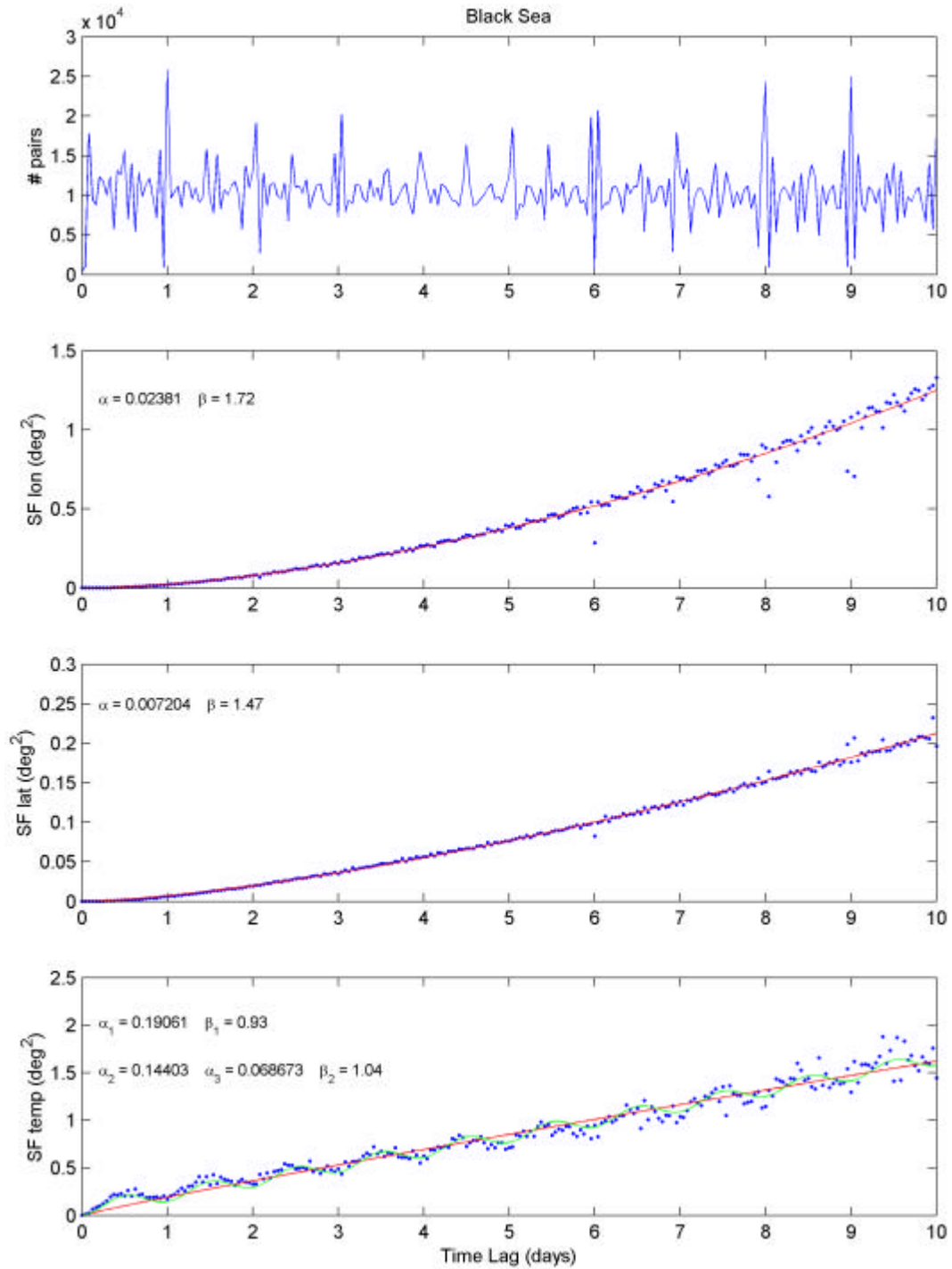


Figura 7. Quantità di paia, "structure function" e "fitting" dei dati di latitudine, di longitudine e di temperatura per il Mar Nero.

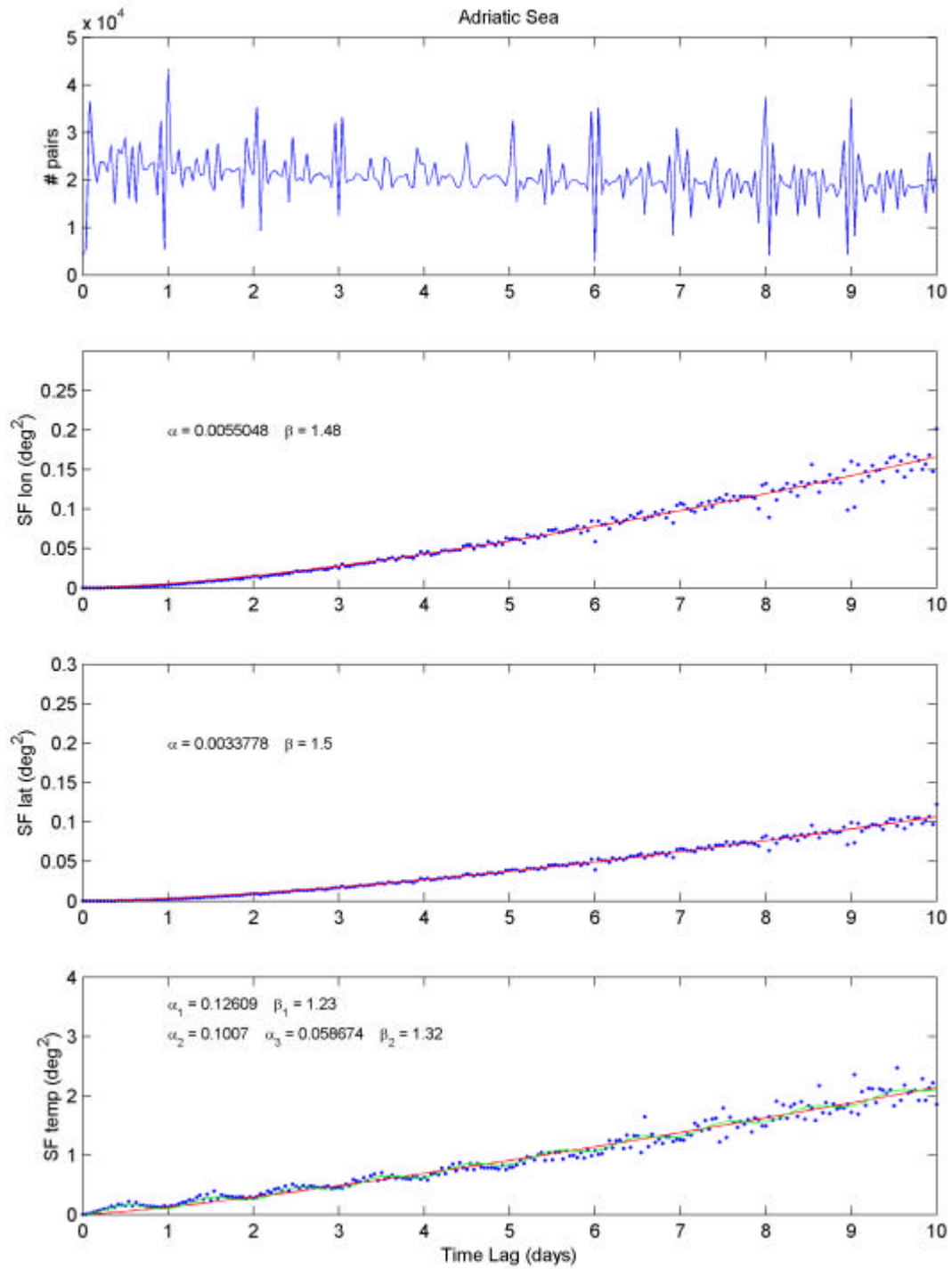


Figura 8. Come in figura 3 ma per il Mar Adriatico (dati Argos).

