

I fondamentali del GPS

Il libro vuole analizzare e illustrare le problematiche, le metodologie e le accuratèzze del posizionamento satellitare nelle applicazioni che vanno dalla navigazione in tempo reale al monitoraggio di alta accuratezza per reti geodetiche locali; il testo discute il problema dei sistemi di riferimento e l'elaborazione delle osservazioni nei suoi diversi passi, dal dato grezzo alla compensazione di reti: il livello della trattazione è quello della didattica universitaria nel corpo della Geomatica.

Il volume è stato scritto nell'ambito del Progetto *Galileo e il posizionamento satellitare modernizzato*, finanziato nel 2006 dal Ministero dell'Università e della Ricerca come Progetto di Rilevante Interesse Nazionale. Le copie stampate del volume verranno distribuite a colleghi e biblioteche; i pdf del testo sono viceversa scaricabili liberamente dal sito <http://geomatica.como.polimi.it/fondamentaligps/>

Ludovico Biagi è ricercatore presso il Politecnico di Milano, DIIAR, dove è docente dei corsi di *Positioning, Geodetic Monitoring and Navigation* e di *Remote Sensing*, per la laurea specialistica in Environmental and Geomatic Engineering. Negli ultimi anni la sua attività prevalente è stata nelle tematiche delle reti GNSS permanenti per la realizzazione del sistema di riferimento, per la stima dei campi di spostamento e per la distribuzione di servizi di posizionamento; è membro dello Study Group sul sistema di riferimento della Commissione EUREF ed è coautore di circa 70 pubblicazioni e di 2 libri.

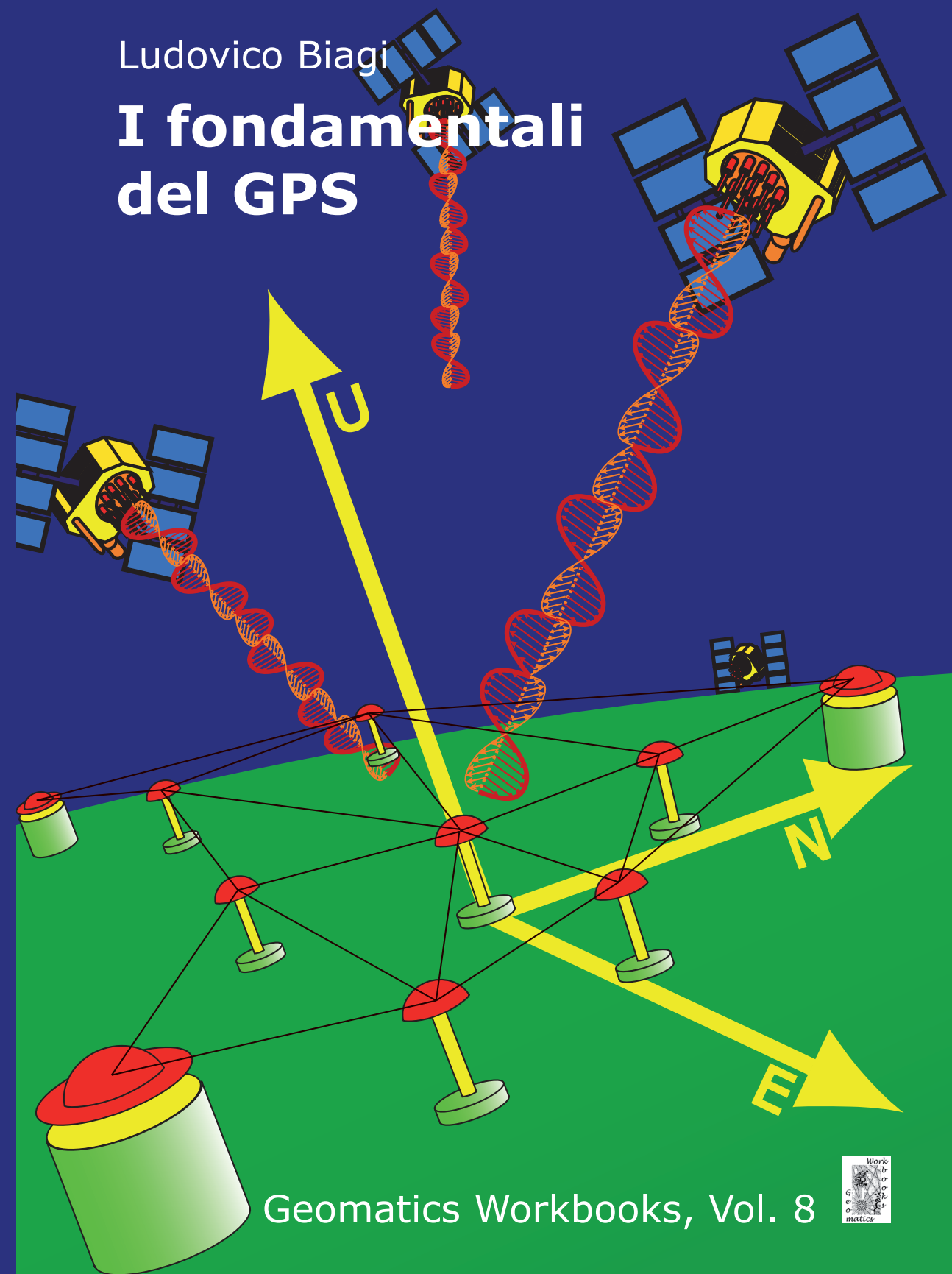


Geomatics Workbooks Vol. 8, ISSN: 1591-092X

Ludovico Biagi

I fondamentali del GPS

I fondamentali del GPS



Geomatics Workbooks, Vol. 8



Geomatics Workbooks, Volume 8

I fondamentali del GPS

Un volume scritto e distribuito nell'ambito del
Progetto di Rilevante Interesse Nazionale
Galileo e il posizionamento satellitare modernizzato,
coordinatore nazionale Prof. Fernando Sansò,
cofinanziato nel 2006 dal Ministero dell'Università e della Ricerca,

di

Ludovico Biagi

Ludovico Biagi
Politecnico di Milano,
DIIAR c/o Polo Regionale di Como,
via Valleggio 11, 22100 Como
ludovico.biagi@polimi.it

Disegno di copertina dell'autore

I fondamentali del GPS
Geomatics Workbooks, Volume 8, 2009
ISSN 1591-092X

<http://geomatica.como.polimi.it/workbooks/>
<http://geomatica.como.polimi.it/fondamentalgps/>

geowb@geomatica.como.polimi.it

Editorial Board: M. A. Brovelli, L. Biagi, M. Negretti

Geomatics Laboratory,
<http://geomatica.como.polimi.it>,
Politecnico di Milano, Polo Regionale di Como,
via Valleggio 11, 22100 Como, Italy

Indice

	Prefazione	vii
1.	Introduzione: il GNSS e le tecniche di posizionamento	1
2.	I sistemi di riferimento e di coordinate	9
2.1	La trasformazione fra sistemi di riferimento	10
2.1.1	Il caso di piccole rotazioni e fattore di scala	13
2.2	Introduzione ai sistemi di riferimento in geodesia	14
2.2.1	I moti del pianeta visti dallo spazio	14
2.2.2	Gli altri moti rigidi del pianeta	15
2.2.3	I moti non rigidi della massa terrestre: i fenomeni periodici	17
2.2.4	La rimozione dei fenomeni periodici	18
2.2.5	I fenomeni non periodici	19
2.3	I diversi sistemi di riferimento	22
2.4	Il sistema di riferimento globale: International Terrestrial Reference System	24
2.4.1	Le reti per la realizzazione di ITRS	26
2.4.1.1	La rete globale VLBI: l'International VLBI Service, IVS	26
2.4.1.2	La rete globale SLR: l'International SLR Service, ILRS	28
2.4.1.3	La rete globale GNSS	29
2.4.2	La realizzazione ITRFyy	30
2.4.3	La realizzazione IGSyy	33
2.5	I Sistemi di Riferimento locali	33
2.5.1	Il sistema di riferimento Europeo: ETRS89	33
2.5.2	La trasformazione fra ITRS e ETRS89: la codificazione internazionale	35
2.5.3	La realizzazione italiana di ETRS89: il vecchio inquadramento di IGM95	36
2.5.4	La realizzazione italiana di ETRS89: RDN e il nuovo inquadramento di IGM95	38
2.6	Le coordinate geodetiche in un SR 3D: l'ellissoide di riferimento	40
2.6.1	Una nota sulle unità di misura degli angoli	43
2.7	Gli altri sistemi di coordinate e le coordinate cartesiane locali	44
2.8	I sistemi di riferimento locali storici (topografici)	45
2.9	Le scale temporali	46
2.9.1	Le scale di tempo dinamico	46
2.9.2	Le scale di tempo siderale e solare	46
2.9.3	Le scale di tempo atomico	47
3.	Il sistema GPS	49
3.1	Le orbite e le effemeridi dei satelliti	49

3.2.1	L'orbita Kepleriana	50
3.2.2	Gli effetti delle altre forze	52
3.3	Le effemeridi del GPS	53
3.3.1	Le effemeridi trasmesse e la relativa procedura di calcolo	53
3.3.2	Le effemeridi calcolate a posteriori	55
3.4	Il satellite GPS	56
3.5	La costellazione GPS	57
3.6	Il segnale GPS	60
3.6.1	Le portanti	60
3.6.2	I codici binari	63
3.6.3	La modulazione dei segnali	64
4.	Le osservazioni, la propagazione e i disturbi atmosferici	67
4.1	La frequenza, la fase e il tempo di un oscillatore	67
4.1.1	Il termine relativistico nell'offset d'orologio del satellite	69
4.2	Le osservazioni di codice (o pseudorange, pseudo distanza)	69
4.2.1	L'ambiguità delle osservazioni C/A	71
4.3	Le osservazioni di fase	71
4.4	L'errore di osservazione, i disturbi e gli errori di modellizzazione	74
4.5	Il disturbo atmosferico	74
4.5.1	Il disturbo ionosferico	76
4.5.2	Il disturbo troposferico (T)	79
4.6	L'equazione di osservazione finale	81
4.7	Le osservazioni complessivamente disponibili da un ricevitore a un satellite a un'epoca e le loro combinazioni	85
4.7.1	La combinazione Geometry Free	87
4.7.2	La combinazione Wide Lane	88
4.7.3	La combinazione Narrow lane	89
4.7.4	La combinazione Ionospheric Free	90
4.8	L'andamento nel tempo delle osservazioni da un ricevitore a un satellite	93
4.9	Il filtro di Hatch	94
5.	Il posizionamento assoluto	97
5.1	La soluzione di Bancroft	97
5.2	La soluzione ridondante in singola epoca mediante osservazioni di codice	100
5.3	Un'analisi degli errori di modello	101
5.4	La soluzione statica mediante osservazioni di codice	108
5.5	Il posizionamento assoluto mediante osservazioni di fase	108
5.6	Il sistema di riferimento in cui viene fornita la stima	112
5.7	Osservazioni sulla linearizzazione della distanza fra satellite e ricevitore	113

6.	Il posizionamento relativo	115
6.1	Le differenze singole	115
6.2	Le differenze doppie (DD)	117
6.3	Il contenuto geometrico delle differenze singole e doppie	118
6.4	Discussione dei termini presenti nella differenza doppia	121
6.5	L'offset d'antenna	122
6.6	La costruzione del vettore di differenze doppie relative a un'epoca di osservazione	125
6.6.1	La propagazione del rumore di misura: la matrice di covarianza delle osservazioni di differenza doppia	126
6.7	Introduzione all'elaborazione delle differenze doppie	128
6.8	La soluzione di solo codice in singola epoca	130
6.9	L'elaborazione di un rilievo statico in doppia frequenza	131
6.9.1	La preelaborazione dei dati	131
6.9.2	L'identificazione e la rimozione dei cycle slip	132
6.9.3	La stima della ionosfera	134
6.9.4	La stima delle ambiguità	135
6.9.5	Il fissaggio delle ambiguità	138
6.10	La stima finale della base	142
6.10.1	La stima finale per le sessioni brevi	143
6.10.2	La stima finale per le sessioni lunghe	143
6.10.3	Le accuratezze tipicamente ottenibili	146
6.11	L'elaborazione di un rilievo cinematico	147
6.12	Il controllo dei risultati di una postelaborazione di base	150
6.13	Il rilievo in tempo reale mediante osservazioni di fase	150
6.14	Il sistema di riferimento per il posizionamento relativo	152
7.	I servizi di posizionamento	155
7.1	Le stazioni permanenti	156
7.2	I servizi di rete per il tempo reale	157
7.3	L'approccio di rete alla modellizzazione degli errori	157
7.3.1	La riduzione delle incognite di stazione.	159
7.3.2	La modellizzazione e la stima degli errori e dei disturbi	160
7.3.3	La distribuzione dei modelli	161
7.4	Il sistema di riferimento per un servizio di posizionamento	164
7.5	Riepilogo finale dei prodotti tipicamente distribuiti da un servizio di posizionamento	166
7.5.1	Il formato di distribuzione dei dati per la postelaborazione	167
7.5.2	Il formato di distribuzione dei dati e dei prodotti per il tempo reale	167
7.6	I servizi di posizionamento in Italia e la sperimentazione del 2005	168
7.7	Alcune indicazioni all'utilizzo dei prodotti in tempo reale	170
8.	Il rilievo e l'inquadramento di reti	173
8.1.1	Lo schema a un solo ricevitore, o a ricevitori indipendenti	175

8.1.2	Lo schema a più ricevitori, a stella gerarchica	178
8.1.3	Altri schemi a più ricevitori	179
8.2	Il rilievo della rete	179
8.3	La postelaborazione delle basi	189
8.4	Introduzione all'analisi di rete	190
8.5	L'analisi delle ripetibilità e delle chiusure	190
8.6	La compensazione di rete	193
8.7	Le regioni di confidenza per le stime dei punti	197
8.7.1	L'ellissi di errore in planimetria	198
Appendice 1. I minimi quadrati		201
A1.1	I minimi quadrati per il caso non lineare	202
A1.1.1	La linearizzazione delle osservazioni di distanza	203
A1.2	La deficienza di rango: il problema e le soluzioni	205
A1.2.1	L'identificazione del nucleo della matrice disegno	207
A1.2.2	La soluzione di un problema con deficienza di rango	208
A1.2.2.1	La soluzione a minimi vincoli deterministici	209
A1.2.2.2	La soluzione a minimi vincoli stocastici	210
A1.2.2.3	La soluzione di Tykonov	210
A1.2.2.4	Applicazione delle soluzioni al triangolo di livellazione	211
A1.2.3	La soluzione sovravincolata, ovvero l'inquadramento di sottoreti	214
A1.2.4	Due definizioni e una nota	215
A1.3	La stima dei minimi quadrati	215
A1.3.1	La stima sequenziale	216
A1.4	La verifica del modello	217
A1.4.1	Il test del χ^2 o test globale sul modello	220
A1.4.2	Il test locale sulla singola osservazione	221
A1.4.3	L'affidabilità interna del test sulla singola osservazione	223
A1.4.4	L'affidabilità esterna del test sulla singola osservazione	224
A1.4.5	La prassi di esecuzione del test locale sulla singola osservazione	225
A1.5	La ripesatura delle osservazioni	226
A1.6	L'accuratezza dei parametri stimati	227
Bibliografia, acronimi e siti		229

Prefazione

I fondamentali del GPS vuole analizzare e illustrare le problematiche, le metodologie e le accuratèzze del posizionamento satellitare nelle applicazioni che vanno dalla navigazione in tempo reale al monitoraggio di alta accuratezza per reti geodetiche locali; il testo si concentra sul problema dei sistemi di riferimento e sull'elaborazione delle osservazioni nei suoi diversi passi, dal dato grezzo alla compensazione di reti: il livello della trattazione è quello della Geomatica per l'ambito universitario. Il volume è stato scritto nell'ambito del Progetto *Galileo e il posizionamento satellitare modernizzato*, finanziato nel 2006 dal Ministero dell'Università e della Ricerca come Progetto di Rilevante Interesse Nazionale. Insieme ai prodotti di ricerca, il progetto prevedeva la produzione di documentazione sul GNSS da distribuire liberamente: grazie ai fondi allocati sul finanziamento, sono state stampate 150 copie cartacee del presente volume, che verranno distribuite a colleghi e biblioteche. I pdf del testo sono viceversa scaricabili liberamente dal sito

<http://geomatica.como.polimi.it/fondamentalgps>

Per effettuare lo scaricamento dei file è semplicemente richiesta l'iscrizione al sito: ciò serve innanzitutto per disporre di una statistica sul volume degli interessati e dei lettori, ma anche per poter informare questi ultimi sugli aggiornamenti che effettuerò al materiale e ciò tocca un aspetto rilevante del presente progetto di distribuzione via web. La lettura su materiale cartaceo, la possibilità di annotare ed evidenziare a margine sono elementi insostituibili dello studio; però vi sono tematiche molto settoriali e moderne, quindi intrinsecamente con poca diffusione e in evoluzione molto rapida: la loro cristallizzazione su libro rischia di invecchiare rapidamente, e la scarsa tiratura impedisce un frequente aggiornamento del volume. Almeno allo stato attuale e in un ambito nazionale il posizionamento satellitare è uno di tali argomenti e perciò mi ha attratto l'ipotesi di pubblicare il materiale via web: in tal modo mi sarà possibile effettuarne un aggiornamento almeno ogni due anni, perlomeno al passo con gli sviluppi tecnologici dell'argomento.

Como, 9 giugno 2009,
Ludovico Biagi

Capitolo 1.

Introduzione: il GNSS e le tecniche di posizionamento

L'acronimo GNSS, ovvero Global Navigation Satellite Systems, indica l'insieme di tutte le costellazioni di satelliti artificiali terrestri per la navigazione degli utenti, ovvero quelle applicazioni in cui è necessario stimare la posizione con elevata frequenza di aggiornamento e in tempo reale: l'ausilio di rotta per la guida di autoveicoli rappresenta la più diffusa applicazione; meno diffuse, ma naturalmente molto più delicate in termini di accuratezza e affidabilità richieste, sono la navigazione degli aerei e il controllo di instradamento sulle linee ferroviarie per i treni; esistono poi anche numerose applicazioni molto specialistiche: ad esempio il controllo di mezzi in cantiere o il tracciamento di reti tecnologiche. In termini di mercato, la navigazione in tempo reale con requisiti di accuratezza di alcuni metri costituisce la prima applicazione del GNSS, può essere effettuata con ricevitori economici e richiede metodi di elaborazione del dato semplici; ben più articolato è il caso delle applicazioni che abbiano requisiti superiori, diciamo di alcuni decimetri. Esistono poi altre applicazioni, di tipo topografico e geodetico, le quali richiedono accuratezze molto maggiori, dal decimetro a meglio del centimetro: fra queste, a titolo d'esempio citiamo:

- il rilievo di punti per la cartografia a grande scala,
- la determinazione di punti d'appoggio per l'aerofotogrammetria,
- il monitoraggio nel tempo di strutture, come le dighe, o di fenomeni naturali, come le frane,
- il monitoraggio delle deformazioni geodinamiche alla scala regionale,
- il monitoraggio globale della tettonica a placche.

