

POLITECNICO DI MILANO

Corso di Laurea in Ingegneria per l'Ambiente e il
Territorio



INFERENZA STATISTICA CON MISURE SAR E GPS APPLICATE A MODELLI DI INVERSIONE GEOFISICA

Relatore: Prof. Riccardo Barzaghi

Tesi di laurea di:

Marco Motta
Matr. 739529

Anno accademico 2010/2011

Indice

Indice delle figure	III
Abstract	VI
Introduzione	1
Capitolo 1	
Acquisizione, trattamento e filtraggio dei dati	
1.1 L'utilizzo del GPS per fini geodinamici: le reti di stazioni permanenti e non permanenti	3
1.2 Cenni sulle modalità di rilievo con strumentazione GPS.....	10
1.3 Il caso in esame: le campagne nell'area del Pollino	14
1.4 Analisi statistica e trattamento preliminare dei dati GPS	15
1.5 Compensazione delle reti	17
1.6 Stima delle velocità per le stazioni della rete del Pollino.....	19
1.7 Procedura di filtraggio COMM (Common-mode error).....	24
1.7.1 Individuazione del trend	25
1.7.2 Individuazione dell'errore medio giornaliero	25
1.7.3 Filtraggio delle osservazioni.....	26
1.7.4 Risultati della procedura COMM.....	26
1.8 Procedura di filtraggio COMM modificata	31
1.9 Considerazioni finali sui risultati del filtraggio	36

1.10	Acquisizione e analisi dei dati SAR	37
1.11	Dati sintetici per la valutazione di integrabilità tra misure SAR e GPS	40
 Capitolo 2		
<i>Il modello di faglia e i metodi di inversione</i>		
2.1	Il modello di Okada.....	41
2.2	Parametri e caratteristiche della faglia di Castrovillari	45
2.3	Il problema dell'inversione	50
2.4	Minimi quadrati regolarizzati	52
 Capitolo 3		
<i>Confronto incrociato dei risultati e test statistici</i>		
3.1	Il confronto incrociato	58
3.2	Schema logico e passi della simulazione	59
3.3	Generazione della matrice di varianza e covarianza nel caso del confronto incrociato	61
3.4	I test statistici	65
3.4.1	Il test classico	65
3.4.2	Il test classico con matrice di covarianza modello.....	67
3.4.3	Il test Bayesiano	68
3.5	Analisi su dati simulati.....	73
 Conclusione		83
 Appendice		85
 Bibliografia		92

Indice delle figure

Capitolo 1

Acquisizione, trattamento e filtraggio dei dati

1.1	Alcuni esempi di antenne e ricevitori di varie marche produttrici	4
1.2	Schema ed esempio reale di una stazione permanente realizzata con colonna in cemento	5
1.3	Schemi di realizzazione delle stazioni di tipo short e long drilled braced	5
1.4	Schema ed esempio reale di una stazione permanente realizzata tramite sostegno in acciaio inossidabile	6
1.5	Variazione del centro di fase dell'antenna e definizione dell'angolo di cut-off	7
1.6	La basetta prima e dopo l'installazione sul punto di misura	8
1.7	Il sostegno dell'antenna ricevente e l'antenna montata	9
1.8	Schema di un posizionamento assoluto	12
1.9	Schema di un posizionamento relativo	12
1.10	La rete di stazioni non-permanenti del Pollino e la faglia di Castrovillari ...	14
1.11	Serie temporale di misura originale	16
1.12	Serie temporale dopo la rimozione degli outliers	16
1.13	Serie definitivamente corretta con la rimozione delle discontinuità	17
1.14	Stelle su un punto di riferimento	17
1.15	Basi a minor lunghezza	17
1.16	Posizione ed orientamento della faglia rispetto alle stazioni della rete	23

1.17	Spostamenti e relativi scarti delle componenti planimetriche prima e dopo l'applicazione della procedura COMM nella stazione di Cassano allo Ionio	30
1.18	Spostamenti e relativi scarti della componente altimetrica prima e dopo l'applicazione della procedura COMM nella stazione di Cassano allo Ionio	31
1.19	Spostamenti e relativi scarti delle componenti planimetriche prima e dopo l'applicazione della procedura COMM modificata nella stazione di Cassano allo Ionio	35
1.20	Spostamenti e relativi scarti della componente altimetrica prima e dopo l'applicazione della procedura COMM modificata nella stazione di Cassano allo Ionio.....	36
1.21	Interferogramma ottenuto nella zona dell'Aquila dopo il sisma del 6 aprile 2009 che evidenzia la forte depressione nella zona urbana	38
1.22	Posizioni dei 661 punti di rilievo SAR in relazione alla faglia di Castrovillari.....	39

Capitolo 2

Il modello di faglia e i metodi di inversione

2.1	Definizione delle due placche, degli angoli di Rake λ e Dip δ e del punto di misura della profondità nel modello di Okada	42
2.2	Metodo di misura delle dimensioni caratteristiche della faglia W ed L e determinazione dell'angolo θ	43
2.3	La faglia di Castrovillari inquadrata nel suo contesto geografico	46
2.4	La faglia discretizzata con una configurazione sintetica delle velocità calcolata per ogni singola patch	47
2.5	Visione in pianta dei punti di controllo superficiali (in verde i punti del rilievo SAR, in rosso le stazioni non-permanenti della rete) con la posizione relativa della faglia, così come mostrata nel suo contesto in figura 2.3	48
2.6	Dettaglio della faglia di Castrovillari che si localizza al di sotto della scarpata sul terreno (le frecce evidenziano la porzione di terreno sollevata) in prossimità della catena del Monte Pollino	49

2.7	Esempio di approssimazione per la serie di dati misurati nella stazione di Cassano allo Ionio relativi alla coordinata Nord; vengono evidenziati in blu i valori osservati ed in rosso i valori predetti dal modello. L'errore e_i è mostrato in figura per la prima osservazione	51
2.8	La curva di trade-off: χ^2 in funzione di $1/\beta$	56

Capitolo 3

Confronto incrociato dei risultati e test statistici

3.1	Funzione modello normale	62
3.2	Funzione modello esponenziale.....	63
3.3	Funzione modello standard di Bessel	63
3.4	Schematizzazione della suddivisione in patch della faglia con il relativo progressivo numerico	74
3.5	Confronto tra valori del vettore sintetico dei movimenti di faglia e i valori ottenuti dall'inversione dei dati SAR simulati per le singole patch all'aumentare del rumore applicato	75
3.6	Andamenti delle covarianze empiriche e modello ottenute dall'elaborazione di dati SAR con rumore pari a 0,4 mm/anno	76
3.7	Andamenti delle covarianze empiriche e modello per la componente NORD al variare del rumore	77
3.8	Andamenti delle covarianze empiriche e modello per la componente EST al variare del rumore	78
3.9	Risultati delle statistiche test ottenuti sulle componenti NORD con ϵ_{SAR} pari a 0,2 mm/anno per valori di ϵ_{GPS} pari ad 1 e a 4 mm	80
3.10	Spostamenti stimati per la stazione dorm con ϵ_{GPS} pari ad 1 mm	81
3.11	Spostamenti stimati per la stazione dorm con ϵ_{GPS} pari ad 4 mm	81
3.12	Valori delle statistiche test ottenuti con ϵ_{GPS} pari ad 2 mm ed ϵ_{SAR} pari a 4 mm/anno	82

Abstract

Nell'ambito del monitoraggio dei fenomeni geodinamici è fondamentale valutare la possibilità di integrazione tra dati di tipo SAR e serie temporali GPS nell'applicazione di modelli di inversione geofisica. Nella tesi si considera il caso particolare della faglia di Castrovillari e, dopo uno studio preliminare sul filtraggio di misure reali GPS con la procedura del Common Mode Error, viene effettuata un'analisi su dati simulati al variare dei livelli di rumore applicato su entrambe le tipologie di dati.

L'applicazione del modello di Okada, in forma diretta e inversa, e lo svolgimento di opportuni test statistici, dimostrano l'effettiva possibilità di integrare i due tipi di dati ed evidenziano la maggiore rapidità dei test che considerano anche la variabilità del modello nel convergere in significatività all'aumentare delle campagne di misura considerate. In particolare, i risultati mostrano che il test di tipo Bayesiano è sempre il più efficiente e la sua applicazione consente di cogliere la congruenza tra i dati e il modello di faglia con un numero inferiore di campagne di misura.

Introduzione

Il monitoraggio dei fenomeni geodinamici è in continua evoluzione, grazie all'uso di tecniche e strumentazioni sempre più raffinate. La crescente disponibilità di dati richiede però una loro verifica, in termini di confronto e integrazione, volta al loro corretto utilizzo nei modelli specifici in grado di simulare la realtà in esame. E' ciò che ci si propone con quest'analisi, intesa a valutare l'effettiva integrabilità dei dati di tipo SAR con dati di tipo GPS, acquisiti utilizzando reti non-permanenti, nell'ottica di descrivere il fenomeno dei movimenti crostali relativi al caso specifico della regione del Pollino.

E' importante considerare che le serie temporali acquisite contengono sia le informazioni significative, legate direttamente alla tettonica della regione considerata, sia componenti che non dipendono dalle modifiche morfologiche della superficie.

Alla luce di questo fatto, verrà svolta un'analisi preliminare per valutare l'efficacia di alcune tecniche di miglioramento della qualità dei dati.

Si è scelto di applicare la procedura di filtraggio COMM (Common-mode error) ad una serie di misure reali ottenute sulla rete dal 2003 al 2009, per identificare e rimuovere le caratteristiche delle serie temporali che non siano legate al ciclo di deformazione della faglia in esame.

Il primo capitolo sarà interamente dedicato al trattamento del dato GPS; dopo una breve descrizione introduttiva sulle reti di rilevamento non-permanenti, verrà dato ampio spazio all'analisi dei dati acquisiti nelle singole stazioni e ai risultati del loro filtraggio.

Il dato, una volta trattato, deve essere quindi inserito nel contesto di un modello che definisca il legame tra le deformazioni profonde sul piano di faglia e le evoluzioni in superficie. In letteratura è disponibile una vasta gamma di modelli di questo tipo e nel secondo capitolo verrà preso in esame uno dei più importanti, il modello di Okada, descrivendone le caratteristiche principali.

Si affronterà poi, dal punto di vista teorico, il delicato problema della sua inversione, introducendo il metodo dei minimi quadrati regolarizzati.

Dopo aver definito il modello da usare e il metodo della sua inversione, nel terzo capitolo verrà effettuata una simulazione partendo da dati sintetici di tipo GPS e SAR, generati da una distribuzione artificiale di velocità sul piano di faglia.

Allo scopo di dimostrare l'integrabilità dei dati di tipo GPS con quelli SAR, questi ultimi verranno utilizzati per ricavare una seconda distribuzione di velocità sul piano di faglia, applicando l'inversione del modello di Okada. Dalla nuova distribuzione di velocità si ricaveranno nuove serie temporali GPS da comparare con quelle create in precedenza.

Le problematiche di confronto verranno affrontate attraverso tre diversi tipi di test statistici: quello classico del chi-quadrato, quello classico con matrice di covarianza modello ed infine quello Bayesiano.

L'ingresso in significatività dei test dimostrerà, oltre alla robustezza del modello e della sua inversione, l'effettiva possibilità di integrare dati di diverso tipo nell'ambito dello stesso modello.

Ci si attende inoltre che i risultati dei test classico con matrice di covarianza e Bayesiano tendano a raggiungere più rapidamente la significatività, in quanto la loro maggiore stabilità in questo tipo di applicazioni è nota in letteratura.

Lo studio verrà effettuato applicando un rumore variabile ai dati sintetici e verrà ripetuto al progressivo aumentare dei dati nella successione delle campagne di misura.

Capitolo 1

Acquisizione, trattamento e filtraggio dei dati

1.1 L'utilizzo del GPS per fini geodinamici: le reti di stazioni permanenti e non-permanenti

L'utilizzo della tecnologia GPS per scopi geodinamici ha come obiettivo il monitoraggio delle deformazioni crostali a diverse scale, che vanno da quella continentale o globale, alla scala regionale, che riguarda aree di estensione pari a qualche grado di latitudine e di longitudine, fino ad arrivare alla scala locale per fenomeni molto circoscritti, compresi nell'arco di pochi chilometri.

Dal monitoraggio geodinamico ci si aspetta di determinare valori deformativi che variano dai millimetri ai centimetri per anno; questo è possibile solo grazie a particolari accorgimenti, che devono essere tenuti nella dovuta considerazione già nelle fasi di realizzazione della stazione permanente.

E' innanzitutto essenziale servirsi di ricevitori di alta precisione a doppia frequenza, ovvero che possano sfruttare entrambe le portanti L1 ed L2 per l'acquisizione dei dati. Questi strumenti dovranno a loro volta essere collegati ad antenne geodetiche esterne di tipo Choke ring, con elevata immunità al multipath, immuni quindi da potenziali riflessioni multiple del segnale causate dalla configurazione topografica del punto di rilievo.

Oltre a scegliere strumenti adeguati, è necessario considerare un sito adatto per l'installazione, che garantisca una copertura ottimale del segnale GPS e che non sia soggetto ad interferenze elettromagnetiche significative. Il sito di installazione deve essere inoltre di interesse geofisico e deve trovarsi su terreno assolutamente stabile, non coinvolto da possibili dissesti locali.



Figura (1.1) Alcuni esempi di antenne e ricevitori di varie marche produttrici

La realizzazione materiale della stazione permanente deve quindi prevedere una struttura esterna resistente, che può essere eseguita secondo diversi standard costruttivi:

- I pilastri in cemento, che possono essere rinforzati con diverse tecniche, sono particolarmente adatti per installazioni non direttamente eseguite sul bedrock ed hanno una grande durata in termini temporali, oltre che un costo contenuto.
- Le stazioni short drilled braced a treppiede con fondamenta superficiali vengono principalmente usate per installazioni sul bedrock in località remote e sono relativamente economiche.
- Le stazioni deep drilled braced a treppiede con fondamenta profonde sono invece usate sia per installazioni su bedrock, sia per installazioni su terreno e hanno costi più elevati di realizzazione.
- Sostegni in acciaio inossidabile, che garantiscono la minima interazione con il riscaldamento solare ed una stabilità a lungo termine.
- Sostegni in INVAR fissati in basi di cemento, che uniscono ad una buona stabilità un'espansione termica minima ed un'elevata immunità al multipath.

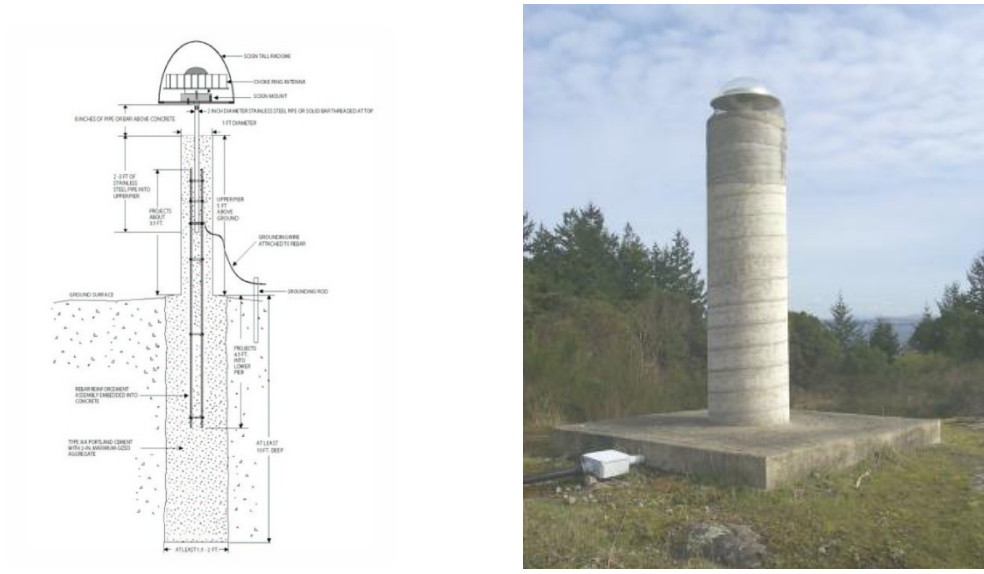


Figura (1.2) Schema ed esempio reale di una stazione permanente realizzata con colonna in cemento

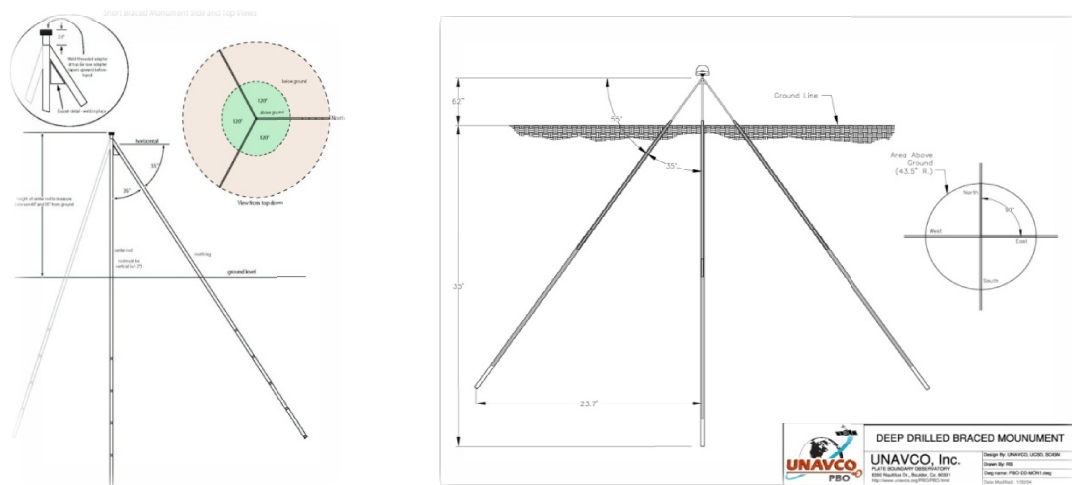


Figura (1.3) Schemi di realizzazione delle stazioni di tipo short e long drilled braced

Le stazioni permanenti devono essere equipaggiate con un orologio esterno al rubidio, con prestazioni superiori ai normali orologi interni dei ricevitori e sono generalmente affiancate da una centralina meteo che fornisce con continuità dati di pressione, temperatura e umidità relativa.

All'interno della struttura hanno inoltre una unità di smistamento in grado di inviare i dati di correzione e i dati di osservazione (raw data) su più canali radio-modem e su più canali GSM, ed un computer che ha il compito, mediante apposito software, di gestire la stazione svolgendo le funzioni di colloquio e di scambio dati tra le varie componenti del sistema e di monitorare la visibilità dei satelliti.

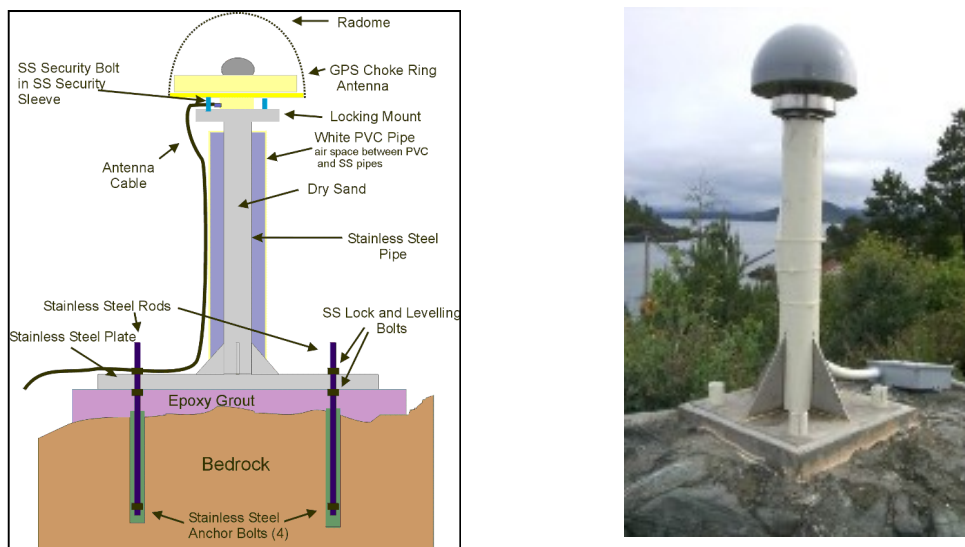


Figura (1.4) Schema ed esempio reale di una stazione permanente realizzata tramite sostegno in acciaio inossidabile

Il computer deve inoltre gestire la traduzione e l'archiviazione dei dati acquisiti, generando allarmi in caso di malfunzionamento di alcune parti, e controllare il collegamento Internet.

E' fondamentale che tutto il sistema sia alimentato mediante un gruppo di continuità che garantisca il funzionamento anche per lunghi periodi di assenza di energia elettrica della normale rete di distribuzione. E' inoltre importante la presenza di dispositivi di protezione dalle scariche elettriche.

E' necessario specificare che l'affidabilità della stazione permanente appena realizzata debba essere verificata con un'analisi sulle coordinate ottenute dalle serie temporali nei primi periodi di funzionamento; per quest'analisi vengono

spesso utilizzate le stazioni permanenti adiacenti a quella in esame, che abbiano delle serie temporali di misura abbastanza consistenti.

Nell'ottica di ottenere stime corrette in campo geodinamico sono necessari alcuni accorgimenti specifici nella fase di acquisizione e di processamento dei dati.

Per prima cosa si deve scegliere una corretta frequenza di campionamento, che può essere ragionevolmente dell'ordine di una acquisizione ogni 15-30 secondi.

Si deve inoltre impostare un adeguato valore del cut-off, definito come l'angolo α a partire dall'orizzonte entro il quale si evita di considerare i segnali provenienti dai satelliti perché si ritiene che possano essere influenzati maggiormente da riflessioni o deviazioni atmosferiche. La scelta del cut-off può dipendere anche dal contesto topografico della singola stazione, ma generalmente si sceglie un angolo compreso tra 0 e 15 gradi di inclinazione dall'orizzontale.

Una volta ottenuti, i dati vengono processati con software scientifici (ad esempio GAMIT o il Bernese) effettuando anche una corretta modellazione delle variazioni del centro di fase dell'antenna utilizzata.

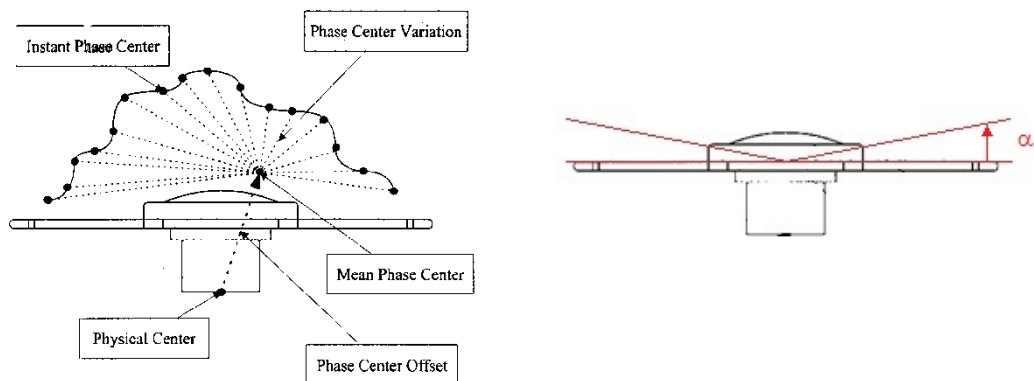


Figura (1.5) Variazione del centro di fase dell'antenna e definizione dell'angolo di cut-off

E' indispensabile, al fine di ottenere risultati consistenti, riferirsi alla corretta realizzazione del sistema di riferimento ITRF (International Terrestrial Reference Frame); le misure effettuate dalle stazioni permanenti infatti possono anche contribuire al mantenimento della realizzazione ITRF nel sistema internazionale di riferimento convenzionale e in particolare, nel caso dell'Europa, al

mantenimento del sistema ETRF2000, ultima realizzazione del sistema internazionale Europeo EUREF.

Le operazioni di analisi statistica dei dati ottenuti partono dalla semplice eliminazione degli outliers e dalla stima delle velocità, che verranno riprese più avanti, alla possibile considerazione della correlazione spaziale e temporale nelle varie serie di misura ottenute dalle stazioni permanenti.

Le stazioni permanenti rappresentano la soluzione migliore per le stime di deformazione in geodinamica, ma per lo studio delle deformazioni in piccole realtà locali è possibile utilizzare reti di stazioni non-permanenti sulle quali effettuare campagne di misura a intervalli prestabiliti.

Raggiungere risultati di alta qualità è più difficile rispetto al caso delle stazioni permanenti, sia a causa del fatto che non disporremo di serie temporali così lunghe come nel caso delle stazioni permanenti, sia per via di alcuni problemi dovuti all'effettiva ripetibilità delle misure.

La monumentazione del punto richiede soluzioni più speditive rispetto alle stazioni permanenti: i punti vengono materializzati con basette cementate nel bedrock come mostrato nelle figure seguenti.



Figura (1.6) La basetta prima e dopo l'installazione sul punto di misura

L'installazione dell'antenna sulla base, effettuata attraverso un sostegno in acciaio, è un altro fattore che può comportare ulteriori incertezze nelle varie campagne di misura.

La necessità di rimontare i sostegni ad ogni campagna di misura impone di posizionare in modo accurato la strumentazione per non rischiare di causare errori sistematici. L'utilizzo delle livelle toriche consente in questi casi di raggiungere precisioni di livello sub-millimetrico nella determinazione della verticale.