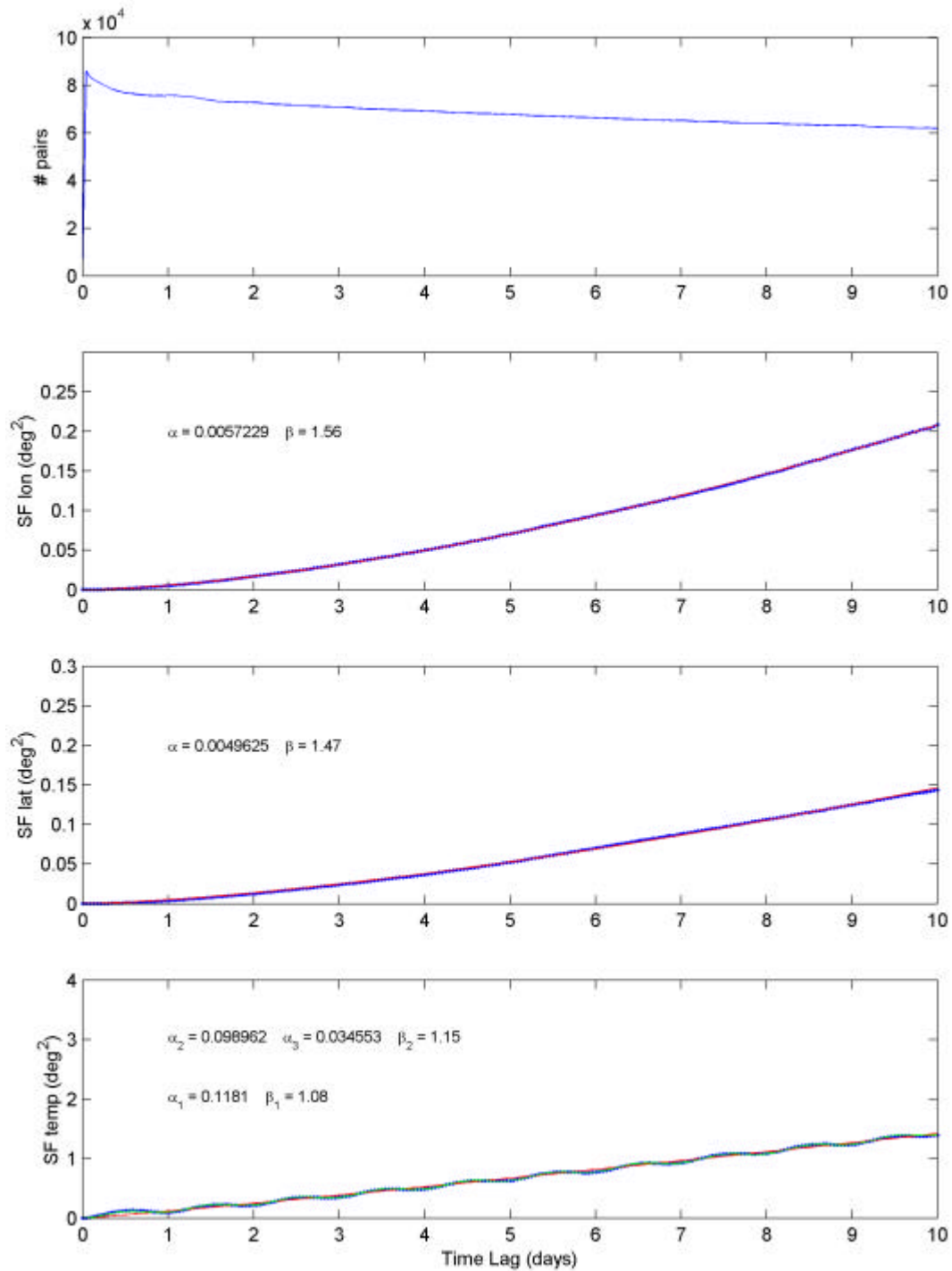


Figura 9. Come in figura 8 ma utilizzando i soli dati GPS.

Di seguito è presentato un esempio di pessime serie temporali di dati di latitudine e temperatura interpolata per il Mar Nero, con le relative barre di errore (figure 9 – 10). Il programma di interpolazione considera le 5 osservazioni precedenti e le 5 consecutive rispetto ai tempi nei quali si interpola.

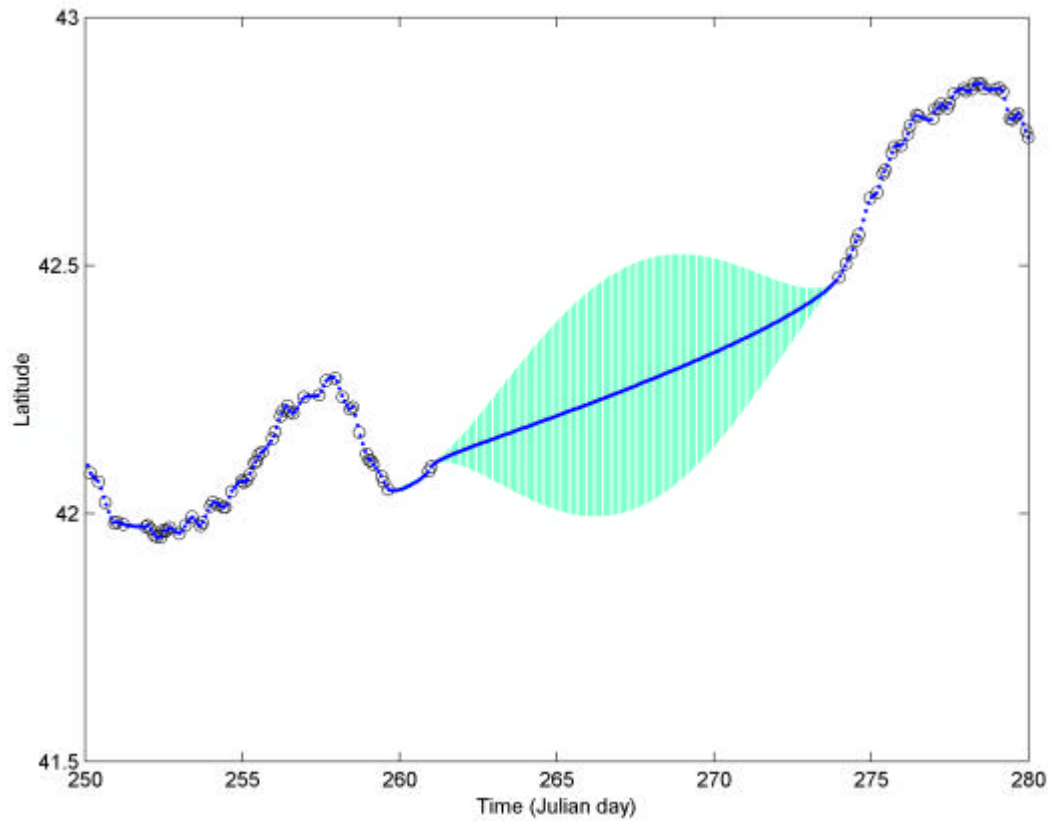


Figura 9. Esempio di interpolazione della latitudine del drifter 16332. Osservazioni (circoli), valori interpolati filtrati (punti) e barre di errore in verde.