Tali applicazioni costituiscono evidentemente un settore molto ristretto per il mercato GNSS, si stima circa del 1%, rispetto alle applicazioni di navigazione; peraltro sono quelle verso le quali si è mossa una cospicua attività di ricerca, di base e applicata, degli ultimi venti anni: ciò per alcuni motivi. Innanzitutto tali applicazioni spingono al massimo le potenzialità del GNSS, e la ricerca finalizzata ad esse ha ricadute anche in discipline eterogenee, ad esempio la fisica delle antenne per la descrizione dei centri di fase, il comportamento dell'atmosfera per la propagazione del segnale, la statistica per l'analisi delle serie temporali delle stazioni permanenti. Inoltre tali applicazioni muovono unitariamente più economia rispetto al mercato di massa, proprio perché sono di nicchia, legate a ben specifiche attività professionali e con requisiti di elevata accuratezza; per alcune di esse le tecniche di osservazione tradizionali o comunque concorrenti al GNSS hanno costo elevatissimo, e dunque si giustifica la spinta e lo studio per l'introduzione di questo nuovo metodo.

Lo scopo sintetico di questo volume è quello di fornire un supporto teorico e pratico all'utente del GNSS: si è perciò tentato di descrivere con un ragionevole livello di dettaglio la teoria, il metodo e le elaborazioni che portano dalle osservazioni alla

determinazione delle posizioni, per le diverse classi di accuratezza a cui si può lavorare, ovvero da alcuni metri a meglio del centimetro. Poiché non lo si riteneva rilevante in questo contesto, si è scelto di escludere quasi completamente dalla discussione tutta la teoria e le metodologie sottostanti alla messa in orbita dei satelliti, al monitoraggio delle loro orbite, al loro funzionamento, alla generazione del segnale e alla sua lettura da parte dei ricevitori. Inoltre si consideri che i modi di utilizzo e le tecniche di elaborazione delle osservazioni sono molteplici: si è scelto di illustrare solo i metodi con rilevanza per un libro di testo, escludendo quelli più marginali. Per fornire una prima idea della molteplicità delle tecniche di rilievo GPS e delle relative accuratezze, la tabella al termine dell'introduzione fornisce alcuni elementi di sintesi.

Il progetto NAVSTAR GPS (NAVigation Satellite Timing And Ranging Global Positioning System, tipicamente abbreviato con GPS) è il primo nato, negli anni '70, ad opera del Dipartimento della Difesa degli Stati Uniti d'America; il primo satellite è stato lanciato nel 1978 e il sistema si è gradualmente sviluppato sino al 1995, anno in cui è stata ufficialmente dichiarata la sua piena operatività: con tale termine si intende che la navigazione in tempo reale è garantita 24 ore su 24 (H24) ovunque sulla Terra. Maggiore dettaglio alla storia del GPS, al suo stato attuale e al suo futuro viene dedicato nel Capitolo 3; in tale capitolo si spiegano inoltre le principali caratteristiche dei satelliti, le loro orbite, la struttura e il contenuto del loro segnale.

Nel presente volume si discute unicamente la costellazione GPS, perché è l'unica oggi pienamente operativa per la navigazione; inoltre il suo utilizzo per applicazioni topografiche e geodetiche è ormai un argomento ben consolidato, sul quale è possibile appoggiare alcune nozioni didattiche: perciò si presta a una trattazione in un libro di testo.

In questa introduzione si vogliono però perlomeno ricordare gli altri sistemi GNSS che in parte già ora e completamente in un prossimo futuro interopereranno con il GPS. Il progetto della costellazione russa GLONASS (Global'naja Navigacionnaja Sputnikovaja Sistema) partì quasi contemporaneamente al GPS e fino agli anni '90 il lancio dei satelliti proseguì regolarmente; a causa della seguente crisi sovietica, dapprima venne abbandonato il completamento del progetto e poi anche la manutenzione di quanto già esistente. Solo nel 2001 il governo russo ha dichiarato ufficialmente la volontà di riportare GLONASS alla piena operatività entro il 2010-2011; da allora lo sviluppo è proseguito con regolarità e oggi la costellazione conta 17 dei 24 satelliti previsti: si può quindi affermare che il GLONASS, pur non essendo ancora pienamente operativo H24, garantisce già ausilio al GPS.

Nel 2003 l'Unione Europea ha formalmente approvato il progetto della costellazione Europea, chiamata GALILEO; la piena operatività era inizialmente stimata per il 2008 ma tale scadenza è stata ripetutamente rinviata: attualmente sono in orbita due satelliti di test (Giove-A e Giove-B) e il completamento del sistema, con 30 satelliti in orbita, è ufficialmente previsto nel 2013.

Esistono poi sistemi satellitari di navigazione finalizzati alla copertura di aree specifiche del globo. La Cina ha una sua costellazione, chiamata Beidou Satellite Navigation and

Positioning System, che attualmente è operativa H24 sul rettangolo geografico compreso fra 70° Est, 5° Nord e 140° Est, 55° Nord; è annunciata l'evoluzione della copertura del sistema da locale a globale, con il progetto COMPASS di cui si prevede il completamento prima del 2020. Il Giappone ha la costellazione QZSS che attualmente offre interoperabilità al GPS sul Giappone e sulle regioni limitrofe. L'India infine ha avviato il progetto della costellazione Indian Regional Navigational Satellite System (IRNSS).

La navigazione GNSS in tempo reale avviene mediante un semplice processo. All'accensione il ricevitore riconosce tutti i satelliti visibili e acquisisce da essi un certo numero di segnali e informazioni, che servono per calcolare le posizioni dei satelliti stessi e il tempo di tragitto dei segnali dai satelliti al ricevitore. Moltiplicando i tempi di tragitto per la velocità di propagazione dei segnali elettromagnetici nel vuoto si derivano le misure di distanza dai satelliti al ricevitore: date le posizioni dei satelliti e le misure di distanza da questi al ricevitore, è finalmente possibile stimare la posizione del ricevitore. Una volta effettuate le operazioni di accensione e di decodifica iniziale dei segnali, l'acquisizione dei segnali è un processo a flusso continuo e le elaborazioni possono potenzialmente avvenire in tempo reale, con altissima frequenza di aggiornamento, tipicamente da 1 a 10 Hz.

La generazione del segnale nei satelliti e le equazioni di osservazione dei ricevitori sono discusse e ricavate nel capitolo 4. In seguito, il capitolo 5 tratta il cosiddetto posizionamento assoluto: si noti che il nome del metodo è dovuto al fatto che il processo di stima è autonomo e ad una via, nel senso che il ricevitore stima la propria posizione semplicemente utilizzando i segnali ricevuti dai satelliti e non ha necessità di comunicazioni attive, né verso i satelliti stessi, né verso altre stazioni a terra. Ad ogni passo dell'elaborazione intervengono numerose fonti di errore: si può affermare che il posizionamento mediante ricevitori di basso costo attualmente ha accuratezza dell'ordine dei 10 metri, che possono essere migliorati attraverso due vie. Innanzitutto si possono costruire opportune combinazioni di osservazioni che permettono di abbattere l'effetto delle principali fonti di errore, sino ad arrivare ad accuratezze dell'ordine del metro, ma che richiedono ricevitori di costo più elevato: per il caso del GPS tali combinazioni sono discusse nel presente volume.

Al fine di migliorarne le stime, il GNSS può anche essere integrato da altri sensori o informazioni esterne: ad esempio in navigazione si possono utilizzare sensori inerziali per filtrare il rumore dalle stime di traiettoria. Oppure, nel caso di navigazione di automezzi, possono essere utilizzate informazioni cartografiche ausiliarie che vincolino la stima della quota a un modello digitale della superficie o la posizione orizzontale alle linee definite da un grafo stradale. L'integrazione del GNSS con altre fonti di informazione riguarda un ambito ben più esteso di quello cui si rivolge il presente volume e quindi non viene qui discusso.

Il capitolo 6 discute il posizionamento GNSS relativo, tecnica che ha costituito il punto di svolta per ottenere le massime accuratezze. Si parte dalla costruzione di particolari combinazioni di osservazioni fra due ricevitori, le cosiddette differenze doppie, che

permette l'eliminazione di alcune incognite di disturbo e una significativa riduzione degli errori di osservazione. La riduzione degli errori aumenta ovviamente l'accuratezza dei risultati, e questo è un primo vantaggio delle differenze doppie, ma è l'eliminazione delle incognite di disturbo a rappresentare il vero punto cardinale: grazie ad essa è infatti possibile il cosiddetto fissaggio delle ambiguità, che permette di ottenere accuratezze di alcuni centimetri con poche epoche di osservazione. Come detto le differenze doppie sono costruite dalle osservazioni di due ricevitori: la loro elaborazione permette la stima del vettore di base (in seguito semplicemente base) fra i due ricevitori e richiede che uno dei due, detto di riferimento, occupi un punto di posizione nota: perciò si dice che l'elaborazione delle differenze doppie porta al posizionamento relativo, ovvero alla stima di posizioni di punti incogniti rispetto alle coordinate note di un punto di riferimento.

In funzione delle accuratezze richieste nel posizionamento relativo, si hanno molteplici metodi di rilievo e di elaborazione: innanzitutto si può operare in modalità cinematica, ovvero con il ricevitore incognito in moto continuo e la stima di una base per ogni epoca, oppure statico, ovvero con la stima di un'unica base per tutte le epoche di osservazione; un rilievo statico può prevedere un'occupazione di pochi minuti, e si dice statico rapido, oppure arrivare ad occupazioni che durino anche intere giornate. Inoltre si deve porre un'ulteriore distinzione: tradizionalmente il posizionamento relativo veniva effettuato mediante postelaborazione dei dati, e questa è ancora la regola per le applicazioni di massima accuratezza; oggi però nelle applicazioni topocartografiche si sta affermando la prassi del cosiddetto rilievo in tempo reale: i due ricevitori scambiano reciprocamente le osservazioni, ad esempio via GSM, e la stima della base viene fatta epoca per epoca dal firmware di uno dei due. Per sintetizzare l'ampio spettro delle accuratezze ottenibili, l'elaborazione cinematica delle differenze doppie di codice in singola frequenza e in tempo reale garantisce accuratezze dell'ordine del metro per ricevitori entro i 50 km di distanza; sulle stesse distanze il posizionamento statico mediante postelaborazione di fasi in doppia frequenza permette accuratezze anche migliori del centimetro per sessioni di una giornata.

Tali aspetti sono discussi nel capitolo 6, con l'obiettivo primario di discutere nel massimo dettaglio i due casi principali di elaborazione e di fornire alcune indicazioni generali per tutti gli altri. Nel capitolo non sono presenti la descrizione analitica degli algoritmi di elaborazione per il rilievo cinematico e per il rilievo in tempo reale, sui quali si fornisce comunque una discussione qualitativa; manca inoltre la discussione della postelaborazione di reti in modalità multibase alla scala regionale, ovvero di reti per applicazioni geodinamiche con basi di lunghezza superiore ai 100 km. Tali aspetti sarebbero evidentemente rilevanti per un discussione completa, ma ne mancava il tempo e lo spazio: saranno sicuramente inclusi nella prossima revisione del volume.

I problemi fondamentali del rilievo delle basi sono essenzialmente di costo e complessità. Quanto al costo, per stimare le coordinate di un punto è richiesto l'utilizzo di due ricevitori e per minimizzare i tempi di rilievo i due ricevitori devono essere reciprocamente quanto più vicini possibile: quindi, in assenza di supporto esterno, si devono possedere, installare e presidiare sia il ricevitore di riferimento sia quello

mobile. Quanto alle complessità, queste riguardano le scelte di postelaborazione in ufficio per le applicazioni di maggiore accuratezza, come anche la soluzione sul campo di eventuali problemi tecnici del tempo reale per le applicazioni topocartografiche. Tali problemi sono particolarmente limitanti nel rilievo topocartografico alla scala locale, ove gli strumenti topografici tradizionali pongono meno problemi di costo e complessità; per ovviarvi è nata e si è sviluppata in molte nazioni l'idea di reti di stazioni permanenti per servizi di posizionamento. Una rete di stazioni permanenti si compone di un insieme di ricevitori GNSS stabilmente monumentati in acquisizione continua; le stazioni sono coordinate da un centro di elaborazione, che a un primo livello ne effettua la manutenzione, il monitoraggio di qualità e delle coordinate. Le reti di stazioni permanenti sono nate quasi 20 anni fa per scopi di monitoraggio, dalla scala spaziale globale del sistema di riferimento a quella locale del controllo; le reti per i servizi di posizionamento naturalmente possono assolvere anche a scopi di monitoraggio, ma sono soprattutto finalizzate a distribuire coordinate, dati e prodotti di rete agli utenti, che permettano di utilizzare il GNSS per scopi di rilievo, tipicamente nella modalità del posizionamento relativo in tempo reale, abbattendone i costi e cercando di semplificarne le complessità. I principi di funzionamento e di calcolo dei servizi di posizionamento e i tipi di prodotti che essi erogano sono discussi nel capitolo 7, che viene chiuso da alcune indicazioni operative sull'utilizzo di tali prodotti.

Ricordiamo che l'inquadramento di una rete locale consiste nella stima delle coordinate dei suoi punti rispetto a una rete di ordine superiore scelta come riferimento. Il problema ha radici storiche ma non ha perso di attualità con la nascita dei servizi di posizionamento: questi infatti nascono in primis per diffondere l'utilizzo del GNSS nel rilievo in tempo reale per applicazioni topocartografiche; però, se correttamente monitorati, costituiscono anche un tramite chiarissimo fra la rete permanente globale e le reti locali, poiché facilitano il loro inquadramento e soprattutto il monitoraggio dei loro punti di riferimento locali, che altrimenti debbono essere considerati stabili apriori. Nel capitolo 8 viene quindi trattato l'inquadramento di reti GNSS, anche con attenzione alla nascita dei servizi di posizionamento: in particolare si danno alcune indicazioni sulla progettazione di un rilievo di rete, sull'esecuzione del rilievo stesso e su come condurre la postelaborazione delle basi; infine viene trattata la teoria e il metodo per condurre la compensazione delle basi. Si noti che il capitolo non tratta il caso delle reti miste, ovvero delle reti rilevate sia mediante GNSS sia mediante strumenti topografici tradizionali, poiché questo riguarda un ambito di integrazione fra strumenti e informazioni eterogenee che, come già detto, è più esteso rispetto all'obiettivo presente del volume.

Sin qui si è sviluppato un discorso sui vari aspetti del posizionamento GNSS: su questo si è cercato di costruire un percorso logico nelle sue varie sfaccettature metodologiche e tecniche; si è ommesso un concetto fondamentale, per riassumere finalmente e in modo definitivo la sua importanza. Il posizionamento, in qualunque modalità esso venga effettuato, è stima di quantità fisiche e percepibili, ovvero le posizioni, per le quali si ricorre però a una descrizione matematica e numerica, ovvero le coordinate; il passaggio

logico non è banale e richiede innanzitutto la definizione, la formalizzazione e la realizzazione di quelli che chiamiamo sistemi di riferimento che per la loro rilevanza vengono discussi all'inizio del libro. La definizione di un sistema di riferimento, la descrizione matematica delle coordinate di un punto e del vettore fra due punti sono operazioni semplici finché si rimane in un ambito geometrico; l'applicazione di tali concetti alla descrizione della posizione di punti sulla superficie terrestre si scontra però con due problemi principali, intrinseci alla natura del nostro pianeta e delle osservazioni geodetiche utilizzate. Innanzitutto la superficie del pianeta è in continuo moto, dovuto alla geodinamica della crosta terrestre e quindi la posizione dei punti localmente stabili cambia continuamente in senso assoluto: le stime delle posizioni devono anche essere temporeferenziate o, in alternativa, devono essere stimati i loro modelli nel tempo, ovvero le loro equazioni caratteristiche. Inoltre il GPS permette l'osservazione di differenze di posizione e non di posizioni assolute: come già introdotto, si osservano distanze fra satelliti e ricevitori, oppure basi fra ricevitori: ciò implica una deficienza di rango nella stima delle posizioni, che richiede l'adozione di regole e misure per definire e realizzare il sistema di riferimento. Lo scopo del capitolo 2 è quello di introdurre il problema dei sistemi di riferimento, cercando di porlo in modo semplice ma completo; innanzitutto si discutono le regole formali; si introduce quindi il concetto di sistema di riferimento geodetico come rete di punti di coordinate già stimate, rispetto ai quali è possibile stimare ulteriori sottoreti: in particolare viene posta la distinzione fra sistemi di riferimento statici e sistemi di riferimento permanenti con attenzione alle relative accuratezze. Vengono quindi descritti i casi rilevanti all'utente italiano del GNSS, ovvero i sistemi di riferimento globale, quello europeo nella sua realizzazione continentale e italiana e le relazioni che intercorrono fra essi.

I minimi quadrati costituiscono il tassello fondamentale nell'elaborazione dei dati ma ovviamente la loro discussione fuoriesce dall'ambito di questo volume. Peraltro vi sono alcuni aspetti particolarmente rilevanti al nostro interesse, ovvero quelli relativi alla deficienza di rango, per il problema dei sistemi di riferimento, quello della stima sequenziale, per tutte le applicazioni in tempo reale, e quello del controllo di affidabilità e di rilevazione di errori grossolani per l'analisi dei risultati di una compensazione di rete. Si è scelto quindi di dedicare un'appendice alla discussione di tali aspetti.

Tipo rilievo	Osservabili	Requisiti e note	Ordine di accuratezza
Posizionamento cinematico assoluto in tempo reale: Capitolo 5	Codici in singola frequenza	Generica navigazione con ricevitori di diverse fasce di costo	dieci metri
	Codici in doppia frequenza		circa un metro
	Codici in doppia frequenza lisciati con le fasi (filtro di Hatch)		meglio del metro
Posizionamento statico assoluto in postelaborazione: Capitolo 5	Fasi in doppia frequenza	Metodi ancora oggetto di sperimentazione universitaria	alcuni centimetri
Posizionamento cinematico relativo in tempo reale: Capitoli 6 e 7	Codici in singola frequenza	Interdistanza fra stazione di riferimento e mobile indicativamente entro 100 km	uno o due metri
	Codici in doppia frequenza		meglio del metro
	Codici in doppia frequenza lisciati con le fasi (filtro di Hatch)		alcuni decimetri
	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 20 km oppure presenza di un servizio di posizionamento	meglio del decimetro
Posizionamento cinematico relativo mediante postelaborazione: Capitolo 6	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 20 km oppure presenza di un servizio di posizionamento	alcuni centimetri
Posizionamento statico rapido (pochi minuti) relativo in tempo reale: Capitoli 6 e 7			

(La tabella continua ed è descritta in pagina seguente)

Tipo rilievo	Osservabili	Requisiti	Ordine di accuratezza
Posizionamento statico rapido (pochi minuti) relativo mediante postelaborazione: Capitoli 6 e 8	Fasi in singola frequenza	Interdistanza fra stazione di riferimento e mobile entro 5-10 km	meglio del decimetro
	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 20 km	alcuni centimetri
Posizionamento statico (occupazione di un'ora) relativo mediante postelaborazione: Capitoli 6 e 8	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 100 km	alcuni centimetri
		Interdistanza fra stazione di riferimento e mobile entro 10 km	uno/due centimetri
Posizionamento statico prolungato (occupazione permanente) relativo mediante postelaborazione: Capitoli 6 e 8	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile anche superiore a 100 km	meglio del centimetro

Tabella 1.1. Elenco schematico dei metodi di rilievo del GNSS, delle osservabili utilizzate, dei requisiti e delle relative accuratezze attualmente ottenibili. L'elenco non vuole avere carattere di completezza ma solo esemplificativo; le accuratezze sono quelle relative a discrete condizioni di osservazione e corretta elaborazione delle osservazioni stesse.