Figura (1.7) Il sostegno dell'antenna ricevente e l'antenna montata

Altre incertezze potrebbero derivare dal possibile cambiamento delle antenne e dei ricevitori da una campagna all'altra; è bene quindi mantenere il più possibile inalterato l'assetto iniziale della strumentazione per ottenere risultati migliori.

Soluzioni più precise possono essere ottenute inserendo nella rete non permanente anche un collegamento alle stazioni permanenti vicine e utilizzando un periodo di misura più lungo, che può arrivare fino a 24 ore.

In sintesi, i limiti di applicazione e le modalità di acquisizione nelle reti non permanenti, che ben si applicano al nostro caso specifico del Pollino, sono i seguenti:

- Copertura areale della rete di 20-30 chilometri al massimo
- Campagne di misura ripetute annualmente nello stesso periodo dell'anno
- Acquisizioni dei dati per 8 ore consecutive ripetute per 4 giorni con una frequenza di campionamento di 15 secondi e un cut-off pari a 15 gradi sull'orizzontale
- Mantenimento degli stessi strumenti sugli stessi punti in tutte le campagne
- Collegamento, se possibile, a stazioni permanenti vicine alla rete in esame
- Ripetibilità delle misure dell'ordine di qualche millimetro

1.2 Cenni sulle modalità di rilievo con strumentazione GPS

Nell'utilizzo del GPS esistono vari tipi di ricevitori che consentono diverse modalità di determinazione della posizione dei punti, che può essere relativa o assoluta, sul singolo punto, come nel caso delle stazioni permanenti.

A queste diverse tecniche corrispondono livelli più o meno elevati di precisione, come vedremo nel seguito del paragrafo.

Le modalità di funzionamento possono essere classificate a seconda delle osservabili acquisite, del funzionamento contemporaneo o meno di più strumenti, del movimento o meno del ricevitore e della necessità di post elaborazione.

Il **posizionamento assoluto** (Point Positioning) consente di ottenere la posizione di un singolo ricevitore nel sistema WGS84. Per ogni posizione del ricevitore le incognite geometriche sono le tre coordinate del punto di stazione; ad esse si aggiunge, per ogni epoca, l'incognita dovuta all'instabilità dell'orologio del ricevitore, che induce un offset rispetto al tempo GPS. Per ogni epoca devono essere eseguite quindi almeno quattro misure su altrettanti satelliti per risolvere

le incognite in tempo reale. Nel caso si utilizzino le misure di codice C/A la precisione è dell'ordine di dieci metri in planimetria e venti metri in altimetria, ad un livello del 95%.

Nel caso si utilizzino le fasi invece ci si riferisce al **posizionamento assoluto di precisione** (Precise Point Positioning); in questo caso vengono effettuate le misure sulla fase utilizzando entrambe le frequenze portanti del segnale GPS con ricevitori idonei.

Con opportuni e complessi software è possibile modellare efficacemente gli errori sistematici in post-processamento e, utilizzando orbite più precise di quelle trasmesse, si possono raggiungere precisioni nell'ordine della decina di centimetri per la singola misurazione.

Nel **posizionamento relativo** invece, grazie allo stazionamento contemporaneo di due o più ricevitori durante una sessione di rilievo, è possibile ottenere il posizionamento dalla misura della "BASE GPS", ovvero il segmento compreso tra due ricevitori. La precisione della base è decisamente più elevata di quella garantita dal posizionamento assoluto: lo stazionamento contemporaneo infatti consente di eliminare gli errori sistematici perché, facendo uso di misure di fase, è possibile effettuare le differenze sulle osservazioni in post-processamento.

La precisione raggiungibile nelle misurazioni dipende dalla lunghezza della base e, definendo con L la lunghezza della base, è dell'ordine di circa:

- $(5 + 1 * L[\text{km}])$ mm in planimetria
- $(10 + 2 * L[\text{km}])$ mm per il dislivello tra i due ricevitori

Nel posizionamento relativo la base tra due ricevitori è ottenuta determinando, con le misure acquisite nella sessione di rilievo, le tre incognite geometriche (le tre differenze di coordinate) e le ambiguità di fase.

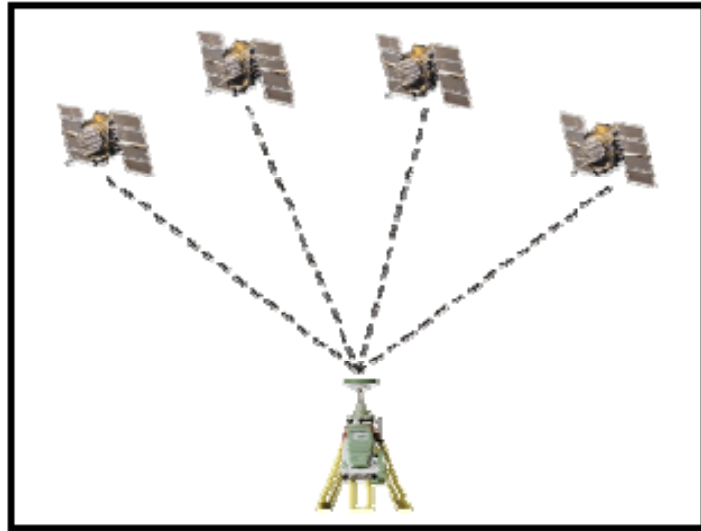


Figura (1.8) Schema di un posizionamento assoluto

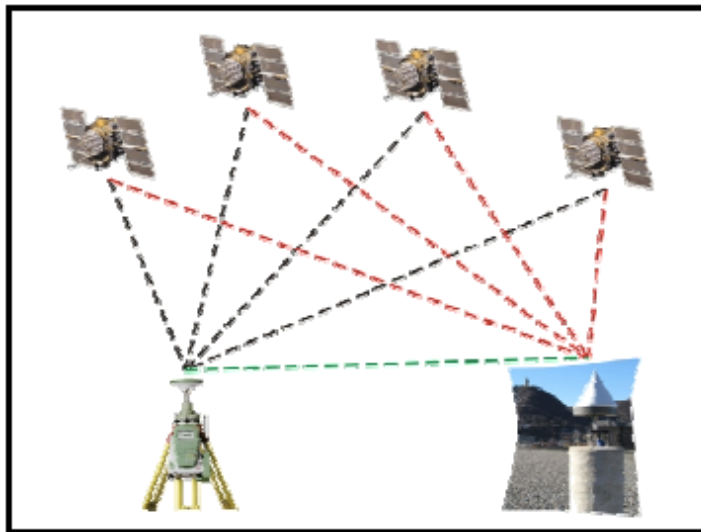


Figura (1.9) Schema di un posizionamento relativo

Un'ulteriore metodologia è rappresentata dal **posizionamento differenziale** (DGPS). Le misure di posizionamento assoluto eseguite da un ricevitore mobile (Rover) possono essere corrette tramite analoghe misure eseguite da un secondo ricevitore (Master) mantenuto su un punto di posizione nota.

La contemporaneità delle ricezioni e la correlazione spaziale degli effetti sistematici consente di ridurre o addirittura eliminare alcuni errori. Per distanze ridotte tra Rover e Master si raggiungono precisioni dell'ordine del metro o di pochi metri. Tale risultato può essere ottenuto in tempo reale se le correzioni vengono trasmesse al Rover con un modesto ritardo (tempo di latenza).

Per ogni posizione del ricevitore le incognite geometriche sono le 3 coordinate del punto di stazione.

Sempre nell'ottica di fornire una panoramica generale delle tipologie di rilievi, si possono definire alcune ulteriori distinzioni:

- considerando la possibilità di muovere il ricevitore i rilievi possono essere:
 - statici, se i ricevitori stazionano sempre sullo stesso punto
 - cinematici, se un ricevitore si muove durante l'acquisizione dei dati

- considerando il momento di elaborazione dei dati le correzioni sono:
 - in tempo reale, quando il processamento dei dati per determinare le incognite viene eseguito contemporaneamente all'acquisizione degli stessi
 - in post-processamento, quando le elaborazioni sono eseguite in un momento successivo al rilievo, utilizzando correzioni più accurate e ed effemeridi più precise di quelle trasmesse.

Nel caso di stazioni permanenti e non-permanenti ci si trova ad eseguire dei posizionamenti che raggiungono un'altissima precisione grazie alle modalità di installazione viste in precedenza e soprattutto grazie alla continuità nell'acquisizione dei dati sul punto di stazionamento.

1.3 Il caso in esame: le campagne nell'area del Pollino

I dati utilizzati nella nostra analisi riguardano sei campagne eseguite tra il 2003 ed il 2009 nella regione del Pollino su dieci stazioni non-permanenti distribuite attorno alla faglia locale come mostrato nella figura seguente.

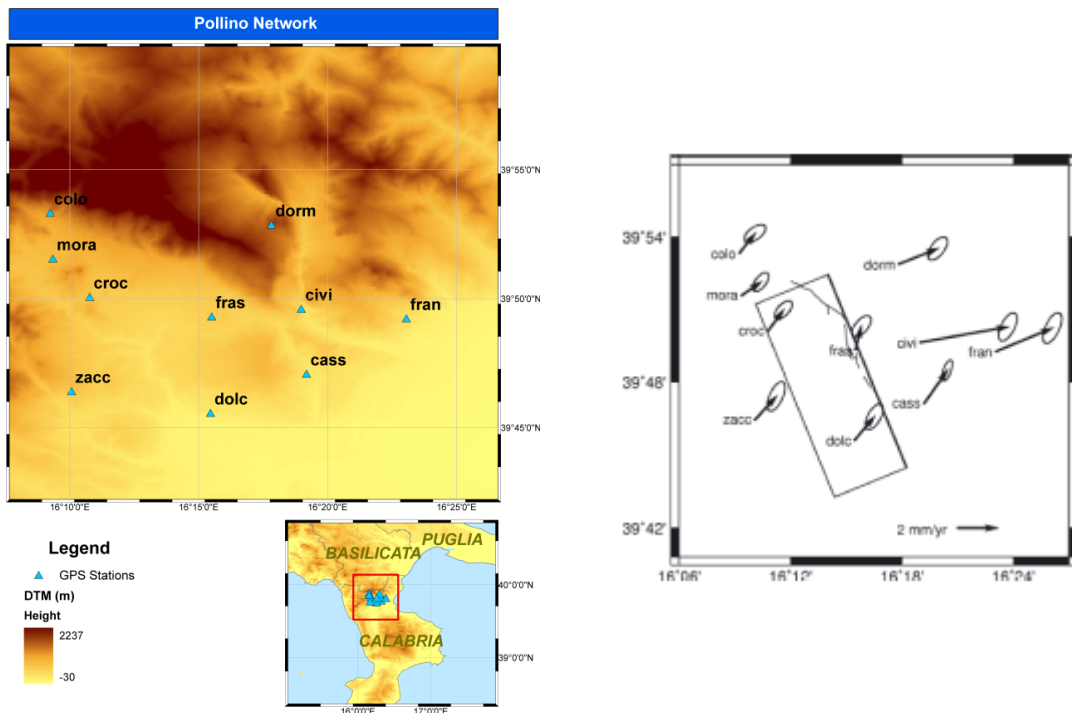


Figura (1.10) La rete di stazioni non-permanenti del Pollino e la faglia di Castrovillari

In particolare, si tratta di rilievi di tre giorni per le prime due campagne, e di quattro giorni per le successive, durante i quali sono state effettuate misurazioni continue di 8 ore in contemporanea su tutte e dieci le stazioni considerate per lo studio. La tabella seguente mostra i periodi dell'anno relativi alle varie campagne effettuate.

<i>TABELLA (1.1)</i>		
Anno di campagna	Giorni giuliani	Mese
2003	253-255	settembre
2004	125-128	aprile
2005	249-252	settembre
2007	282-285	ottobre
2008	253-256	settembre
2009	258-261	settembre

I rilievi sono stati tutti eseguiti nel periodo autunnale, all'incirca negli stessi giorni dell'anno ad eccezione del 2004, in cui le misure sono state svolte in aprile, e del 2006, anno che non ha visto lo svolgimento dei rilievi.

1.4 Analisi statistica e trattamento preliminare dei dati GPS

Nel trattamento delle serie temporali è importante per prima cosa identificare e rimuovere eventuali discontinuità e trend che non siano effettivamente dovute al ciclo di deformazione della faglia in esame.

Il primo passo consiste nell'individuare ed eliminare le discontinuità nelle misure che solitamente sono legate a:

- cambi di sistema di riferimento
- sostituzione dell'antenna
- modifiche all'hardware o al software

Le discontinuità devono essere individuate nel tempo e devono essere stimate per poi essere rimosse dalle misure.

E' prevista in ogni caso una documentazione che tenga conto dell'impatto di questi fattori sulle serie temporali.

Un ulteriore passo per migliorare i dati è certamente l'eliminazione degli outliers, ovvero le misure da considerare errate, che vengono individuate attraverso dei test sullo scarto quadratico medio σ .

Con il metodo più semplice si individua un trend lineare dei dati e si eliminano le osservazioni al di fuori dell'intervallo $\pm 3\sigma$ centrato sulla regressione lineare calcolata. Altri metodi prevedono il calcolo degli scarti normalizzati, che possono essere ulteriormente analizzati attraverso test di ipotesi basati sulla distribuzione normale.

Si riporta in seguito una serie di grafici che mostra la procedura per la stazione permanente di Bolzano.

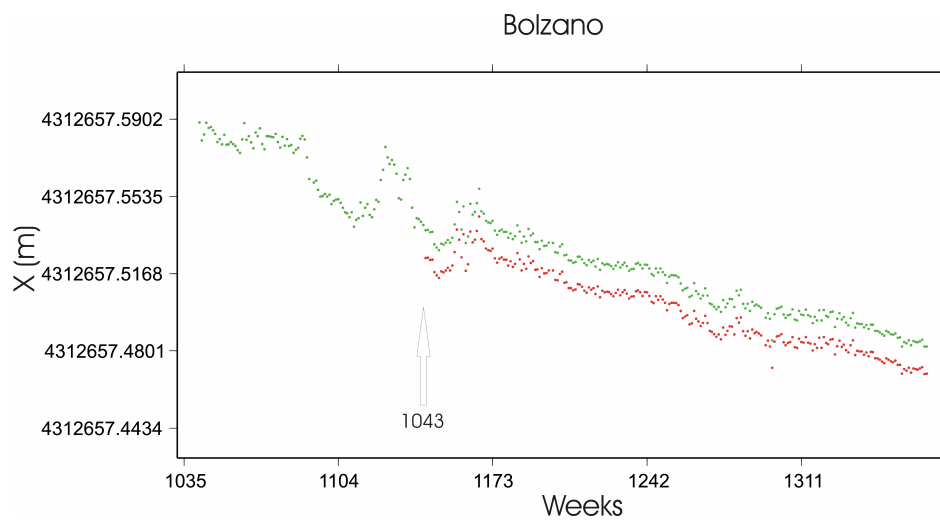


Figura (1.11) Serie temporale di misura originale

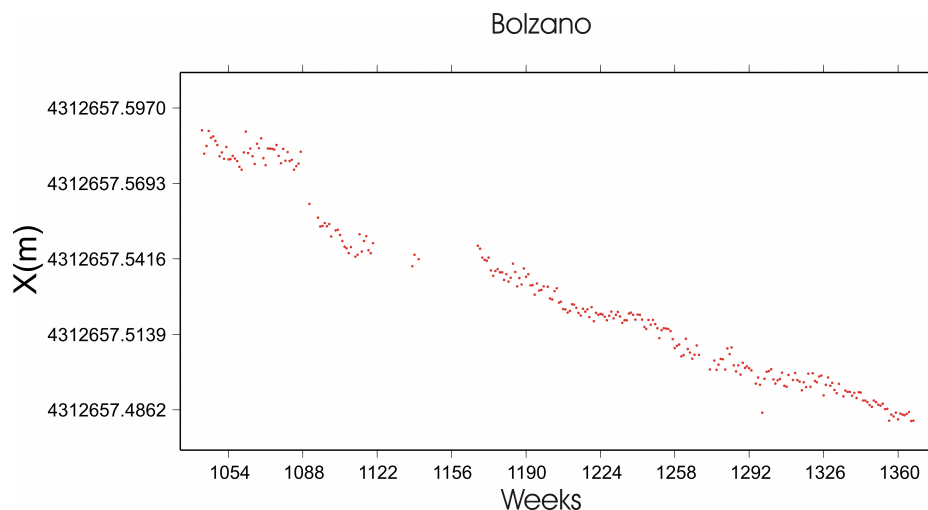


Figura (1.12) Serie temporale dopo la rimozione degli outliers

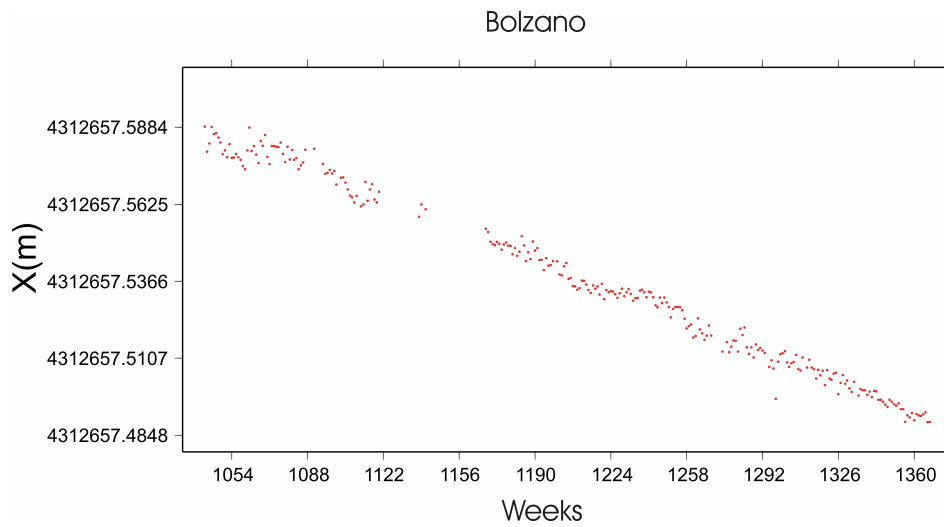


Figura (1.13) Serie definitivamente corretta con la rimozione delle discontinuità

1.5 Compensazione delle reti

L'elaborazione dei dati GPS ottenuti da una rete di rilievo classica impone di eseguire le operazioni di compensazione e di calcolo delle basi comprese tra i vertici della rete stessa.

Con n ricevitori è teoricamente possibile la costruzione e la stima di $n(n-1)/2$ basi, delle quali però solo $n-1$ saranno indipendenti tra loro. In genere si utilizzano schemi simili a quelli riportati in seguito

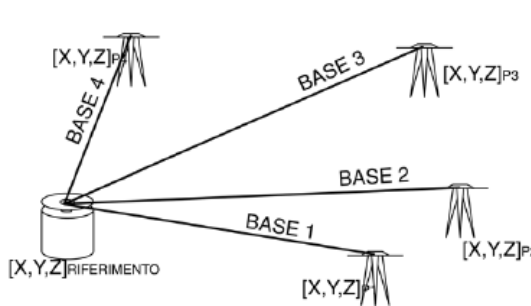


Figura (1.14) Stelle su un punto di riferimento

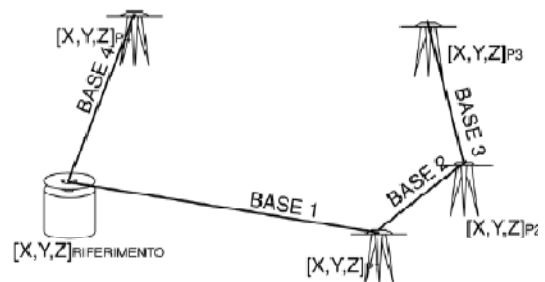


Figura (1.15) Basi a minor lunghezza

Nel rilievo completo di una rete con N vertici si devono connettere fra loro tutti i punti, ovvero rilevare $N-1$ basi.

Se ad esempio disponiamo di due ricevitori sarà necessario effettuare almeno $N-1$ sessioni (misurando una sola base per sessione); all'aumentare dei ricevitori disponibili il numero di sessioni strettamente necessarie diminuisce fino ad arrivare ad una singola sessione nel caso delle reti permanenti o non-permanenti, in cui su tutti i punti si ha un ricevitore attivo durante il rilievo.

Nell'analisi delle reti è importante il concetto di ridondanza, ovvero la differenza fra il numero di osservazioni e il numero di incognite della rete. Le reti rilevate a ridondanza nulla sono, per ovvie ragioni, poco affidabili e si cerca in genere di realizzare schemi di rilievo a ridondanza maggiore di zero, che permettono un controllo preventivo degli errori grossolani grazie all'analisi delle chiusure e consentono di effettuare la compensazione rigorosa della rete con la determinazione delle coordinate e delle precisioni di stima dei punti della rete.

Su di un percorso chiuso la somma delle variazioni di ciascuna componente delle basi deve essere nulla teoricamente ed il calcolo degli errori di chiusura permette di valutare anche a posteriori la qualità delle stime delle basi e la presenza degli eventuali errori grossolani.

Dopo aver eliminato gli errori grossolani dalle serie di misura, è possibile effettuare la compensazione sulla rete ai minimi quadrati; i dati in ingresso sono:

- la posizione del o dei punti di riferimento della rete
- le stime delle basi indipendenti date dall'elaborazione delle osservazioni
- le matrici di covarianza dal modello stocastico

I parametri incogniti invece saranno:

- le coordinate dei rimanenti punti della rete
- le precisioni delle stime e la loro matrice di covarianza
- le regioni di confidenza

Nel processo di compensazione il primo passo è la costruzione del modello globale; vengono quindi definiti i vettori delle incognite, delle osservazioni, dei

termini noti e la matrice disegno della rete e viene costruita la matrice di covarianza delle osservazioni. Con queste premesse è poi possibile ottenere le stime ai minimi quadrati dei valori cercati e valutare l'affidabilità delle analisi attraverso test statistici come quello del chi-quadro.

Nell'analisi delle reti permanenti e non-permanenti, le singole osservazioni che vengono trattate hanno di per sé una grande precisione e questo permette di ottenere risultati soddisfacenti a prescindere dalla eventuale compensazione.

1.6 Stima delle velocità per le stazioni della rete del Pollino

Dalle serie temporali ottenute durante le misure, grazie all'analisi statistica ed alla compensazione di rete viste in precedenza, sono state determinate, per ogni stazione non-permanente, una serie di soluzioni giornaliere delle coordinate geografiche (N, E e h) con la relativa matrice di covarianza delle stime.

Per comodità interpretativa i valori delle coordinate sono stati espressi in termini relativi rispetto al primo giorno di osservazione della prima campagna eseguita.

A partire da questi dati è possibile valutare la velocità media di deriva delle singole stazioni sull'arco temporale di tutte le campagne svolte, e determinare le tre componenti separate v_N , v_E e v_h .

Le velocità sono rappresentate dalle pendenze delle rette che interpolano le serie nel piano spazio-tempo e sono facilmente ricavabili grazie alla teoria dei minimi quadrati.

Nel nostro caso è stata utilizzata la seguente formulazione del metodo; vogliamo stimare un opportuno modello lineare

$$y = ax + b \quad (1)$$

e più precisamente definendo con y_i l'osservazione i -esima della serie avremo

$$ax + b - y_i = v_i \quad (2)$$

Dove:

- a è la pendenza della retta interpolante (e quindi la velocità)
- b è l'intercetta
- v_i è l' i -esimo scarto tra il valore osservato y_i e la previsione data dal modello

L'obiettivo del metodo è quello di minimizzare la somma degli scarti quadratici v_i^2 ottenuti per ogni osservazione della serie. Senza riportare i passaggi matematici che conducono al risultato, si ricorda che la teoria dei minimi quadrati permette di definire il vettore delle incognite, a e b , attraverso la seguente relazione matriciale:

$$\begin{bmatrix} a \\ b \end{bmatrix} = -(A^T P A)^{-1} A^T P l \quad (3)$$

Dove:

$$P = \begin{bmatrix} \frac{\sigma_0^2}{\sigma_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{\sigma_0^2}{\sigma_n^2} \end{bmatrix} \quad A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ \vdots & \vdots \\ x_i & 1 \\ \vdots & \vdots \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \quad l = \begin{bmatrix} -y_1 \\ -y_2 \\ \vdots \\ \vdots \\ -y_i \\ \vdots \\ \vdots \\ -y_n \end{bmatrix} \quad (4)$$

P è la matrice dei pesi che permette di considerare nella formulazione le singole varianze relative ad ogni osservazione y_i , A è la matrice disegno mentre l è il vettore delle osservazioni.