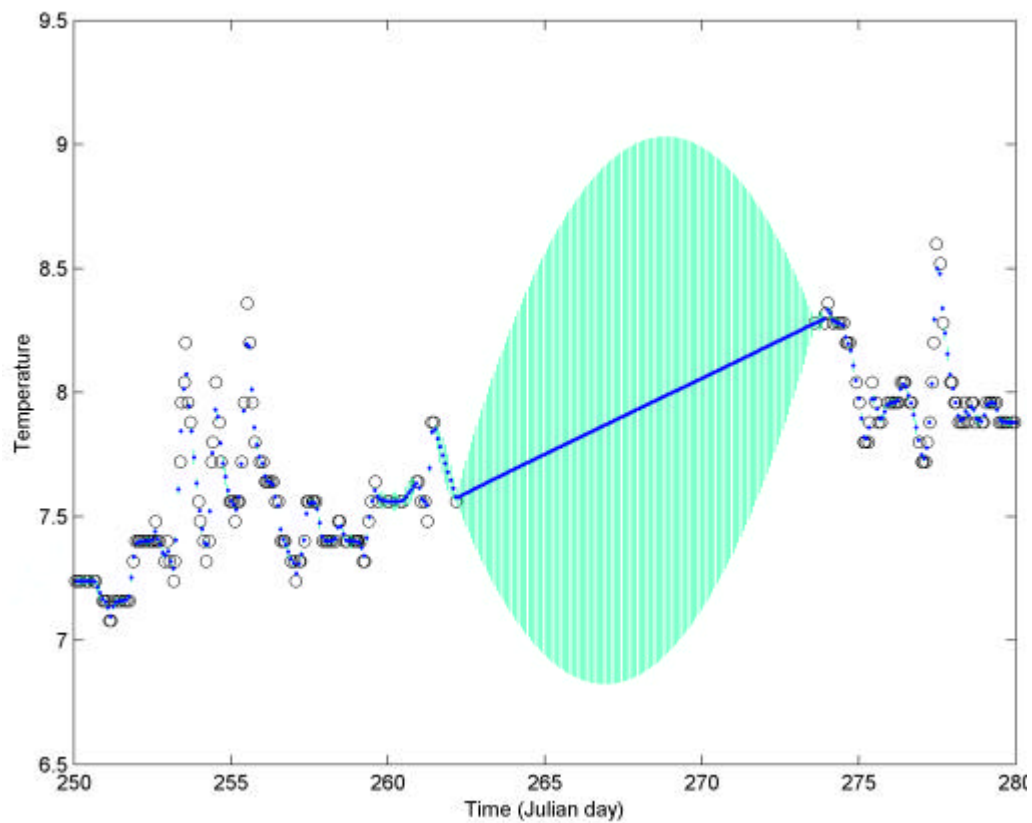


Figura 10. Esempio di interpolazione di una serie temporale di temperatura superficiale (SST) del drifter 16332. Osservazioni (circoli), valori interpolati (punti) e barre di errore in verde

Occorre particolare attenzione qualora il kriging calcoli valori in presenza di grossi buchi nei dati. Infatti, l'errore aumenta con l'aumentare dell'intervallo temporale tra osservazioni successive, inoltre il kriging estrapola (questo non dovrebbe accadere) anche in mancanza di osservazioni all'inizio o alla fine di una serie temporale, attribuendo un valore costante pari al dato iniziale o finale alla serie temporale (figura 11). Quindi in alcuni casi si sono rese necessarie separazioni in più "file" di serie temporali che presentavano buchi (>13 gg.) dovuti ad esempio ad una non buona trasmissione. Nel caso di estrapolazione si è proceduto ad un editing manuale dei dati di temperatura estrapolati non corretti.

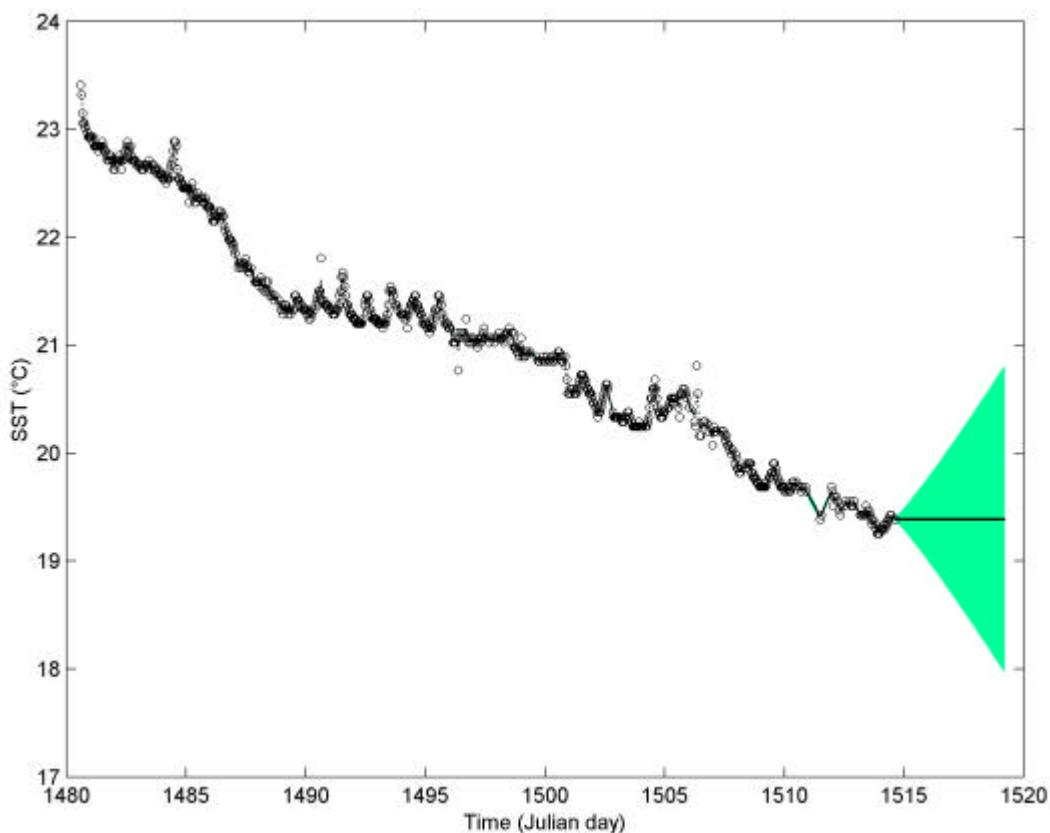


Figura 11. Esempio di interpolazione di una serie temporale in cui nella parte finale non sono presenti osservazioni di temperatura del drifter b37678.