Capitolo 2. I sistemi di riferimento e di coordinate

Innanzitutto viene presentata una definizione di sistema di riferimento tridimensionale e viene discussa la modellizzazione matematica generale della trasformazione fra sistemi di riferimento: si presentano quindi la formulazione linearizzata usualmente adottata in geodesia e il metodo di stima della trasformazione a partire da punti doppi. Il formalismo geometrico viene quindi collegato al problema della stima di posizioni di punti sulla superficie terrestre, ovvero al problema geodetico e topografico: ciò richiede innanzitutto una breve elencazione dei moti del pianeta, soprattutto per introdurre l'idea che non esistono punti della superficie terrestre la cui posizione sia costante nel tempo. I sistemi di riferimento geodetici sono realizzati mediante reti di punti fondamentali, e tali reti possono essere categorizzate in reti permanenti e reti statiche: la descrizione delle loro caratteristiche occupa la parte centrale del capitolo e quindi segue l'elenco dei principali sistemi di riferimento di interesse per il GPS, dalla scala globale sino allo specifico caso italiano. Infine vengono presentati i diversi metodi di rappresentazione numerica delle posizioni di un punto all'interno di un sistema di riferimento e le trasformazioni di coordinate. Il capitolo si chiude con una breve descrizione dei sistemi di riferimento per le scale temporali.

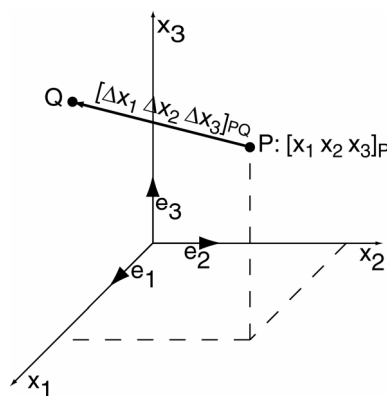


Figura 2.1. Versori ortogonali, posizione di un punto e vettore fra due punti

Un sistema di riferimento viene definito mediante tre versori $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ con origine comune, ortogonali e tali da formare una terna destrorsa:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}, \quad \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$$

Si noti che la lunghezza dei versori implica intrinsecamente la definizione della scala del sistema di riferimento. Le coordinate di un punto P sono rappresentate dalle

lunghezze delle sue proiezioni ortogonali sui tre assi e vengono usualmente indicate con la terna

$$\mathbf{x}_P = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_P, \quad \|\mathbf{x}\| = \sqrt{x_{1_P}^2 + x_{2_P}^2 + x_{3_P}^2} \quad (2.1)$$

Il vettore fra due punti P e Q è dato dalla

$$\Delta \mathbf{x}_{PQ} = \begin{bmatrix} x_{1_Q} - x_{1_P} \\ x_{2_Q} - x_{2_P} \\ x_{3_Q} - x_{3_P} \end{bmatrix}, \quad \Delta \mathbf{x}_{PQ} = -\Delta \mathbf{x}_{QP}, \quad (2.2)$$

$$\|\Delta \mathbf{x}_{PQ}\| = \sqrt{(x_{1_Q} - x_{1_P})^2 + (x_{2_Q} - x_{2_P})^2 + (x_{3_Q} - x_{3_P})^2}$$

2.1 La trasformazione fra sistemi di riferimento

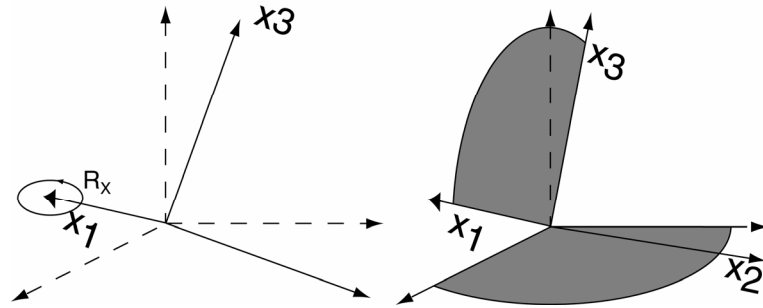
Siano SRI ($[\mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3]_I$) e SRII ($[\mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3]_{II}$) due sistemi di riferimento con diversa origine, diverso orientamento degli assi e diversa unità di misura di lunghezza; siano $\mathbf{t} = [t_1 \ t_2 \ t_3]^T$ le coordinate dell'origine di SRI rispetto a SRII, $\mathbf{R}$ la matrice di rotazione da SRI a SRII, λ il fattore di scala, ovvero il rapporto fra le unità di misura di SRI e SRII.

Dato un punto P , di coordinate $[x_1 \ x_2 \ x_3]_{P,I}^T$ nel SRI, le sue coordinate in SRII sono date dalla

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{P,II} = \mathbf{t} + \lambda \mathbf{R} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{P,I} \quad (2.3)$$

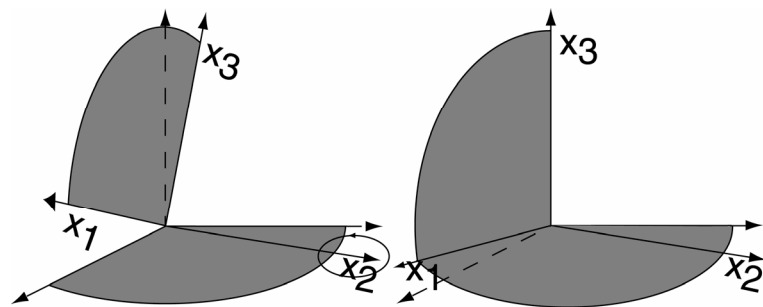
La rotazione può essere realizzata mediante la composizione di tre opportune rotazioni piane rispetto ai tre assi; questo non è l'unico modo possibile, ma è la scelta convenzionale nell'ambito geodetico ed è il metodo qui descritto. La prima rotazione avviene rispetto all'asse x_1 , ed è tale da portare x_2 e x_3 rispettivamente sui piani x_1x_2 e x_1x_3 finali:

$$\mathbf{R}_1(r_1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos r_1 & \sin r_1 \\ 0 & -\sin r_1 & \cos r_1 \end{bmatrix}$$


 Figura 2.2 La rotazione piana intorno all'asse x_1

La seconda rotazione avviene rispetto all'asse x_2 ed è tale da portare x_3 nella direzione finale, x_1 sul piano x_1x_2 finale:

$$\mathbf{R}_2(r_2) = \begin{bmatrix} \cos r_2 & 0 & -\sin r_2 \\ 0 & 1 & 0 \\ \sin r_2 & 0 & \cos r_2 \end{bmatrix}$$


 Figura 2.3 La rotazione piana intorno all'asse x_2

La terza e ultima rotazione è rispetto all'asse x_3 , tale da portare x_1 e x_2 nelle loro direzioni finali:

$$\mathbf{R}_3(r_3) = \begin{bmatrix} \cos r_3 & \sin r_3 & 0 \\ -\sin r_3 & \cos r_3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

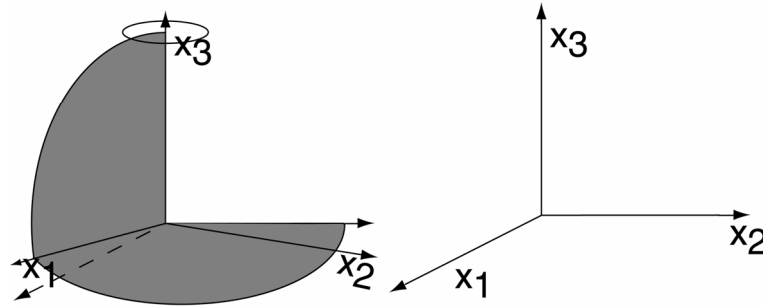


Figura 2.4 La rotazione piana intorno all'asse x_3

La matrice di rotazione finale è data dal prodotto:

$$\mathbf{R}(r_1, r_2, r_3) = \mathbf{R}_3(r_3)\mathbf{R}_2(r_2)\mathbf{R}_1(r_1)$$

ovvero dalla

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (2.4)$$

ove

$$\begin{aligned} r_{11} &= \cos(r_2)\cos(r_3) \\ r_{12} &= \cos(r_2)\sin(r_3) + \sin(r_1)\sin(r_2)\cos(r_3) \\ r_{13} &= \sin(r_1)\sin(r_3) - \cos(r_1)\sin(r_2)\cos(r_3) \\ r_{21} &= -\cos(r_2)\sin(r_3) \\ r_{22} &= \cos(r_1)\cos(r_3) - \sin(r_1)\sin(r_2)\sin(r_3) \\ r_{23} &= \sin(r_1)\cos(r_3) + \cos(r_1)\sin(r_2)\sin(r_3) \\ r_{31} &= \sin(r_2) \\ r_{32} &= -\sin(r_1)\cos(r_2) \\ r_{33} &= \cos(r_1)\cos(r_2) \end{aligned}$$

Poiché la rotazione dipende da 3 angoli la trasformazione dipende in totale da 7 parametri. La rototraslazione con fattore di scala prende anche nome di trasformazione di Helmert oppure trasformazione di similarità.

2.1.1 Il caso di piccole rotazioni e fattore di scala

I principali sistemi di riferimento geodetici sono definiti in modo molto simile (Par. 2.3 e seguenti): geocentrici, con asse x_3 diretto secondo l'asse di rotazione terrestre e asse x_1 giacente in un meridiano convenzionale; essi differiscono fra loro per differenti definizioni di geocentro e asse di rotazione e differenti insiemi di misure con cui sono realizzati, ovvero differiscono di quantità molto piccole: in tal caso la precedente trasformazione può essere semplificata.

Per piccole rotazioni e fattore di scala prossimo all'unità ($r_i \cong 1''$, $\cos r_i \cong 1$, $\sin r_i \cong r_i$, $\lambda \cong 1.0$) la trasformazione può essere linearizzata in

$$\mathbf{R} \cong \begin{bmatrix} 1 & r_3 & -r_2 \\ -r_3 & 1 & r_1 \\ r_2 & -r_1 & 1 \end{bmatrix} = \mathbf{I} + \delta \mathbf{R}; \lambda = 1 + \mu, \mu \delta \mathbf{R} \cong 0$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{SR II} \cong \mathbf{t} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{SR I} + \delta \mathbf{R} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{SR I} + \mu \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{SR I} \quad (2.5)$$

In tal caso, la stima dei parametri di trasformazione fra due sistemi di riferimento è particolarmente semplice; è necessario disporre di punti le cui coordinate siano state misurate in entrambi i sistemi di riferimento, che sono detti punti doppi; per ogni punto doppio si può scrivere l'equazione

$$\mathbf{x}_{i,SR II} = \mathbf{t} + \mathbf{x}_{i,SR I} + \delta \mathbf{R} \mathbf{x}_{i,SR I} + \mu \mathbf{x}_{i,SR I}$$

La precedente può essere riscritta come

$$\mathbf{x}_{i,SR II} - \mathbf{x}_{i,SR I} = \begin{bmatrix} \mathbf{I} & [\mathbf{x}_{i,SR I} \times] & \mathbf{x}_{i,SR I} \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \mathbf{r} \\ \mu \end{bmatrix} = \mathbf{T}_i \mathbf{q}$$

ove

$$\mathbf{I} \text{ rappresenta la matrice identità } [3 \times 3], [\mathbf{x}_{i,SR I} \times] = \begin{bmatrix} 0 & -x_{3i_i} & x_{2i_i} \\ x_{3i_i} & 0 & -x_{1i_i} \\ -x_{2i_i} & x_{1i_i} & 0 \end{bmatrix}, \mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

Nel caso di n punti doppi si può scrivere la

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \dots \\ \mathbf{x}_n \end{bmatrix}_{II} - \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \dots \\ \mathbf{x}_n \end{bmatrix}_I = \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \\ \dots \\ \mathbf{T}_n \end{bmatrix} \mathbf{q}$$

Si ha quindi un sistema in 7 incognite: note le coordinate di n punti doppi in entrambi i sistemi di riferimento, è possibile scrivere $3n$ equazioni in 7 incognite; se $n \geq 3$ il sistema è caratterizzato da più equazioni che incognite e può essere risolto mediante minimi quadrati (App. 1).

2.2 Introduzione ai sistemi di riferimento in geodesia

Un sistema di riferimento in geodesia è un insieme di regole e misure per rispondere ai quesiti

Dove si trova un punto?

Quando è avvenuto un evento?

Le regole sono necessarie per eliminare la deficienza di rango intrinseca (App. 1) e naturalmente sono derivate dalla definizione geometrica; le misure applicano le regole e servono per distribuire coordinate.

Si noti che vi sono esigenze di spazio-referenziazione e di tempo-referenziazione e tradizionalmente si adottava una distinzione fra sistemi di riferimento nello spazio e sistemi di riferimento nel tempo. La distinzione non ha più senso: con i moderni strumenti si osservano tempi, dalle osservazioni di tempo si derivano osservazioni di distanza: quindi la scala dei tempi non è separabile dalla scala delle coordinate. Inoltre la forma della Terra cambia: si hanno sia fenomeni periodici come le maree terrestri, sia fenomeni irreversibili, dalla scala globale a quella locale, come ad esempio la geodinamica e la subsidenza. Rispetto ai fenomeni periodici si può definire una posizione media; considerando però i fenomeni irreversibili si può affermare che, in generale, qualunque punto del pianeta occupa una certa posizione solo per un istante infinitesimo: ovvero la spazio-referenziazione di un punto deve essere tempo-referenzata. Nel seguito vengono presentati i principali fenomeni cui è soggetto il nostro pianeta, dai suoi moti nello spazio alle sue deformazioni interne.

2.2.1 I moti del pianeta visti dallo spazio

Il nostro pianeta compie un'orbita ellittica intorno al Sole, che giace in un piano detto piano dell'eclittica: il moto è detto di rivoluzione e richiede 365.25 giorni solari. Il pianeta inoltre ruota intorno al proprio asse: il moto è detto di rotazione e richiede 1 giorno solare; l'asse di rotazione terrestre e l'asse ortogonale al piano dell'eclittica formano un angolo di circa $23^\circ 27'$. L'asse di rotazione terrestre compie due moti conici

nello spazio; il primo è detto moto di precessione, ha angolo al centro di $23^{\circ}27'$ e periodo di 25800 anni; il secondo è detto moto di nutazione: è un moto perturbativo intorno al moto di precessione, con ampiezza angolare di $9.2''$ e con periodo di 18.6 anni.

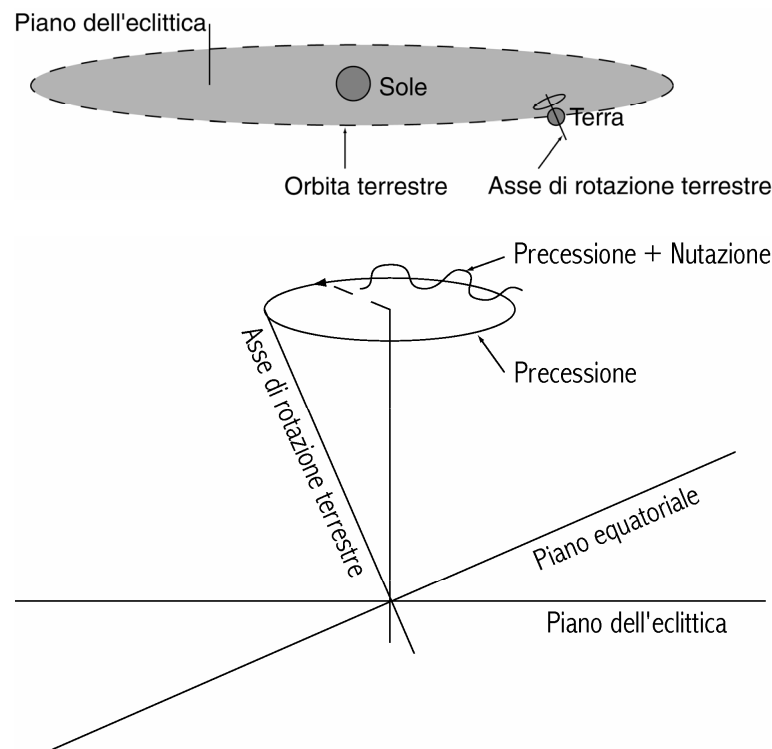


Figura 2.5. I moti della Terra vista dallo spazio

2.2.2 Gli altri moti rigidi del pianeta

Il pianeta “oscilla” rispetto al proprio asse di rotazione, ovvero la posizione del polo di rotazione rispetto alla superficie terrestre cambia nel tempo; tale moto, che prende nome di Chandler Wobble o Polar Motion, è approssimativamente conico, intorno a un punto convenzionale medio, detto Polo Convenzionale Terrestre (CTP), con periodo di circa 435 giorni e ampiezza di circa $0.1''$ – $0.2''$ (si noti che $0.1'' \cong 3m$).

La velocità angolare di rotazione terrestre non è costante nel tempo (Par. 2.9): con il termine LOD (Length Of Day) si definisce la differenza fra la lunghezza del giorno solare, legata alla rotazione terrestre, e la scala del tempo atomica, considerata stabile.

Col termine EOP (Earth Orientation Parameters), vengono accorpati i tre parametri che variano nel tempo: le due direzioni dell’asse istantaneo di rotazione rispetto al CTP e la LOD. Gli EOP vengono stimati da diversi organismi internazionali: ad esempio IERS e IGS (Par. 2.4).

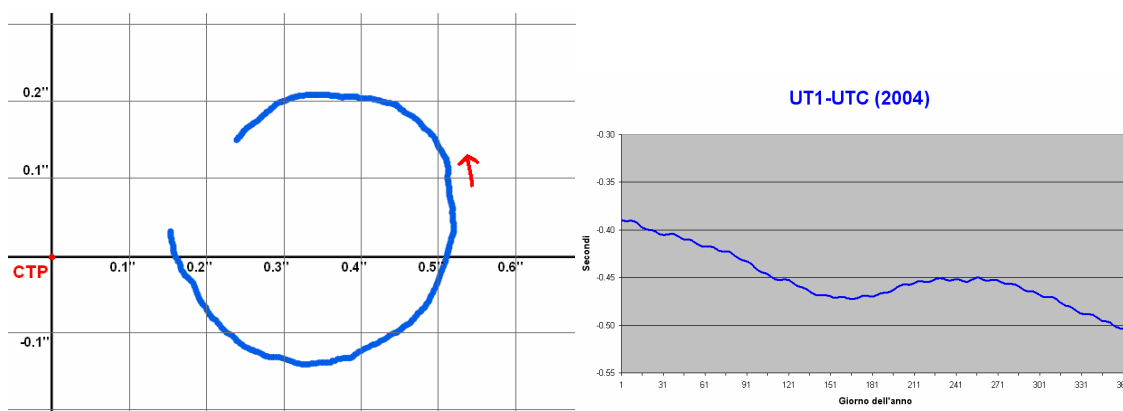


Figura 2.6. EOP per l'anno 2004. A sinistra, posizione dell'asse istantaneo di rotazione rispetto al polo convenzionale terrestre: $0.1'' \cong 3m$. A destra, LOD espressa come differenza fra durata del giorno solare (UT1) e scala dei tempi atomica (UTC): la durata del giorno solare aumenta.