Nell'ottica di valutare la consistenza dei valori stimati di velocità è necessario ricavare anche la varianza dei parametri ottenuti; $\hat{\sigma}_0^2$ può essere calcolato grazie alla relazione

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^T \mathbf{P} \mathbf{v}}{r} \quad \text{con} \quad r = m - n \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_i \\ \cdot \\ \cdot \\ v_n \end{bmatrix} \quad (5)$$

Dove:

r è la ridondanza del sistema

m è il numero di osservazioni

n è il numero di incognite

\mathbf{v} è il vettore degli scarti

Dalla stima di $\hat{\sigma}_0^2$ si possono quindi determinare σ_a^2 e σ_b^2 dalle relazioni

$$(\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \cdot \hat{\sigma}_0^2 \quad \begin{cases} \sigma^2(\mathbf{a}) = \hat{\sigma}_0^2 \cdot (\text{elem. 1,1 di } (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1}) \\ \sigma^2(\mathbf{b}) = \hat{\sigma}_0^2 \cdot (\text{elem. 2,2 di } (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1}) \end{cases} \quad (6)$$

Ai fini della caratterizzazione fisica del problema, i valori dell'intercetta b con la relativa varianza di stima non sono significativi e ci si è quindi concentrati solo sulla determinazione delle velocità medie; i risultati ottenuti sono riportati nelle seguenti tabelle

TABELLA (1.2)			
Stazione	v_N	v_E	v_h
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,019736428	0,023867313	0,001563675
<i>civi</i>	0,019103371	0,026365217	8,68738E-05
<i>colo</i>	0,019372900	0,023710359	0,001726423
<i>croc</i>	0,018753113	0,024020289	0,002718445
<i>dolc</i>	0,019195277	0,023803558	0,000854133
<i>dorm</i>	0,018680730	0,023928722	0,003541370
<i>fran</i>	0,019149339	0,024855284	-0,002001685
<i>fras</i>	0,019493457	0,023040769	0,003923387
<i>mora</i>	0,018778936	0,023925422	0,001354879
<i>zacc</i>	0,019972352	0,024911188	0,003527951

TABELLA (1.3)			
Stazione	σv_N	σv_E	σv_h
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,000253774	0,000363624	0,000692372
<i>civi</i>	0,000346213	0,000349527	0,000699818
<i>colo</i>	0,000347918	0,000221077	0,000754617
<i>croc</i>	0,000287092	0,000308333	0,000793290
<i>dolc</i>	0,000287305	0,000334235	0,000809011
<i>dorm</i>	0,000315083	0,000383816	0,000749656
<i>fran</i>	0,000281499	0,000445456	0,001203920
<i>fras</i>	0,000276288	0,000409360	0,001727236
<i>mora</i>	0,000309162	0,000302309	0,001433283
<i>zacc</i>	0,000468350	0,000493758	0,001017418

Dai risultati ottenuti emerge chiaramente che le stime sulle velocità per la coordinata h sono più disperse a causa dell'intrinseca carenza di precisione del sistema GPS nel determinare spostamenti verticali rispetto a quanto possibile per le componenti orizzontali.

E' anche interessante notare che le varianze sulle velocità lungo la coordinata Est sono generalmente maggiori rispetto a quelle che si ottengono sulla coordinata Nord; questa differenza è molto probabilmente legata all'orientamento geometrico ed alla direzione di scorrimento della faglia.

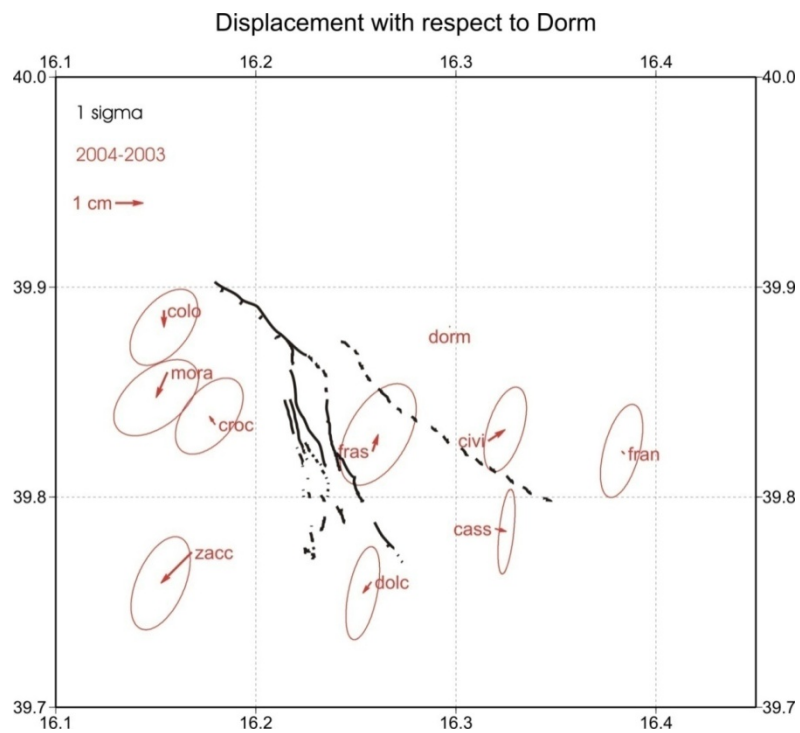


Figura (1.16) Posizione ed orientamento della faglia rispetto alle stazioni della rete

1.7 Procedura di filtraggio COMM (Common-mode error)

Abbiamo visto nei paragrafi precedenti come sia importante nel trattamento dei dati identificare e rimuovere tutte le caratteristiche insite nelle serie temporali che non siano dovute al ciclo di deformazione della faglia in esame.

Oltre agli errori di misura ed alle discontinuità è frequente che vi siano altri fattori ambientali non facilmente modellabili, che possono influenzare la stima delle grandezze caratteristiche di interesse.

Il modello lineare introdotto per la stima delle velocità può essere riespresso nella forma più generale

$$Y = Ax + B + s + n \quad (7)$$

Dove:

- s è un segnale correlato col tempo non legato alla geodinamica del problema
- n è un rumore bianco

Oltre ai parametri già descritti in precedenza ed espressi in forma matriciale generale, nella formula compaiono s ed n che possono essere considerate come grandezze scorrelate.

E' lecito ipotizzare che s sia un disturbo comune a tutte le stazioni nello stesso istante di tempo ed è su questa affermazione che si basa la procedura COMM (Common-mode error).

Il COMM è un metodo che permette in genere di attenuare le componenti non direttamente legate ai movimenti di faglia e si basa su tre passi principali che vedremo in dettaglio nei prossimi paragrafi.

1.7.1 Individuazione del trend

Come già visto nel paragrafo precedente, attraverso la regressione lineare si determina la retta che meglio si adatta ai dati originali e si stima la velocità dei singoli punti della rete. Ottenuto il modello si procede calcolando per ogni giorno d e per ogni stazione s lo scarto, cioè la differenza tra la coordinata originale e quella prevista dal modello.

$$v_s(d) = (a_s d + b_s) - y_s(d) \quad (8)$$

Dove:

- $v_s(d)$ è lo scarto al giorno d per la stazione s
- $(a_s d + b_s)$ è il valore predetto dal modello al giorno d per la stazione s
- $y_s(d)$ è il valore originale della coordinata al giorno d per la stazione s

1.7.2 Individuazione dell'errore medio giornaliero

Ottenuto il trend, si calcola quindi la media degli scarti ottenuti in un singolo giorno per tutte le stazioni ripetendo l'operazione per ogni giornata di misura

$$v(d) = \frac{\sum_{s=1}^S v_s(d)}{S} \quad (9)$$

Dove:

- S è il numero totale delle stazioni

La precisione nella determinazione dell'errore medio giornaliero tende generalmente a migliorare con l'aumento del numero di siti coinvolti nell'analisi.

1.7.3 Filtraggio delle osservazioni

Una volta ottenuti i valori di $v(d)$ per tutti i giorni di misura è possibile ottenere nuove serie di osservazioni filtrate sottraendo l'errore medio ai valori originali delle coordinate

$$\hat{y}_s(d) = y_s(d) - v(d) \quad (10)$$

Dove:

- $\hat{y}_s(d)$ è il valore della nuova coordinata filtrata al giorno d per la stazione s
- $v(d)$ è l'errore medio al giorno d
- $y_s(d)$ è il valore originale della coordinata al giorno d per la stazione s

Sulle nuove serie ottenute si deve quindi eseguire nuovamente la stima delle velocità medie su ciascun vertice della rete determinando anche le nuove varianze di stima.

1.7.4 Risultati della procedura COMM

Il metodo COMM è stato applicato alle nostre osservazioni. La stima degli errori medi giornalieri è stata effettuata per ciascuna delle tre coordinate e ha permesso di determinare i valori riportati nella tabella (1.4)

TABELLA (1.4)			
Tempo	Errore medio giornaliero		
	N	E	h
[anni]	[m/anno]	[m/anno]	[m/anno]
2003,692	7,61934E-05	-0,006137709	0,000619489
2003,695	-0,002276136	0,000985020	-0,006445322
2003,697	0,000372311	0,002703505	-0,007211863
2004,343	0,001370751	-0,001685638	-0,004344577
2004,346	-0,000411579	-0,001332910	-0,003369389
2004,348	0,000256869	0,000365576	0,002504071
2005,681	0,000712847	-0,002409062	0,003437273
2005,684	0,000929585	0,000793973	0,000324743
2005,686	0,001298033	0,003062458	0,005388202
2005,689	0,001945703	0,001635187	0,004573391
2007,771	-0,000700782	0,000198722	-0,000565697
2007,774	-9,31109E-05	-0,002548550	0,003239492
2007,777	0,000234560	0,000574178	0,001594681
2007,779	-0,002196993	-0,002617336	0,003518140
2008,690	-0,002261374	-0,000356327	-0,004644373
2008,693	-0,001656631	0,000850594	0,002728944
2008,695	-0,002068184	0,000169080	0,002852403
2008,698	-0,000950513	0,002421808	-0,000962408
2009,705	0,001837642	-0,000775680	-0,003110756
2009,708	0,001975313	-0,001252952	-0,002105568
2009,711	0,001132984	0,000739777	0,000359621
2009,714	0,003100654	0,001412505	-0,004745190

Trovati i valori dell'errore medio giornaliero è stato effettuato il filtraggio delle coordinate e si sono ristimate le velocità a partire dai dati filtrati.

I risultati sono riassunti nelle tabelle riportate in seguito; la (1.5) contiene le nuove stime delle velocità e la (1.6) riporta le differenze tra i valori ottenuti dai dati filtrati e quelli ottenuti dai dati originali.

Un confronto analogo è stato fatto con le varianze di stima ed è riportato nelle tabelle (1.7) e (1.8).

TABELLA (1.5)			
Stazione	v_N	v_E	v_h
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,019768857	0,023810748	0,001316206
<i>civi</i>	0,018976057	0,026336301	0,000231570
<i>colo</i>	0,019361558	0,023752999	0,001838054
<i>croc</i>	0,018649442	0,023997817	0,002812170
<i>dolc</i>	0,019182291	0,023749488	0,000602181
<i>dorm</i>	0,018719350	0,023920036	0,003632410
<i>fran</i>	0,019156856	0,024786770	-0,002241708
<i>fras</i>	0,019481758	0,022986030	0,003780811
<i>mora</i>	0,018651881	0,023917329	0,001303173
<i>zacc</i>	0,020028041	0,024867242	0,003244915

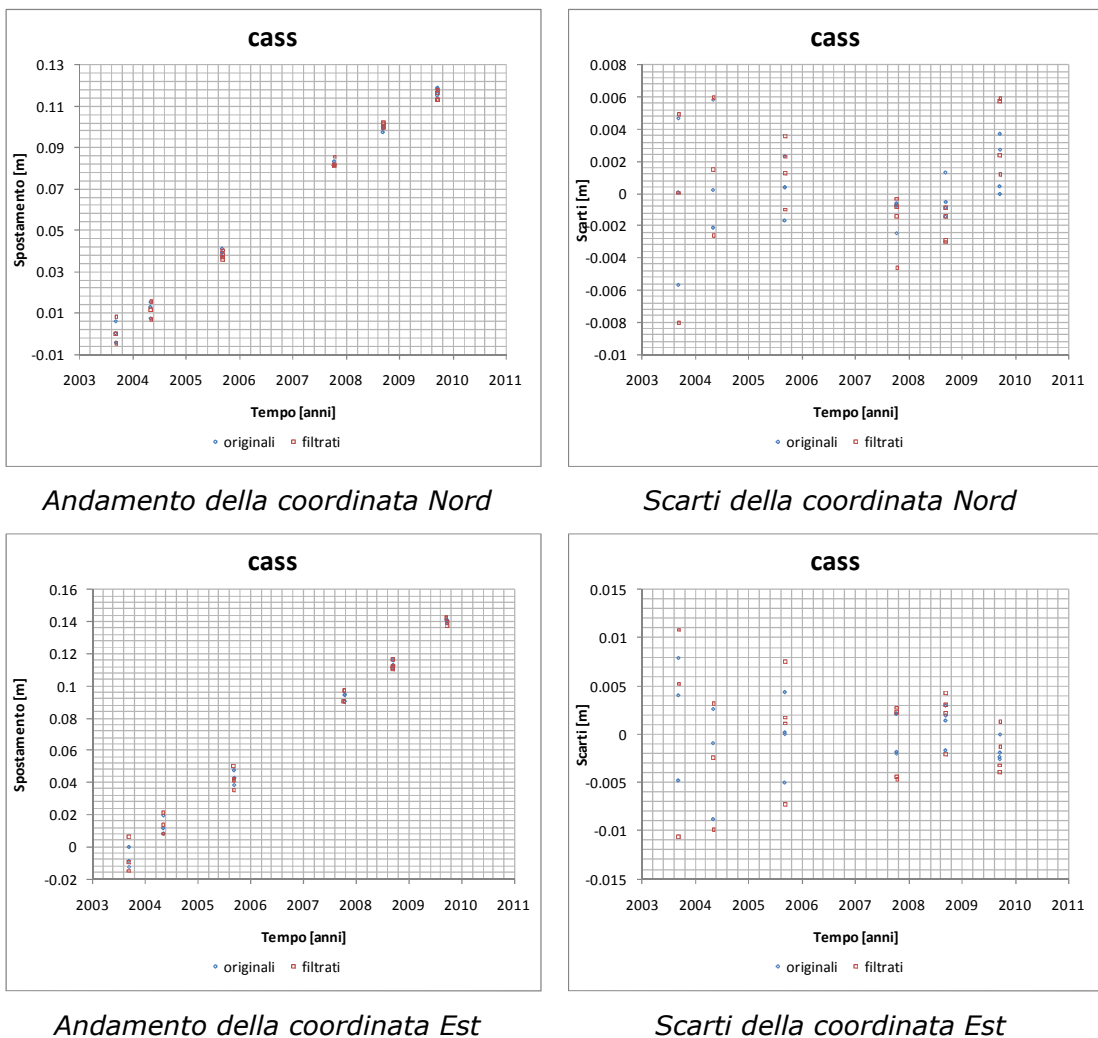
TABELLA (1.6)			
Stazione	dv_N	dv_E	dv_h
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	3,24288E-05	-5,65644E-05	-0,000247469
<i>civi</i>	-0,000127315	-2,89160E-05	0,000144697
<i>colo</i>	-1,13419E-05	4,26403E-05	0,000111631
<i>croc</i>	-0,000103671	-2,24725E-05	9,37251E-05
<i>dolc</i>	-1,29857E-05	-5,40708E-05	-0,000251951
<i>dorm</i>	3,86198E-05	-8,68579E-06	9,10402E-05
<i>fran</i>	7,51723E-06	-6,85137E-05	-0,000240023
<i>fras</i>	-1,16994E-05	-5,47387E-05	-0,000142576
<i>mora</i>	-0,000127054	-8,09289E-06	-5,17051E-05
<i>zacc</i>	5,56891E-05	-4,39460E-05	-0,000283036

TABELLA (1.7)			
Stazione	σ_{v_N}	σ_{v_E}	σ_{v_h}
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,000376195	0,000515047	0,000975953
<i>civi</i>	0,000466097	0,000475654	0,000962887
<i>colo</i>	0,000506948	0,000395682	0,000930508
<i>croc</i>	0,000436537	0,000426826	0,000977696
<i>dolc</i>	0,000407217	0,000441922	0,001101461
<i>dorm</i>	0,000392367	0,000522934	0,000929123
<i>fran</i>	0,000431839	0,000576651	0,001397511
<i>fras</i>	0,000382773	0,000568144	0,001781848
<i>mora</i>	0,000466717	0,000469678	0,001684904
<i>zacc</i>	0,000547679	0,000559146	0,001187030

TABELLA (1.8)			
Stazione	$d\sigma_{v_N}$	$d\sigma_{v_E}$	$d\sigma_{v_h}$
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,000122421	0,000151423	0,000283580
<i>civi</i>	0,000119884	0,000126127	0,000263068
<i>colo</i>	0,000159030	0,000174605	0,000175891
<i>croc</i>	0,000149445	0,000118493	0,000184406
<i>dolc</i>	0,000119911	0,000107686	0,000292450
<i>dorm</i>	7,72848E-05	0,000139118	0,000179467
<i>fran</i>	0,000150340	0,000131195	0,000193590
<i>fras</i>	0,000106485	0,000158784	5,46113E-05
<i>mora</i>	0,000157555	0,000167369	0,000251621
<i>zacc</i>	7,93290E-05	6,53875E-05	0,000169612

Come si può notare dalla tabella (1.8) le differenze tra le varianze ottenute dai dati filtrati e quelle ottenute dai dati originali sono positive e questo significa che la procedura ha causato un minimo aumento del rumore nelle osservazioni piuttosto che una sua riduzione.

Ulteriori considerazioni possono essere fatte analizzando i grafici comparati delle singole coordinate e dei relativi scarti; si riportano qui in seguito, a titolo di esempio, i risultati ottenuti per la stazione di Cassano allo Ionio (cass). Considerazioni del tutto analoghe possono essere fatte per tutte le altre stazioni.

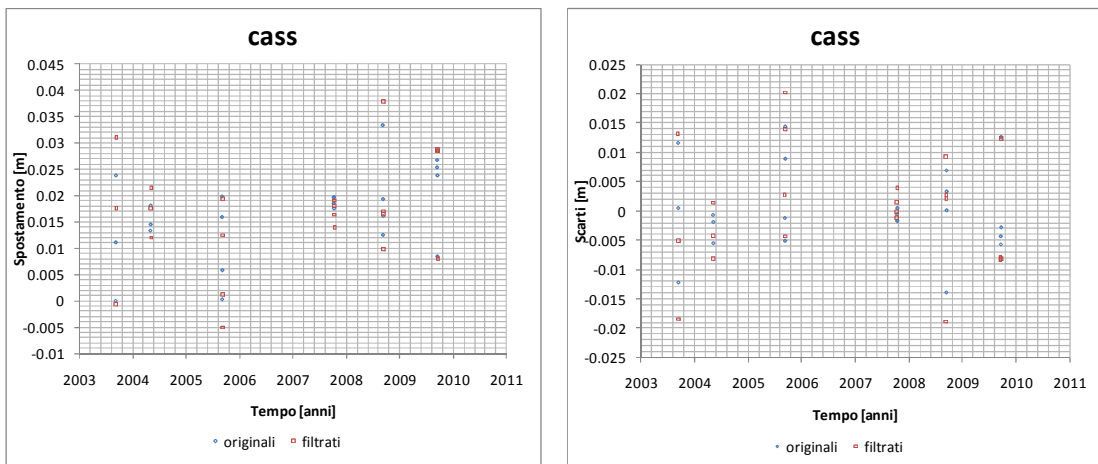


Andamento della coordinata Nord Scarti della coordinata Nord

Andamento della coordinata Est Scarti della coordinata Est

Figura (1.17) Spostamenti e relativi scarti delle componenti planimetriche prima e dopo l'applicazione della procedura COMM nella stazione di Cassano allo Ionio

In questo caso specifico, che rispecchia comunque i dati ottenuti anche per le altre stazioni, l'andamento delle coordinate planimetriche Nord e Est sembra non risentire in maniera efficace dell'operazione di filtraggio; le variabilità dei dati filtrati ad un esame più dettagliato sembrano assolutamente comparabili con quelle ottenute dai dati originali.



Andamento della coordinata h

Scarti della coordinata h

Figura (1.18) Spostamenti e relativi scarti della componente altimetrica prima e dopo l'applicazione della procedura COMM nella stazione di Cassano allo Ionio

Per la coordinata h, sempre relativa alla stazione di Cassano allo Ionio, è evidente la maggiore dispersione del dato sia in termini di coordinate, sia per quanto riguarda gli scarti.

1.8 Procedura di filtraggio COMM modificata

I fenomeni fisici che governano la variabilità del segnale s potrebbero essere effettivamente legati anche a fattori atmosferici e alla condizione della ionosfera nel periodo di misura. Appare ragionevole supporre che durante i quattro giorni di rilievo questi fattori rimangano relativamente costanti e che il segnale s possa essere rappresentato meglio da un errore medio complessivo di campagna $\bar{v}(c)$ calcolato sull'intera campagna di misura c .

Riadattando ulteriormente la formula vista in precedenza si ottiene la relazione

$$\bar{v}(c) = \frac{\sum_{d=1}^D v(d)}{D} \quad \text{dove:} \quad v(d) = \frac{\sum_{s=1}^S v_s(d)}{S} \quad (11)$$

Dove:

- D è il numero totale di giorni di misurazione per la campagna c

Applicando quindi la nuova correzione abbiamo ottenuto una seconda serie di valori di correzione coi quali è stata effettuata una nuova operazione di filtraggio dei dati.

Sulle serie così ricavate sono quindi state ripetute le stime di velocità per ogni stazione e i risultati vengono riportati nelle tabelle in seguito.

I valori ottenuti sono stati ancora una volta confrontati con le stime fatte sui dati originali, sia per quanto riguarda le coordinate, sia per le varianze di stima.

TABELLA (1.9)			
Campagna	Errore medio complessivo di campagna		
	N	E	h
[anni]	[m/anno]	[m/anno]	[m/anno]
2003	-0,000609210	-0,000816395	-0,004345899
2004	0,000405347	-0,000884324	-0,001736632
2005	0,001221542	0,000770639	0,003430902
2007	-0,000689081	-0,001098247	0,001946654
2008	-0,001734176	0,000771289	-6,35861E-06
2009	0,000443564	0,000823064	-0,002036582

TABELLA (1.10)

Stazione	v_N	v_E	v_h
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,019938846	0,023668256	0,001417793
<i>civi</i>	0,019206723	0,026144623	0,000243510
<i>colo</i>	0,019572200	0,023542015	0,001934335
<i>croc</i>	0,018874756	0,023792905	0,002838658
<i>dolc</i>	0,019362047	0,023597937	0,000686895
<i>dorm</i>	0,018907705	0,023756573	0,003722994
<i>fran</i>	0,019351693	0,024649613	-0,002145231
<i>fras</i>	0,019657324	0,022829424	0,003822730
<i>mora</i>	0,018879002	0,023683166	0,001421596
<i>zacc</i>	0,020184275	0,024718508	0,003353775

TABELLA (1.11)

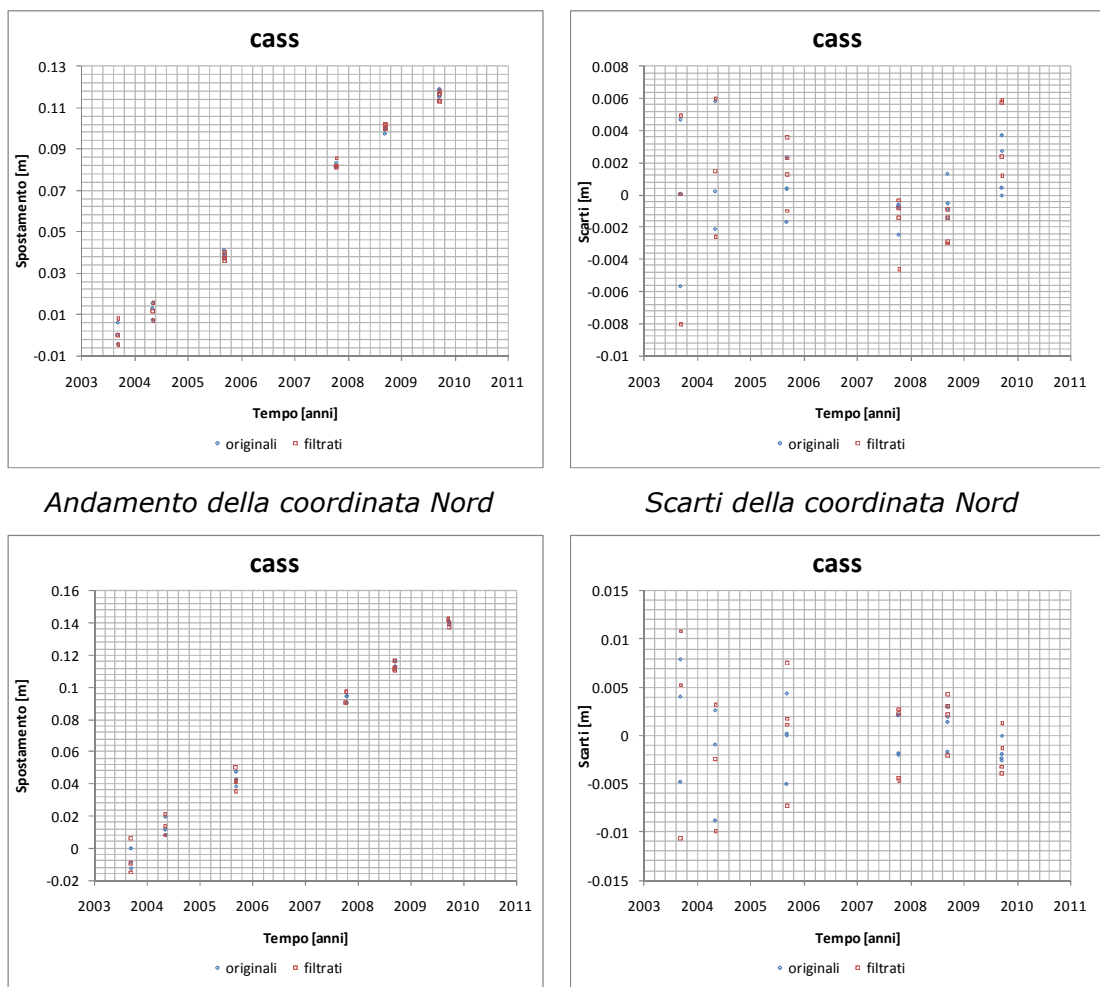
Stazione	dv_N	dv_E	dv_h
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,000202418	-0,000199056	-0,000145882
<i>civi</i>	0,000103352	-0,000220594	0,000156636
<i>colo</i>	0,000199300	-0,000168344	0,000207912
<i>croc</i>	0,000121643	-0,000227385	0,000120213
<i>dolc</i>	0,000166770	-0,000205622	-0,000167238
<i>dorm</i>	0,000226975	-0,000172149	0,000181624
<i>fran</i>	0,000202354	-0,000205671	-0,000143546
<i>fras</i>	0,000163867	-0,000211345	-0,000100657
<i>mora</i>	0,000100066	-0,000242256	6,67176E-05
<i>zacc</i>	0,000211923	-0,000192680	-0,000174176

TABELLA (1.12)			
Stazione	σ_{v_N}	σ_{v_E}	σ_{v_h}
	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	0,000304321	0,000375590	0,000804007
<i>civi</i>	0,000409829	0,000378219	0,000807962
<i>colo</i>	0,000440864	0,000284481	0,000810645
<i>croc</i>	0,000362575	0,000336973	0,000947324
<i>dolc</i>	0,000352485	0,000360350	0,000979623
<i>dorm</i>	0,000364826	0,000442942	0,000853308
<i>fran</i>	0,000356438	0,000475544	0,001335011
<i>fras</i>	0,000321679	0,000429805	0,001653047
<i>mora</i>	0,000395750	0,000361606	0,001594232
<i>zacc</i>	0,000506711	0,000507513	0,001169470

TABELLA (1.13)			
	[m/anno]	[m/anno]	[m/anno]
Stazione	$d\sigma_{v_N}$	$d\sigma_{v_E}$	$d\sigma_{v_h}$
<i>cass</i>	5,05469E-05	1,19654E-05	0,000111635
<i>civi</i>	6,36157E-05	2,86926E-05	0,000108143
<i>colo</i>	9,29459E-05	6,34044E-05	5,60283E-05
<i>croc</i>	7,54831E-05	2,86405E-05	0,000154035
<i>dolc</i>	6,51795E-05	2,61148E-05	0,000170612
<i>dorm</i>	4,97439E-05	5,91256E-05	0,000103652
<i>fran</i>	7,49393E-05	3,00880E-05	0,000131090
<i>fras</i>	4,53910E-05	2,04451E-05	-7,41889E-05
<i>mora</i>	8,65885E-05	5,92969E-05	0,000160949
<i>zacc</i>	3,83605E-05	1,37543E-05	0,000152051

Nella tabella (1.13) si deduce che il secondo filtraggio ha dato ancora un aumento del rumore, ma è importante notare come l'adattamento del metodo, da un errore medio giornaliero ad un errore medio di campagna, abbia ridotto per tutte le stazioni le differenze tra le varianze di stima.