Per alcuni drifter aventi il sistema GPS si è riscontrato, per brevi periodi, il blocco dell'acquisizione della posizione, e conseguentemente il medesimo dato per tutto l'intervallo temporale; questo porta ad una erronea ricostruzione della traiettoria durante l'interpolazione come è evidente in un esempio in figura 12. Per ovviare a tale inconveniente, in fase precedente all'interpolazione, viene effettuata una automatica eliminazione delle posizioni GPS ripetute.

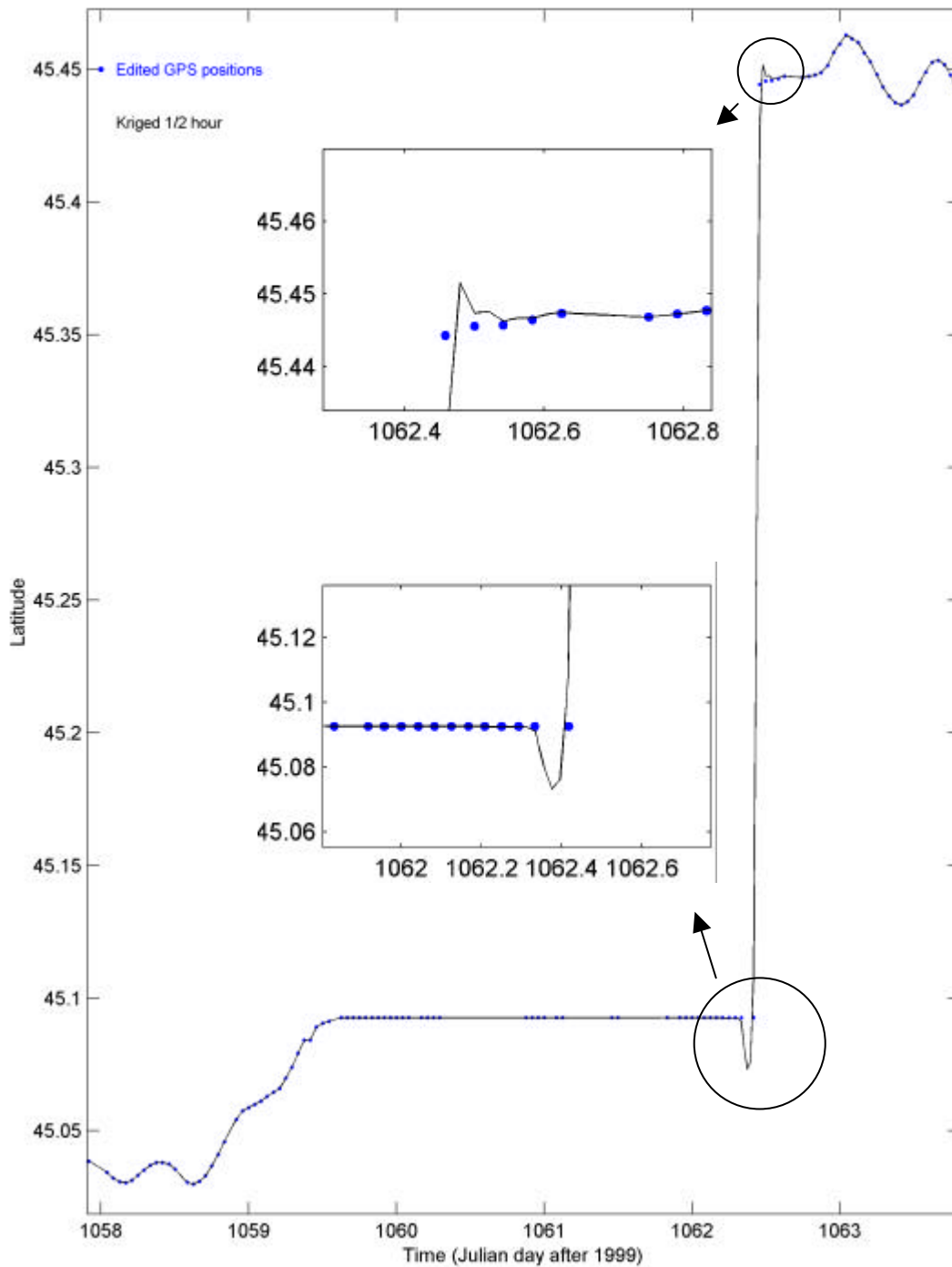


Figura 12. Esempio di interpolazione di una serie temporale di latitudine (linea continua) e osservazioni (punti) in cui si verifica il “blocco” del GPS del drifter b06952.

Nelle figure 13 e 14 vengono mostrati due esempi riguardanti il Mar Nero e il Mar Adriatico nelle quali vengono graficate le traiettorie delle boe, editate, interpolate, filtrate (filtro passa basso con taglio a 36 ore) e sotto-campionate ogni 6 ore.