IERS C04 POLE										15-JUN-06 18:31															
NUTATION MODEL : IAU2000										SUBDAILY POLE MODEL: IERS2000															
DATE	TIME	X-POLE	Y-POLE	UT1-UTC	GPS-UTC	RMS XP	RMS YP	RMS DT	DE-CPO	DP-CPO	RMS EP	RMS PS	DATE	TIME	X-POLE	Y-POLE	UT1-UTC	GPS-UTC	RMS XP	RMS YP	RMS DT	DE-CPO	DP-CPO	RMS EP	RMS PS
YYYY	MM	DD	HH	MM	(")	(")	(S)	(S)	REM	(")	(")	(S)	(")	(")	(")	(")	(S)	(S)	REM	(")	(")	(S)	(")	(")	(S)
2006	1	1	0	0	0.05271	0.38335	0.338834	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	2	0	0	0.05179	0.38302	0.338597	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	3	0	0	0.05097	0.38276	0.338113	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	4	0	0	0.05018	0.38244	0.337398	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	5	0	0	0.04961	0.38208	0.336521	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	6	0	0	0.04950	0.38178	0.335577	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	7	0	0	0.04957	0.38166	0.334679	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	8	0	0	0.04951	0.38160	0.333951	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	9	0	0	0.04951	0.38131	0.333476	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	10	0	0	0.04958	0.38117	0.333274	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	11	0	0	0.04927	0.38105	0.333310	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	12	0	0	0.04877	0.38071	0.333506	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	13	0	0	0.04854	0.38036	0.333768	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	14	0	0	0.04868	0.38002	0.334007	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	15	0	0	0.04921	0.37977	0.334109	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	16	0	0	0.04995	0.37969	0.334092	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	17	0	0	0.05040	0.37997	0.333954	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	18	0	0	0.05063	0.38042	0.333636	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	19	0	0	0.05072	0.38075	0.333103	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	20	0	0	0.05050	0.38098	0.332531	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	21	0	0	0.05024	0.38121	0.331863	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	22	0	0	0.05017	0.38126	0.331179	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	23	0	0	0.05024	0.38072	0.330539	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	24	0	0	0.05031	0.38024	0.329980	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	25	0	0	0.05020	0.38035	0.329508	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	26	0	0	0.05012	0.38058	0.329102	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	27	0	0	0.05031	0.38071	0.328710	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	28	0	0	0.05070	0.38098	0.328278	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	29	0	0	0.05115	0.38135	0.327725	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	30	0	0	0.05149	0.38170	0.326896	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	
2006	1	31	0	0	0.05161	0.38230	0.325730	14	C04	0.00000	0.00000	0.000000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	

Figura 2.7. Esempio di bollettino EOP rilasciato dallo IERS. Contiene le stime degli EOP e delle relative accuratze, oltre a altri parametri qui non discussi: le stime sono fornite a intervalli di 24 ore, per le ore 00.00 UTC di ogni giorno.

2.2.3 I moti non rigidi della massa terrestre: i fenomeni periodici

Come la superficie degli oceani, anche la superficie della parte solida del pianeta si deforma in accordo al continuo ciclo di variazione dell'attrazione gravitazionale di Sole e Luna: il fenomeno viene chiamato di marea terrestre e ha entità di alcuni decimetri. Vi è inoltre il fenomeno di carico oceanico, ovvero un fenomeno ciclico di deformazione della crosta terrestre dovuto a sua volta al ciclo di carico delle maree oceaniche: in prossimità delle coste questo può raggiungere valori di alcuni centimetri. I periodi fondamentali di questi fenomeni sono sottomultipli dei periodi di orbita apparente di Sole e Luna rispetto alla Terra: 12 ore, 24 ore, 28 giorni, 365 giorni.

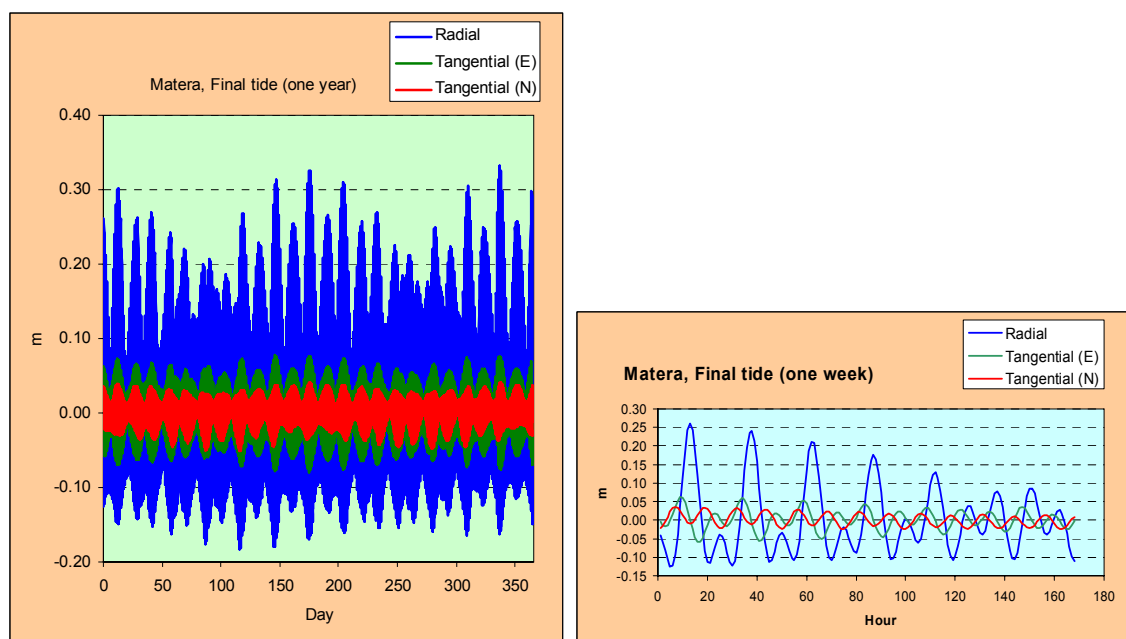


Figura 2.8. Marea terrestre presso la SP del Centro di Geodesia Spaziale di Matera: a sinistra campionamento, ogni ora, per un anno; a destra una magnificazione su una settimana; valori in metri.

	Nord	Est	Quota
Media (m)	0.00	0.00	0.00
SQM (m)	0.02	0.03	0.10
Min/Max (m)	-0.05/0.04	-0.08/0.08	-0.18/0.33

Tabella 2.1. Statistiche sull'effetto di marea presso la SP del Centro di Geodesia Spaziale di Matera: 2 anni, intervallo di campionamento di 1 ora: 17520 epoche.

2.2.4 La rimozione dei fenomeni periodici

I fenomeni periodici sono la risposta della Terra a forze periodiche: costituiscono un segnale geofisico, la cui analisi permette lo studio del nostro pianeta. Per le analisi finalizzate al monitoraggio della geodinamica a lungo termine, i fenomeni periodici costituiscono però un rumore sulle serie di coordinate: istante per istante è necessario rimuoverne l'effetto dalle coordinate stimate. A titolo di esempio, nella figura seguente viene presentata una simulazione grafica realistica del moto di un punto geodetico.

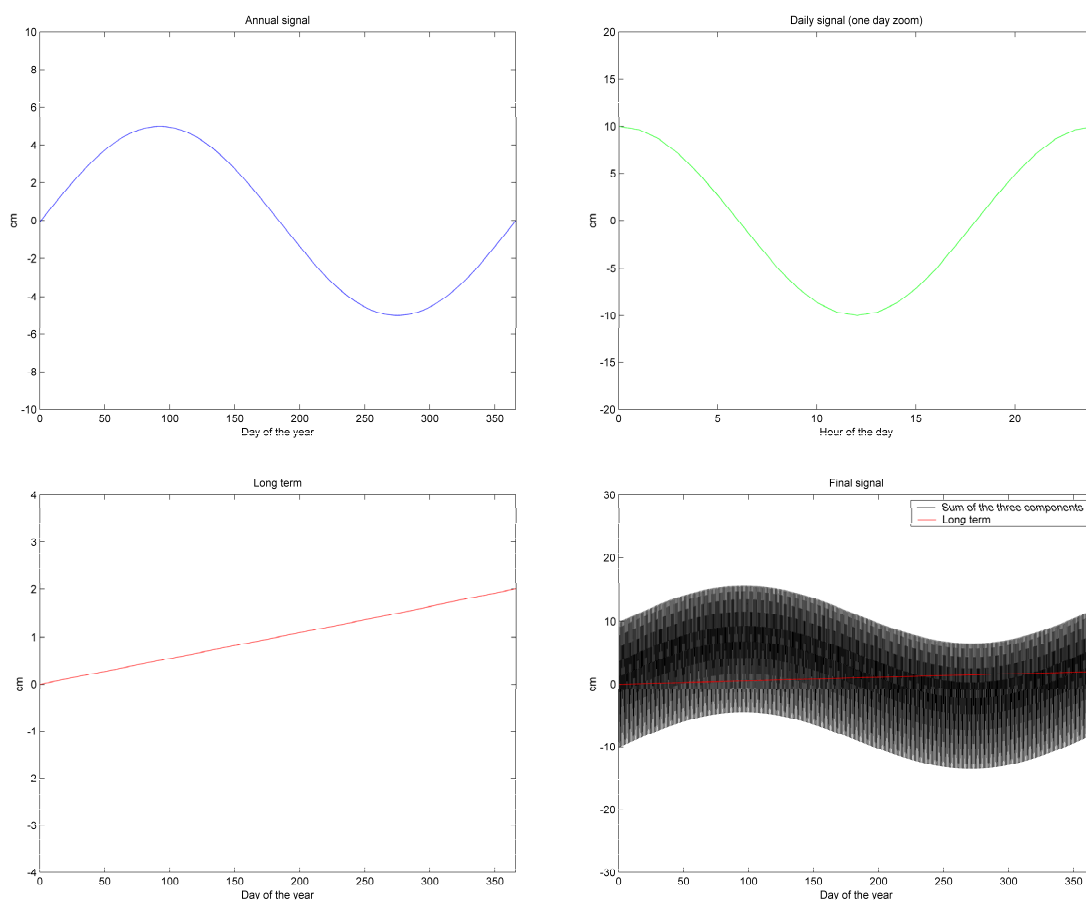


Figura 2.9. In alto a sinistra è rappresentata una simulazione di marea terrestre, nel suo periodo annuale, con ampiezza di 5 cm; in alto a destra è rappresentato il periodo diurno della marea terrestre, con ampiezza 10 cm. In basso a sinistra, un tipico moto dovuto alla geodinamica a lungo termine, lineare di 2 cm/anno. In basso a destra, la composizione dei tre moti su un anno: la marea terrestre maschera completamente il moto a lungo termine.

La rimozione degli effetti periodici dalle posizioni istantanee è peraltro abbastanza banale perché oggi disponiamo di accurati modelli matematici per il loro calcolo. Per un sito di osservazione geodetica la posizione istantanea $\xi(t)$ può essere scritta come

$$\xi(t) = \mathbf{x}(t) + \delta\mathbf{x}(\mathbf{x}, t)$$

ove $\delta\mathbf{x}(t)$ è il contributo dei fenomeni periodici nel sito all'epoca t . A partire da modelli $\delta\mathbf{x}_M(\mathbf{x}, t)$, la posizione media del sito viene calcolata mediante la

$$\mathbf{x}(t) = \xi(t) - \delta\mathbf{x}_M(\mathbf{x}, t) \quad (2.6)$$

2.2.5 I fenomeni non periodici

Qualunque punto della superficie terrestre è tipicamente soggetto a moti orizzontali, se non altro per la geodinamica: in Italia questa ha entità di 2-3 cm/anno verso Nord Est, con piccole differenze fra le varie regioni; può inoltre esservi presenza di moto verticale, ad esempio per la subsidenza, con entità che possono arrivare in casi estremi a 1-10 cm/anno. Per descrivere la geodinamica globale esistono diversi modelli studiati da geofisici e geologi, basati anche su serie temporali di osservazioni geodetiche: quelli più tipicamente citati in letteratura sono il NUVEL-1A e il NNR-NUVEL-1A. Per modellizzare matematicamente l'effetto della geodinamica a lungo termine si adotta l'ipotesi di moto lineare:

$$\mathbf{x}(t) = \mathbf{x}_0 + (t - t_0)\dot{\mathbf{x}}$$

ove $\mathbf{x}_0$ è la posizione del punto ad un'epoca di riferimento t_0 , $\dot{\mathbf{x}}$ è la sua velocità, ipotizzata costante. Questa è a sua volta un'approssimazione del moto delle grandi placche di cui è composta la crosta terrestre, che ruotano sulla superficie rispetto a un sistema terrestre.

Si noti che l'ipotesi di linearità deve naturalmente essere abbandonata in presenza di eventi particolari e bruscamente discontinui come gli eventi sismici oppure i cedimenti strutturali: in tali casi il modello da adottare è discontinuo, ovvero del tipo

$$\mathbf{x}(t) = \begin{cases} \mathbf{x}_0 + (t - t_0)\dot{\mathbf{x}} & t \leq \bar{t} \\ \mathbf{x}_0 + \delta\mathbf{x} + (t - t_0)(\dot{\mathbf{x}}_2 + \delta\dot{\mathbf{x}}) & t > \bar{t} \end{cases} \quad (2.7)$$

ove $\bar{t}$ è l'epoca della discontinuità.

Inoltre in natura possono accadere anche fenomeni locali non lineari, come ad esempio un moto franoso o le deformazioni strutturali lente: anche in tal caso il semplice modello lineare deve essere abbandonato, per modelli più rispondenti alla realtà locale.

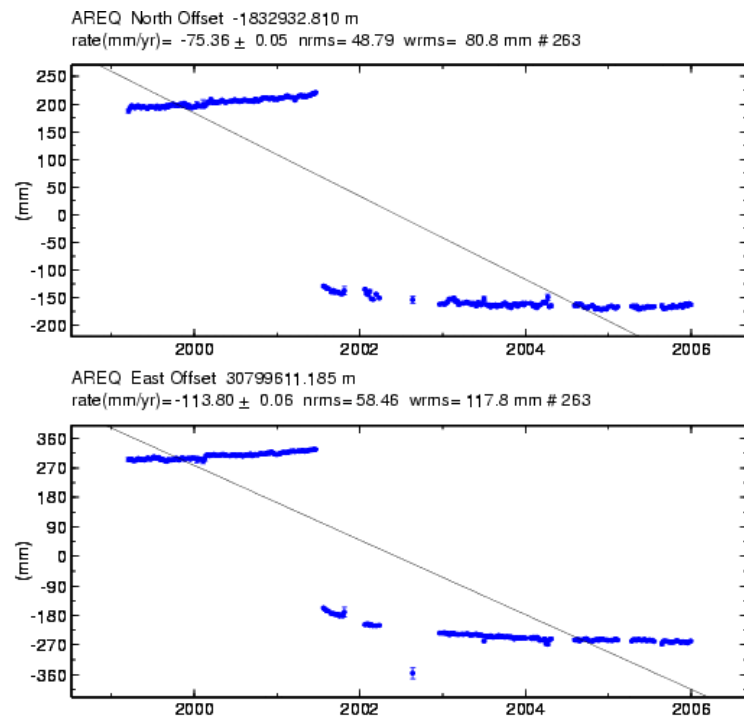


Figura 2.10. Sito di Arequipa, esempio di serie Nord e Est stimate da IGS. Prima del sisma del giugno 2001 l'andamento è lineare; il sisma causa una discontinuità di circa 35 cm. Dopo il sisma si nota in entrambe le coordinate una fase di rilassamento post-sismico non lineare. Dal sito IGS.

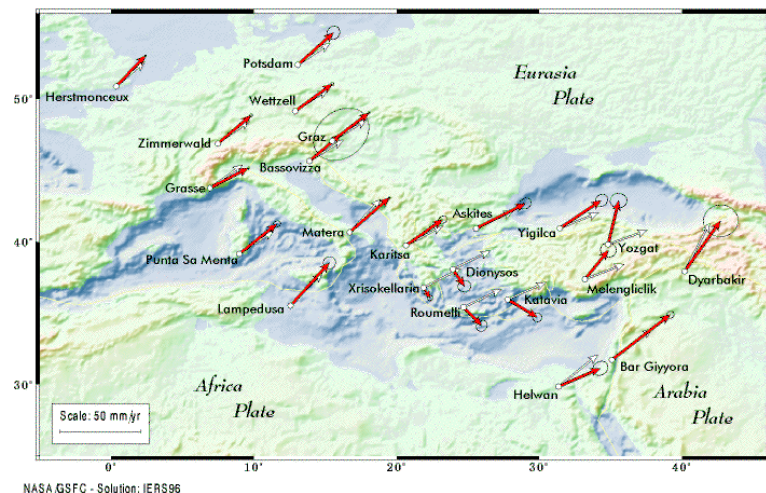


Figura 2.11. Esempio di stima delle velocità geodinamiche: stime del CDDIS della NASA per alcuni siti europei. Dal sito CDDIS

NNR-NUVEL-1A (No-net-rotation)

DeMets et al.,

Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions,
Geophys. Res. Lett. Vol. 21 No. 20, p. 2191-2194, 1994

Plate Abbr.	Rot. Vector (geogr.)		Rot. Vector (cart.)				Plate Name
	PHI [deg]	LAM	OMEGA [deg/Ma]	omega(x) [rad/Ma]	omega(y)	omega(z)	
AFRC	50.569	-73.978	0.2909	0.000891	-0.003099	0.003922	Africa
ANTA	62.986	244.264	0.2383	-0.000821	-0.001701	0.003706	Antarctica
ARAB	45.233	-4.464	0.5455	0.006685	-0.000521	0.006760	Arabia
AUST	33.852	33.175	0.6461	0.007839	0.005124	0.006282	Australia
CARB	25.014	266.989	0.2143	-0.000178	-0.003385	0.001581	Caribbea
COCO	24.487	244.242	1.5103	-0.010425	-0.021605	0.010925	Cocos
EURA	50.631	247.725	0.2337	-0.000981	-0.002395	0.003153	Eurasia
INDI	45.505	0.345	0.5453	0.006670	0.000040	0.006790	India
NOAM	2.438	-85.895	0.2069	0.000258	-0.003599	-0.000153	N. America
NAZC	-47.804	259.870	0.7432	-0.001532	-0.008577	0.009609	Nazca
PCFC	-63.045	107.325	0.6408	-0.001510	0.004840	-0.009970	Pacific
SOAM	-25.325	235.570	0.1164	-0.001038	-0.001515	-0.000870	S. America
JUFU	-30.054	58.870	0.6658	0.005200	0.008610	-0.005820	Juan de Fuca
PHIL	-38.011	-35.360	0.8997	0.010090	-0.007160	-0.009670	Philippine
RIVR	20.428	253.128	1.9781	-0.009390	-0.030960	0.012050	Rivera
SCOT	-25.273	261.234	0.1705	-0.000410	-0.002660	-0.001270	Scotia

<http://dgfi2.dgfi.baw-muenchen.de/geodis/GDYN/platemodels.html>

```

dphi/dt(phi,lam) = OMEGA*cos(PHI)*sin(lam-LAM)
dlam/dt(phi,lam) = OMEGA*(sin(PHI) - cos(lam-LAM)*tan(phi)*cos(PHI))
dx/dt(x,y,z) = omega_x*z - omega_y*y
dy/dt(x,y,z) = omega_x*x - omega_z*z
dz/dt(x,y,z) = omega_y*y - omega_x*x

```

Figura 2.12. Esempio di modello geodinamico: parametri numerici del modello NNR-NUVEL-1A. Dal sito DGFI.