La diminuzione delle dov_i dà forza all'ipotesi introdotta nell'applicazione del metodo modificato, anche se questa non si è rivelata sufficiente a risolvere il problema.



Andamento della coordinata Nord

Scarti della coordinata Nord

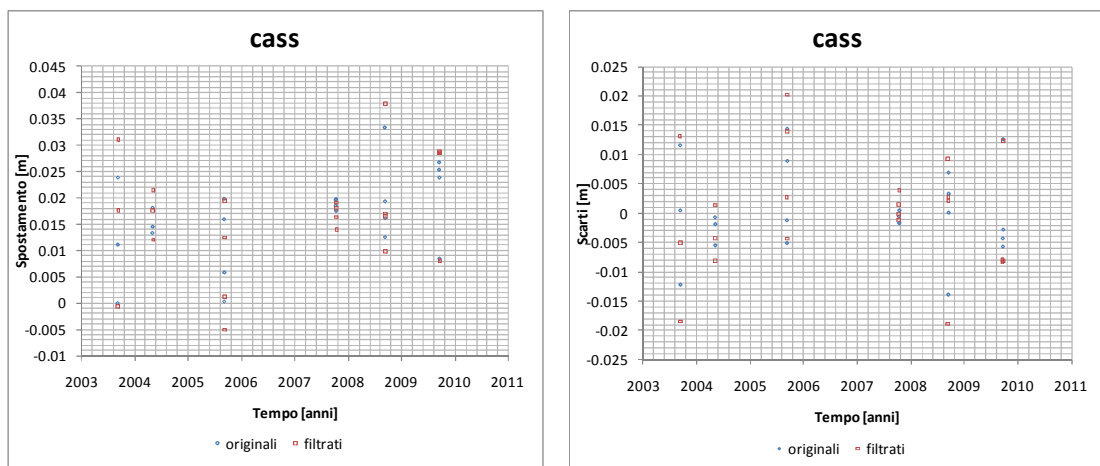
Andamento della coordinata Est

Scarti della coordinata Est

Figura (1.19) Spostamenti e relativi scarti delle componenti planimetriche prima e dopo l'applicazione della procedura COMM modificata nella stazione di Cassano allo Ionio

Anche nell'analisi dei grafici comparati delle singole coordinate e dei rispettivi scarti, dove qui sono riportati sempre i risultati relativi alla stazione di Cassano allo Ionio, è evidente che i punti della serie filtrata risultano meno dispersi rispetto al caso precedente per tutte e tre le coordinate.

Il risultato ottenuto è simile in tutte le altre stazioni come già ci si poteva attendere dalle tabelle riassuntive riportate in precedenza.



Andamento della coordinata h

Scarti della coordinata h

Figura (1.20) Spostamenti e relativi scarti della componente altimetrica prima e dopo l'applicazione della procedura COMM modificata nella stazione di Cassano allo Ionio

1.9 Considerazioni finali sui risultati del filtraggio

Nel nostro caso specifico dobbiamo concludere che la procedura COMM non è efficace nel tentativo di isolare il segnale s e di ridurre ulteriormente il rumore nei dati originali; questo risultato può essere dovuto ad alcune considerazioni generali sul metodo:

1. Il metodo COMM è un metodo piuttosto semplificato che non fa nessuna considerazione sulla propagazione della varianza nell'operazione del filtraggio delle coordinate; a rigore nell'analisi delle serie filtrate si dovrebbe utilizzare una matrice dei pesi più complessa che tenga conto delle correlazioni tra i nuovi dati ottenuti
2. La procedura è pensata per regioni di grande estensione e, in particolar modo, è stata applicata con successo nella California meridionale; l'area occupata dalla rete non-permanente del Pollino è di gran lunga inferiore e questo può certamente influire sull'efficacia del metodo, in quanto intervengono evidentemente alcune variabilità locali che non consentono di individuare in maniera corretta il segnale s

1.10 Acquisizione e analisi dei dati SAR

Nelle analisi dell'evoluzione deformativa sul territorio, oltre ai dati GPS, si ricorre spesso anche ai dati SAR.

Il SAR (Syntetic Aperture Radar) è una delle tecnologie che permettono di acquisire immagini della superficie terrestre grazie all'emissione di onde radar da dispositivi installati su aerei o satelliti artificiali. Si tratta di un sistema di telerilevamento attivo che consente di produrre delle mappe particolari, gli interferogrammi, dove, grazie al confronto tra due immagini ottenute con il medesimo angolo visivo in due istanti di tempo differenti, si possono stimare i cedimenti e le deformazioni superficiali.

La particolarità della tecnologia SAR sta nell'uso di antenne cosiddette "sintetiche", realizzate grazie all'uso di un'antenna reale di piccole dimensioni che trasmette e riceve un fascio di impulsi ad intervalli regolari di tempo.

Dall'unione dei segnali di ritorno, memorizzati durante la scansione, si ottiene un risultato analogo a quello che si avrebbe con un'antenna fisica più grande.

L'emissione di una serie di impulsi però fa sì che ogni soggetto produca un gran numero di riflessioni, a differenza di quanto non si avrebbe con l'emissione di un singolo impulso (tipico dei radar terrestri). Nella successiva ricostruzione

dell'immagine questo deve essere tenuto in considerazione e i segnali vengono quindi sintetizzati nello spazio per ottenere l'immagine finale.

Con la tecnologia SAR vengono diminuiti gli ingombri fisici dell'antenna e si raggiunge un alto livello di risoluzione al suolo, che non dipende dalla distanza degli oggetti dal punto di rilievo.

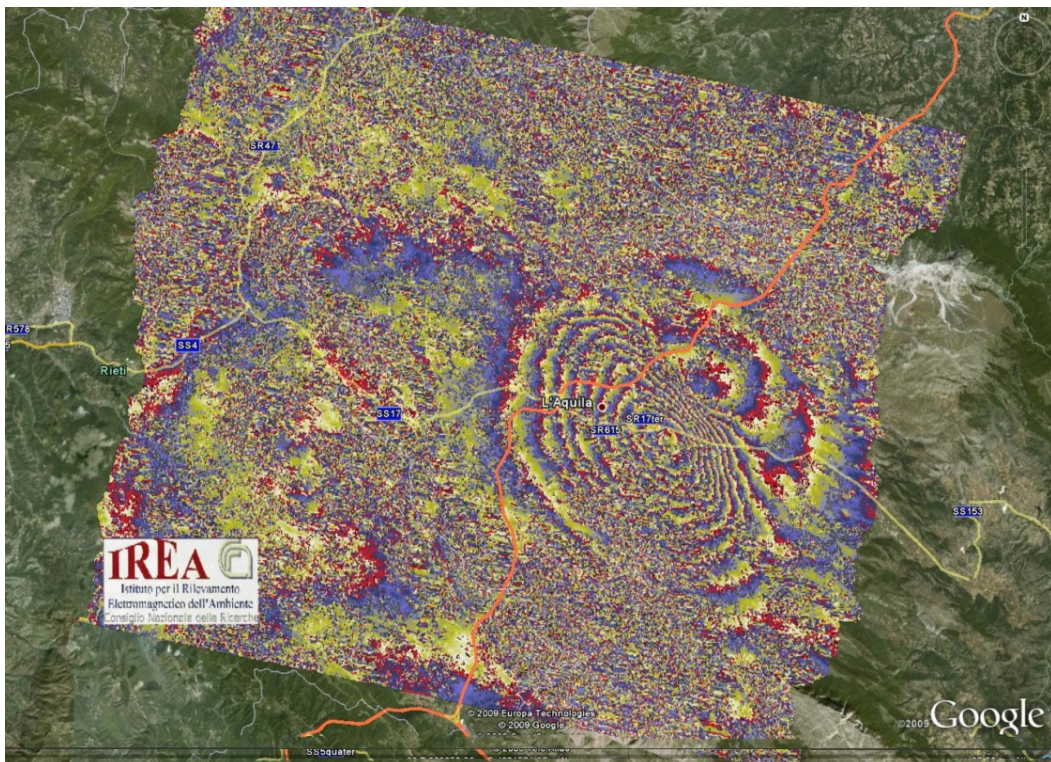


Figura (1.21) Interferogramma ottenuto nella zona dell'Aquila dopo il sisma del 6 aprile 2009 che evidenzia la forte depressione nella zona urbana

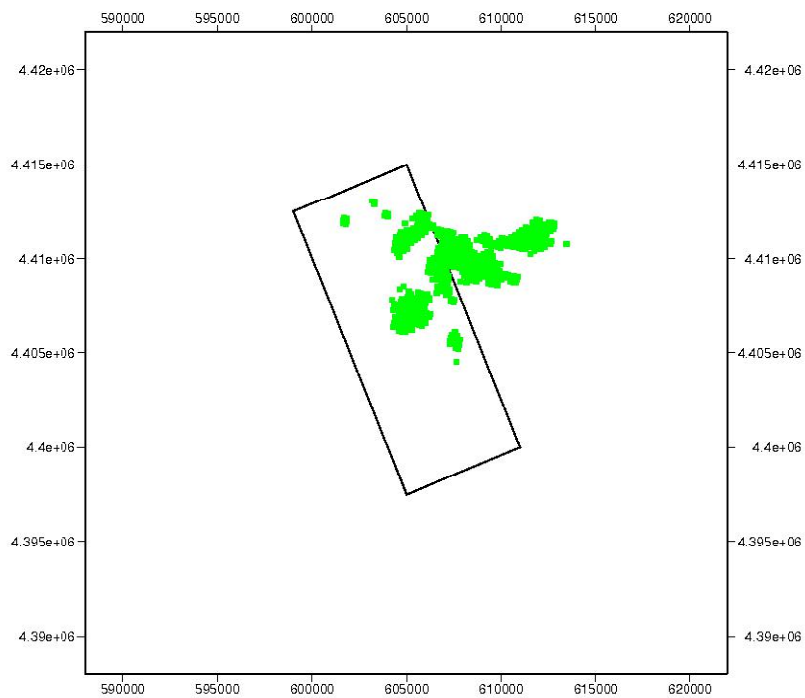
L'antenna sintetica genera infatti una serie di fasci di onde radar molto più stretti di quanto possa fare un'antenna fisica di grandi dimensioni e per questo produce immagini a risoluzioni molto più elevate.

L'uso delle onde radar consente di effettuare i rilievi anche durante la notte o con cielo nuvoloso grazie al fatto che l'acquisizione è totalmente indipendente dalle

condizioni naturali di illuminazione e che l'indice di assorbimento atmosferico alle lunghezze d'onda tipiche del radar è molto basso.

Numerose applicazioni SAR utilizzano l'ampiezza del segnale di ritorno e non considerano i dati di fase, ma, per quanto riguarda il caso dell'interferometria, il dato che viene analizzato è proprio la fase del segnale riflesso. Se consideriamo l'onda emessa dall'antenna, questa sarà a fase nota, mentre la fase dell'onda di ritorno dipenderà dalla lunghezza del percorso effettuato.

Il percorso conterrà un certo numero intero di lunghezze d'onda più uno sfasamento che può essere misurato con grande precisione.



*Figura (1.22) Posizioni dei 661 punti di rilievo SAR
in relazione alla faglia di Castrovillari*

Una volta rimossi alcuni fattori che disturbano il segnale, è possibile quindi generare l'interferogramma operando la differenza di fase tra le due immagini prese in tempi diversi dallo stesso punto di vista.

Per la regione del Pollino sono state create 38 immagini dalle quali si sono ricavati i dati relativi all'evoluzione delle quote in corrispondenza di 661 punti appartenenti alla stessa zona della rete non-permanente GPS.

La loro distribuzione spaziale, in relazione al piano di scorrimento della faglia, è riportata in figura (1.22).

1.11 Dati sintetici per la valutazione di integrabilità tra misure SAR e GPS

Nei paragrafi precedenti abbiamo analizzato due possibili tipologie di misure che possono essere acquisite nell'ambito del monitoraggio geodinamico.

Il nostro obiettivo è una valutazione generale sulla possibilità di integrare informazioni di tipo SAR, legate ai soli spostamenti verticali della superficie terrestre, con misure di tipo GPS passando dall'applicazione di modelli che descrivono il legame tra gli spostamenti del piano di faglia e le deformazioni in superficie.

A questo scopo verranno create alcune serie temporali di misure sintetiche strutturate come quelle reali, che presentino un livello di rumore variabile sia sul dato SAR sia sul dato GPS. Grazie al loro confronto incrociato, si potranno fare anche considerazioni più approfondite al variare del numero di campagne svolte. Tutto questo verrà sviluppato in dettaglio nei prossimi capitoli.

Capitolo 2

Il modello di faglia e i metodi di inversione

2.1 Il modello di Okada

Nella sua formulazione originale, il modello di Okada permette di determinare le evidenze superficiali, in termini di spostamento, velocità e deformazione, dovute ad una dislocazione nota sul piano di faglia in profondità, attraverso la formalizzazione di una serie di relazioni analitiche.

Il modello, presentato per la prima volta nel 1985, ha infatti due obiettivi principali:

1. Fornire una serie di formule che esprimano il legame tra gli spostamenti e gli sforzi ottenuti in profondità e in superficie e ne consentano la stima a partire dai dati specifici della faglia
2. Evitare e gestire attraverso metodi pratici le instabilità numeriche e le singolarità matematiche che sono intrinsecamente presenti nelle relazioni analitiche

Okada sceglie di definire la faglia come il piano di interfaccia lungo il quale avviene l'accumulo di deformazione; questo è compreso tra la placca di sovrascorrimento, definita tetto, e la placca di subduzione, detta piede.

Nell'analisi numerica ci si riferisce proprio al piede che appare tratteggiato nella figura seguente

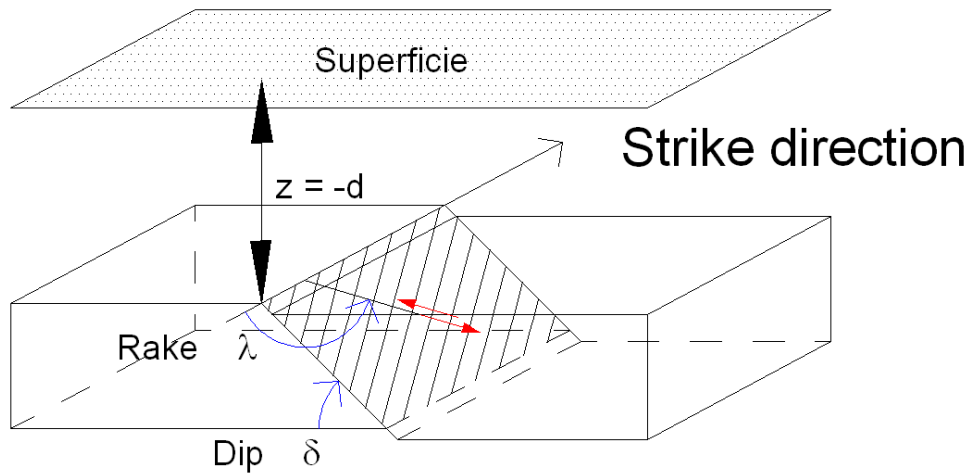


Figura (2.1) Definizione delle due placche, degli angoli di Rake (λ) e Dip (δ) e del punto di misura della profondità nel modello di Okada

Per poter determinare in modo completo la faglia è necessario introdurre alcuni parametri geometrici fondamentali:

- La posizione dell'estremo sinistro del bordo superiore della faglia in termini di coordinate planimetriche Nord ed Est e di profondità rispetto al piano campagna z
- La lunghezza L e la larghezza W della faglia misurati sul suo piano di scorrimento
- L'angolo di strike θ , o angolo di orientamento, che dipende dalla direzione del vettore uscente dal bordo superiore della faglia e orientato secondo il suo sviluppo in lunghezza. L'angolo θ sarà quindi misurato in senso orario tra la direttrice verso Nord e lo strike come mostrato nella figura (2.2). Ovviamente i valori limite dell'angolo di strike saranno compresi nell'intervallo $0^\circ \leq \theta \leq 360^\circ$
- L'angolo di dip δ , o angolo di immersione del piano di faglia, è l'angolo misurato in senso antiorario tra il piano orizzontale e il piano di scorrimento come evidenziato dalla figura (2.1). Per l'angolo di dip i valori limite saranno compresi nell'intervallo $0^\circ \leq \delta \leq 90^\circ$

- L'angolo di rake λ , che definisce la direzione di accumulo delle deformazioni della faglia, viene misurato in senso antiorario sul piano di scorrimento tra la direzione di strike e la direzione lungo la quale la placca di sovrascorrimento scivola rispetto a quella di subduzione. Anche per l'angolo di rake i limiti ammissibili sono di $0^\circ \leq \lambda \leq 360^\circ$

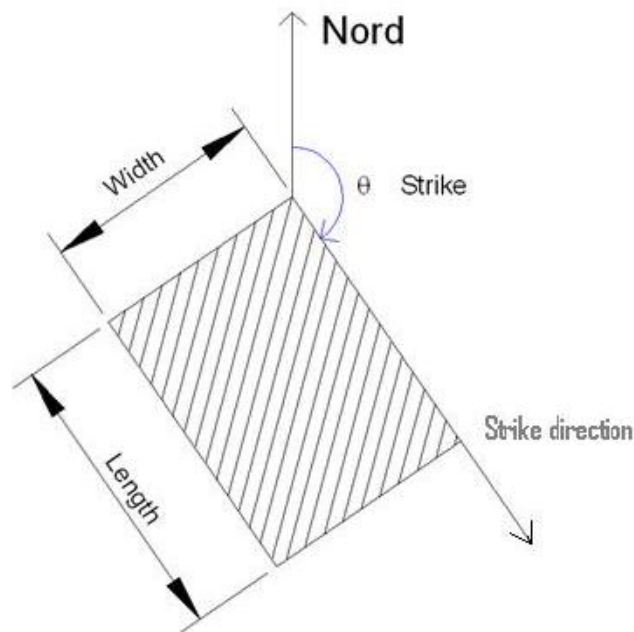


Figura (2.2) Metodo di misura delle dimensioni caratteristiche della faglia W ed L e determinazione dell'angolo θ

E' possibile effettuare in generale una classificazione dei meccanismi di faglia a partire dai valori dell'angolo di rake come mostrato nella tabella seguente

TABELLA (2.1)	
Angolo di rake (λ)	Meccanismo di faglia prevalente
0°	Faglia trascorrente
90°	Faglia inversa
180°	Faglia trascorrente
270°	Faglia normale

Ovviamente in natura ci si trova spesso ad affrontare casi in cui i valori di rake siano compresi tra questi valori limite. Assisteremo in questo caso a dei comportamenti complessi dati dall'unione di più componenti; ad esempio nel caso che il rake sia compreso tra 0° e 90° la faglia sarà di tipo inverso con componente trascorrente.

Per concludere la definizione geometrica, l'ultimo passo è la discretizzazione della faglia grazie alla suddivisione in patch della sua superficie di scorrimento; questo permette di ottenere i risultati numerici e consente di legare le velocità su ogni elemento di faglia agli spostamenti superficiali.

A livello analitico il modello diretto di Okada può essere sinteticamente espresso dalla relazione matriciale (1)

$$d = G \cdot m + \varepsilon \quad (1)$$

Dove, definendo con n il numero di punti di controllo in superficie e con p il numero di patch in cui viene suddivisa la faglia, avremo che:

- $d_{[n \times 1]}$ è il vettore che descrive gli spostamenti superficiali per ogni singolo punto di misura
- $m_{[p \times 1]}$ è un vettore che contiene il modulo dell'ampiezza di dislocazione per ogni singola patch del piano di faglia
- $G_{[n \times p]}$ è la matrice di Green che definisce il legame tra le dislocazioni delle singole patch e gli spostamenti sui punti di misura in termini unitari. Questa matrice è funzione dei parametri geometrici noti della faglia introdotti in precedenza
- $\varepsilon_{[n \times 1]}$ è il vettore degli errori che si ipotizza possano essere distribuiti come una normale a media nulla e con matrice di covarianza pari a Σ_d . Utilizzando una notazione contratta avremo quindi

$$\varepsilon \sim N(0, \Sigma_d) \quad (2)$$

Nelle analisi svolte per il nostro studio si è reso necessario anche l'uso del metodo indiretto, che permette di stimare il vettore delle velocità sulla faglia $m_{[p \times 1]}$ a partire da un vettore di osservazioni superficiali note $d_{[n \times 1]}$.

Questa operazione impone di introdurre un opportuno metodo di inversione da applicare alle relazioni della teoria originale di Okada, come vedremo in dettaglio nei prossimi paragrafi.

2.2 Parametri e caratteristiche della faglia di Castrovillari

Come riportato nella più recente pubblicazione del Database of Individual Seismogenetic Sources (DISS Working Group 2007), la faglia di Castrovillari, situata nella regione del Pollino, presenta un meccanismo puramente normale.

I suoi parametri principali sono riassunti nella seguente tabella

Coordinata Est	E	605000 [m]
Coordinata Nord	N	441500 [m]
Profondità	Z	1000 [m]
Lunghezza	L	16 [km]
Larghezza	W	12 [km]
Angolo di strike	θ	157°
Angolo di dip	δ	60°
Angolo di rake	λ	270°

Il valore dell'angolo di immersione pari a 60° è caratteristico di tutte le faglie presenti nella regione meridionale degli Appennini, a partire dalla faglia di Melandro-Pergola, situata più a Nord, fino ad arrivare alla faglia di Castiglione-Agri, a Sud. L'unica eccezione è rappresentata dalla faglia di Luzzi che ha un valore di δ di poco superiore, pari a 65°.