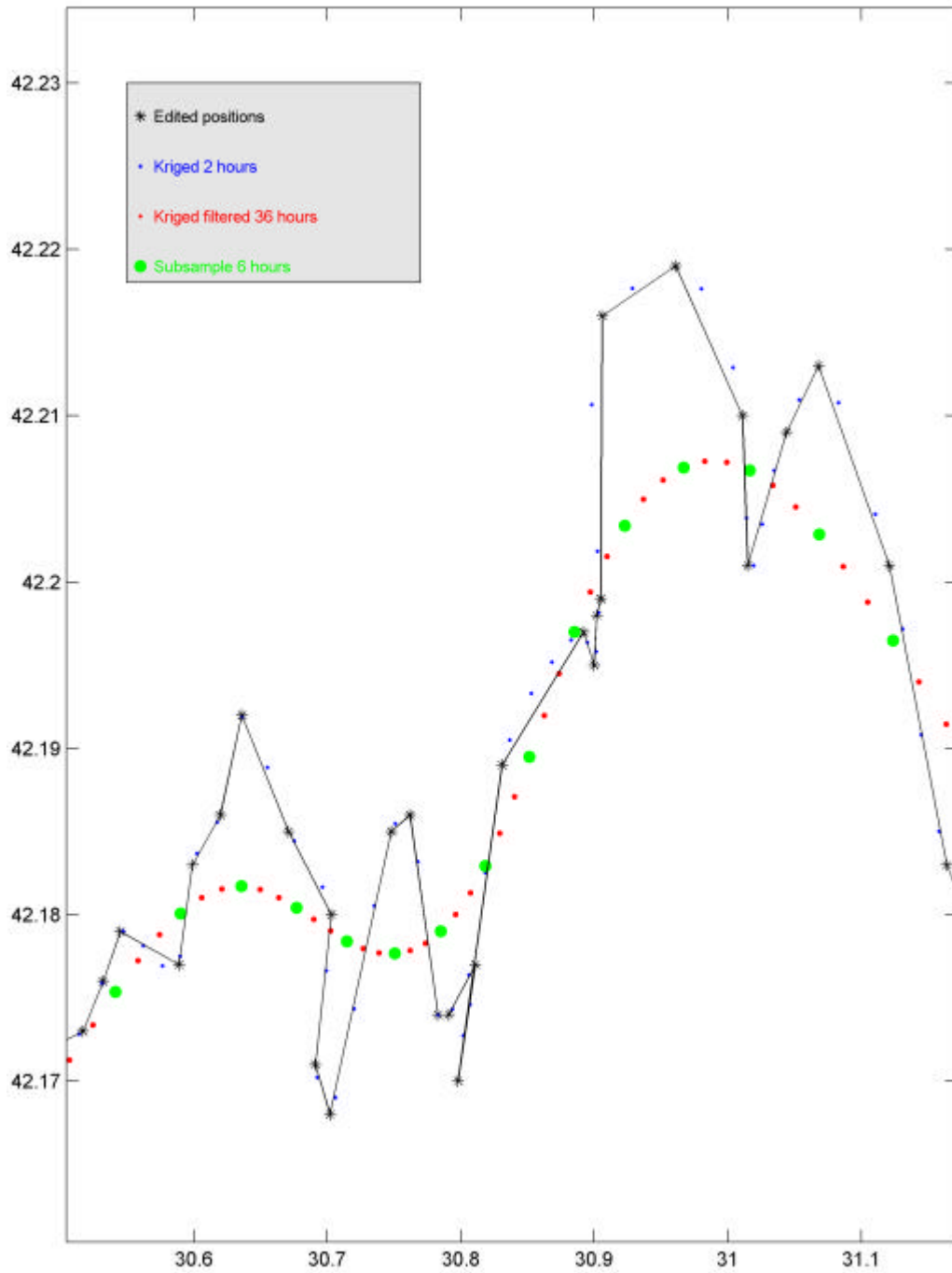


Figura 13. Esempio di interpolazione di una timeseries del drifter b16330 del Mar Nero in cui sono rappresentati: 1) nero - valori editati, 2) blu - interpolati ogni 2 ore, 3) rosso - filtraggio a 36 ore, 4) verde - sottocampionamento dei valori filtrati ogni 6 ore.

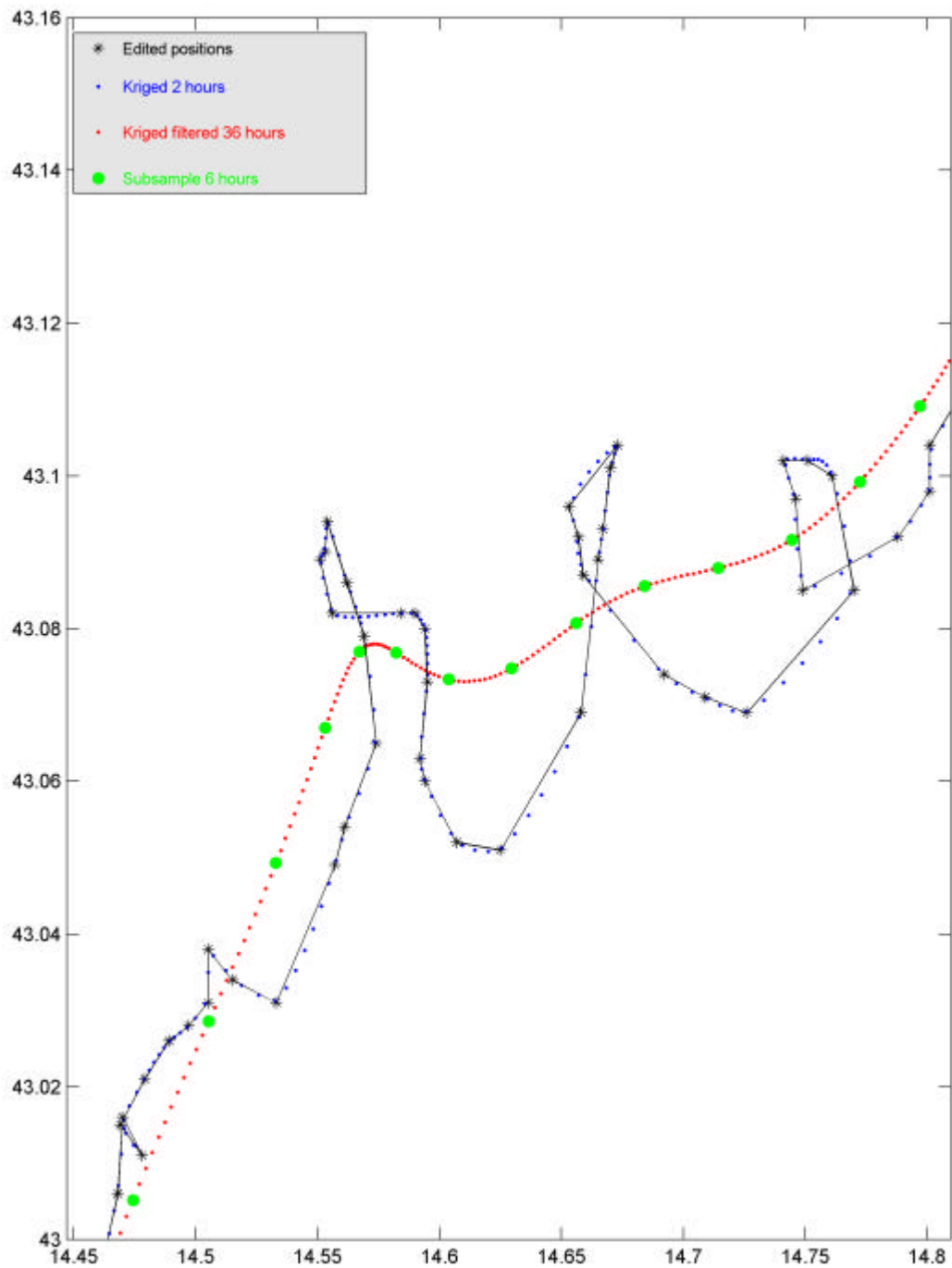


Figura 14. Esempio di interpolazione di una timeseries del drifter b37701 del Mar Adriatico in cui sono rappresentati: 1) nero - valori editati, 2) blu - interpolati ogni  $\frac{1}{2}$  ora, 3) rosso - filtraggio a 36 ore, 5) verde - sottocampionamento dei valori filtrati ogni 6 ore.



## 5. Conclusioni

Utilizzando dei dati di galleggianti appartenenti ai database del Mar Nero e del Mar Adriatico abbiamo studiato la tecnica d'interpolazione delle traiettorie dei drifter e delle serie temporali di temperatura superficiale attraverso il metodo di analisi oggettiva chiamato kriging. Abbiamo inoltre calcolato la “structure function” dai dati, e mediato come essa viene parametrizzata e come viene calcolato l'errore di interpolazione.

Abbiamo mostrato assieme ad alcune statistiche che qualificano i due insiemi di dati utilizzati per questo studio, anche esempi di “structure function” di latitudine, longitudine e temperatura ed esempi di dati interpolati e filtrati per il Mar Nero e per il Mar Adriatico.