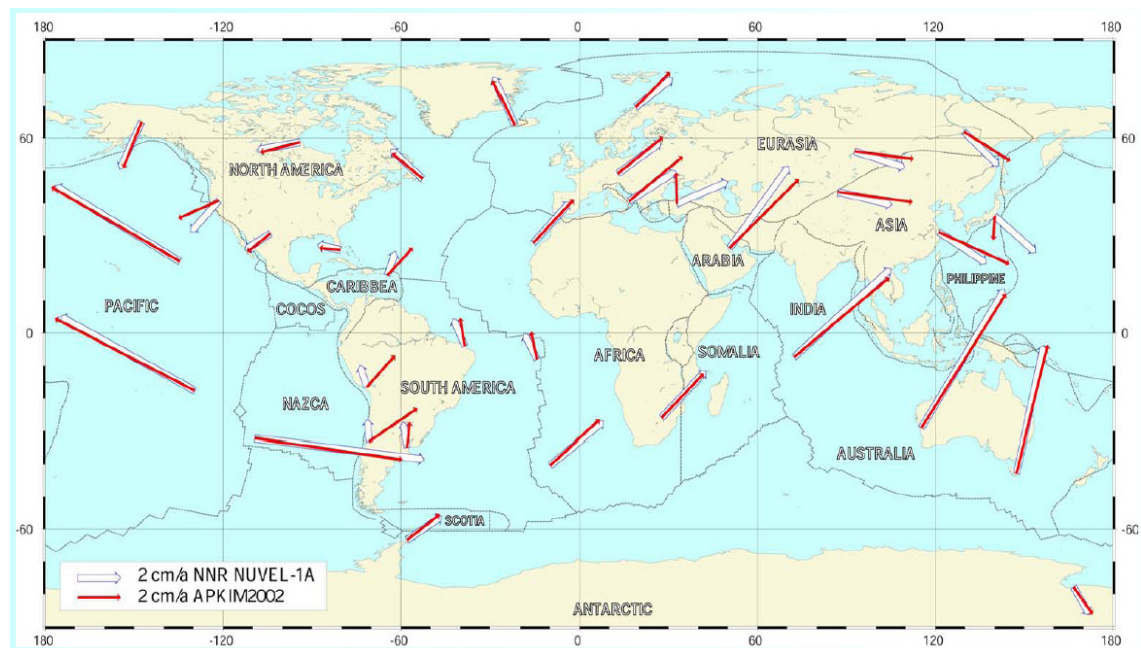


Figura 2.13. Esempio di confronto fra modelli geodinamici: mappa delle placche tettoniche e loro velocità secondo i modelli NNR-NUVEL-1A e APKIM2002. Sostanziale accordo ma anche alcune differenze significative. Dal sito DGFI.

2.3 I diversi sistemi di riferimento

Mediante le regole (rimozione della deficienza di rango) e le misure, (osservazioni fisiche e geometriche, al suolo e dallo spazio), viene realizzato un Sistema di Riferimento. In geodesia la realizzazione di un sistema di riferimento consiste in un catalogo di coordinate stimate per un insieme, detto rete, di punti fondamentali: gli utenti del sistema di riferimento utilizzano i punti fondamentali e le loro coordinate come riferimento per determinare le coordinate di nuove sottoreti, dette di raffittimento. Nella storia si sono avute differenti esigenze di referenziazione e si sono via via rese disponibili nuove metodologie di osservazione: per tale motivo si hanno differenti sistemi di riferimento, tipicamente di accuratezza crescente nel tempo. In particolare riteniamo opportuno discutere le seguenti distinzioni:

- sistemi di riferimento celesti e terrestri;
- sistemi di riferimento terrestri permanenti e statici;
- sistemi di riferimento terrestri globali e locali.

I sistemi di riferimento celesti servono per il calcolo di posizioni di galassie, di orbite di pianeti e di satelliti e non vengono qui trattati. I sistemi di riferimento terrestri servono invece per lo studio del pianeta, ovvero della sua forma, della sua dinamica e delle sue deformazioni; servono inoltre per la georeferenziazione di informazioni, ad esempio a fini cartografici e di navigazione: costituiscono quindi l'argomento di nostro interesse. Intendiamo con il termine di sistemi di riferimento permanenti quelli realizzati mediante reti di stazioni in osservazione continua (reti permanenti), le cui coordinate vengono stimate e monitorate con continuità; evidentemente i sistemi di riferimento permanenti possono garantire la massima accuratezza e coerenza, poiché le coordinate pubblicate sono continuamente controllate (si veda l'esempio di Fig. 2.14).

I sistemi di riferimento statici sono realizzati mediante reti di caposaldi osservati in una singola campagna di rilievo, le cui coordinate vengono determinate e pubblicate una sola volta. I sistemi di riferimento statici forniscono coordinate cristallizzate e possono accumulare negli anni deformazioni alla scala regionale, dovute agli spostamenti differenziali delle varie regioni interessate; possono presentare anche, e soprattutto, errori locali dovuti a una molteplicità di cause, come ad esempio deformazioni locali, cedimenti strutturali, manomissione dei punti: si veda l'esempio in Fig. 2.15; infine, nel caso del metodo GPS, le reti statiche nazionali sono state realizzate nei decenni passati, quando ancora la tecnologia e i metodi di elaborazione dei dati non erano arrivati alla piena maturità: presentano quindi accuratezze intrinseche molto inferiori a quelle oggi ottenibili. I sistemi di riferimento statici sono legati a esigenze della topografia storica ma risultano oggi inattuali per le tecniche GNSS perché queste permettono realizzazioni mediante reti permanenti.

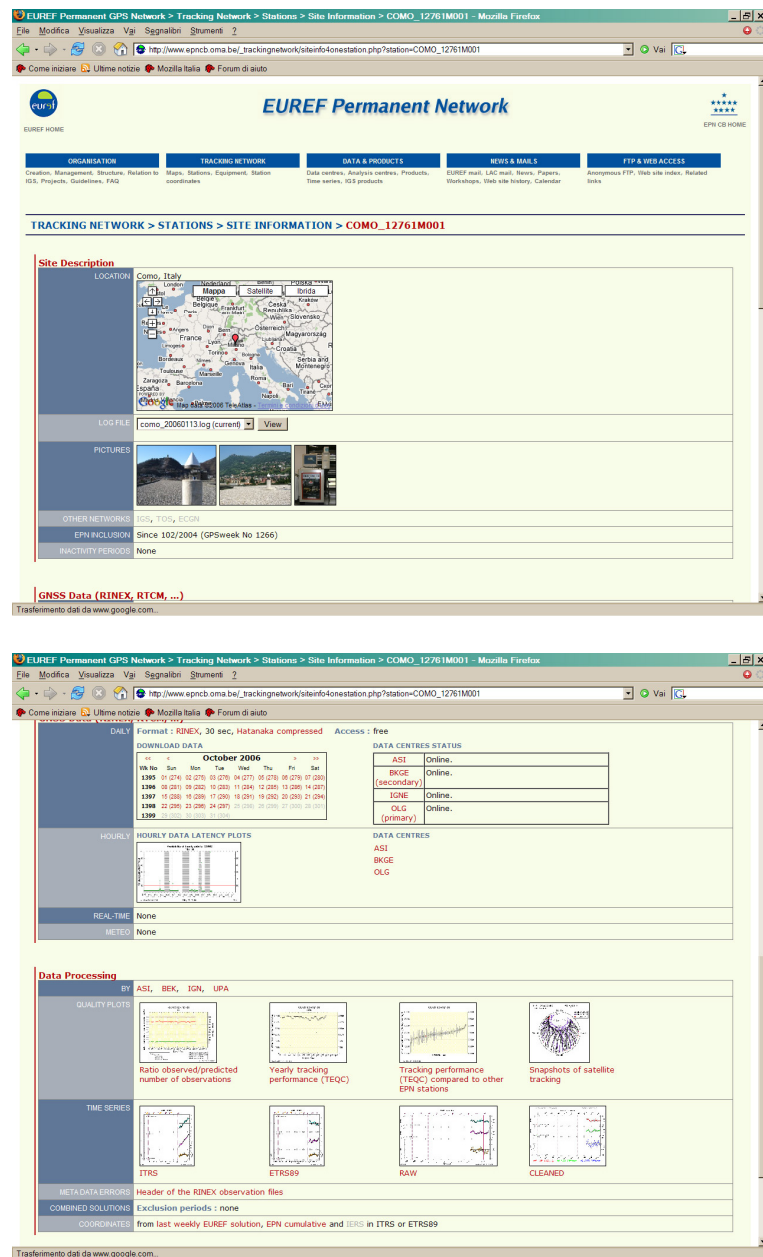


Figura 2.14. Esempio di stazione appartenente a rete permanente; sito di EPN, pagina della SP di Como: gli ultimi grafici riportano le stime di coordinate

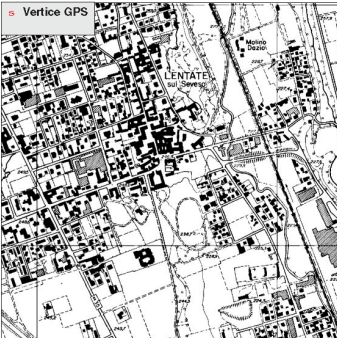
 Provincia di Milano	LENTATE SUL SEVESO		CODICE	4089
	Proprietà: SUOLO PUBBLICO Comune: Lentate sul Seveso			
	Rilievo effettuato da: GPS works - Nizza Monferrato (AT) - anno 2000			
Percorrendo via Garibaldi da Lazzate in direzione di Meda, giunti in piazza S.Vito e oltrepassata la chiesa, proseguire fino a trovare una cabina sulla sinistra prima della discesa. Note: livellazione effettuata davanti la porta della cabina.		Vertice GPS di riferimento Accesso: Il vertice è situato davanti alla cabina. Materializzazione: Centrino di ottone sul muretto in cls di piazza S.Vito dal lato opposto della chiesa.		
				
		Coordinate del vertice GPS: Geografiche (Roma 40) ϕ : 45°40'40,736" λ : - 03°19'47,209" Alt.: 246,43 m		
		Piane (Gauss-Boaga) N: 5.058.372,59 m E: 1.509.544,87 m		
		Geografiche (WGS84) ϕ : 45°40'43,176" λ : 09°07'19,984" Alt.: 291,34 m		

Figura 2.15. Esempio di rete statica: monografia di un punto della rete di raffittimento IGM95 in Lombardia. Il punto è stato rilevato e monografato nel 2000; negli anni immediatamente successivi il muretto su cui è materializzato il punto ha subito lavori di manutenzione, con un conseguente spostamento del caposaldo di circa 20 cm: il fatto è stato fortuitamente rilevato in un test strumentale. Monografia della Provincia di Milano.

2.4 Il sistema di riferimento globale: International Terrestrial Reference System

I sistemi di riferimento globali sono realizzati mediante reti di stazioni permanenti distribuite su tutto il globo, le cui posizioni sono stimate mediante le diverse tecniche della geodesia satellitare: discuteremo nel seguito la definizione di ITRS e le sue realizzazioni ITRF e IGS.

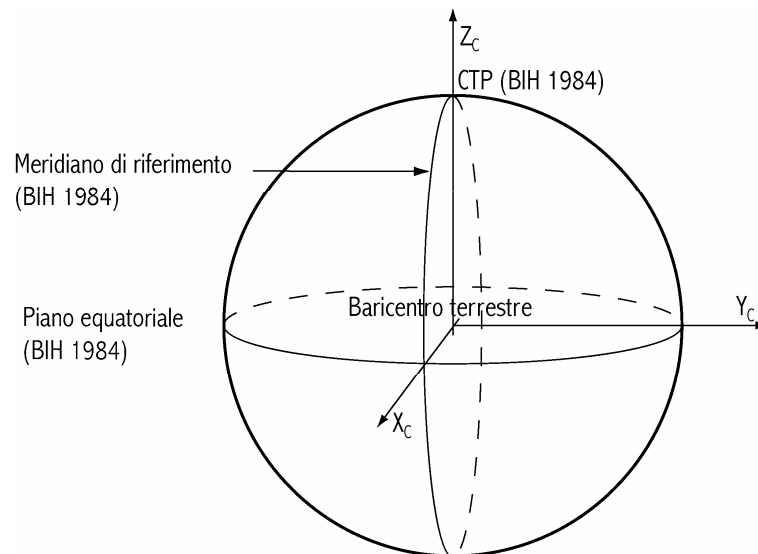


Figura 2.16. Gli assi e l'origine di ITRS.

Nella tradizione e nella maggioranza della letteratura geodetica gli assi x_1, x_2, x_3 vengono rispettivamente indicati con X, Y, Z : nel seguito ci atterremo quindi a questa convenzione. ITRS è definito con origine nel baricentro convenzionale della Terra; l'orientamento iniziale degli assi è in accordo alla definizione del BIH 1984:

- asse Z passante per il Polo Convenzionale Terrestre;
- asse X definito dall'intersezione fra il piano meridiano di riferimento (Greenwich) e il Piano Equatoriale Convenzionale Terrestre;
- asse Y tale da completare la terna destrorsa.



Figura 2.17. L'osservatorio astronomico di Greenwich. Da Wikipedia.

Gli assi X , Y , Z non vengono ancorati nel tempo a punti fisici o convenzionali della superficie fisica del pianeta: questi infatti si muovono reciprocamente, sia per la geodinamica globale, sia per eventuali fenomeni locali, e non hanno quindi alcuna relazione con la dinamica globale del pianeta. Viceversa l'evoluzione dell'orientamento degli assi, ferma restando la loro reciproca ortogonalità, è definita in modo da annullare il momento angolare globale della Terra. La scala delle lunghezze è definita coerente con la scala temporale Geocentric Coordinate Time (Par. 2.9).

2.4.1 Le reti per la realizzazione di ITRS

Le reti di stazioni permanenti per la realizzazione di ITRS sono sostanzialmente quelle dei tre metodi principali di osservazione geodetica:

- VLBI: Very Long Baseline Interferometry,
- SLR: Satellite Laser Ranging,
- GNSS: Global Navigation Satellite System.

Le tre reti sono realizzate da punti fondamentali monumentati in modo stabile, con strumenti di osservazione che operano pressoché continuamente nel tempo: ovvero sono reti permanenti.

2.4.1.1 La rete globale VLBI: l'International VLBI Service, IVS

La rete globale VLBI è coordinata dall'International VLBI Service (IVS): si compone di circa 30 stazioni (radiotelescopi) che effettuano misure contemporanee di differenze di angoli e distanze a quasar: dalle misure, mediante complessi metodi di elaborazione dei dati (ovviamente qui non discussi) si stimano le basi (vettori) fra le stazioni.



Figura 2.18. La rete VLBI, dal sito di IVS.



Figura 2.19. Un esempio di stazione VLBI. Dal sito IVS.

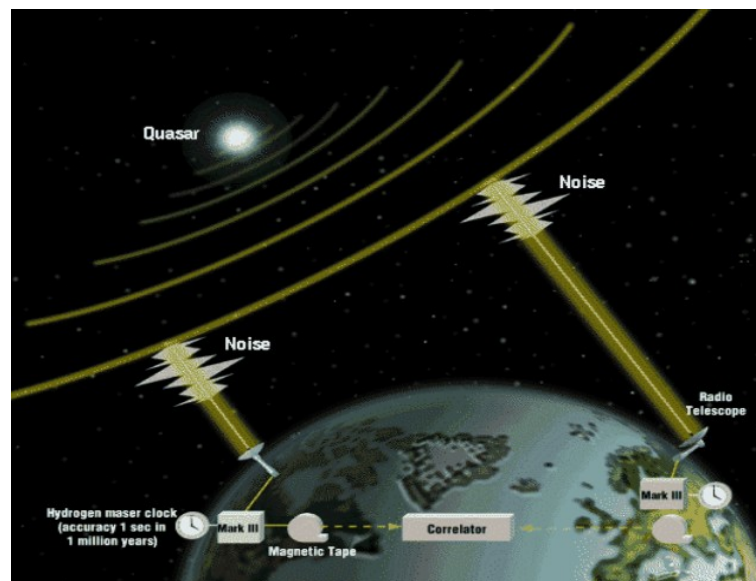


Figura 2.20. Rappresentazione pittorica del processo di stima di una base VLBI. Dal sito IVS.

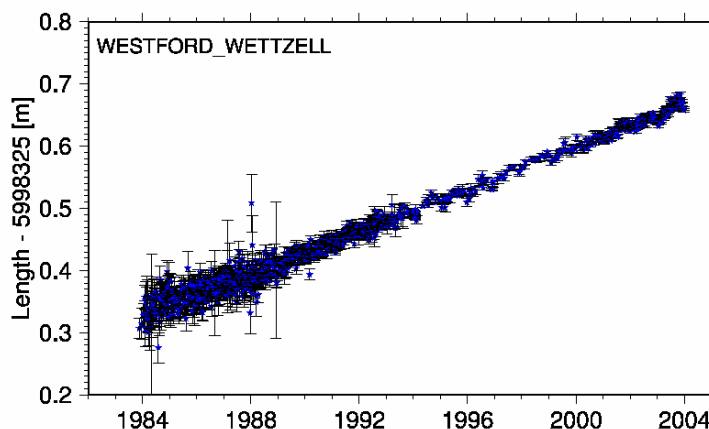


Figura 2.21. Un esempio di stima di lunghezza di base: stazioni VLBI di Westford (USA) e Wettzell (Germania): si nota il miglioramento delle ripetibilità con gli anni, in corrispondenza all'affinarsi delle tecniche di osservazione e elaborazione. Dal sito IVS.

2.4.1.2 La rete globale SLR: l'International SLR Service, ILRS

La rete SLR è coordinata dall'International Satellite Laser Ranging Service (ILRS) e si compone di circa 90 stazioni, realizzate mediante cannoni laser orientabili: sostanzialmente una stazione osserva il tempo di andata e ritorno di un fascio laser inviato dal cannone e riflesso da appositi satelliti artificiali in orbita intorno al pianeta: le osservazioni permettono la modellizzazione delle orbite dei satelliti e la stima della posizione delle stazioni.