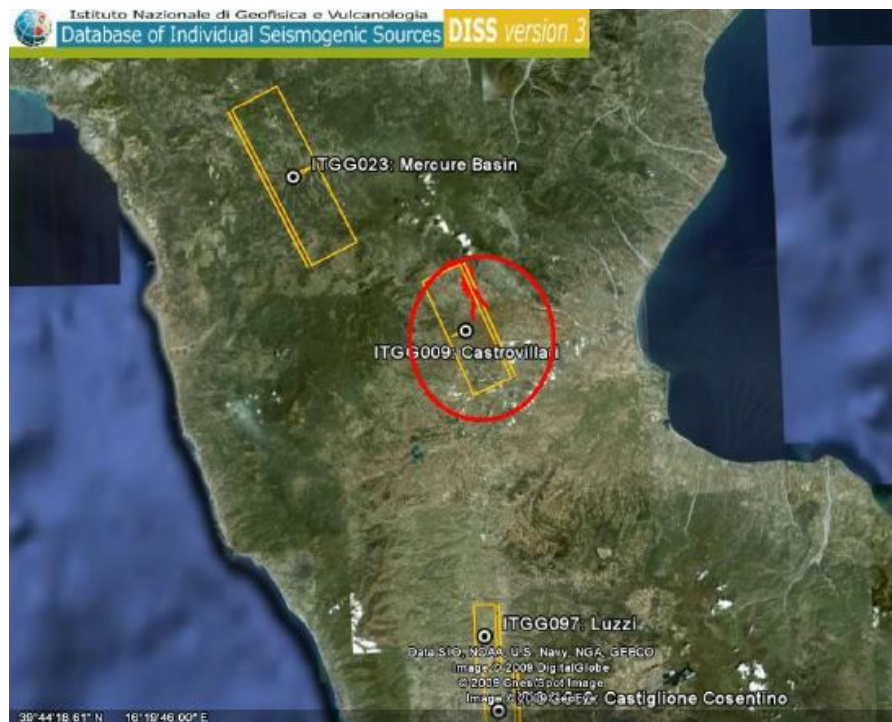


Figura (2.3) La faglia di Castrovillari inquadrata nel suo contesto geografico

Considerare un meccanismo puramente normale è nel nostro caso una semplificazione, poiché la distribuzione del rake appare certamente più complessa nella realtà. Tuttavia, dato che non si dispone di modelli in letteratura che possano descrivere un valore variabile di rake e che il comportamento della faglia è a dominanza estensionale, l'ipotesi di adottare un meccanismo puramente normale nelle analisi successive è più che ragionevole.

Per la discretizzazione del piano di scorrimento si è scelto di suddividere l'area della faglia in 12 patch quadrate di 4 km di lato, 4 in lunghezza e 3 in larghezza, come mostrato nella figura (2.4).

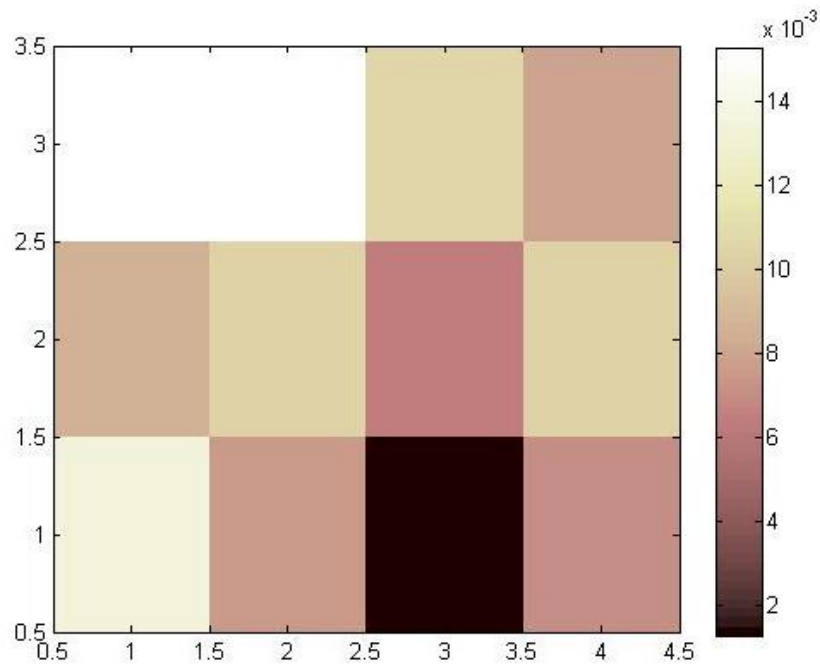


Figura (2.4) La faglia discretizzata con una configurazione sintetica delle velocità calcolata per ogni singola patch

Nella scelta dei punti di controllo in superficie sono state considerate le due serie di misurazioni effettuate nella regione del Pollino:

1. I 661 punti sparsi del rilievo SAR, utilizzati anche come base per realizzare l'inversione del modello
2. I 10 punti corrispondenti alle stazioni non-permanenti della rete, introdotta nel capitolo 1, scelti come punti di riferimento per le successive operazioni di confronto tra i dati di partenza e i dati ottenuti dall'applicazione del modello diretto

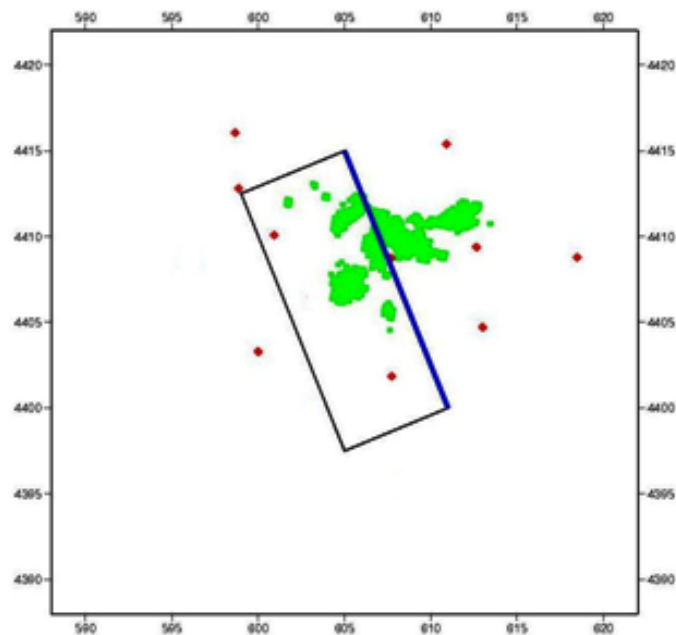


Figura (2.5) Visione in pianta dei punti di controllo superficiali (in verde i punti del rilievo SAR, in rosso le stazioni non-permanenti della rete) con la posizione relativa della faglia così come mostrata nel suo contesto in figura (2.3)

Una volta determinati i parametri del modello, definite le posizioni dei punti di controllo in superficie ed impostato il numero di patch nella discretizzazione della faglia, si hanno tutte le informazioni necessarie per costruire le matrici di Green di modello che verranno utilizzate nelle successive analisi numeriche.

Il modello di Okada permette di ricavare tre differenti matrici per le tre componenti spaziali delle coordinate Nord, Est e altezza z .

Sono quindi state calcolate:

- G_N per la componente Nord e G_E per la componente Est relative agli spostamenti sul piano orizzontale
- G_z per gli spostamenti lungo l'asse z delle quote, e quindi legata alle osservazioni SAR

Le matrici così ottenute potranno essere utilizzate in tutte le applicazioni numeriche del modello affrontate in seguito, sia nel caso diretto, sia nel caso inverso. La matrice G ha infatti la proprietà di condensare le informazioni geometriche ed i parametri relativi alla dinamica della faglia e non dipende in alcun modo dall'entità degli spostamenti misurati.

Ammettendo di non cambiare i parametri della faglia introdotti in precedenza, si potranno quindi utilizzare in ogni calcolo sempre le stesse matrici G_N , G_E e G_Z .

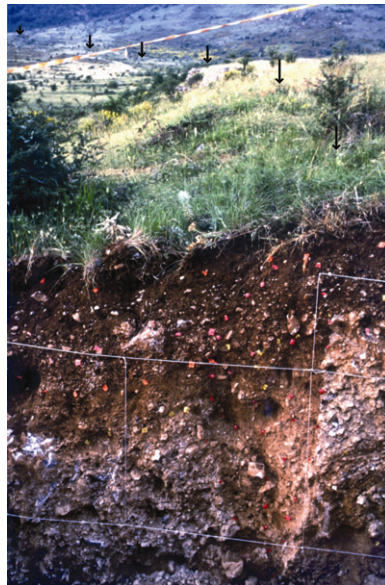


Figura (2.6) Dettaglio della faglia di Castrovillari che si localizza al di sotto della scarpata sul terreno (le frecce evidenziano la porzione di terreno sollevata) in prossimità della catena del Monte Pollino

2.3 Il problema dell'inversione

L'inversione di un modello, come quello di Okada, può generare delle instabilità numeriche, specialmente se si vuole stimare la distribuzione delle velocità con un elevato grado di dettaglio e di conseguenza si è costretti a discretizzare il piano di scorrimento con un grande numero di patch.

Nell'analisi è inoltre importante trovare una soluzione in grado allo stesso tempo di minimizzare l'errore di predizione in funzione dei dati iniziali e di descrivere con regolarità le velocità sul piano di faglia. Un risultato coerente dal punto di vista fisico dovrà quindi prevedere che la velocità ottenuta sulla singola patch non si discosti eccessivamente dalle velocità delle patch adiacenti.

I minimi quadrati sono senza dubbio il più semplice tra i metodi di risoluzione del problema lineare inverso, esprimibile nel nostro caso dalla relazione

$$G \cdot m = d \quad (3)$$

Lo scopo della procedura è quello di minimizzare correttamente, in funzione delle misure di partenza, l'errore di predizione o scarto quadratico medio, definito come

$$E = e^t e = \sum_{i=1}^N e_i^2 \quad \text{con:} \quad e_i = d_i^{\text{calc}} - d_i^{\text{oss}} \quad (4)$$

Dove:

- e_i rappresenta l'errore di predizione
- d^{calc} è il vettore dei dati predetti
- d^{oss} è il vettore dei dati osservati

Questa procedura è del tutto analoga a quella già affrontata nel capitolo 1 per l'approssimazione dei dati osservati mediante una retta del tipo $y = mx + q$ nella stima delle velocità per le stazioni non-permanenti della rete del Pollino.

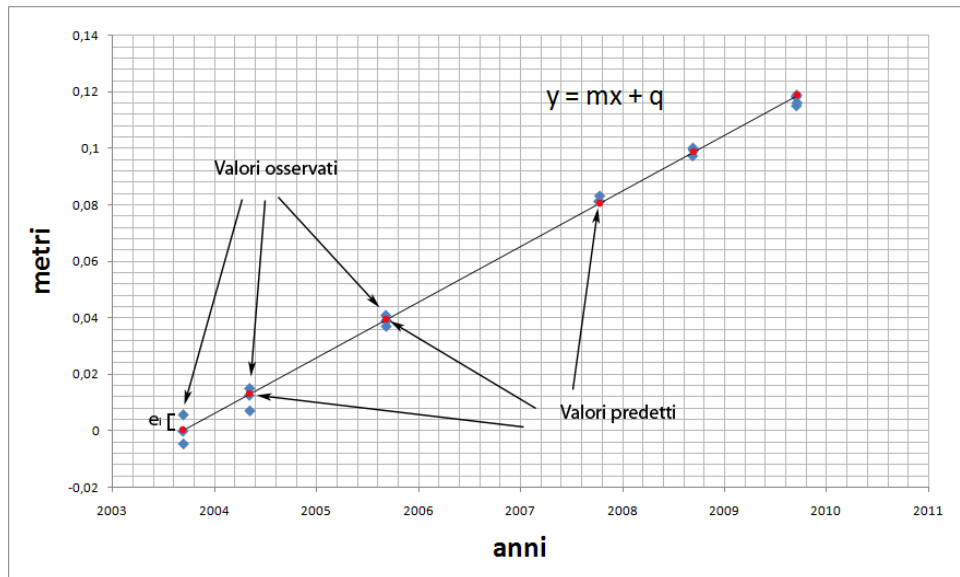


Figura (2.7) Esempio di approssimazione per la serie di dati misurati nella stazione di Cassano allo Ionio relativi alla coordinata Nord; vengono evidenziati in blu i valori osservati ed in rosso i valori predetti dal modello. L'errore e_i è mostrato in figura per la prima osservazione

Come si può notare dalla prima parte dell'espressione (4), l'errore E rappresenta esattamente il quadrato della lunghezza euclidea del vettore e .

Estendendo il metodo al caso generale del problema inverso, si può dimostrare che la soluzione ai minimi quadrati, ammettendo che la matrice G^tG sia invertibile, sarà data da

$$m^{est} = [G^tG]^{-1}G^td \quad (5)$$

Come sovente accade per il caso di problemi inversi, la soluzione risulta essere instabile. Ciò è inerente al problema e non può essere risolto con un ulteriore aumento dei dati. Si può ricorrere però a metodi di stabilizzazione della soluzione imponendo opportuni vincoli di regolarizzazione come verrà descritto nel prossimo paragrafo.

2.4 Minimi quadrati regolarizzati

Il metodo dei minimi quadrati regolarizzati è un valido strumento per ottenere l'inversione del modello di Okada e per arrivare alla stima corretta delle velocità sul piano di scorrimento a partire da misurazioni in superficie.

La regolarizzazione consiste nell'introduzione di un'informazione a priori che permette di inserire nuovi elementi nell'espressione generale del modello, e consente quindi di risolvere le instabilità numeriche del processo di inversione. La scelta del tipo di informazione da introdurre dipende strettamente dal suo ruolo nella stima dei parametri; nel caso si voglia ottenere un semplice esempio di soluzione del problema, la scelta non è determinante, mentre se si vogliono considerare argomenti dipendenti dall'unicità delle stime, le assunzioni a priori saranno di fondamentale importanza.

In generale la "semplicità" della soluzione del problema inverso può essere quantificata con una misura della lunghezza della soluzione; nel caso più semplice si considera la lunghezza euclidea del vettore soluzione m , data da

$$L = m^t m \quad (6)$$

In molti casi è però opportuno ritenere che la soluzione migliore non sia quella che si avvicina di più allo zero, ma quella che tende a dei valori dati a priori; generalizzando la (6) otterremo

$$L = (m - \langle m \rangle)^t (m - \langle m \rangle) \quad (7)$$

Dove:

- $\langle m \rangle$ è il valore a priori dei parametri del modello

Quando i parametri di modello descrivono una funzione continua discretizzata, come nel nostro caso, il concetto di semplicità della soluzione non è legato propriamente al concetto di lunghezza, ma piuttosto a quello di regolarità.

La fisica del problema impone quindi che la scelta relativa all'informazione a priori per il modello di Okada ricada sulla gradualità delle variazioni di velocità sul piano di faglia, e quindi sul fatto che la soluzione debba essere sufficientemente liscia. Occorre pertanto introdurre un'ulteriore generalizzazione che coinvolga le differenze tra i parametri su patch fisicamente adiacenti, approssimandone l'andamento alle derivate. Introduciamo così l'omogeneità l di un vettore m , data da

$$l = \begin{bmatrix} -1 & 1 & & & \\ & 0 & \ddots & & \\ & 0 & & \ddots & \\ & & & & -1 & 1 \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_M \end{bmatrix} = Dm \quad (8)$$

Dove:

- D è la matrice dei gradienti

E' possibile quindi ricavare la nuova espressione di L in seguito all'introduzione della (8) ottenendo

$$L = l^t l = [Dm]^t [Dm] = m^t D^t D m = m^t W_m m \quad (9)$$

Dove:

- $W_m = D^t D$ è una matrice dei pesi che entra nel calcolo della lunghezza del vettore m .

E' doveroso precisare che la norma pesata di m non è una vera e propria norma poiché non rispetta la condizione di stretta positività; si annulla infatti anche per vettori costanti non nulli.

In conclusione la misura di regolarità della soluzione può essere ulteriormente generalizzata dalla relazione

$$L = (m - \langle m \rangle)^t W_m (m - \langle m \rangle) \quad (10)$$

E analogamente per l'errore di predizione si avrà

$$E = e^t W_e e \quad (11)$$

Dove:

- la matrice W_e è una matrice che esprime il contributo relativo di ogni singolo errore all'errore di predizione totale.

A differenza del metodo classico, nei minimi quadrati regolarizzati la soluzione dell'equazione generale (3) viene determinata minimizzando non il solo errore di predizione E , ma la sua somma con il prodotto della misura di regolarità della soluzione (10) appena introdotta con un parametro β , detto parametro di regolarizzazione; l'espressione generale da minimizzare diventa quindi:

$$E + \beta L \quad (12)$$

Il problema regolarizzato si riduce quindi alla ricerca del minimo di un funzionale $\Phi(m)$ definito dalla seguente relazione

$$\Phi(m) = E + \beta L = e^t W_e e + \beta (m - \langle m \rangle)^t W_m (m - \langle m \rangle) \quad (13)$$

Dove il misfit è espresso da

$$e = G \cdot m_{calc} - G \cdot m_{oss} \quad (14)$$

Si dimostra che l'operazione di minimizzazione del funzionale $\Phi(m)$ porta alla determinazione di uno stimatore espresso dalla seguente espressione

$$m^{est} = \langle m \rangle + [G^t W_e G + \beta W_m]^{-1} G^t W_e (d - G \langle m \rangle) \quad (15)$$

Dove si ricorda che:

- $\langle m \rangle$ è la soluzione di guess o soluzione approssimata dei parametri
- W_e è la matrice dei pesi delle osservazioni
- W_m è la matrice dei gradienti
- β è il fattore che indica il peso dell'equazione di regolarità della soluzione

La formulazione introdotta finora impone pertanto di introdurre una stima appropriata del parametro β .

In primissima approssimazione si sarebbe potuto considerare un valore fisso di β sulla base di ipotesi geofisiche, ma ai fini dell'analisi questa scelta non è risultata soddisfacente. È stato quindi calcolato un indice statistico appropriato per poter scegliere un valore ottimale di β .

La procedura per la determinazione del valore ottimale prevede che, al variare del parametro β , venga calcolata volta per volta la soluzione ai minimi quadrati regolarizzati e successivamente il valore del χ^2 dato dalla formula seguente

$$\chi^2 = \frac{\hat{\sigma}_0^2}{\sigma_0^2} \quad (16)$$

Il numeratore dell'espressione, $\hat{\sigma}_0^2$, rappresenta la varianza stimata con la procedura dei minimi quadrati, data da

$$\hat{\sigma}_0^2 = \frac{v^t C_{dd} v}{n - m} \quad \text{con:} \quad v = G \cdot \hat{m} - d \quad (17)$$

Dove:

- n è il numero delle osservazioni
- m è il numero delle incognite
- v è il vettore degli scarti

Il denominatore dell'espressione di χ^2 rappresenta invece la varianza a priori delle misure di partenza legata alla tolleranza strumentale.

Tutti i valori di χ^2 ottenuti dalla relazione (16) sono stati quindi confrontati con i valori teorici corrispondenti ad un livello di significatività del 5%.

Nelle successive considerazioni è stato necessario introdurre un parametro indicatore di disomogeneità; per comodità si è scelto di considerare $1/\beta$ come indicatore di riferimento

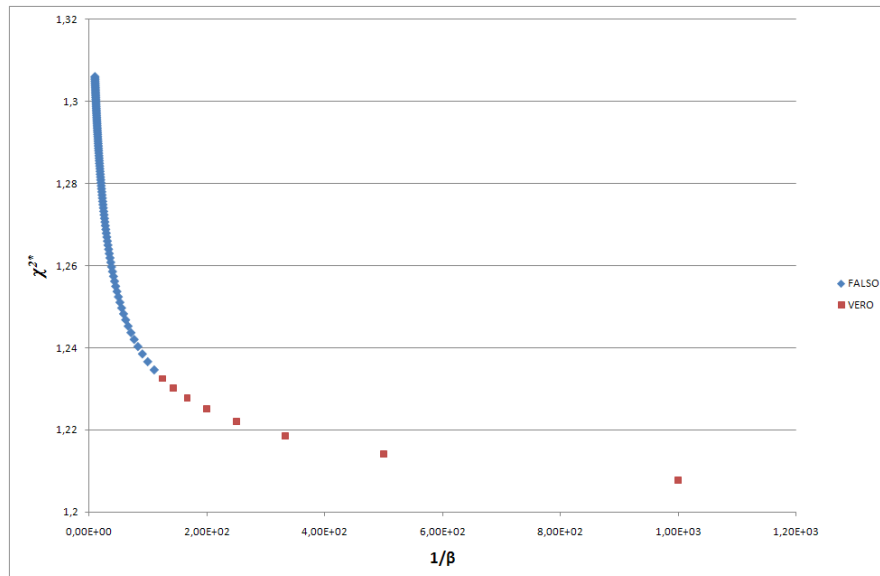


Figura (2.8) La curva di trade-off: χ^2 in funzione di $1/\beta$

La regolarità della soluzione dipende direttamente dal valore di $1/\beta$; per valori prossimi allo zero del parametro indicatore di disomogeneità si ottengono soluzioni molto regolarizzate, mentre valori elevati di $1/\beta$ consentono una maggiore libertà di variazione ai rapporti incrementali della soluzione conducendo a soluzioni più disomogenee.

Per la modalità con la quale è stato introdotto, il χ^2 può essere anche considerato come un indice di adattamento tra le misure e le velocità calcolate in superficie; la scelta di β deve quindi soddisfare due esigenze fondamentali:

- Garantire un buon adattamento in superficie, e quindi portare a valori di χ^2 relativamente bassi
- Contenere la disomogeneità della soluzione sul piano di faglia limitando i gradienti di velocità tra le singole patch

In definitiva la determinazione del valore ottimale di β è stata effettuata ripetendo il calcolo dei minimi quadrati variando con passo fissato il valore di β in un intervallo compreso tra 0,001 e 0,1. Confrontando di volta in volta i valori ottenuti per l'indicatore χ^2 con quelli tabulati corrispondenti ad una chi-quadrato con $(n - m)$ gradi di libertà si sono determinati i valori per i quali il test dà esito positivo con un livello di significatività α pari al 5% (evidenziati in rosso nella figura (2.8)).

Si può inoltre verificare, tramite il test di Fisher, che per tutti i valori di $1/\beta$ superiori al primo che entra in significatività, i modelli che si ottengono sono statisticamente equivalenti. Questo significa che, pur aumentando la disomogeneità delle velocità sul piano di scorrimento, non si ottengono miglioramenti significativi dell'adattamento in superficie.

La scelta del valore ottimale ricade quindi sul primo valore ottenuto che porta ad adottare un \hat{m} il più regolare possibile compatibilmente con il livello di significatività del test assegnato al 5%.

Capitolo 3

Confronto incrociato dei risultati e test statistici

3.1 Il confronto incrociato

L'esecuzione del confronto nella nostra analisi ha come obiettivo principale la valutazione dell'effettiva integrabilità tra i dati verticali e i dati planimetrici nel contesto del modello di faglia e della sua modalità di inversione.

Il confronto deve essere quindi di tipo incrociato e nella sua realizzazione deve essere tenuta in conto la diversa tipologia delle osservazioni; in particolare avremo:

- I dati di tipo SAR, che descrivono la componente verticale della deformazione superficiale
- I dati GPS relativi agli spostamenti orizzontali dei punti di stazione. Le serie GPS danno ovviamente anche la componente verticale z che però viene trascurata vista la sua intrinseca rumorosità

Le due serie di dati sono chiaramente indipendenti tra loro perché vengono generalmente acquisite in tempi diversi, con strumenti e tecnologie diverse e su punti di misura diversi ed è quindi corretto considerarle come statisticamente non correlate.

Il confronto è stato quindi applicato ad un set di dati simulati partendo da un pattern di velocità sul piano di faglia generato artificialmente.

3.2 Schema logico e passi della simulazione

Il primo passo nell'affrontare il confronto su dati sintetici è la generazione delle serie di osservazioni superficiali di partenza.

Viene introdotto un pattern di velocità sulle patch del piano di faglia generato artificialmente che chiameremo m_{synt} ; a partire da questo vettore si possono calcolare, grazie all'applicazione diretta del metodo di Okada, le serie temporali espresse dalle seguenti relazioni:

$$d_{GPS(N)} = G_N \cdot m_{synt} \quad s_{GPS(N)} = d_{GPS(N)} \cdot \Delta t + \varepsilon_{GPS} \quad (1)$$

$$d_{GPS(E)} = G_E \cdot m_{synt} \quad s_{GPS(E)} = d_{GPS(E)} \cdot \Delta t + \varepsilon_{GPS} \quad (2)$$

$$d_{SAR(z)} = G_z \cdot m_{synt} + \varepsilon_{SAR} \quad (3)$$

Dove:

- G_N , G_E e G_z sono le matrici di Green già ottenute secondo la formulazione di Okada nel capitolo 2
- ε_{SAR} rappresenta il rumore artificiale, che viene aggiunto direttamente in termini di velocità alla sola stima del vettore delle velocità verticali $d_{SAR(z)}$
- ε_{GPS} è il rumore artificiale applicato alle serie di spostamenti GPS $s_{GPS(N)}$ ed $s_{GPS(E)}$ ottenute dalle relazioni (1) e (2) a partire dai vettori delle velocità planimetriche $d_{GPS(N)}$ e $d_{GPS(E)}$

Il modello con il quale vengono generati i dati superficiali è considerato noto e questo permetterà, come vedremo in seguito, di valutare il comportamento dei successivi test di confronto al variare dei rumori applicati.

Le serie $s_{GPS(N)}$, $s_{GPS(E)}$ e $d_{SAR(z)}$ appena determinate saranno quindi considerate al pari di un set di dati osservati reali. A partire dai dati di deformazione verticale in superficie $d_{SAR(z)}$ verrà quindi applicato il modello inverso di Okada, grazie ai minimi quadrati regolarizzati trattati in dettaglio nel capitolo 2, per ricavare un nuovo vettore degli spostamenti sul piano di faglia m_{calc} .