## 6. Appendici

### *Appendice A: dimostrazione del calcolo dell'errore quadratico medio di interpolazione.*

La varianza è definita come:

$$\mathbf{s}_k^2 = \left\langle (\hat{x}_0 - \hat{x}_0^*)^2 \right\rangle. \quad (\text{A1})$$

Dobbiamo dimostrare che la varianza può essere scritta nel seguente modo:

$$\mathbf{s}_k^2 = 2 \sum_{i=1}^n w_i \hat{S}_{0i} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij} + \sum_{i=1}^n w_i^2 \langle e^2 \rangle, \quad (\text{A2})$$

che è equivalente alla (A1). Sviluppando la (A1) otteniamo:

$$\mathbf{s}_k^2 = \left\langle (\hat{x}_0 - \hat{x}_0^*)^2 \right\rangle = \left\langle \hat{x}_0^2 + \hat{x}_0^{*2} - 2\hat{x}_0 \hat{x}_0^* \right\rangle = \left\langle \hat{x}_0^2 \right\rangle + \left\langle \hat{x}_0^{*2} \right\rangle - \left\langle 2\hat{x}_0 \hat{x}_0^* \right\rangle, \quad (\text{A3})$$

ma  $\hat{x}_0^* = \sum_{i=1}^n w_i \cdot x_i$  e sostituendo questa ultima nella (A1) si ha:

$$\begin{aligned} \mathbf{s}_k^2 &= \left\langle \hat{x}_0^2 \right\rangle + \left\langle \left( \sum_{i=1}^n w_i x_i \right)^2 \right\rangle - \left\langle 2\hat{x}_0 \sum_{i=1}^n w_i x_i \right\rangle = \\ &= \hat{x}_0^2 + \sum_{i=1}^n w_i^2 \langle x_i^2 \rangle + \sum_{i \neq j} w_i w_j \langle x_i \cdot x_j \rangle - 2\hat{x}_0 \sum_{i=1}^n w_i \langle x_i \rangle. \end{aligned} \quad (\text{A4})$$

Ora consideriamo la definizione della “structure function” e la sviluppiamo per  $S_{0i}$  e  $S_{ij}$ :

$$\hat{S}_{0i} = \frac{1}{2} \left\langle (\hat{x}_0 - \hat{x}_i)^2 \right\rangle = \frac{1}{2} \left\langle \hat{x}_0^2 - 2\hat{x}_0 \hat{x}_i + \hat{x}_i^2 \right\rangle = \frac{1}{2} \hat{x}_0^2 + \frac{1}{2} \langle \hat{x}_i^2 \rangle - \hat{x}_0 \langle \hat{x}_i \rangle \quad (\text{A5})$$

$$\hat{S}_{ij} = \frac{1}{2} \left\langle (\hat{x}_i - \hat{x}_j)^2 \right\rangle = \frac{1}{2} \left\langle \hat{x}_i^2 - 2\hat{x}_i \hat{x}_j + \hat{x}_j^2 \right\rangle = \frac{1}{2} \langle \hat{x}_i^2 \rangle + \frac{1}{2} \langle \hat{x}_j^2 \rangle - \langle \hat{x}_i \hat{x}_j \rangle. \quad (\text{A6})$$

Calcoliamo ora la quantità:

$$2 \sum_{i=1}^n w_i \hat{S}_{0i}, \quad (\text{A7})$$

utilizzando la (A5):

$$2 \sum_{i=1}^n w_i \hat{S}_{0i} = 2 \sum_{i=1}^n w_i \left( \frac{1}{2} \hat{x}_0^2 + \frac{1}{2} \langle \hat{x}_i^2 \rangle - \hat{x}_0 \langle \hat{x}_i \rangle \right) = \sum_{i=1}^n w_i \hat{x}_0^2 + \sum_{i=1}^n w_i \langle \hat{x}_i^2 \rangle - 2 \sum_{i=1}^n w_i \hat{x}_0 \langle \hat{x}_i \rangle, \quad (\text{A8})$$

ma  $\sum_{i=1}^n w_i = 1$  pertanto:

$$2 \sum_{i=1}^n w_i \hat{S}_{0i} = \hat{x}_0^2 + \sum_{i=1}^n w_i \langle \hat{x}_i^2 \rangle - 2 \sum_{i=1}^n w_i \hat{x}_0 \langle \hat{x}_i \rangle. \quad (\text{A9})$$

Sviluppiamo ora il termine  $-\sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij}$  e utilizzando la (A6) :

$$\begin{aligned} & - \sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij} = - \sum_{i=1}^n \sum_{j=1}^n w_i w_j \left( \frac{1}{2} \langle \hat{x}_i^2 \rangle + \frac{1}{2} \langle \hat{x}_j^2 \rangle - \langle \hat{x}_i \hat{x}_j \rangle \right) = \\ & - \sum_{i=1}^n \sum_{j=1}^n w_i w_j \left( \frac{1}{2} \langle \hat{x}_i^2 \rangle \right) - \sum_{i=1}^n \sum_{j=1}^n w_i w_j \left( \frac{1}{2} \langle \hat{x}_j^2 \rangle \right) + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle = \\ & = - \frac{1}{2} \sum_{i=1}^n w_i \langle \hat{x}_i^2 \rangle \sum_{j=1}^n w_j - \frac{1}{2} \sum_{j=1}^n w_j \langle \hat{x}_j^2 \rangle \sum_{i=1}^n w_i + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle, \end{aligned}$$

ma  $\sum_{i=1}^n w_i = 1$  e poiché,

$$- \frac{1}{2} \sum_{i=1}^n w_i \langle \hat{x}_i^2 \rangle = - \frac{1}{2} \sum_{j=1}^n w_j \langle \hat{x}_j^2 \rangle,$$

si ha :

$$- \sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij} = - \sum_{j=1}^n w_j \langle \hat{x}_j^2 \rangle + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle. \quad (\text{A10})$$

Sostituendo nella (A2) le espressioni (A9) e (A10) otteniamo:

$$\begin{aligned} \mathbf{s}_k^2 &= \hat{x}_0^2 + \sum_{i=1}^n w_i \langle \hat{x}_i^2 \rangle - 2 \sum_{i=1}^n w_i \hat{x}_0 \langle \hat{x}_i \rangle - \sum_{j=1}^n w_j \langle \hat{x}_j^2 \rangle + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle + \sum_{i=1}^n w_i^2 \langle e^2 \rangle, \\ \mathbf{s}_k^2 &= \hat{x}_0^2 - 2 \hat{x}_0 \sum_{i=1}^n w_i \langle \hat{x}_i \rangle + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle + \sum_{i=1}^n w_i^2 \langle e^2 \rangle. \end{aligned} \quad (\text{A11})$$

Sviluppiamo il seguente termine:

$$\sum_{i=1}^n w_i^2 \langle x_i^2 \rangle,$$

appartenente alla (A4) e tenendo conto che  $x_i = \hat{x}_i + e_i$  si ha:

$$\begin{aligned} \sum_{i=1}^n w_i^2 \langle x_i^2 \rangle &= \sum_{i=1}^n w_i^2 \langle (\hat{x}_i + e_i)^2 \rangle = \sum_{i=1}^n w_i^2 \langle \hat{x}_i^2 + e_i^2 + 2\hat{x}_i e_i \rangle = \\ &= \sum_{i=1}^n w_i^2 \langle \hat{x}_i^2 \rangle + \sum_{i=1}^n w_i^2 \langle e_i^2 \rangle + \sum_{i=1}^n w_i^2 \langle 2\hat{x}_i e_i \rangle, \end{aligned}$$

nella quale l'ultimo termine è uguale a zero poiché:

$$\langle e_i \rangle = \langle \hat{x}_i - \hat{x}_i^* \rangle = 0,$$

per cui possiamo scrivere:

$$\sum_{i=1}^n w_i^2 \langle x_i^2 \rangle = \sum_{i=1}^n w_i^2 \langle \hat{x}_i^2 \rangle + \sum_{i=1}^n w_i^2 \langle e_i^2 \rangle. \quad (\text{A12})$$