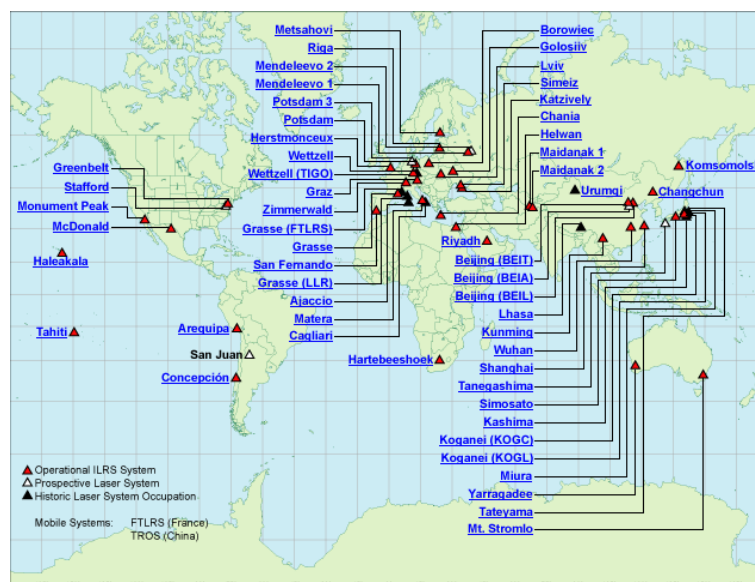


Figura 2.22. La rete SLR, dal sito ILRS.



Figura 2.23. Un esempio di stazione SLR: il sito di Metsahovi.
Dal sito ILRS.



Figura 2.24. Rappresentazione pittorica del processo di misura
di una stazione SLR. Dal sito ILRS.

2.4.1.3 La rete globale GNSS

L'International GNSS Service, IGS. L'International GNSS Service (IGS) è un servizio istituito nel 1993 i cui scopi sono di contribuire alla realizzazione e distribuzione di ITRS, fornire prodotti GNSS (efemeridi, EOP, ...), definire gli standard per reti di SP GNSS e sostenere la ricerca sul GNSS. La rete IGS si compone di circa 380 stazioni; per il controllo, il monitoraggio, l'analisi della rete sono attivi diversi Centri di analisi, Gruppi di Lavoro, Progetti Pilota, oltre a un Ufficio Centrale di Coordinamento.

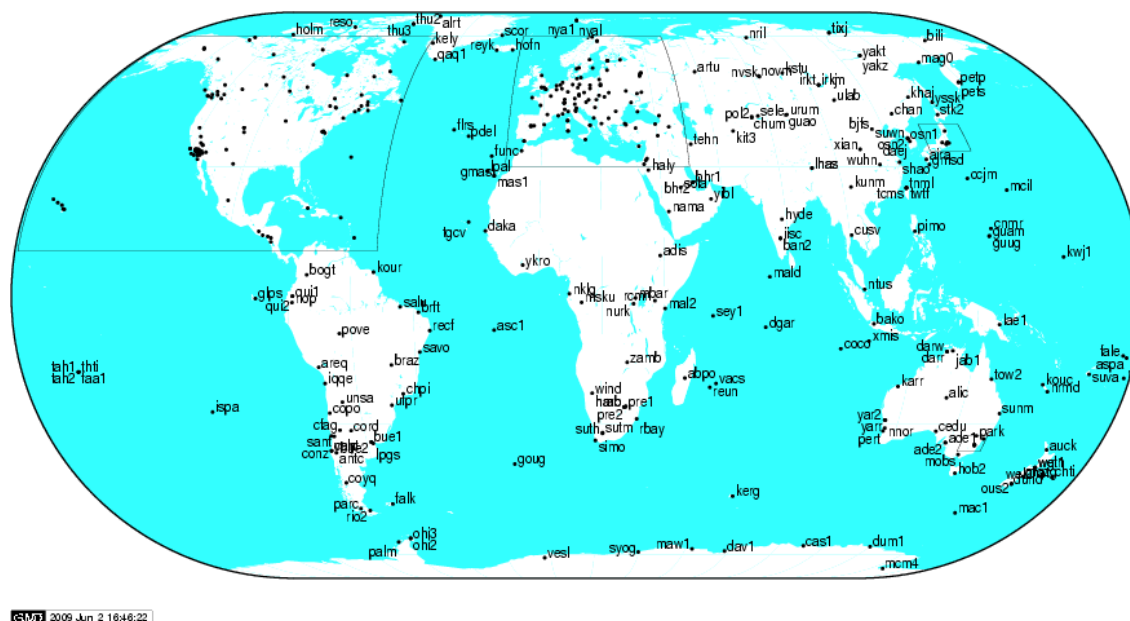


Figura 2.25. La rete globale GNSS, dal sito IGS.

2.4.2 La realizzazione ITRFyy

L'International Earth Rotation Service (IERS) è una sottocommissione di IAG (International Association of Geodesy) e IAU (International Astronomical Union), nata per:

- monitorare i parametri di orientamento e di rotazione terrestre (EOP),
- realizzare il Sistema di Riferimento Internazionale Celeste,
- realizzare il Sistema di Riferimento Internazionale Terrestre,
- monitorare le deformazioni in atto sulla crosta terrestre.

In particolare IERS realizza l'unico sistema di riferimento globale multitecnica di cui viene qui brevemente ricordato il processo di stima. Dal 1988 i risultati forniti dai centri di analisi VLBI, SLR e GPS vengono afferiti a IERS che periodicamente ne effettua la validazione e la compensazione congiunta: ad ogni compensazione viene aggiornata e pubblicata una nuova realizzazione di ITRS, che prende nome di ITRFyy. ITRF sta per International Terrestrial Reference FRAME, mentre yy indica l'anno di pubblicazione della stima. Gli ITRF si evolvono con il numero di stazioni e gli algoritmi di stima; in particolare si sono avuti ITRF89, ITRF94, ITRF97, ITRF2000, ITRF2005: si ha quindi una periodicità di aggiornamento di circa 5 anni.

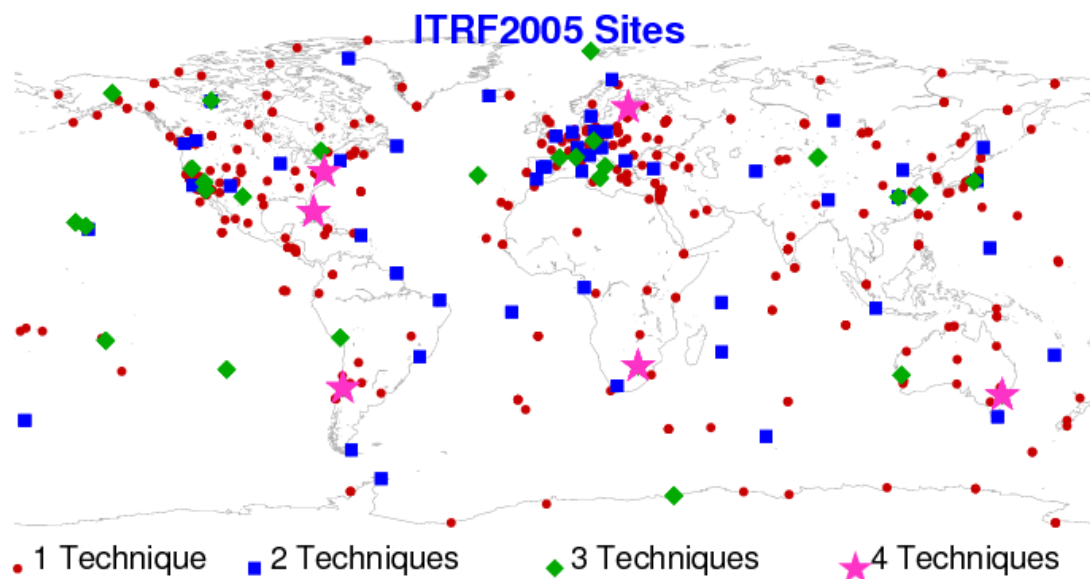


Figura 2.26. Le 338 stazioni usate per ITRF2005, dal sito ITRF. Per ogni sito è indicato il numero di stazioni, ovvero tecniche di osservazione, presenti.

Nello specifico caso dell'ultima realizzazione, ITRF2005, sono stati acquisiti 10 anni di soluzioni settimanali (dal 1995 al 2005) delle reti IVS, ILRS, IGS; inoltre sono state utilizzate le soluzioni di sei stazioni DORIS (tecnica qui non descritta) provenienti da due centri di calcolo. La soluzione coinvolge 338 siti dei quali 254 dispongono di una sola tecnica di osservazione (tipicamente GPS); 84 sono collocati, ovvero presentano 2 o più stazioni, relative ad altrettante tecniche. Nella stima vengono imposti i necessari vincoli sul sistema di riferimento, in modo da allineare la realizzazione ITRF alla definizione ITRS; in particolare:

- il geocentro e il suo moto nel tempo sono vincolati alle stime ottenute dall'analisi delle orbite dei satelliti ILRS;
- la scala e la sua evoluzione nel tempo, sono ottenute dalle soluzioni IVS;
- l'orientamento degli assi X, Y, Z nel tempo soddisfa una condizione di non rotazione di ITRF2005 rispetto a ITRF2000, vincolata su un insieme di 70 siti; a sua volta l'orientamento nel tempo degli assi di ITRF2000 era stato vincolato in modo da annullare il momento angolare della rete rispetto al modello geodinamico NNR-NUVEL-1A.

Posti i necessari vincoli, le soluzioni sulle tre reti sono state compensate congiuntamente ai minimi quadrati, per ottenere le stime di coordinate e velocità di tutte le stazioni. Le stime e le relative covarianze, nonché i dettagli sul processo di stima, sono pubblicate in un catalogo scaricabile dal sito. Si noti che $t_0 = 2000.0$, ovvero coincide con l'epoca centrale delle serie analizzate: il calcolo della posizione della stazione a una generica epoca $t \neq t_0$, è ottenuta mediante propagazione lineare:

$$x_i(t) = x_i(t_0) + \dot{x}_i \cdot (t - t_0), \quad \sigma_i^2(t) = \sigma_{i_0}^2 + \sigma_{i_v}^2 \cdot (t - t_0)^2, \quad x_i = X, Y, Z \quad (2.8)$$

Nome	$X(m)$	$Y(m)$	$Z(m)$	$\dot{X}(m/y)$	$\dot{Y}(m/y)$	$\dot{Z}(m/y)$
Lampedusa	5073164.830	1134512.480	3683181.068	-.0147	.0168	.0153
Matera	4641949.648	1393045.330	4133287.386	-.0179	.0188	.0155
Cagliari	4893378.891	772649.688	4004182.102	-.0132	.0197	.0127
Bologna	4461400.834	919593.484	4449504.712	-.0182	.0190	.0110
Padova	4389531.226	923253.699	4519256.380	-.0172	.0177	.0113

Tabella 2.2. Posizione e velocità ITRF2005 di alcuni siti italiani.

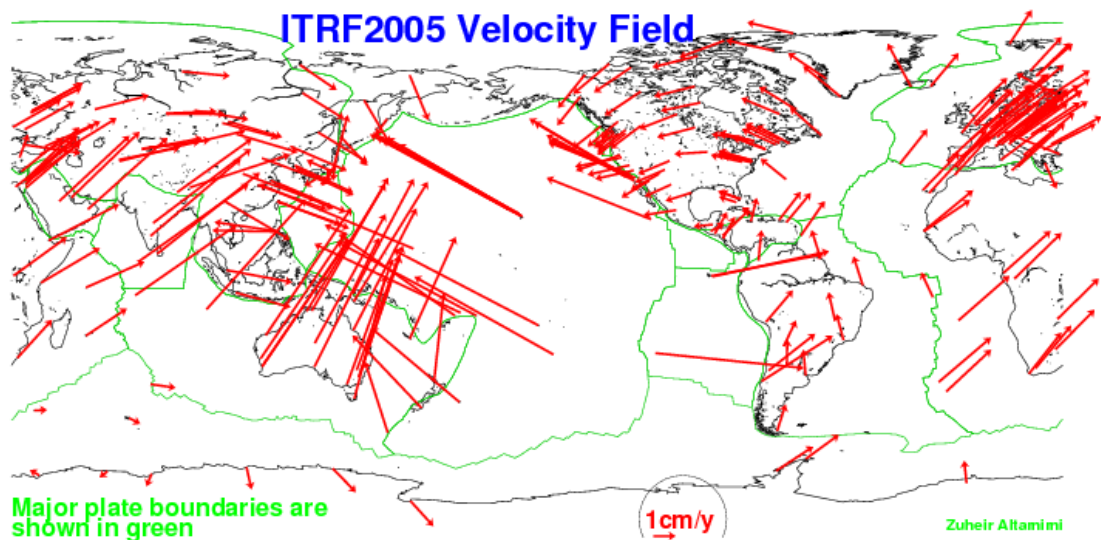


Figura 2.27. Carta delle velocità stimate delle stazioni ITRF. Dal sito ITRF.

ITRF è realizzato mediante una compensazione multitecnica e quindi minimizza l'effetto di eventuali sistematismi dovuti a una particolare tecnica di osservazione; inoltre ITRF si basa su serie molto lunghe di dati: in tal senso è ragionevole sostenere che le velocità geodinamiche stimate nella soluzione ITRF siano caratterizzate dalle massime accuratèzze oggi ottenibili. Però ITRF viene aggiornato con bassa frequenza, ovvero circa ogni 5 anni: questa è una caratteristica intrinseca, perché nasce dalla complessità delle operazioni di preparazione, omogeneizzazione e validazione reciproca dei dati delle diverse tecniche; il problema non è superabile a breve, poiché richiederebbe un significativo aumento di personale del centro di calcolo, che non è al momento previsto.

La scarsa frequenza di aggiornamento comporta due problemi operativi; per quanto riguarda le stazioni già esistenti, abbiamo già discusso il fatto che l'ipotesi di linearità sul moto possa essere eccessivamente rigida a lungo termine: in altri termini, l'utilizzo

della soluzione per propagare all'epoca corrente le coordinate di una stazione può fornire risultati inaccurati a diversi anni dal rilascio della soluzione stessa; inoltre una nuova stazione entra nel catalogo della soluzione solo dopo molti anni dal suo avvio e sino ad allora risulta quindi inutilizzabile come riferimento per ulteriori rilievi locali. Per tale motivo si presentano nel seguito le soluzioni settimanali IGS, che a tutti gli effetti costituiscono un'altra realizzazione di ITRS.

2.4.3 La realizzazione IGSy

IGS stima e distribuisce una propria soluzione, basata sulla compensazione della sola rete IGS, ovvero utilizzando solo la tecnica GPS: attualmente la realizzazione ha nome IGS05; la rete IGS viene compensata ogni settimana, e per ogni settimana viene prodotta una soluzione con la stima delle coordinate di tutte le stazioni al momento centrale della settimana stessa.

IGS05 viene distribuito in due modi:

- le soluzioni settimanali;
- le serie a lungo termine, ovvero la stima delle coordinate e delle velocità ($t_0 = 2005.0$), aggiornata incrementalmente a partire dalle soluzioni settimanali.

La coerenza fra ITRF e IGS è tipicamente di alcuni millimetri: per alcune stazioni le differenze arrivano a 1-2 cm e sono tipicamente dovute a discontinuità intercorse dopo il rilascio di ITRF. Pur ritenendo che l'approccio di stima di ITRF sia più completo di quello di IGS, riteniamo che per calcolare le coordinate di una stazione a qualche anno dal rilascio della soluzione ITRF, sia preferibile l'uso delle soluzioni aggiornate settimanalmente da IGS. E' necessario puntualizzare che, naturalmente, tali considerazioni sono essenziali solo per applicazioni di altissima precisione: per applicazioni cartografiche la differenza è assolutamente non significativa.

2.5 I Sistemi di Riferimento locali

I sistemi di riferimento locali sono definiti su scala locale, ovvero continentale, nazionale,..., sino a particolari applicazioni di controllo di strutture: essi possono essere realizzati mediante tecniche GPS, ovvero mediante raffittimento della rete globale IGS; in tal senso discuteremo il sistema di riferimento europeo ETRS89 e più in particolare quanto di interesse per l'Italia: la rete permanente europea EPN, la vecchia rete statica italiana IGM95, la istituenda rete permanente italiana RDN.

Ricorderemo anche, senza però discuterlo nel dettaglio, il sistema di riferimento nazionale Italiano, realizzato mediante rilievo topografico tradizionale: la rete Roma40.

2.5.1 Il sistema di riferimento Europeo: ETRS89

ETRS89 (European Terrestrial Reference System 1989.0) è definito coincidente con ITRS nel 1989.0, ma solidale al moto Europeo: origine e orientamento degli assi si

muovono nel tempo insieme alla placca media Europea. Si noti che la definizione di ETRS89 nasce da esigenze cartografiche: un punto europeo infatti si muove di 2-3 cm/anno rispetto a ITRS e questo, già sulla scala temporale del decennio, causa spostamenti significativi per la cartografia a grande scala; rispetto a ETRS89 lo stesso punto si muove molto meno, dell'ordine di 2-3 mm/anno, e quindi non pone tale problema.

La definizione e la realizzazione a livello internazionale di ETRS89 sono affidate alla commissione IAG European Reference Frame (EUREF), e alla rete permanente GPS European Permanent Network (EPN), che si compone di circa 200 stazioni, diversi centri di archiviazione, controllo e analisi dati, con una gerarchia simile a IGS. Per ogni realizzazione ITRF_y viene calcolata anche una realizzazione ETRF_y, ovvero:

- la stima delle coordinate e delle velocità delle stazioni europee in ETRS89,
- la stima dei parametri di trasformazione da ITRS a ETRS89,
- la distribuzione delle coordinate e dei parametri di trasformazione.

Si noti che ETRF_y è distribuito anche mediante il sito di EPN, ove sono disponibili le serie di soluzioni settimanali per tutte le stazioni della rete e viene settimanalmente riaggiornata una stima a lungo termine.

Evidentemente ETRF non è un sistema di riferimento globale, poiché la stima ETRF è limitata alle sole stazioni Europee ma è monitorato permanentemente. Si noti che, per tecnicismi piuttosto complessi e qui non discussi, il nome della versione attuale non è ETRF2005, bensì ETRF2000(R05).