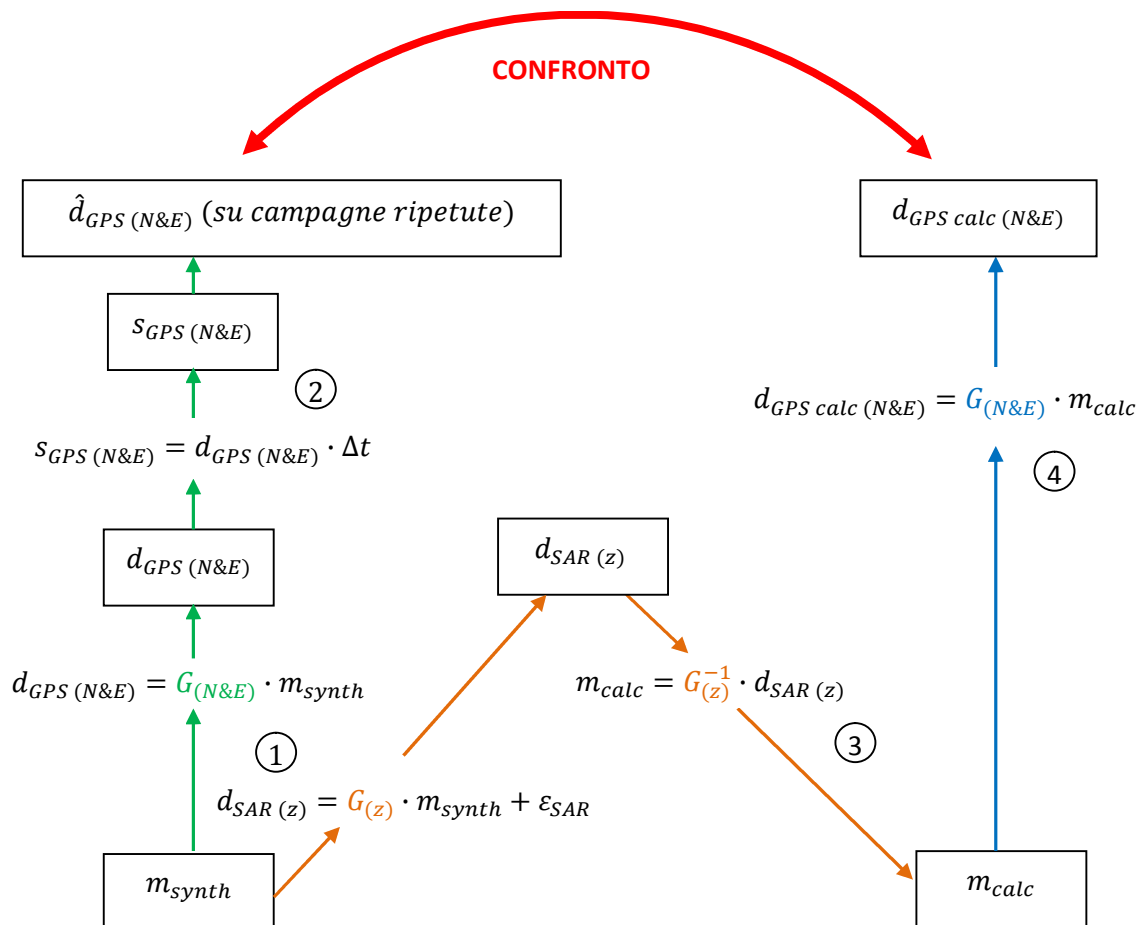
Da questo secondo vettore saranno calcolate, applicando ancora una volta il modello diretto, due nuove serie di deformazione sul piano espresse dalle relazioni:

$$d_{GPS\ calc\ (N)} = G_N \cdot m_{calc} \quad (4)$$

$$d_{GPS\ calc\ (E)} = G_E \cdot m_{calc} \quad (5)$$

I confronti verranno quindi eseguiti tra le osservazioni planimetriche appena determinate e le diverse velocità stimate $\hat{d}_{GPS\ (N\&E)}$ dalle serie di spostamenti ottenute con le formule (1) e (2) all'aumentare del numero di campagne considerate.

Nello schema seguente è riassunta la procedura appena descritta



3.3 Generazione della matrice di varianza e covarianza nel caso del confronto incrociato

La generazione di una matrice di covarianza relativa al modello in superficie è fondamentale in alcuni dei test statistici che verranno utilizzati in seguito per eseguire i confronti tra le serie di dati in superficie.

Il calcolo delle velocità sul piano di faglia è stato eseguito utilizzando dati di deformazione puramente verticale; si è infatti partiti da dati SAR per ricavare m_{calc} . Nel calcolo delle deformazioni superficiali ci si è invece serviti delle matrici di Green per la proiezione delle componenti orizzontali; sono state usate quindi matrici differenti che possono essere considerate indipendenti tra loro.

L'inversione del modello di Okada è stata effettuata con il metodo dei minimi quadrati regolarizzati; questo fa sì che si abbia a disposizione una struttura di covarianza delle velocità sul piano di faglia data dalla matrice $C_{\hat{m}\hat{m}}$.

Se consideriamo ancora una volta il legame funzionale dato dall'espressione

$$d = G \cdot m \quad (6)$$

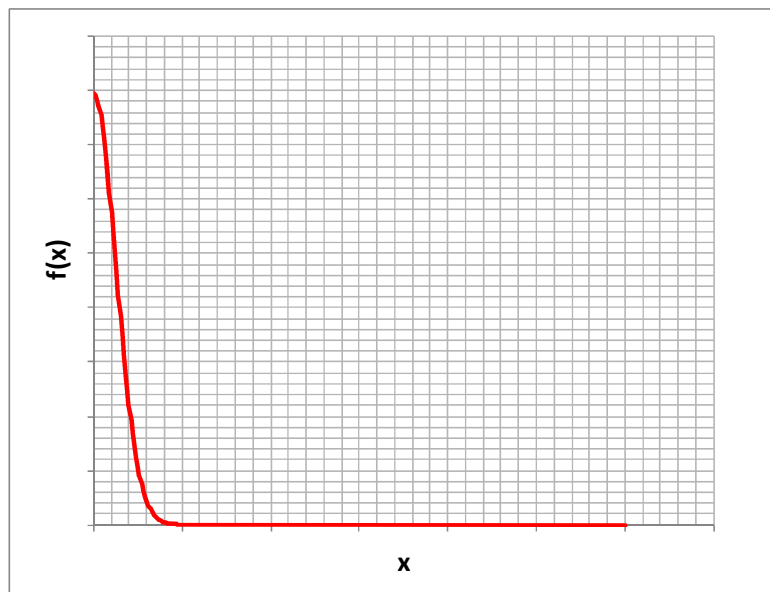
ipotizzando che la matrice di Green sia assimilabile ad una costante, è quindi possibile determinare analiticamente in forma chiusa la matrice di covarianza del modello per i dati in superficie grazie all'applicazione della legge di propagazione della covarianza. Si otterrà quindi l'espressione seguente

$$C_{\bar{a}\bar{a}} = G \cdot C_{\hat{m}\hat{m}} \cdot G^t \quad (7)$$

Nonostante l'applicazione della relazione (7) per il calcolo della matrice $C_{\bar{a}\bar{a}}$ sia perfettamente lecito dal punto di vista analitico, le stime che si ottengono sono affette da un rumore non trascurabile.

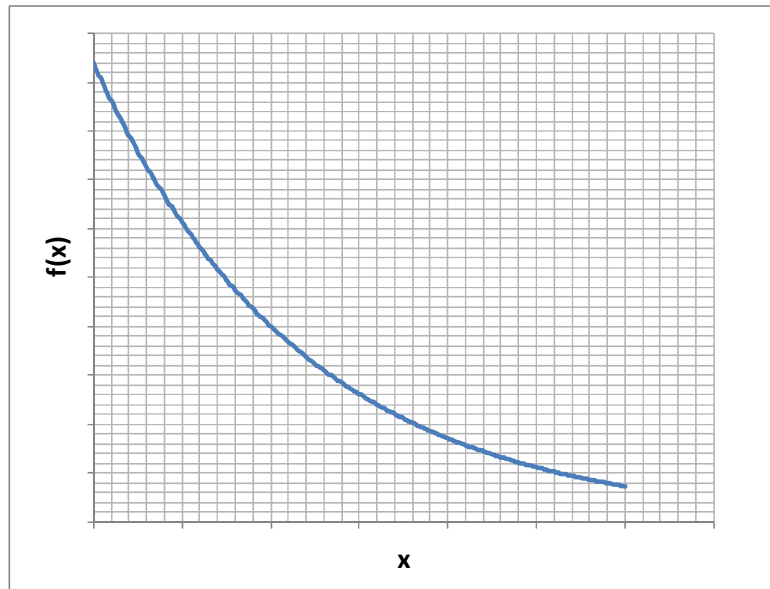
Si è scelto quindi di calcolare la covarianza empirica, ricavata dalla media dei valori della matrice di covarianza $C_{\bar{a}\bar{a}}$ ad una determinata distanza p dai punti di stazione.

Al fine di garantire le caratteristiche di definita positività e di dominanza diagonale proprie delle matrici di covarianza è stata effettuata l'interpolazione dei valori di covarianza propagata utilizzando funzioni modello di covarianza di tipo esponenziale, di tipo normale o di Bessel. Si riportano nelle tabelle seguenti gli andamenti delle funzioni modello appena citate a titolo di esempio



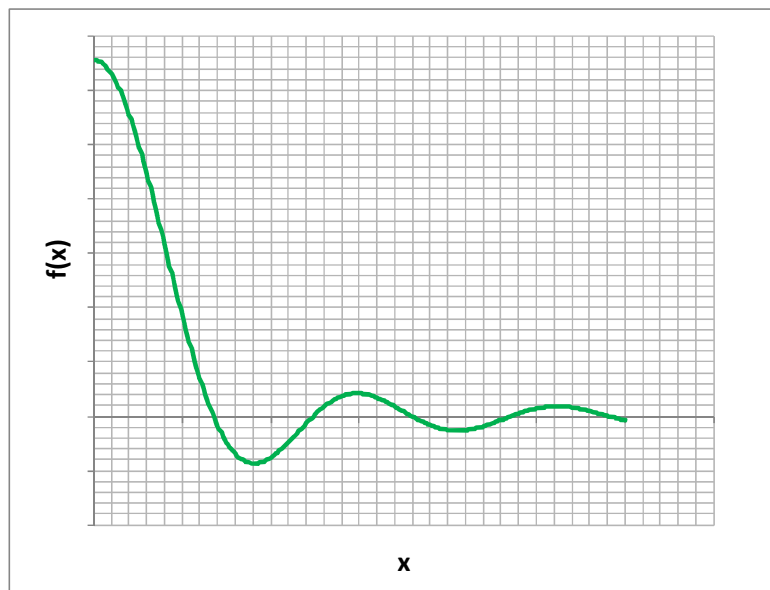
$$y = A \cdot e^{-Bx^2}$$

Figura (3.1) Funzione modello normale



$$y = A \cdot e^{-B|x|}$$

Figura (3.2) Funzione modello esponenziale



$$y = A \cdot \frac{J_1(Bx)}{Bx}$$

Figura (3.3) Funzione modello standard di Bessel

La scelta tra le possibili tipologie di funzione è stata affrontata seguendo il criterio AIC (Akaike Information Criteria).

Quando si definisce un modello si deve accettare che non esistono modelli veri, ma solo modelli che approssimano la realtà. L'obiettivo è quindi di scegliere tra tutte le soluzioni possibili quella più vicina al caso reale che stiamo descrivendo in funzione dei dati in nostro possesso.

Questo coincide con la minimizzazione della perdita di informazione contenuta nei dati e con questo scopo Kullback e Leibler (1951) definirono una distanza capace di rappresentare l'informazione persa in seguito all'approssimazione della realtà tramite modelli.

Akaike propose successivamente di utilizzare questa distanza nella scelta del modello. Il criterio si basa infatti sul legame tra la massima verosimiglianza, concetto che viene spesso utilizzato in molte applicazioni statistiche, e la distanza di Kullback e Leibler introdotta in precedenza ed è espresso dalla seguente relazione

$$AIC = n \cdot \log(\hat{\sigma}^2) + 2K \quad (8)$$

Dove:

- $\hat{\sigma}^2$ è la varianza calcolata con i minimi quadrati
- n è la dimensione del campione
- K è il numero dei parametri stimati del modello

La scelta del modello migliore viene effettuata minimizzando il valore di AIC e ricade quindi sul tipo di funzione che fornisce l'AIC più basso nel nostro caso.

3.4 I test statistici

I confronti tra le diverse serie di dati ottenute nelle analisi fin qui svolte possono essere eseguiti in maniera rigorosa attraverso alcuni test statistici.

In particolare nel nostro caso sono stati scelti i tre test seguenti:

- Test classico
- Test classico con matrice di covarianza modello
- Test Bayesiano

Lo scopo dei test è quello di verificare che la media di una determinata variabile casuale e le medie di alcune variabili statistiche date, supposte come campioni rappresentativi estratti da una popolazione costituita dalla stessa variabile casuale, siano sufficientemente vicine, ovvero che il valore osservato del test sia contenuto entro un intervallo di significatività di livello α dato.

Grazie agli stessi test è possibile inoltre verificare l'uguaglianza tra le differenze delle medie che, essendo ancora medie di campioni estratti da un'unica popolazione, costituiscono una nuova variabile casuale avente questa volta una media unica, nulla in caso di uguaglianza. Nel secondo caso avremo un aumento di un grado di libertà della varianza spiegata ed una diminuzione di un grado di libertà della varianza residua.

3.4.1 Il test classico

Il primo test utilizzato per eseguire i confronti parte dai dati di deformazione superficiale considerandoli come valori osservati distribuiti normalmente come riportato nella seguente notazione

$$\hat{x} \sim N[\bar{x}, C_{\hat{x}\hat{x}}] \quad (9)$$

Dove:

- \hat{x} sono i valori osservati
- \bar{x} la media della distribuzione, pari al valore del modello
- $C_{\hat{x}\hat{x}}$ è la varianza della distribuzione, assunta pari alla matrice di covarianza delle osservazioni

Per ricavare l'espressione della matrice di covarianza viene utilizzata la seguente relazione

$$C_{\hat{x}\hat{x}} = \sigma_0^2 N^{-1} \quad (10)$$

Dove:

- σ_0^2 rappresenta un valore di varianza a priori
- N^{-1} è la matrice dei cofattori

Lo svolgimento del test impone di considerare i valori osservati \hat{x} , i valori del modello di riferimento \bar{x} e la matrice dei cofattori $C_{\hat{x}\hat{x}}$ appena calcolata come dati in ingresso dai quali calcolare il valore osservato con l'espressione

$$\chi_e^2 = \frac{(\hat{x} - \bar{x})^t N (\hat{x} - \bar{x})}{\sigma_0^2} \quad (11)$$

Dopo aver fissato un opportuno livello di significatività α , si potrà quindi procedere al confronto del valore osservato, ottenuto con i limiti dell'intervallo di confidenza corrispondenti ad una distribuzione chi-quadrato con $(n - 1)$ gradi di libertà, dove n è la numerosità del campione analizzato.

Il test darà esito positivo se sarà verificata la condizione

$$\chi_t^{2(-)} \leq \chi_e^2 \leq \chi_t^{2(+)} \quad (12)$$

Dove:

- $\chi_t^{2(-)}$ e $\chi_t^{2(+)}$ sono i limiti dell'intervallo di confidenza

3.4.2 Il test classico con matrice di covarianza modello

In modo analogo a quanto fatto per il test precedente, partiamo nuovamente dai dati di deformazione superficiale considerandoli ancora una volta come valori osservati distribuiti normalmente secondo la relazione (9) dove la media è pari al valore del modello e la varianza è data dalla classe di rumore in esame, per la quale viene ancora considerata la matrice di covarianza $C_{\hat{x}\hat{x}}$ così come definita in precedenza.

La differenza rispetto al test precedente sta nell'introduzione della matrice di covarianza del modello $C_{\bar{x}\bar{x}}$ con cui si vogliono confrontare i valori osservati; $C_{\bar{x}\bar{x}}$ può essere costruita con le procedure di interpolazione con funzioni modello già viste nel paragrafo 3.3 e, analogamente a quanto fatto per la matrice di covarianza delle osservazioni $C_{\hat{x}\hat{x}}$, può essere considerata come il seguente prodotto

$$C_{\bar{x}\bar{x}} = \sigma_0^2 M^{-1} \quad (13)$$

Dove:

- σ_0^2 rappresenta un valore di varianza a priori del modello
- M^{-1} è la matrice dei cofattori del modello

In questo secondo test, oltre ai valori osservati \hat{x} , ai valori calcolati col modello di riferimento \bar{x} e alla matrice dei cofattori $C_{\hat{x}\hat{x}}$, avremo tra i dati in ingresso anche la matrice dei cofattori del modello $C_{\bar{x}\bar{x}}$ determinata in precedenza.

Dai dati appena elencati è quindi possibile risalire al valore osservato con l'espressione

$$\chi_e^2 = \frac{(\hat{x} - \bar{x})^t (N + M)^{-1} (\hat{x} - \bar{x})}{\sigma_0^2} \quad (14)$$

Sempre considerando un opportuno livello di significatività α , si potrà quindi procedere nuovamente al confronto del valore osservato ottenuto con i limiti dell'intervallo di confidenza corrispondenti ad una distribuzione chi-quadrato con $(n - 1)$ gradi di libertà, verificando ancora la condizione (12) con il nuovo valore osservato appena ottenuto.

3.4.3 Il test Bayesiano

Come ultimo test di confronto è stato utilizzato un metodo più complesso basato sull'omonimo teorema. Definito un vettore delle osservazioni y_0 ed un vettore dei k parametri incogniti θ del tipo

$$y_0 = (y_1, y_2, y_3, \dots, y_N) \quad (15)$$

$$\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_k) \quad (16)$$

Si assume che la densità di probabilità condizionata del vettore y_0 , dipendente dai valori dei k parametri incogniti θ ,

$$P(y|\theta) \quad (17)$$

sia nota; in base a queste premesse il teorema di Bayes afferma che la probabilità condizionata dei k parametri incogniti θ , date le osservazioni y , risulterà dalla formula

$$P(\theta|y) = \frac{P(y|\theta) \cdot P(\theta)}{\int P(y|\theta) \cdot P(\theta) d(\theta)} \quad (18)$$

Dove:

- y è il vettore delle osservabili
- θ è il vettore dei parametri della funzione densità di probabilità di y , sul quale si vuole fare inferenza
- $P(\theta)$ è la funzione densità di probabilità dei k parametri θ e viene chiamata "prior" poiché rappresenta la distribuzione a priori di θ , cioè lo stato di conoscenza dei parametri nella fase che precede le osservazioni
- $P(\theta|y)$ è la probabilità condizionata di θ , dato y , e rappresenta la distribuzione a posteriori del parametro θ , ovvero dopo l'introduzione delle osservazioni y ; viene anche detta "posterior"
- Il denominatore $\int P(y|\theta) \cdot P(\theta) d(\theta)$ funge da costante di normalizzazione per ricondurre l'espressione ad una densità di probabilità

Se si utilizza il teorema di Bayes per determinare la distribuzione a posteriori di un parametro, risulta possibile calcolare anche la probabilità che tale parametro sia compreso in un sottoinsieme dello spazio dei parametri.

E' bene rimarcare che spesso ci si trova nella situazione di voler trovare il sottoinsieme in cui è maggiormente concentrata la probabilità, ad esempio ad un livello del 95%.

La determinazione di una regione di questo tipo può essere effettuata con innumerevoli modalità differenti.

Per ottenere una soluzione che sia unica, posta la condizione che la funzione di densità di probabilità abbia un solo valore modale, come ad esempio un solo massimo, la regione dovrebbe essere definita in modo che la densità di ogni punto interno ad essa sia grande almeno quanto quella di tutti i punti all'esterno della regione stessa.

Una regione caratterizzata da queste proprietà è chiamata regione di massima probabilità a posteriori (HPDR, Highest Posterior Density Region) o regione di confidenza Bayesiana.

Una regione HPDR ha inoltre la caratteristica di occupare, per un determinato livello di probabilità, il minor volume possibile nello spazio dei parametri.

La definizione dei valori osservati nel test bayesiano avviene attraverso il seguente modello matematico

$$\hat{x} = x_0 + \epsilon \quad (19)$$

Dove:

- \hat{x} è il vettore delle osservazioni della variabile x_0
- ϵ è l'errore di osservazione

Indicheremo quindi con $C_{\hat{x}\hat{x}}$ la matrice di covarianza derivata dalle misure geodetiche.

Per ipotesi, la distribuzione a priori della variabile x_0 è considerata come una normale distribuita come riportato nella seguente notazione

$$x_0 \sim N(x_0 | \bar{x}, C_{\bar{x}\bar{x}}) \quad (20)$$

Dove:

- \bar{x} è il vettore dei valori di modello
- $C_{\bar{x}\bar{x}}$ è la matrice di covarianza del modello

Sulla base di quanto affermato, si può dimostrare che la distribuzione a posteriori di x_0 può essere ricondotta ad una nuova normale data da:

$$x_0 \sim N(x_0 | \tilde{x}, W^{-1}) \quad (21)$$

Dove:

$$W = C_{\bar{x}\bar{x}}^{-1} + C_{\hat{x}\hat{x}}^{-1} \quad (22)$$

$$\tilde{x} = W^{-1} (C_{\hat{x}\hat{x}}^{-1} \cdot \hat{x} + C_{\bar{x}\bar{x}}^{-1} \cdot \bar{x}) \quad (23)$$

In modo del tutto analogo a quanto fatto per il test precedente, i dati di ingresso nel test bayesiano sono il vettore dei valori osservati \hat{x} , il vettore dei valori del modello di riferimento \tilde{x} , la matrice dei cofattori $C_{\hat{x}\hat{x}}$ e la matrice dei cofattori del modello $C_{\tilde{x}\tilde{x}}$.

Il valore osservato può essere quindi ricavato dalla relazione

$$\chi_e^2 = (\hat{x} - \tilde{x})^t (C_{vv})^{-1} (\hat{x} - \tilde{x}) \quad (24)$$

In questo caso è necessario svolgere alcuni passaggi ulteriori per determinare la matrice C_{vv} ; partendo dall'espressione che fornisce i gradi di libertà considerata finora

$$v = n - 1 \quad (25)$$

è possibile ricavare l'espressione di C_{vv} riscrivendo la (25) come

$$v = \hat{x} - \tilde{x} \quad (26)$$

Dopo alcuni passaggi è possibile ricavare una nuova espressione di v data da

$$v = [I - W^{-1}C_{\hat{x}\hat{x}}^{-1} \quad -W^{-1}C_{\tilde{x}\tilde{x}}^{-1}] \begin{bmatrix} \hat{x} \\ \tilde{x} \end{bmatrix} \quad (27)$$

dove indicando la prima matrice con

$$A = [I - W^{-1}C_{\hat{x}\hat{x}}^{-1} \quad -W^{-1}C_{\tilde{x}\tilde{x}}^{-1}] \quad (28)$$

e introducendo la matrice di covarianza del vettore $\begin{bmatrix} \hat{x} \\ \tilde{x} \end{bmatrix}$ C_{ss} espressa da

$$C_{ss} = \begin{bmatrix} C_{\hat{x}\hat{x}} & 0 \\ 0 & C_{\tilde{x}\tilde{x}} \end{bmatrix} \quad (29)$$

è possibile ricavare l'espressione di C_{vv} , grazie alla legge di propagazione della varianza, attraverso la seguente relazione

$$C_{vv} = A \cdot C_{ss} \cdot A^t \quad (30)$$

Considerando come sempre un opportuno livello di significatività α , si potrà quindi procedere al confronto del valore osservato ottenuto grazie alla (24) con i limiti dell'intervallo di confidenza sintetizzati dalla relazione

$$\chi_t^{2(-)} \leq \chi_e^2 \leq \chi_t^{2(+)} \quad (31)$$

Quest'ultimo test è a tutti gli effetti una misura della discrepanza pesata tra i dati calcolati e le osservazioni. Esso considera allo stesso tempo la struttura dei dati attraverso il concetto di verosimiglianza o likelihood e la struttura del modello a priori attraverso l'introduzione della funzione densità di probabilità dei parametri ("prior").

3.5 Analisi su dati simulati

In accordo con quanto espresso nei precedenti paragrafi, l'analisi su dati simulati parte dalla generazione delle serie artificiali di osservazioni in superficie grazie alle relazioni

$$\left. \begin{aligned} d_{GPS(N)} &= G_N \cdot m_{synt} \\ d_{GPS(E)} &= G_E \cdot m_{synt} \end{aligned} \right\} + \varepsilon_{GPS} \quad (32)$$

$$d_{SAR(z)} = G_z \cdot m_{synt} + \varepsilon_{SAR}$$

Il vettore m_{synt} e le matrici G_N , G_E e G_z sono quelle già introdotte in precedenza, mentre per il rumore sono state adottate le seguenti condizioni:

- ε_{SAR} è stato considerato variabile con valori pari a 0.2, 0.4 e 0.8 mm/anno ed è stato applicato alle velocità ottenute dai dati simulati SAR
- ε_{GPS} è stato fatto variare su diversi valori medi (1, 2, 3 e 4 mm), creando di fatto quattro serie di rumore completamente scorrelate che sono state applicate ai dati planimetrici artificiali espressi in termini di spostamento

Nella procedura di calcolo utilizzata si è scelto di ricavare dapprima il dato "pulito" in termini di velocità dalla relazione diretta del modello di Okada.