Infine sostituiamo nella (A4) l'espressione (A12):

$$\mathbf{s}_k^2 = \hat{x}_0^2 + \sum_{i=1}^n w_i^2 \langle \hat{x}_i^2 \rangle + \sum_{i=1}^n w_i^2 \langle e_i^2 \rangle + \sum_{i \neq j}^n w_i w_j \langle x_i \cdot x_j \rangle - 2 \hat{x}_0 \sum_{i=1}^n w_i \langle x_i \rangle, \quad (\text{A13})$$

ma,

$$\sum_{i=1}^n w_i^2 \langle \hat{x}_i^2 \rangle + \sum_{i \neq j} w_i w_j \langle x_i \cdot x_j \rangle = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle,$$

allora si ha che l'errore quadratico equivale a:

$$\mathbf{s}_k^2 = \hat{x}_0^2 + \sum_{i=1}^n w_i^2 \langle e_i^2 \rangle + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle \hat{x}_i \hat{x}_j \rangle - 2\hat{x}_0 \sum_{i=1}^n w_i \langle x_i \rangle, \quad (\text{A14})$$

che è l'equazione A2) per la varianza.

**Appendice B: minimizzazione dello scarto quadratico e sviluppo del sistema per il calcolo dei pesi e dell'errore di interpolazione .**

Partendo dalla espressione dello scarto quadratico:

$$\mathbf{s}_k^2 = 2 \sum_{i=1}^n w_i \hat{S}_{0i} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij} + \sum_{i=1}^n w_i^2 \langle e^2 \rangle, \quad (\text{B1})$$

e introducendo il moltiplicatore di lagrange  $\lambda$  si ha:

$$\mathbf{s}_k^2 = 2 \sum_{i=1}^n w_i \hat{S}_{0i} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \hat{S}_{ij} + \sum_{i=1}^n w_i^2 \langle e^2 \rangle + 2 \cdot \mathbf{I} \cdot \left( \sum_{i=1}^n w_i - 1 \right). \quad (\text{B2})$$

Derivando la (B2) rispetto a  $\lambda$  otteniamo:

$$\frac{\partial \mathbf{s}_k^2}{\partial \mathbf{I}} = 2 \cdot \left( \sum_{i=1}^n w_i - 1 \right), \quad (\text{B3})$$

derivando la (B2) rispetto a  $w_i$  otteniamo:

$$\frac{\partial \mathbf{s}_k^2}{\partial w_i} = 2 \cdot \hat{S}_{0i} - \sum_{j=1}^n w_j \hat{S}_{ij} - \sum_{i=1}^n w_i \hat{S}_{ij} + 2 \cdot w_i \langle e^2 \rangle = 2 \cdot \hat{S}_{0i} - 2 \cdot \sum_{i=1}^n w_i \hat{S}_{ij} + 2 \cdot w_i \langle e^2 \rangle. \quad (\text{B4})$$

Poniamo ora la derivata parziale (B4) uguale a zero e troviamo il punto di minimo:

$$2 \cdot \hat{S}_{0i} - 2 \cdot \sum_{i=1}^n w_i \hat{S}_{ij} + 2 \cdot w_i \langle e^2 \rangle - 2 \cdot \mathbf{I} = 0, \quad (\text{B5})$$

da cui:

$$\hat{S}_{0i} - \sum_{i=1}^n w_i \hat{S}_{ij} + w_i \langle e^2 \rangle - \mathbf{I} = 0. \quad (\text{B6})$$

Isoliamo il termine  $\hat{S}_{0i}$  e otteniamo la seguente espressione:

$$\hat{S}_{0i} = \sum_{i=1}^n w_i \hat{S}_{ij} - w_i \langle e^2 \rangle + \mathbf{I}, \quad (\text{B7})$$

che può essere scritta come:

$$\begin{pmatrix} \hat{S}_{11} - \langle e^2 \rangle & \hat{S}_{12} - \langle e^2 \rangle & \dots & \hat{S}_{1n} - \langle e^2 \rangle & 1 \\ \hat{S}_{11} - \langle e^2 \rangle & \hat{S}_{ij} - \langle e^2 \rangle & & & 1 \\ \hat{S}_{n1} - \langle e^2 \rangle & \hat{S}_{11} - \langle e^2 \rangle & \dots & \hat{S}_{nn} - \langle e^2 \rangle & 1 \\ 1 & 1 & & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_n \\ \mathbf{I} \end{pmatrix} = \begin{pmatrix} \hat{S}_{01} \\ \hat{S}_{02} \\ \dots \\ \hat{S}_{0n} \\ 1 \end{pmatrix}. \quad (\text{B8})$$

Quindi possiamo calcolare i pesi come segue:

$$\begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_n \\ \mathbf{I} \end{pmatrix} = \begin{pmatrix} \hat{S}_{11} - \langle e^2 \rangle & \hat{S}_{12} - \langle e^2 \rangle & \dots & \hat{S}_{1n} - \langle e^2 \rangle & 1 \\ \hat{S}_{11} - \langle e^2 \rangle & \hat{S}_{ij} - \langle e^2 \rangle & & & 1 \\ \hat{S}_{n1} - \langle e^2 \rangle & \hat{S}_{11} - \langle e^2 \rangle & \dots & \hat{S}_{nn} - \langle e^2 \rangle & 1 \\ 1 & 1 & & 1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} \hat{S}_{01} \\ \hat{S}_{02} \\ \dots \\ \hat{S}_{0n} \\ 1 \end{pmatrix}. \quad (\text{B9})$$

Moltiplicando ambo i membri della (B7) per  $\sum_{i=1}^n w_i$  si ha:

$$\sum_{i=1}^n w_i \hat{S}_{0i} = \sum_{i=1}^n w_i \sum_{i=1}^n w_i \hat{S}_{ij} - \sum_{i=1}^n w_i^2 \langle e^2 \rangle + \sum_{i=1}^n w_i \mathbf{I}, \quad (\text{B10})$$

ma  $\sum_{i=1}^n w_i = \sum_{j=1}^n w_j = 1$ , e sommando ad ambo i membri la quantità  $\sum_{i=1}^n w_i \hat{S}_{0i}$  si ottiene:

$$\sum_{i=1}^n w_i \hat{S}_{0i} + \sum_{i=1}^n w_i \hat{S}_{0i} = \sum_{j=1}^n \sum_{i=1}^n w_j w_i \hat{S}_{ij} - \sum_{i=1}^n w_i^2 \langle e^2 \rangle + \mathbf{I} + \sum_{i=1}^n w_i \hat{S}_{0i}, \quad (\text{B11})$$

$$2 \cdot \sum_{i=1}^n w_i \hat{S}_{0i} - \sum_{j=1}^n \sum_{i=1}^n w_j w_i \hat{S}_{ij} + \sum_{i=1}^n w_i^2 \langle e^2 \rangle = \sum_{i=1}^n w_i \hat{S}_{0i} + \mathbf{I}, \quad (\text{B12})$$

possiamo verificare che il primo membro della (B12) equivale alla (B1) di conseguenza scriviamo:

$$\mathbf{s}_k^2 = \sum_{i=1}^n w_i \hat{S}_{0i} + \mathbf{I}, \quad (\text{B13})$$

che è l'errore di interpolazione, e scritto in forma matriciale equivale a:

$$\begin{pmatrix} \mathbf{s}_1^2 \\ \mathbf{s}_2^2 \\ \mathbf{s}_n^2 \\ 1 \end{pmatrix} = \begin{pmatrix} \hat{S}_{01} \\ \hat{S}_{02} \\ \hat{S}_{0n} \\ 1 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_n \\ \mathbf{I} \end{pmatrix}. \quad (\text{B14})$$