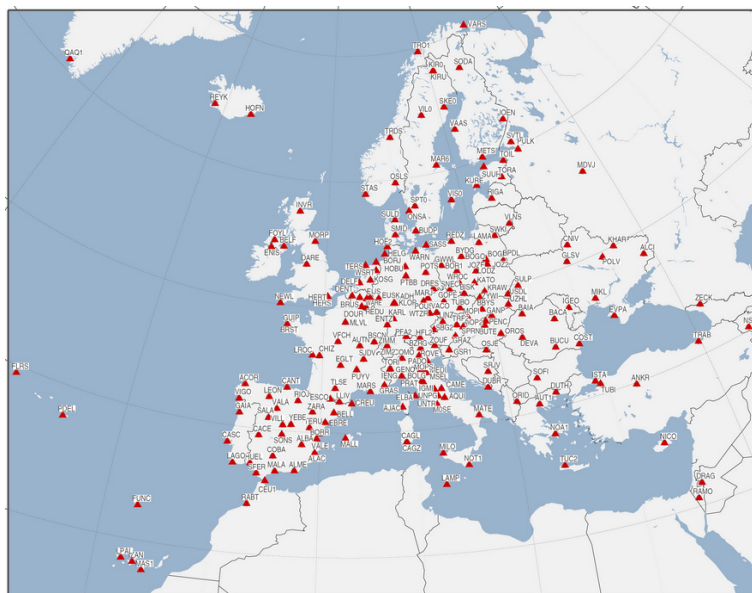


Figura 2. 28. La rete europea. Dal sito EPN

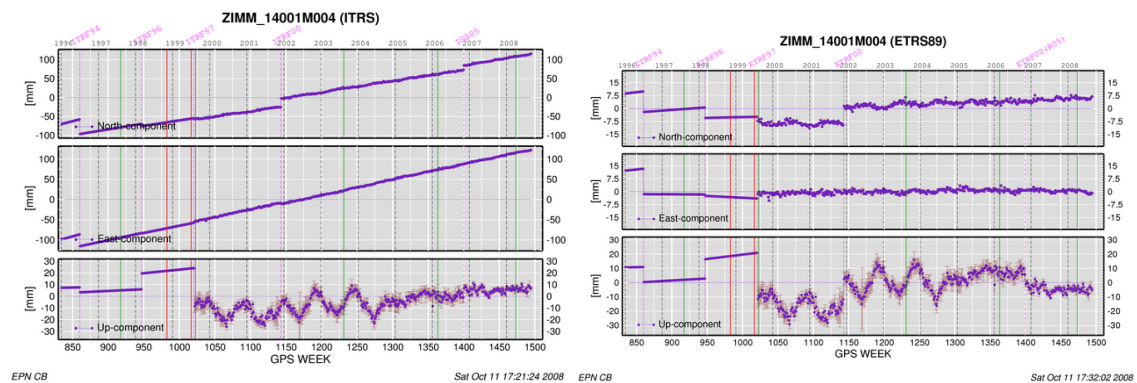


Figura 2.29. Esempio di confronto fra ITRF, IGS e ETRF per una stazione europea: grafici di Zimmerwald (Svizzera). Differenze fra stime settimanali e valor medio, per Nord, Est e quota; a destra ITRS, a sinistra ETRS. Fino a ITRF97, ZIMM ha andamento perfettamente lineare perché nelle soluzioni era una stazione di vincolo; da ITRF97 è una stazione stimata. Si ignorino le discontinuità, dovute ad aggiornamenti del sistema di riferimento e a cambi strumentali. In ITRS, in planimetria, si ha un moto quasi lineare di circa 2 cm/anno; in quota, fino al 2004.5 vi sono apparenti cicli stagionali dovuti a un'errata modellizzazione delle maree terrestri; in seguito rimane rumore di stima debolmente correlato. In ETRS89: la Svizzera è nella parte meridionale della cosiddetta placca “media” europea e ZIMM risulta sostanzialmente ferma.

SR	Est (m)/dE (mm)	Nord (m)/dN (mm)	h (m)/dh (mm)
ITRF2005	383055.4571	5192649.9076	956.3425
IGS05	-22	-21	73
ETRF00(R5)	-294	-331	15

Tabella 2.3. Zimmerwald: coordinate ITRF (in proiezione UTM) e differenze con IGS e ETRF all'epoca 2008.0. Le differenze fra ITRF2005 e IGS05 sono dovute alla necessità di propagare la soluzione ITRF ($t_0 = 2000.0$) per 8 anni; le differenze fra ITRF/IGS e ETRF sono dovute al moto della placca media europea dal 1989.0.

2.5.2 La trasformazione fra ITRS e ETRS89: la codificazione internazionale

Il processo per ottenere le coordinate ETRS89 di un punto, secondo le linee guida di EUREF, è a tre passi:

- stima delle sue coordinate in ITRF corrente, all'epoca corrente,
- trasformazione delle coordinate da ITRF a ETRF, all'epoca corrente,
- eventualmente: propagazione indietro delle coordinate ETRF nel tempo.

Passo 1. Stima in ITRF

Le coordinate del punto vengono determinate in ITRFyy all'epoca corrente ($\mathbf{x}(t)_{P,I}$), mediante inquadramento del punto secondo le linee guida internazionali per l'elaborazione dei dati GPS (si vedano i siti IERS, IGS, EPN).

Passo 2. Trasformazione da ITRF a ETRF

Le coordinate ITRFyy del punto vengono trasformate in ETRFyy (all'epoca corrente), mediante la

$$\begin{aligned}\mathbf{x}(t)_{P,E89} &= \mathbf{t} + \mathbf{x}(t)_{P,I} + \delta\mathbf{R}(t)\mathbf{x}(t)_{P,I}, \\ \delta\mathbf{R}(t) &= -(t-1989.0)[\dot{\mathbf{r}} \times]\end{aligned}\quad (2.9)$$

ove i valori per i parametri $\mathbf{t}, \dot{\mathbf{r}}$ sono tabulati nei cataloghi ufficiali IERS, disponibili sul sito.

Passo 3. Trasporto nel tempo in ETRF

Eventualmente, se è richiesta la stima all'epoca 1989.0, le coordinate ETRFyy devono essere propagate indietro dall'epoca corrente. La formula da adottarsi è

$$\mathbf{x}(1989.0)_{P,E89} = \mathbf{x}(t)_{P,E89} + (1989.0 - t)\dot{\mathbf{x}}_{P,E89} \quad (2.10)$$

ove $\dot{\mathbf{x}}$ è la velocità geodinamica in ETRFyy del punto: per le stazioni IGS e EPN la velocità è pubblicata con le coordinate.

2.5.3 La realizzazione italiana di ETRS89: il vecchio inquadramento di IGM95

La rete fondamentale italiana, monumentata e rilevata dall'Istituto Geografico Militare Italiano (IGM), è denominata IGM95 e costituisce l'attuale sistema di riferimento cartografico ufficiale italiano; IGM95 si compone di circa 1250 caposaldi originali, oltre a circa 800 (in crescita) punti appartenenti a reti di raffittimento regionali. L'interdistanza fra i punti della rete iniziale è di circa 25 km, mentre i raffittimenti a livello regionale hanno interdistanza di circa 7 km. La rete e i suoi successivi raffittimenti vennero inizialmente inquadrati vincolando alcuni punti EPN posti in Italia alle loro coordinate ETRF89, ovvero la prima realizzazione del sistema di riferimento europeo, stimate all'epoca 1989.0: si noti che in Italia le differenze fra ITRF corrente e ETRF89 al 1989.0 sono dell'ordine di 40-50 cm in planimetria; IGM95 era quindi un raffittimento locale e statico di ETRF89 e soffriva di tre problemi, che elenchiamo in ordine crescente di rilevanza:

1. l'Italia non appartiene omogeneamente alla placca Europea: dal 1989.0 a oggi vi sono state deformazioni di alcuni centimetri fra Nord e Sud della nazione;
2. strumenti e metodi di elaborazione adottati negli anni '90 per rilevare IGM95 erano intrinsecamente molto meno accurati di quanto oggi disponibile;

3. in 20 anni, sui singoli caposaldi possono essere avvenuti deformazioni locali, ad esempio cedimenti piuttosto che lavori murari, come mostrato per il caposaldo di Lentate (Fig. 2.15).

In sostanza, gli errori di ETRF89-IGM95 erano in planimetria sino a 5 cm, in altimetria sino a 10 cm: si trattava di errori sparsi ma anche di deformazioni spazialmente correlate, con la presenza anche di valori isolati più alti.



Figura 2.30. I punti IGM95 in Regione Lombardia.

Per analizzare alla scala nazionale le relazioni fra IGS05 e il vecchio inquadramento di IGM95 è stato effettuato il seguente calcolo: sono state stimate con grande accuratezza (migliore di 1-2 cm) le coordinate IGS05 di 28 vertici IGM95 distribuiti sulla nazione, e quindi è stata stimata una trasformazione di similarità; i residui (Tab. 2.4) sugli IGM95 presentano valori elevati, superiori alle accuratezze oggi ottenibili in un rilievo in tempo reale.

(cm)	Est	Nord	Quota
Media	0.0	0.0	0.0
σ	3.9	3.1	5.7
Min	-7.9	-4.3	-12.3
Max	8.0	7.1	11.9

Tabella 2.4. Residui della trasformazione nazionale fra IGS05 e IGM95.

L'analisi spaziale dei residui planimetrici (Fig. 2.31) mostra la presenza di una chiara componente spazialmente correlata ma anche di deformazioni e errori incorrelati. Rispetto alla stima di una trasformazione nazionale si ottengono risultati migliori stimando 3 trasformazioni separate per Nord, Centro e Sud: internamente alle zone di stima i residui della trasformazione assumono sempre valori inferiori ai 5 cm ma in

questo caso vi sono significativi problemi di coerenza nelle due zone di bordo fra Nord e Centro e fra Centro e Sud.

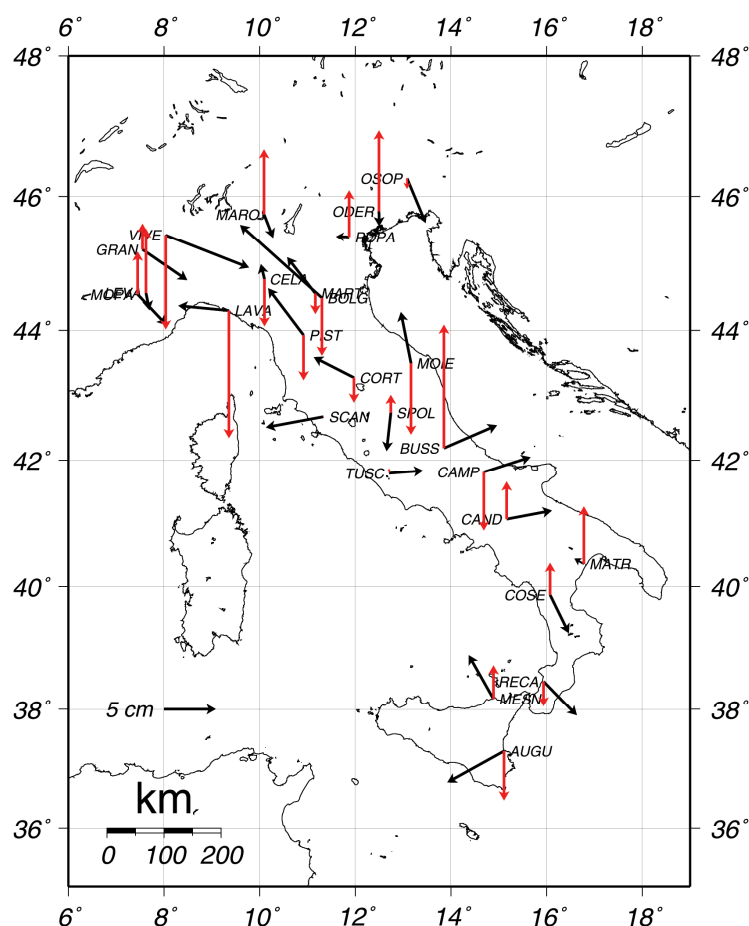


Figura 2.31. Residui della trasformazione nazionale sui punti IGM95.
In rosso residui altimetrici, in nero vettori residui in planimetria.

2.5.4 La realizzazione italiana di ETRS89:

Rete Dinamica Nazionale e il nuovo inquadramento di IGM95

Per ovviare alla obsolescenza del vecchio inquadramento di IGM95, nel 2007 l'Istituto Geografico Militare ha avviato un processo di aggiornamento, che si è ora concluso perlomeno nella sua prima fase. In particolare sono state effettuate le seguenti operazioni:

1. è stata definita una nuova rete di ordine zero nazionale, composta di stazioni permanenti,
2. tale rete, detta Rete Dinamica Nazionale, è stata inquadrata in ITRF05 e in ETRF2000,
3. IGM95 è stata collegata e inquadrata in RDN, in modo da aggiornarne le coordinate.

Più in dettaglio, RDN si compone di 100 stazioni che sono state selezionate fra le stazioni già esistenti e censite in Italia, in base a criteri di omogenea distribuzione e buona qualità: tutte le stazioni sono monumentate permanentemente ma attualmente non tutte hanno una connessione dati continua con IGM.

Erano disponibili e sono state elaborate 4 settimane di dati: dal giorno 357 del 2007 al giorno 19 del 2008. La rete è stata elaborata presso l'IGM e quindi rielaborata presso tre centri di calcolo indipendenti, in modo da permettere il confronto e la crossvalidazione dei risultati: l'Università di Padova, l'Università di Bologna e il Politecnico di Milano. L'inquadramento è avvenuto in ITRF2005, con la stima di 28 sessioni giornaliere; queste sono state quindi compensate per la stima finale delle coordinate, all'epoca centrale della serie, di ogni stazione. Le accuratezze di stima sono dell'ordine, o migliori, del cm: in tabella riportiamo a titolo di esempio le statistiche dei residui delle soluzioni giornaliere rispetto alle soluzioni finali ottenuti presso il Politecnico di Milano.

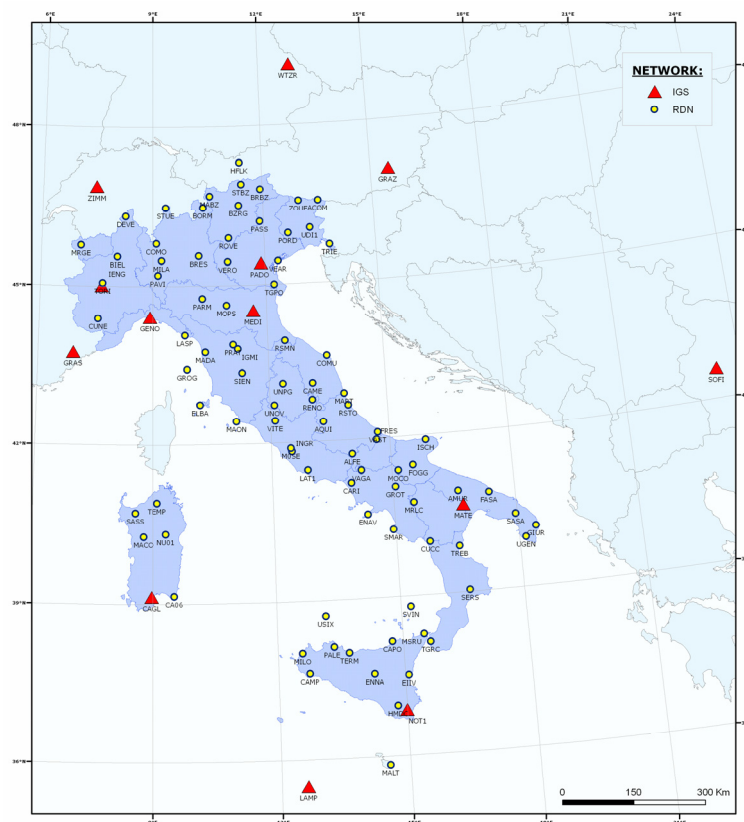


Figura 2.32. La nuova rete di IGM: Rete Dinamica Nazionale.

Le coordinate ITRF2005 sono infine state trasformate a ETRF2000(R5), all'epoca 2008.0, applicando le trasformazioni internazionali (2.9): RDN costituisce quindi il nuovo sistema di riferimento italiano, stimato sia in ITRF2005 sia in ETRF2000(R5) e può essere utilizzata a due livelli.

Innanzitutto RDN è stato utilizzata per inquadrare nuovamente la rete IGM95; a tale riguardo si noti però che IGM95 non è stata completamente rilevata ed rielaborata; sono state invece rilevate le basi da un certo numero di stazioni RDN, circa 45, ai punti IGM95 loro contigui; quindi le “vecchie” basi IGM95 sono state ricompensate utilizzando RDN e le relative coordinate ETRF2000 come vincoli. Si tratta quindi di un aggiornamento parziale da ETRF89 a ETRF2000, che ovvia i problemi dell’obsolescenza del sistema di riferimento ma risente ancora, almeno parzialmente, dei problemi tecnici esposti nel paragrafo precedente: in ogni caso la stima di RDN e la ristima di IGM95 sono state ufficializzate e pubblicate nel marzo del 2009, e si deve ora parlare di ETRF2000-IGM95.

(mm)	Est	Nord	h
σ	1.2	1.2	3.1
Min	-7.0	-2.8	-12.1
Max	6.1	6.1	7.1

Tabella 2.5. Statistiche dei residui giornalieri delle stime RDN.

IGM95 viene distribuito mediante monografie, che devono essere acquistate presso IGM. Naturalmente le monografie distribuite prima dell’ufficializzazione di RDN sono ancora in ETRF89; per trasformarle in ETRF2000 è possibile utilizzare un’apposita macro implementata sul sito web dell’IGM; in alternativa, la trasformazione è implementata anche sulla nuova versione del programmi Verto, commercializzato da IGM e di cui si darà qualche altra informazione nel Par. 2.8.

A un secondo livello si deve ricordare che RDN si compone di stazioni permanenti. Perché l’operazione di ammodernamento del sistema di riferimento sia completa e per ottimizzare lo sforzo già fatto, è necessario che RDN diventi effettivamente una rete permanente, ovvero che:

1. IGM riceva automaticamente e continuamente i suoi dati;
2. i dati e le coordinate siano pubblici e pubblicati tramite interfaccia web (come è il caso delle reti IGS, EPN,...);
3. la rete venga continuamente ricalcolata, in modo da monitorarne la stabilità e la qualità.

In tal modo le reti locali potranno essere inquadrate direttamente in RDN, senza il tramite intrinsecamente meno accurato di IGM95.

2.6 Le coordinate geodetiche in un SR 3D: l’ellissoide di riferimento

Nelle applicazioni di posizionamento terrestre occorre una superficie geometrica di riferimento descritta da pochi parametri che ben approssimi la forma della Terra, rispetto alla quale sia definibile e utilizzabile in pratica un sistema di coordinate: tale superficie è l’ellissoide di rotazione. Definiamo ellissoide di rotazione con centro nell’origine del sistema di riferimento il luogo dei punti tali che

$$\frac{X^2 + Y^2}{a^2} + \frac{Z^2}{b^2} = 1$$

ove a è il semiasse maggiore o equatoriale; b è il semiasse minore o polare.

In funzione dei semiassi si possono definire:

l'eccentricità: $e = \sqrt{1 - \frac{b^2}{a^2}}$

lo schiacciamento: $f = \frac{(a-b)}{a}$

Valgono le:

$$f = 1 - \sqrt{1 - e^2}$$

$$e = f(2 - f)$$

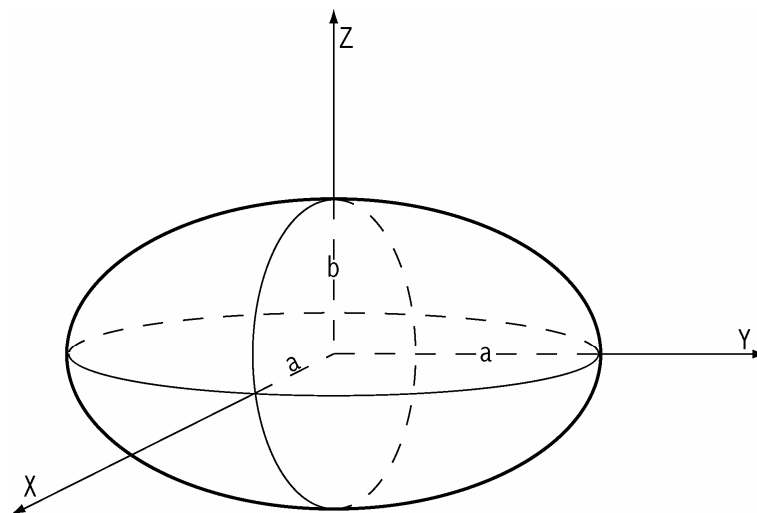


Figura 2.33. L'ellissoide di riferimento.