Si è quindi trasformato in termini di spostamenti il risultato ottenuto per le sole coordinate Nord ed Est e si sono applicati, come ultimo passo, i differenti livelli di rumore ipotizzati.

Per ottenere una simulazione realistica del dato GPS, sono state create delle serie di soluzioni con le seguenti caratteristiche:

- Campagne di misura della durata di 4 giorni consecutivi
- Per ogni campagna sono state ricavate delle soluzioni giornaliere per le componenti Nord ed Est relative alle 10 stazioni

- Le campagne sono state considerate distanziate di un anno esatto (365 giorni) tra loro e svolte per sei anni consecutivi.

Le serie generate con questi accorgimenti rispecchiano la struttura dei dati ottenuti dai veri rilievi e permetteranno nei passi successivi di avere una valutazione più affidabile sulla bontà del metodo utilizzato.

In analogia a quanto fatto nel primo capitolo della tesi sui dati reali, anche per le serie sintetiche sono state stimate le velocità di deriva delle singole stazioni utilizzando ancora una volta il metodo dei minimi quadrati.

Allo scopo di analizzare il comportamento dei test al variare del numero di campagne effettuate, le stime delle velocità e del loro scarto quadratico medio sono state calcolate a partire dai primi due anni di misure. Sono state quindi ripetute in seguito al progressivo aggiungersi, anno per anno, dei dati acquisiti, fino ad arrivare a considerare tutte e sei le campagne generate. L'operazione è stata ripetuta ovviamente per tutte e quattro le classi di rumore introdotte.

1	2	3	4
5	6	7	8
9	10	11	12

Figura (3.4) Schematizzazione della suddivisione in patch della faglia con il relativo progressivo numerico

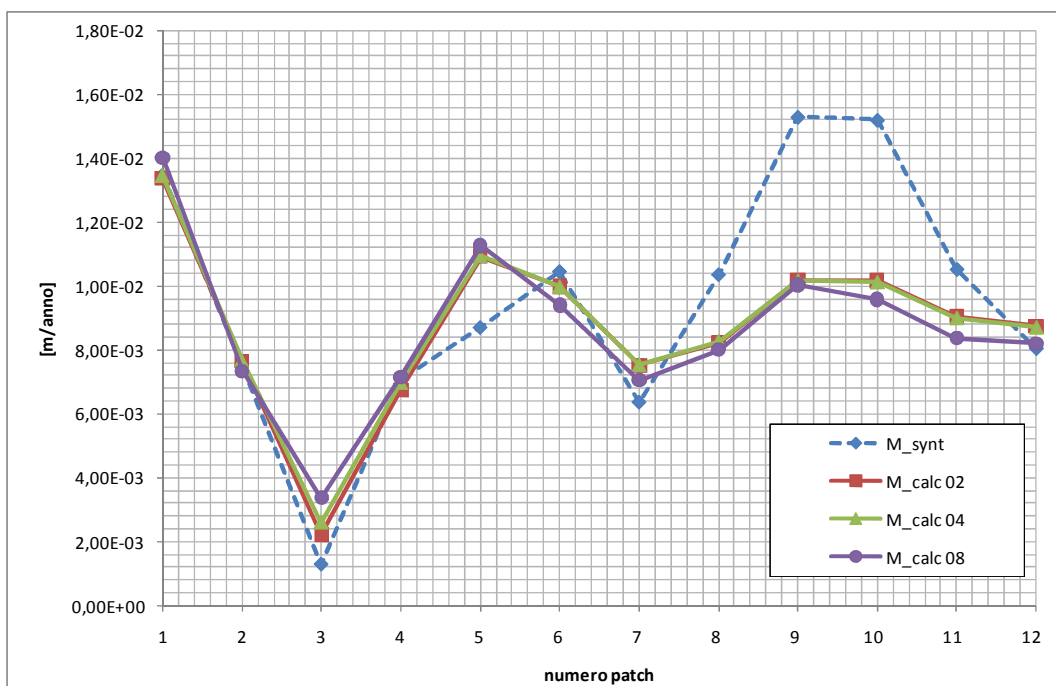


Figura (3.5) Confronto tra valori del vettore sintetico dei movimenti di faglia e i valori ottenuti dall'inversione dei dati SAR simulati per le singole patch all'aumentare del rumore applicato

L'operazione di inversione dei dati è stata fatta a partire dalle serie di spostamento verticale $d_{SAR}(z)$ grazie all'applicazione del metodo dei minimi quadrati regolarizzati visto nel capitolo 2, ed è stata ripetuta per i tre valori di rumore considerati.

Come soluzione approssimata dei parametri $\langle m \rangle$ è stato sempre considerato un valore medio ottenuto da elaborazioni precedenti sulla stessa faglia.

L'elaborazione ha permesso di calcolare tre vettori distinti, m_{calc02} , m_{calc04} e m_{calc08} , con le loro relative matrici di covarianza $C_{\hat{m}\hat{m}}$.

Le tre distribuzioni di velocità, come si può notare dalla figura (3.5), presentano valori di inversione molto buoni per tutte le patch superficiali (dalla 1 alla 4), mentre forniscono un andamento più regolarizzato delle velocità sulle altre patch rispetto a quanto si vede in m_{synt} , soprattutto nella parte più profonda del piano

di faglia. In particolare è evidente una lieve sovrastima delle velocità nella patch numero 5, dovuta probabilmente alla distribuzione non ottimale dei punti di riferimento in superficie.

L'aumento del rumore applicato ai dati SAR di partenza influisce debolmente sul risultato dell'inversione, che restituisce un andamento pressoché invariato delle velocità calcolate per i tre livelli di rumore considerati.

Prima di poter continuare il confronto, è necessario determinare le covarianze relative al modello ed ai dati, indispensabili per i test d'ipotesi.

Dagli scarti quadratici medi delle osservazioni generate si ricavano, per le quattro classi di rumore e per i diversi casi di stima, da due a sei anni, le matrici dei pesi delle osservazioni.

Per il calcolo della covarianza relativa al modello il primo passo consiste nella determinazione della covarianza analitica $C_{\bar{a}\bar{a}}$ dalla semplice propagazione della covarianza espressa dalla formula (7).

Da questa è possibile ricavare la covarianza empirica, che esprime la media delle covarianze comprese entro una distanza fissata dai punti di stazione.

Il valore iniziale a distanza nulla è considerato pari alla media delle varianze delle singole stazioni, mentre i valori delle covarianze empiriche sono stati determinati al variare della distanza e sono stati inseriti nei grafici che seguono per un passo costante di 4 km fino ad arrivare alla distanza tra le due stazioni più lontane, pari a circa 24 km.

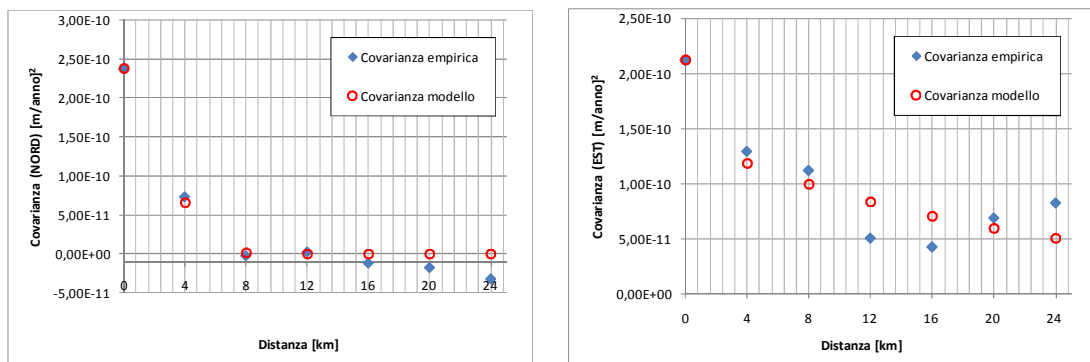


Figura (3.6) Andamenti delle covarianze empiriche e modello ottenute dall'elaborazione di dati SAR con rumore pari a 0,4 mm/anno

Come appare evidente dalla figura (3.6) i valori delle covarianze empiriche sono affetti da un significativo rumore e, in accordo con quanto affermato nei paragrafi precedenti, è stata effettuata l'interpolazione dei dati ottenuti su funzioni modello operando infine la scelta ottimale con il criterio AIC.

Nel nostro caso le covarianze modello, per tutte le classi di rumore applicate ai dati SAR, sono state ottenute utilizzando:

- Una funzione di tipo normale per la componente Nord
- Una funzione di tipo esponenziale per la componente Est

I valori delle covarianze empiriche, e di conseguenza quelli delle covarianze modello, aumentano con l'incremento del rumore applicato ai dati SAR come si può osservare nelle figure seguenti.

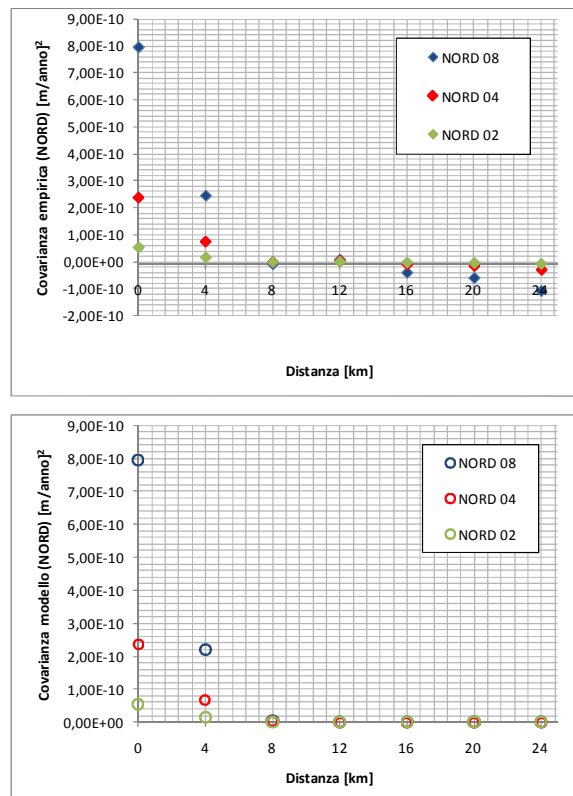


Figura (3.7) Andamenti delle covarianze empiriche e modello per la componente NORD al variare del rumore

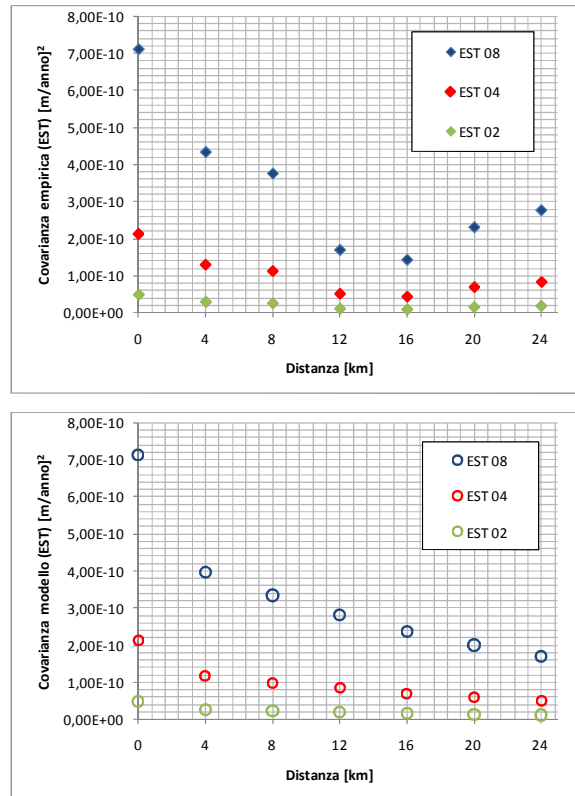


Figura (3.8) Andamenti delle covarianze empiriche e modello per la componente EST al variare del rumore

L'ultimo passo prima di eseguire i test di confronto è stato il calcolo dei valori predetti in superficie dal modello per le componenti Nord ed Est a partire dai vettori m_{calc02} , m_{calc04} e m_{calc08} ottenuti dall'inversione dei dati SAR generati. Applicando ancora una volta il modello diretto grazie alle relazioni

$$d_{GPS\ calc}(N) = G_N \cdot m_{calc} \quad (33)$$

$$d_{GPS\ calc}(E) = G_E \cdot m_{calc} \quad (34)$$

sono state determinate le velocità medie predette dal modello per ogni singola stazione. I risultati ottenuti sono riportati nella tabella seguente

TABELLA (3.1)						
NOISE SAR	$\varepsilon_{SAR} = 0,2 \text{ mm/anno}$		$\varepsilon_{SAR} = 0,4 \text{ mm/anno}$		$\varepsilon_{SAR} = 0,8 \text{ mm/anno}$	
	$d_{GPS \text{ calc } (N)}$	$d_{GPS \text{ calc } (E)}$	$d_{GPS \text{ calc } (N)}$	$d_{GPS \text{ calc } (E)}$	$d_{GPS \text{ calc } (N)}$	$d_{GPS \text{ calc } (E)}$
Stazione	[m/anno]	[m/anno]	[m/anno]	[m/anno]	[m/anno]	[m/anno]
<i>cass</i>	3,99E-04	1,24E-03	4,12E-04	1,28E-03	4,19E-04	1,33E-03
<i>civi</i>	4,69E-04	1,58E-03	4,87E-04	1,60E-03	5,00E-04	1,61E-03
<i>colo</i>	-6,94E-04	4,12E-04	-6,96E-04	4,12E-04	-6,90E-04	4,10E-04
<i>croc</i>	-2,54E-04	8,32E-04	-2,55E-04	8,33E-04	-2,20E-04	8,11E-04
<i>dolc</i>	6,34E-04	1,50E-04	6,56E-04	1,67E-04	6,83E-04	1,98E-04
<i>dorm</i>	9,51E-04	1,48E-03	9,58E-04	1,49E-03	9,59E-04	1,51E-03
<i>fran</i>	2,75E-04	1,08E-03	2,81E-04	1,10E-03	2,80E-04	1,09E-03
<i>fras</i>	8,90E-05	-6,57E-04	7,24E-05	-6,55E-04	5,03E-05	-6,29E-04
<i>mora</i>	-6,20E-04	6,48E-04	-6,21E-04	6,47E-04	-6,03E-04	6,36E-04
<i>zacc</i>	3,08E-04	4,27E-04	3,05E-04	4,19E-04	2,92E-04	3,76E-04

Con gli elementi ricavati fino a questo punto dell'analisi è stato possibile eseguire il confronto delle componenti planimetriche sintetiche $\hat{d}_{GPS(N)}$ e $\hat{d}_{GPS(E)}$, ricavate al variare del numero di campagne considerate, con le velocità predette dal modello riportate in tabella (3.1).

Il test è stato applicato utilizzando i tre metodi introdotti nei paragrafi precedenti e, in tutti i casi, è stato considerato un valore soglia per la statistica del χ^2 corrispondente ad un livello di significatività α del 5%.

Tutti i risultati ottenuti per i tre differenti metodi al variare dei rumori ε_{SAR} ed ε_{GPS} sono riportati nell'appendice.

Dai risultati ottenuti è evidente che, per entrambe le coordinate, i test entrano in significatività mediamente dopo tre anni di misure e in qualche caso addirittura dopo solo due campagne. Questo sembrerebbe un segnale della bontà del modello e del metodo applicato per l'inversione.

Da un'analisi più approfondita, si nota che i valori del χ^2 ottenuti con il test classico sono sempre i più elevati, mentre il test Bayesiano è quello che fornisce in ogni caso i risultati più contenuti della statistica test.

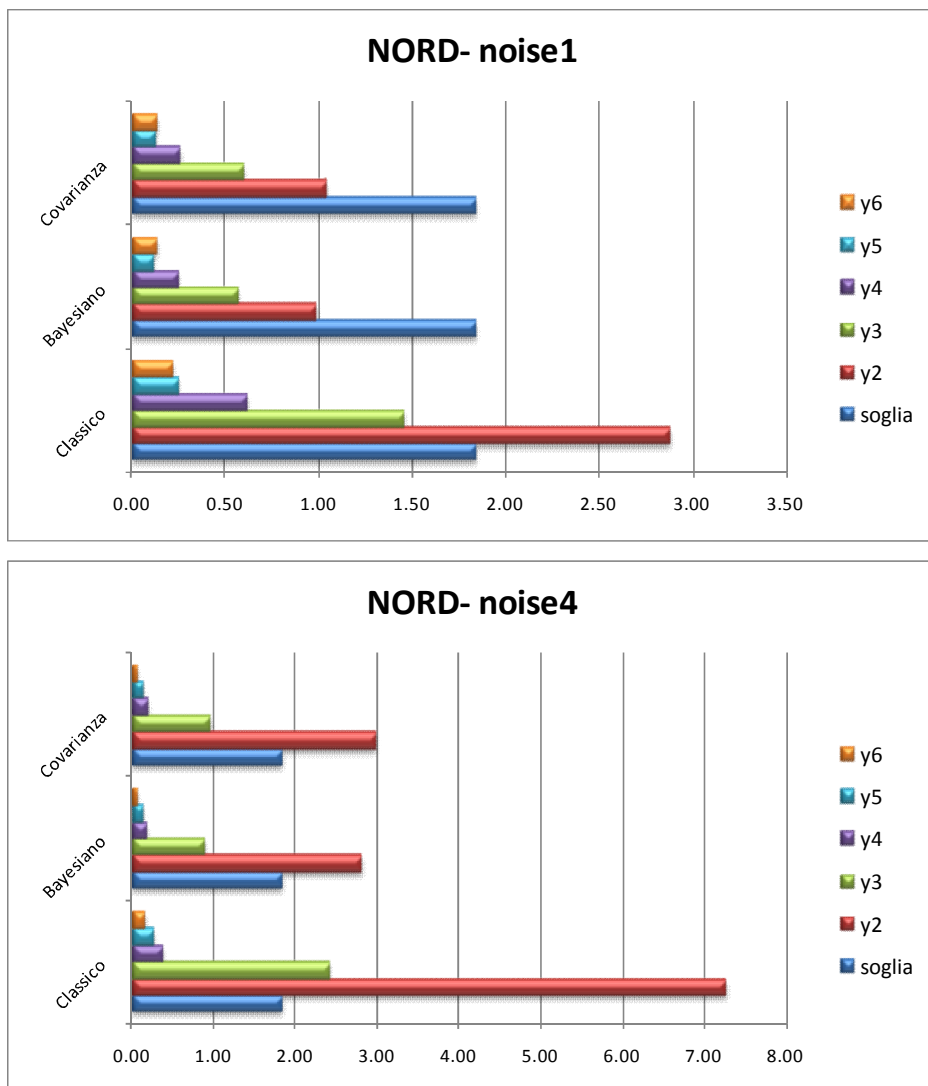


Figura (3.9) Risultati delle statistiche test ottenuti sulle componenti NORD con ε_{SAR} pari a 0,2 mm/anno per valori di ε_{GPS} pari ad 1 e a 4 mm

Il risultato dà forza all'ipotesi che il test Bayesiano sia effettivamente il più robusto in applicazioni di questo tipo e che tenda a convergere verso la significatività in tempi più rapidi degli altri due.

Dalla figura (3.9) si può anche vedere che all'aumentare di ε_{GPS} i valori della statistica test tendono ad aumentare in corrispondenza delle campagne di due e tre anni. Una possibile spiegazione a questa tendenza può essere ricercata negli spostamenti stimati dalle velocità ottenute a diversi valori di ε_{GPS} ; nelle figure (3.10) e (3.11) si riportano i grafici relativi alla stazione dorm ricavati con ε_{SAR} pari a 0,2 mm/anno.

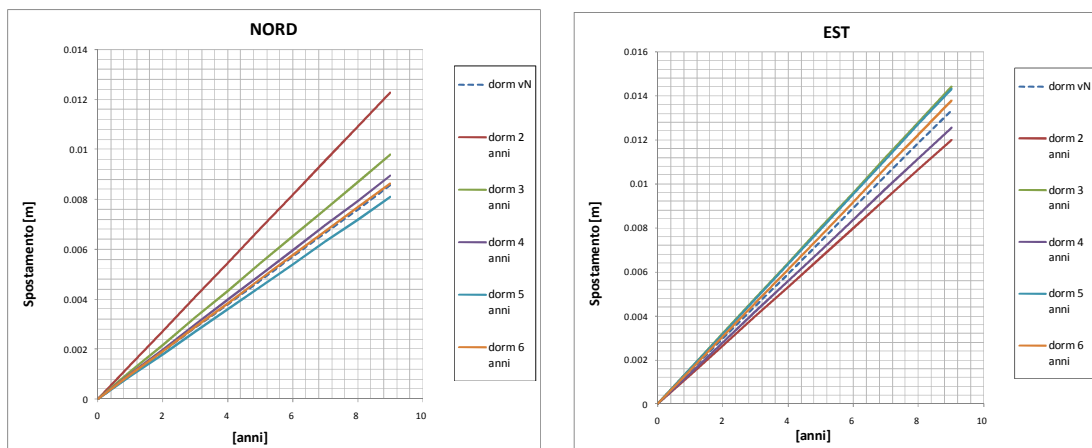


Figura (3.10) Spostamenti stimati per la stazione dorm con ε_{GPS} pari ad 1 mm

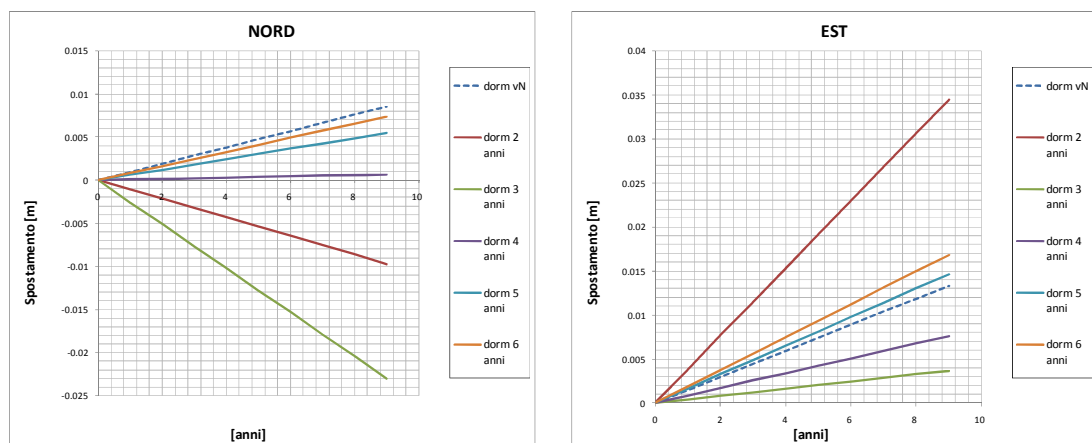


Figura (3.11) Spostamenti stimati per la stazione dorm con ε_{GPS} pari ad 4 mm

L'applicazione di un rumore casuale elevato sulle serie temporali GPS, come nel caso mostrato in figura (3.11), può influenzare in maniera significativa la stima delle velocità e può addirittura indurre un iniziale errore nella valutazione del verso corretto delle stesse. A livelli di rumore più contenuti invece i risultati sono coerenti già dalle prime stime. E' ragionevole supporre che il test, nei casi relativi a rumori elevati, interpreti il fenomeno come una mancata congruenza tra i dati in ingresso ed il modello.

La variazione del rumore applicato ai dati SAR, a differenza di quanto visto nel caso precedente per i dati GPS, non influenza significativamente i valori finali dei test. Un risultato simile è concorde con quanto ottenuto in precedenza nell'ambito dell'inversione del modello di Okada per la determinazione dei vettori m_{calc02} , m_{calc04} e m_{calc08} .

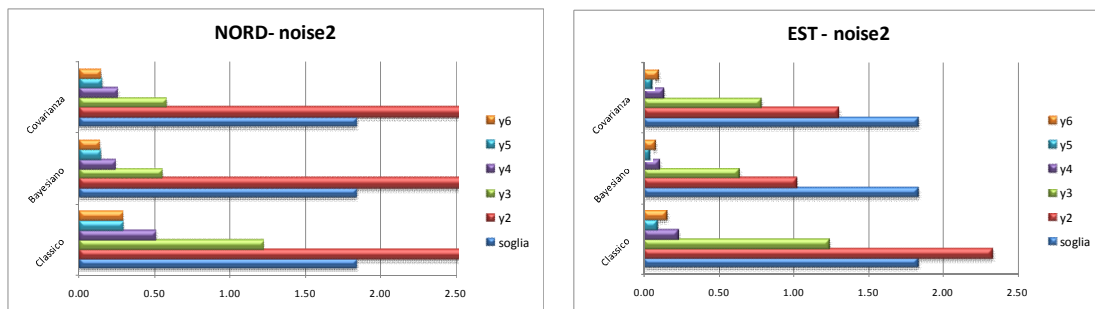


Figura (3.12) Valori delle statistiche test ottenuti con ε_{GPS} pari ad 2 mm ed ε_{SAR} pari a 4 mm/anno

Se si confrontano infine i risultati per le due coordinate a parità di rumore applicato, come nell'esempio mostrato in figura (3.12), si nota che, a partire dal quarto anno di campagna, si hanno valori di χ^2 per la coordinata Est sempre inferiori a quelli relativi alla coordinata Nord.

Come già detto in precedenza, questo fatto può essere ricondotto alla struttura ed alla direzione di scorrimento della faglia.