### Appendice C: relazione tra “structure function” e covariogramma.

La “structure function” è definita come la semi media tra i quadrati delle distanze tra i punti  $x_i$  e  $x_j$ , con  $i = 1, 2, \dots, n-1$  e  $j = 2 \dots n$

$$S_{ij} = \frac{1}{2} \langle (x_i - x_j)^2 \rangle. \quad (C1)$$

Sviluppando il quadrato tra  $\langle \rangle$  si ha:

$$\hat{S}_{ij} = \frac{1}{2} (\langle \hat{x}_i^2 \rangle + \langle \hat{x}_j^2 \rangle) - \langle \hat{x}_i \hat{x}_j \rangle = \langle \hat{x}_i^2 \rangle - \langle \hat{x}_i \hat{x}_j \rangle. \quad (C2)$$

La covarianza è definita:

$$\begin{aligned} \text{cov}_{ij} &= \langle (\hat{x}_i - \langle \hat{x}_i \rangle) \cdot (\hat{x}_j - \langle \hat{x}_j \rangle) \rangle = \langle \hat{x}_i \hat{x}_j \rangle - \langle \hat{x}_i \rangle \langle \hat{x}_j \rangle - \langle \langle \hat{x}_i \rangle \hat{x}_j \rangle + \langle \hat{x}_j \rangle \langle \hat{x}_i \rangle = \\ &= \langle \hat{x}_i \hat{x}_j \rangle - \langle \hat{x}_i \rangle \cdot \langle \hat{x}_j \rangle, \end{aligned}$$

pertanto:

$$\text{cov}_{ij} = \langle \hat{x}_i \hat{x}_j \rangle - \langle \hat{x}_i \rangle \cdot \langle \hat{x}_j \rangle. \quad (C3)$$

Dalla (C2) si ha che:

$$\langle \hat{x}_i \hat{x}_j \rangle = \langle \hat{x}_i^2 \rangle - \hat{S}_{ij}, \quad (C4)$$

che sostituiamo nella (C3) per avere:

$$\text{cov}_{ij} = \langle \hat{x}_i^2 \rangle - \hat{S}_{ij} - \langle \hat{x}_j \rangle \cdot \langle \hat{x}_i \rangle. \quad (C5)$$

Mediando la (C5) sul tempo (cioè per “ensemble” dello medesimo time-lag ( $\tau$ )):

$\mathbf{t} = |t_i - t_j|$  otteniamo che:

$$\text{cov}(\mathbf{t}) = \langle x^2 \rangle_{\mathbf{t}} - \hat{S}(\mathbf{t}) - \langle x \rangle_{\mathbf{t}} \cdot \langle x \rangle_{\mathbf{t}}, \quad (C6)$$

quindi:



$$\boxed{\widehat{cov}(\mathbf{t}) = \langle x^2 \rangle_{\mathbf{t}} - \langle x \rangle_{\mathbf{t}}^2 - \widehat{S}(\mathbf{t})}, \quad (C7)$$

ed infine l'espressione della covarianza sarà:

$$\boxed{\widehat{cov}(\mathbf{t}) = \mathbf{s}^2 - \widehat{S}(\mathbf{t})}. \quad (C8)$$

Sviluppando la (C2) e sostituendo  $\widehat{cov}_{ij} = \langle \hat{x}_i \hat{x}_j \rangle - \langle \hat{x}_i \rangle \cdot \langle \hat{x}_j \rangle$  si ha:

$$\widehat{S}_{ij} = \frac{1}{2} (\langle \hat{x}_i^2 \rangle + \langle \hat{x}_j^2 \rangle) - \langle \hat{x}_i \hat{x}_j \rangle = \langle \hat{x}_i^2 \rangle - \langle \hat{x}_i \hat{x}_j \rangle + \langle x \rangle^2 - \langle x \rangle^2 = \langle \hat{x}_i^2 \rangle - \langle x \rangle^2 - cov(\hat{x}_i \hat{x}_j), \quad (C9)$$

$$x' = x - \langle x \rangle \Rightarrow x = x' + \langle x \rangle, \quad (C10)$$

$$x^2 = x'^2 + \langle x \rangle^2 + 2x'\langle x \rangle,$$

e mediando otteniamo:

$$\langle x^2 \rangle = \langle x'^2 \rangle + \langle x \rangle^2 + 0,$$

ed quindi otteniamo la relazione tra “structure function” e covarianza:

$$S_{ij} = \langle x'^2 \rangle - cov(x_i, x_j) = \mathbf{s}^2 - cov(x_i, x_j), \quad (C11)$$

oppure:

$$\boxed{\widehat{S}(\mathbf{t}) = \mathbf{s}^2 - \widehat{cov}(\mathbf{t})}. \quad (C12)$$

## Referenze

- Bergamasco, A., Teatini, P., e Carbognin, L. "Confronto critico tra kriging e analisi oggettiva", *Il nuovo cemento*, Vol. 16 C, N. 3, 289-303, 1993.
- Bitterman, D. S., e Hansen D. V., "Evaluation of Sea Surface Temperature Measurement from Drifting Buoys", *Journal of Atmospheric and Ocean Technology*, 10, 88-96, 1993.
- Davis, R. E., "Drifter observation of coastal currents during CODE. The method and descriptive view.", *Journal of Geophysical Research*, 90, 4741-4755, 1985.
- Davis, R. E., "Observing the General Circulation with Floats", *Deep-Sea Research*, 38 (Suppl. 1), 531-571, 1991.
- Emery, W. J. e Thomson, R. E. "Data Analysis Methods in Physical Oceanography", edited by-Elsevier Science B.V, Amsterdam, The Netherlands 2001, 638 pp.
- Hansen D. V., e Herman A., "Temporal Sampling Requirements for Surface Drifting Buoys in the Tropical Pacific", *Journal of Atmospheric and Ocean Technology*, 6, 599-607, 1989.
- Hansen, D. V., e Poulain, P. M., "Quality Control and Interpolations of WOCE-TOGA Drifter Data", *Journal of Atmospheric and Ocean Technology*, 13(4), 900-909, 1996.
- Poulain, P. M. "Adriatic Sea Surface Circulation as Derived from Drifter Data Between 1990 and 1999", *Journal of Marine Systems*, 29, 3-32, 2001.