Conclusione

L'evoluzione nel monitoraggio dei fenomeni geodinamici ha comportato un forte incremento della disponibilità di dati, sia nella quantità effettiva di misure, sia nella varietà dei metodi di acquisizione.

L'intento era quello di analizzare le effettive possibilità di migliorare la qualità delle misure stesse e di integrarle nei modelli geofisici di faglia, in particolare nel caso della loro inversione.

L'analisi svolta nel primo capitolo sulle serie temporali ottenute dal 2003 al 2009 per la rete non-permanente del Pollino evidenzia, come già detto, che la procedura COMM (Common-mode error) non è sufficientemente efficace per individuare e rimuovere i segnali correlati col tempo e non legati alla geodinamica della faglia.

Ciò non toglie che, grazie a futuri sviluppi nei metodi matematici di filtraggio, si arrivi a determinare una procedura efficace anche per le reti di modesta estensione.

Alla luce dei risultati ottenuti nella simulazione del terzo capitolo, si può affermare che la possibilità di integrare dati di tipo SAR con serie temporali GPS è concreta e realizzabile nell'ambito dei modelli di inversione.

Il possibile utilizzo di diversi tipi di misure, integrati nella descrizione geodinamica, impone quindi di rivedere il concetto classico di modello fisso e privo di errore.

I dati assumono una forte valenza descrittiva e di conseguenza è necessario prendere in considerazione anche la variabilità del modello nell'applicazione dei test statistici di confronto.

I risultati ottenuti mostrano chiaramente come il test classico con matrice di covarianza modello e soprattutto il test Bayesiano siano in grado di convergere a significatività in tempi più rapidi rispetto al test classico.

L'esito della nostra analisi evidenzia il fatto che questi due tipi di test tendono in genere ad entrare in significatività con un anno d'anticipo rispetto al metodo classico. In particolare il test Bayesiano fornisce sempre il valore di chi-quadrato più basso.

Nell'ottica della programmazione di nuove campagne di misura, l'utilizzo del test Bayesiano garantirebbe quindi una valutazione più rapida della congruenza tra i dati ottenuti e il modello applicato alla zona in studio.

Le positive conseguenze di quest'approccio sono la riduzione dei tempi necessari per ottenere dati significativi e i relativi vantaggi pratici ed economici.

Una conferma di questa conclusione può scaturire da ulteriori analisi che prendano in esame dati reali in casi analoghi.

Appendice

Nelle pagine seguenti si riportano tutti i grafici relativi ai test statistici eseguiti nell'ambito del confronto incrociato del capitolo 3. I valori di chi-quadrato ottenuti sono suddivisi in ogni grafico in base alle tre tipologie di test applicate

- Covarianza
- Bayesiano
- Classico

e sono ordinati in base al numero di anni di campagna considerati nella stima. Sono raggruppati in base alla componente analizzata e presentati al crescere del rumore applicato ai dati.

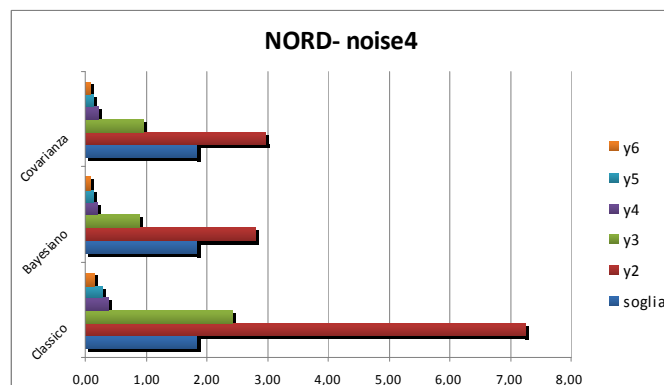
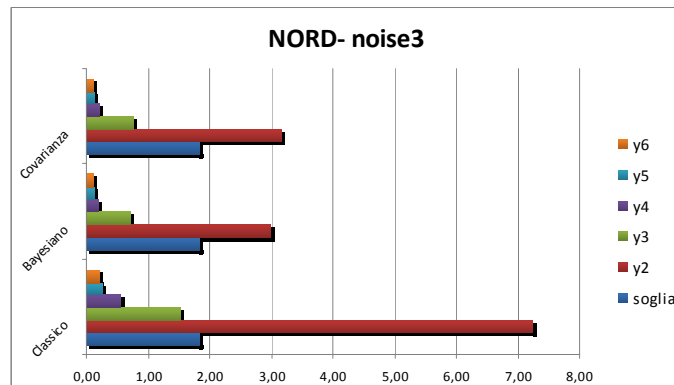
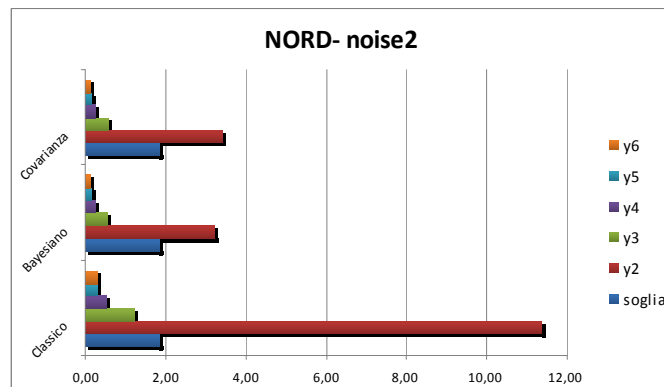
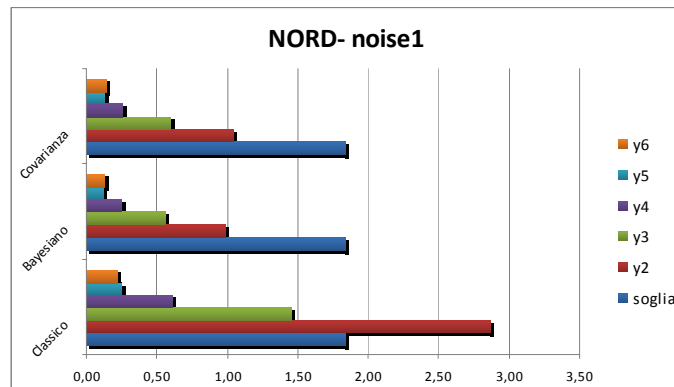
Si ricorda che nelle analisi sono stati applicati i seguenti tre livelli di noise ai dati di tipo SAR:

1. ε_{SAR} pari a 0,2 mm/anno
2. ε_{SAR} pari a 0,4 mm/anno
3. ε_{SAR} pari a 0,8 mm/anno

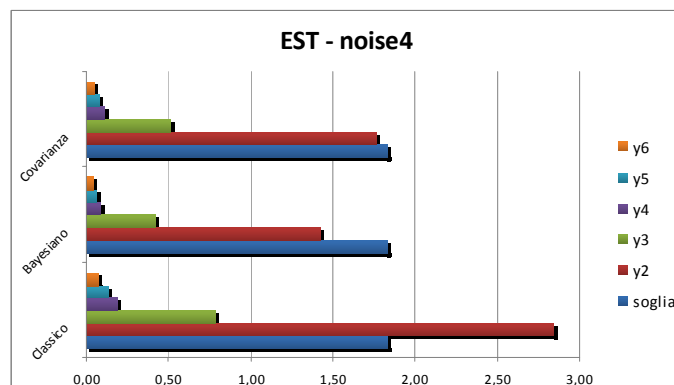
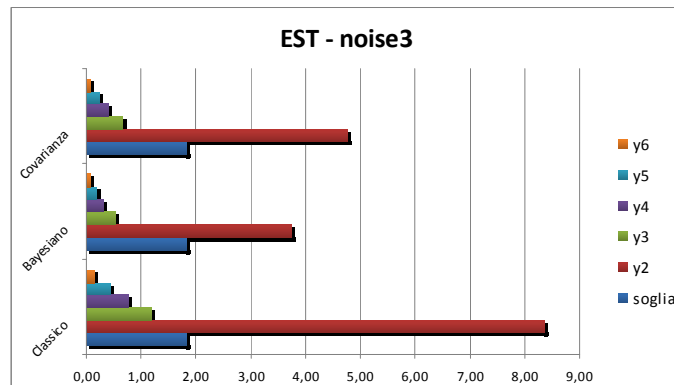
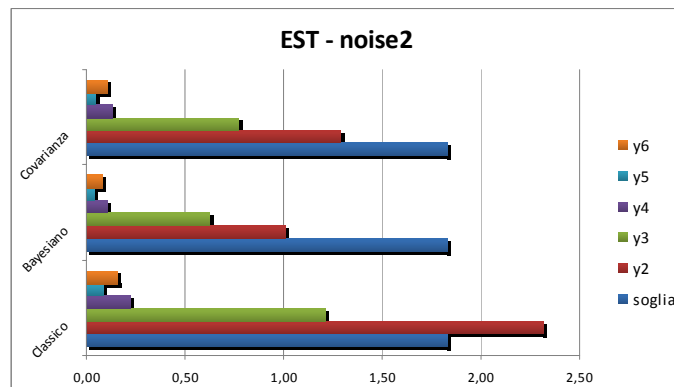
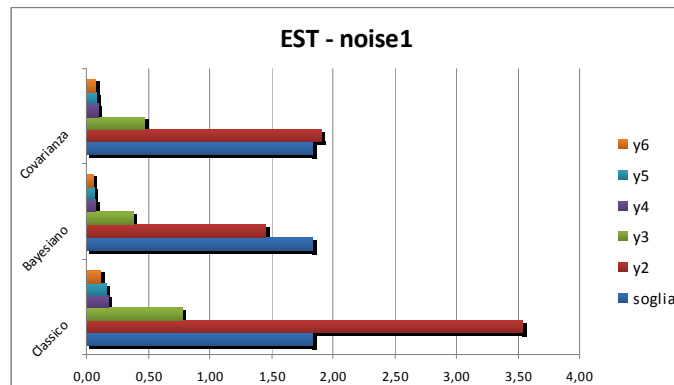
per ognuno dei quali è stato svolto il confronto incrociato considerando quattro livelli di rumore applicato agli spostamenti GPS:

- ε_{GPS} pari ad 1 mm
- ε_{GPS} pari a 2 mm
- ε_{GPS} pari a 3 mm
- ε_{GPS} pari a 4 mm

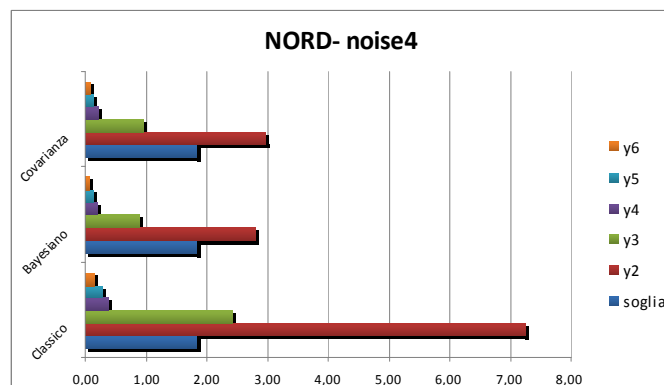
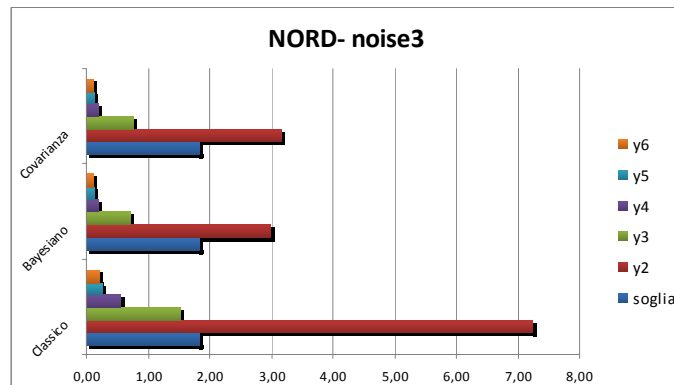
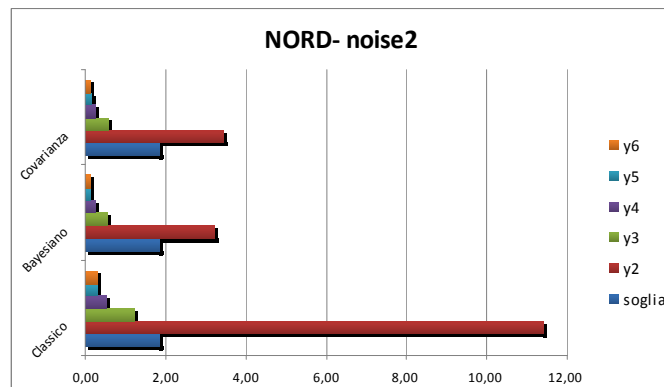
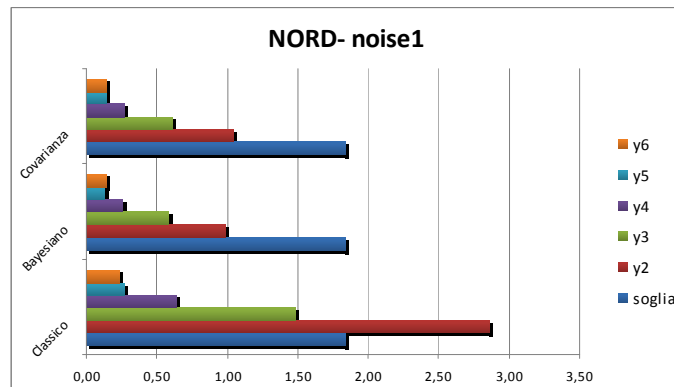
$\varepsilon_{SAR} = 0,2 \text{ mm/anno}$ - componente NORD



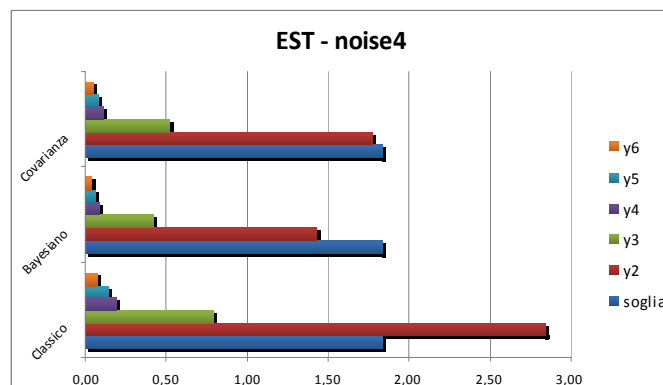
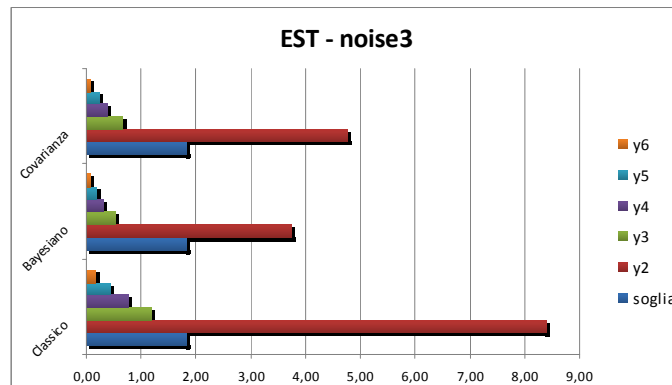
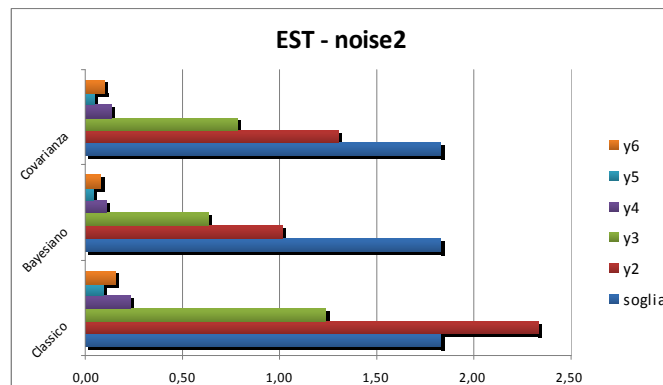
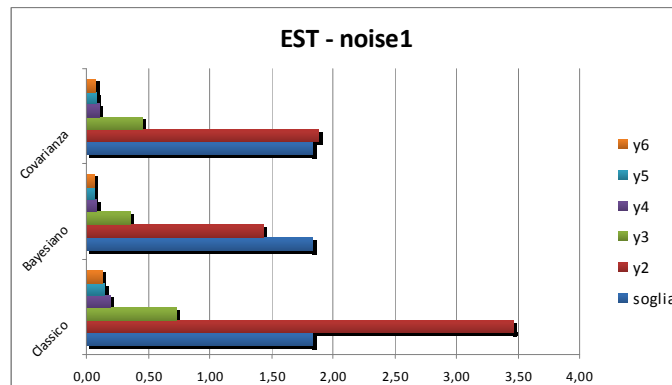
$\varepsilon_{SAR} = 0,2 \text{ mm/anno}$ - componente EST



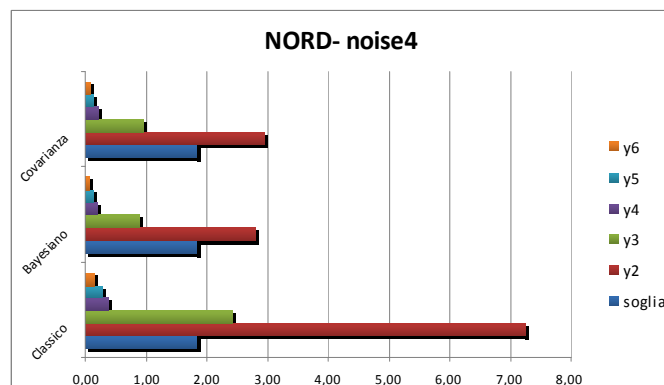
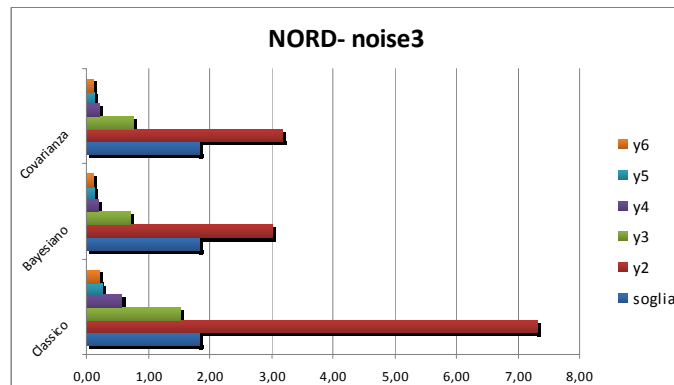
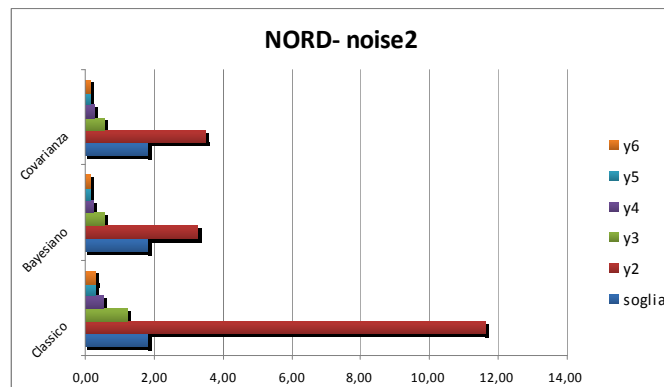
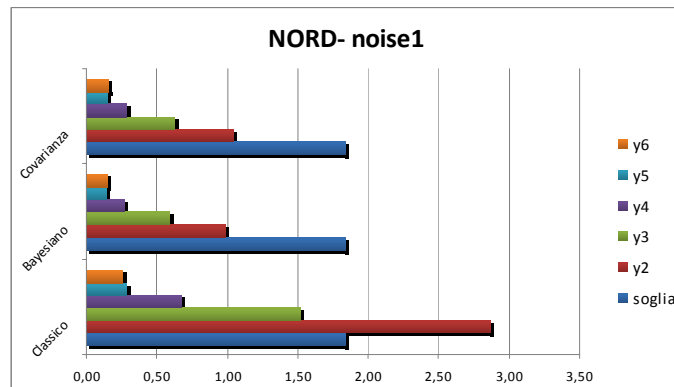
$\varepsilon_{SAR} = 0,4 \text{ mm/anno}$ - componente NORD



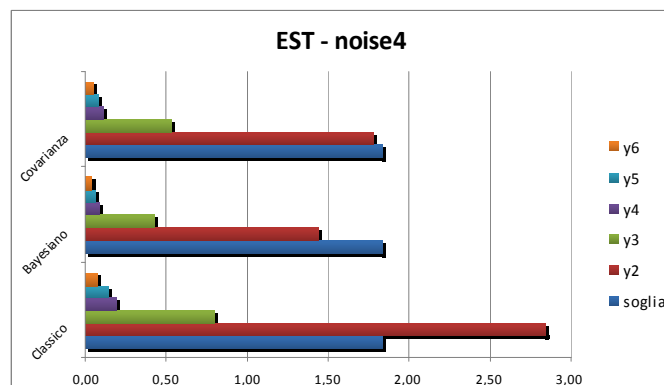
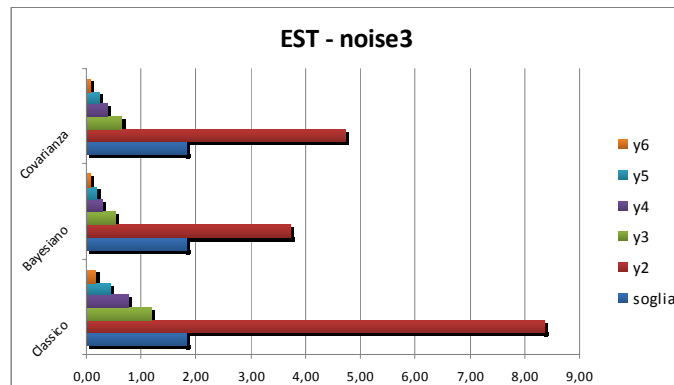
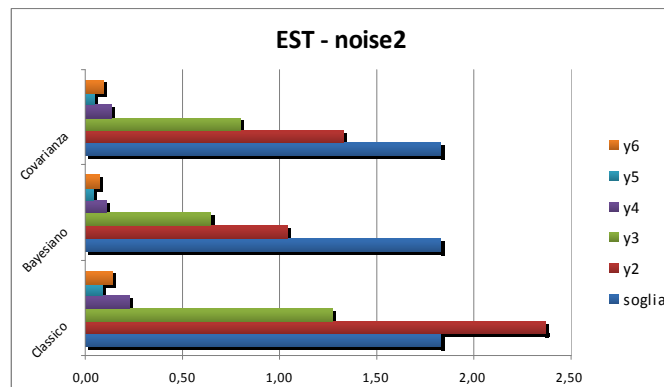
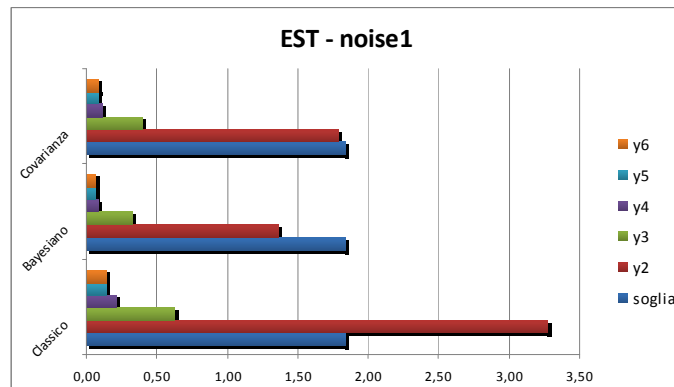
$\varepsilon_{SAR} = 0,4 \text{ mm/anno}$ - componente EST



$\varepsilon_{SAR} = 0,8 \text{ mm/anno}$ - componente NORD



$\varepsilon_{SAR} = 0,8 \text{ mm/anno}$ - componente EST



Bibliografia

- BIAGI L. – *“I fondamentali del GPS”* – Geomatics Workbooks, vol. 8 (2009)
- CINA A. – *“GPS. Principi, modalità, tecniche di posizionamento”* – (2000)
- FAINI S. – *“L’inversione geofisica di osservazioni geodetiche”* – (2010)
- KÚNZLE A.M. – *“Inferenza statistica di misure geodetiche di modelli di inversione geofisica”* – (2010)
- OKADA Y. – *“Internal deformation due to shear and tensile faults in a half-space”*
– Bull. Seism. Soc. Am. Vol. 82 No. 2, pp 1018-1040 (April 1992)
- WDOWINSKI S., BOCK Y., ZHANG J., FANG P., GENRICH J. – *“Southern California Permanent GPS Geodetic Array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake”* – J. Geophys. Res. Vol. 102, No. B8, pp 18.057-18.070 (August 1997)
- SANSÒ F., CLARA DE LACY M. – *“The Bayesian Approach Applied to Significant Deformation Identification”* – IAG Symposium, Jaen, Spain (March 2005)
- SABADINI R., AOUDIA A., BARZAGHI R., CRIPPA B., MAROTTA A.M., BORGHI A., CANNIZZARO L., CALCAGNI L., DALLA VIA G., ROSSI G., SPLENDORE R., CROSETTO M. – *“First evidences of fast creeping on a long-lasting quiescent earthquake normal-fault in the Mediterranean”* – Geophys. J. Int. (2009)