Ellissoide	a (m)	f
Internazionale (Hayford, 1924)	6378388	1/297
WGS84	6378137	1/298.257223563
GRS80	6378137	1/298.257222101

Tabella 2.6. I parametri dei principali ellissoidi adottati in geodesia.

Forma e dimensioni di un ellissoide sono determinate dalla scelta dei suoi parametri che, in applicazioni di posizionamento terrestre, vengono scelti in modo che l'ellissoide approssimi al meglio il geoide terrestre; nella storia, con il miglioramento delle stime della forma del geoide, si sono progressivamente adottati differenti ellissoidi, fra i quali

si ricordano i principali: A titolo di esempio si ricorda che, rispetto all'ellissoide GRS80, il geoide presenta scostamenti a media nulla e con valori massimi di $\cong 100$ m.

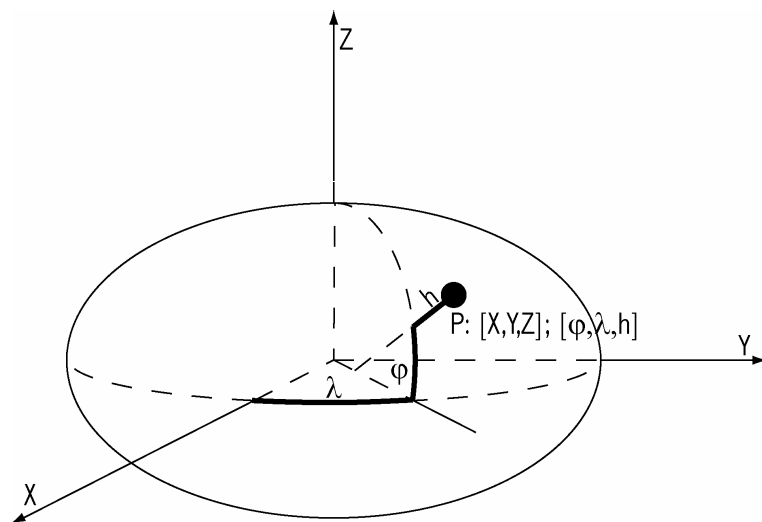


Figura 2.34. Le coordinate geodetiche.

Dato un sistema di riferimento e il relativo ellissoide associato le coordinate geodetiche di P sono definite da:

- φ : (latitudine geodetica) angolo fra la normale all'ellissoide passante per P e il piano equatoriale $[X, Y]$;
- λ : (longitudine geodetica) angolo antiorario fra il piano meridiano per P e il piano meridiano origine $[X, Z]$;
- h : (quota ellissoidica) distanza lungo la normale all'ellissoide fra la superficie dell'ellissoide e P.

La relazione fra coordinate cartesiane e geodetiche di P è data da

$$\begin{aligned} X_p &= (N + h_p) \cos \varphi_p \cos \lambda_p \\ Y_p &= (N + h_p) \cos \varphi_p \sin \lambda_p \\ Z_p &= [N(1 - e^2) + h_p] \sin \varphi_p \end{aligned} \tag{2.11}$$

ove

$$N = \frac{a}{\sqrt{(1 - e^2 \sin^2 \varphi_p)}},$$

N è definito grannormale.

Il passaggio da geodetiche a cartesiane è immediato; il passaggio inverso richiede qualche conto, ma è comunque ben definito. Riportiamo, senza dimostrarla, la seguente procedura: calcolate le seguenti quantità ausiliarie

$$e_b^2 = \frac{a^2 - b^2}{b^2}, \quad \rho = \sqrt{X^2 + Y^2}, \quad \psi = \arctan\left(\frac{Z}{\rho\sqrt{1-e^2}}\right)$$

si calcolano

$$\lambda = \arctan\left(\frac{Y}{X}\right), \quad \varphi = \arctan\left(\frac{(x_3 + e_b^2 b \sin(\psi)^3)}{(\rho - e^2 a \cos(\psi)^3)}\right) \quad (2.12)$$

$$N = \frac{a}{\sqrt{(1 - e^2 \sin^2 \varphi_p)}}, \quad h = \frac{\rho}{\cos(\lambda)} - N$$

2.6.1 Una nota sulle unità di misura degli angoli

Tipicamente gli angoli sono espressi in sessagesimali: gradi (°), primi (′) e secondi (″). Ricordiamo però che vi sono altre unità, sintetizzate nella tabella sottostante.

Tipo	Angolo retto	Angolo giro	Esempio
Sessagesimali	90	360	45° 48′ 7.787822″
Sessadecimali	90	360	45.8021632839
Radiani	$\pi/2$	2π	0.79939855384
Centesimali	100	400	50.8912925376

Tabella 2.7. Le principali unità di misura per angoli. Si noti che talvolta si usano i mas, ovvero i millesimi di arco secondo: $1 \text{ mas} = 1/1000'' = 1/3600000^\circ$

Per fornire alcune indicazioni sugli ordini di grandezza, si consideri un raggio medio della Terra $R \cong 6380 \text{ km}$: a un incremento angolare $\Delta\varphi$ di latitudine (espresso in radianti) corrisponde una distanza $\Delta \cong R\Delta\varphi$ sulla superficie terrestre. Quindi a un incremento di 1° di latitudine corrispondono circa 110 km; $1'$ equivale a 1.8 km e $1''$ a 30 metri; infine $0.0001''$ sono circa 3 mm. La relazione fra incrementi in longitudine e distanze dipende invece dalla latitudine, attraverso la $\Delta \cong R \cos \varphi \Delta\lambda$: a 45° di latitudine (Nord Italia), 1° di longitudine corrisponde a circa 80 km, $1'$ equivale a 1.3 km. e $1''$ a 22 m.

2.7 Gli altri sistemi di coordinate e le coordinate cartesiane locali

Dato un sistema di riferimento, e l'eventuale ellissoide associato, per descrivere la posizione di un punto sono possibili e popolari altri metodi di rappresentazione; ad esempio si ricordano, senza trattarle, le coordinate cartografiche: assegnato un metodo di proiezione piana, le coordinate vengono proiettate sul piano: in IGS, ITRF e ETRF si adotta la proiezione UTM.

Per problemi di controllo locale vengono spesso utilizzate le coordinate locali: definito un punto origine per un'area di interesse, le coordinate degli altri punti sono date come Nord, Est, Up rispetto al punto origine. Si consideri un punto P_0 , di coordinate

$$[\varphi_0 \quad \lambda_0 \quad h_0]^T \leftrightarrow [X_0 \quad Y_0 \quad Z_0]^T$$

In un sistema di coordinate locali avente come origine P_0 , le coordinate di altri punti sono date dalla

$$\begin{aligned} \begin{bmatrix} E \\ N \\ U \end{bmatrix}_P &= \begin{bmatrix} -\sin \lambda_0 & \cos \lambda_0 & 0 \\ -\sin \varphi_0 \cos \lambda_0 & -\sin \varphi_0 \sin \lambda_0 & \cos \varphi_0 \\ \cos \varphi_0 \cos \lambda_0 & \cos \varphi_0 \sin \lambda_0 & \sin \varphi_0 \end{bmatrix} \begin{bmatrix} X_P - X_0 \\ Y_P - Y_0 \\ Z_P - Z_0 \end{bmatrix} \\ &= \mathbf{E}_{P_0} \Delta \mathbf{x}_{P_0, P} \end{aligned} \quad (2.13)$$

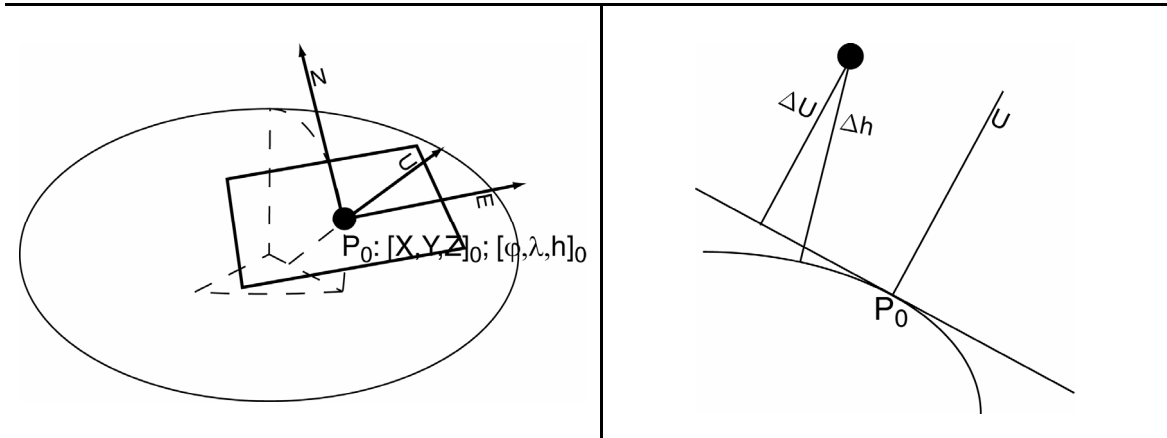


Figura 2.35. Coordinate locali e differenza fra dislivello ellissoidico Δh e componente cartesiana ΔU .

Le coordinate locali costituiscono un sistema di coordinate ortogonali, adatto all'analisi metrica di deformazioni, e risultano utili per il monitoraggio di fenomeni locali; inoltre sono utili per monitorare nel tempo gli spostamenti di un punto P , $P(t)$, rispetto alla sua posizione per un'epoca iniziale $P(t_0)$: infatti permettono una chiara discriminazione

fra movimenti orizzontali e movimenti verticali. La componente U_p non rappresenta effettivi dislivelli, poiché è riferita al piano tangente nel punto origine e non alla differenza di quota ellissoidica: già a 500 m dal punto origine la differenza fra dislivello e U_p è dell'ordine di 2 cm.

Si consideri da ultimo un punto di cui si siano stimate le coordinate geocentriche cartesiane $\hat{\mathbf{x}}_p$ e la relativa matrice di covarianza $\mathbf{C}_{xx}$. Per trasformare la posizione in coordinate geografiche e per far ciò si adotta la (2.12); per derivare la covarianza delle coordinate geografiche, ovvero la matrice di covarianza $\mathbf{C}_{ENU}$ nelle direzioni Est, Nord e Up, si applica la propagazione della covarianza da cartesiane geocentriche a locali, ovvero la

$$\mathbf{C}_{ENU} = \mathbf{E}_p \mathbf{C}_{xx} \mathbf{E}_p^T \quad (2.14)$$

2.8 I sistemi di riferimento locali storici (topografici)

Tipicamente si vincola il sistema di riferimento vincolando l'ellissoide, senza definire assi e coordinate cartesiane: si impongono innanzitutto i suoi parametri di forma a e b ; quindi si impongono 6 condizioni per vincolare i suoi angoli di direzione e il suo centro, dette condizioni di orientamento locale: per fare ciò si considera un punto fondamentale P e si impone che:

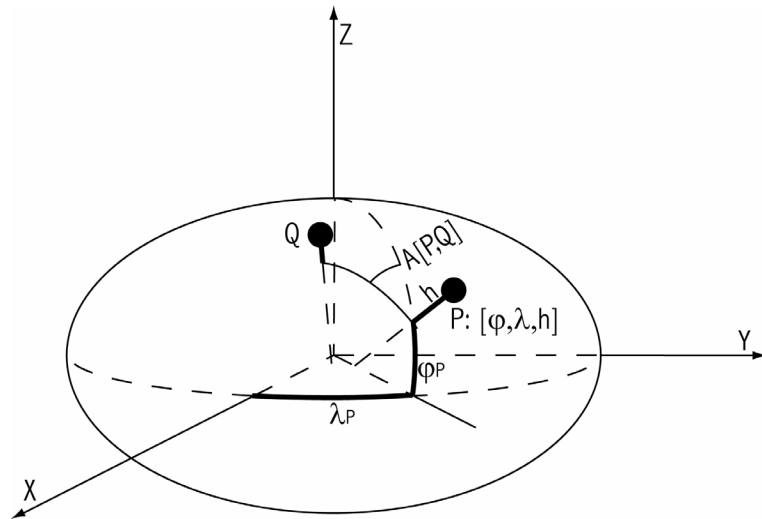


Figura 2.36 Orientamento locale dell'ellissoide.

1. in P le coordinate geodetiche siano uguali a quelle astronomiche, ovvero $\varphi_P = \Phi_P, \lambda_P = \Lambda_P$;
2. la quota ellissoidica di P sia uguale alla sua quota ortometrica;

3. si impone un'ultima condizione di rotazione attorno alla normale in P ; ad esempio imponendo che l'azimut α geodetico da P a un secondo punto Q sia uguale all'azimuth A astronomico: $\alpha_{\widehat{PQ}} = A_{\widehat{PQ}}$.

2.9 Le scale temporali

Le scale temporali servono per tempo-referenziare un evento e si possono raggruppare in tre famiglie principali: per ciascuna famiglia si riassumono qui le scale fondamentali nell'ambito geodetico.

2.9.1 Le scale di tempo dinamico

Si osservano le posizioni reciproche di Sole, Luna e altri pianeti del sistema solare; mediante confronto delle posizioni osservate con le effemeridi del sistema solare si risale alla stima dell'epoca di osservazione. Le scale dinamiche presentano stabilità di circa 10^{-10} , ovvero un errore accumulato di circa $0.05s/10y$. Fra queste si hanno le due scale fondamentali Ephemeris Time (ET) e Geocentric Coordinate Time (TCG); ET venne introdotta nel 1960 e rappresenta la prima scala di tempo dinamico: in effetti è stata poi sostituita con realizzazioni più recenti e accurate, ovvero basate su migliori effemeridi e descrizioni dinamiche del sistema solare; fra queste, la scala TCG è una scala dinamica in cui il tempo osservato da osservatori solidali al moto terrestre viene corretto per l'effetto relativistico indotto dal moto terrestre: si ricorda che TCG è la scala che definisce le lunghezze di ITRS.

2.9.2 Le scale di tempo siderale e solare

Queste sono legate all'osservazione del moto di rotazione terrestre rispetto a una direzione di riferimento nello spazio: le scale siderali utilizzano la direzione dell'equinozio mentre le scale solari adottano la direzione del Sole.

Le varie scale esistenti si distinguono per più aspetti. Innanzitutto la direzione di riferimento può essere quella attuale o una direzione media, ovvero depurata da qualche effetto perturbativo: ad esempio si può utilizzare l'equinozio effettivo oppure quello depurato dell'effetto della nutazione. Inoltre il tempo può essere espresso come angolo orario fra meridiano di Greenwich e direzione di riferimento (Greenwich time) oppure come angolo orario fra meridiano locale nel punto di osservazione e direzione di riferimento (Local Time).

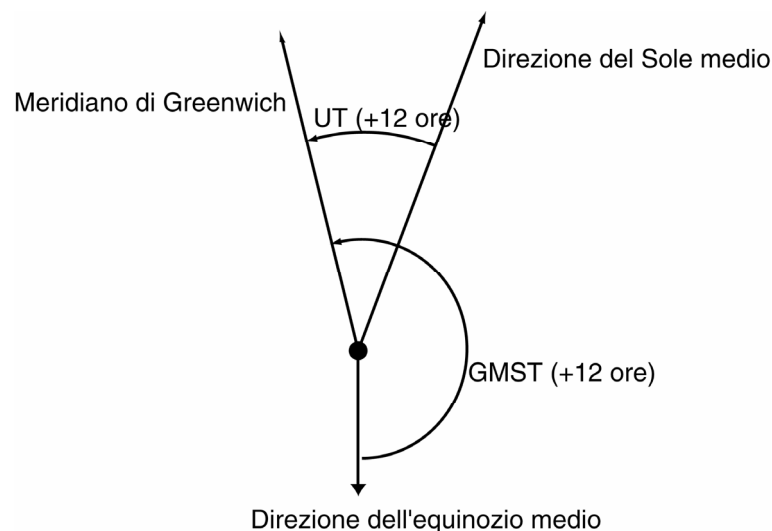


Figura 2.37. Direzioni fondamentali delle scale dei tempi siderale e solare.

Si ricordano in particolare

- UT1 (Universal Time 1), scala di tempo solare: angolo orario fra meridiano di Greenwich e direzione del Sole medio, depurato degli effetti del polar motion, aumentato di 12 ore (in modo che il mezzogiorno coincida con il Sole allo zenit),
- GMST (Greenwich Mean Sidereal Time), scala di tempo siderale: angolo orario fra meridiano di Greenwich e equinozio medio, aumentato di 12 ore.

A causa dell'irregolarità del moto di rotazione terrestre (attualmente la rotazione terrestre rallenta di circa $0.5s/y$) queste scale di tempo presentano irregolarità significative. Si ricorda inoltre che tempo solare e tempo siderale differiscono poiché, a causa del moto di rivoluzione terrestre intorno al Sole, la direzione di quest'ultimo rispetto la Terra retrocede di giorno in giorno: in un anno (tempo tra due passaggi del Sole) vi sono 365.2422 giorni solari e 366.2422 giorni siderali.

2.9.3 Le scale di tempo atomico

Le scale di tempo atomico sono definite da un certo numero di orologi atomici al Cesio 133. Sono le scale di tempo più accurate, con una stabilità stimata di 10^{-14} , che in prospettiva arriverà 10^{-16} .

La scala di tempo fondamentale è la TAI (International Atomic Time), definita dal BIPM di Parigi: TAI ha origine definita in modo da coincidere con UT nel 1958.0, la durata di un secondo è definita uguale a 9192631770 periodi della radiazione generata dalla transizione tra due livelli iperfini dello stato base del cesio 133. Vista l'elevata stabilità di TAI e TCG, è possibile scrivere una relazione matematica "esatta" fra le due scale:

$$\begin{aligned}
 TT &= TAI + 32.184s \\
 TCG &= TT + 0.022s/y \cdot (y - 1977.0)
 \end{aligned}
 \tag{2.15}$$

Il primo passaggio tiene semplicemente conto del disallineamento nelle scelte di origine; il secondo corregge per l'effetto relativistico cui sono soggetti gli orologi atomici operanti sulla superficie terrestre.

UTC (Coordinated Universal Time) ha stessa cadenza del TAI; viene però periodicamente riallineato per mantenerlo sincronizzato entro il secondo con la scala solare UT1: ciò, in questo secolo, significa la sottrazione di un secondo ogni circa due/tre anni. GPST (GPS Time) è il tempo atomico introdotto per il sistema GPS; GPST è definito in modo da coincidere, al 6 gennaio 1980, ore 00.00, con l'UTC e poi da proseguire allineato con il TAI. Perciò la differenza fra GPST e TAI è costante,

$$GPST = TAI - 19s;$$

la differenza fra GPST e UTC viceversa cambia a ogni riallineamento di UTC: per gli ultimi tre anni e per il futuro prossimo si ha

$$\begin{aligned}
 GPST &= UTC + 14s, 2006.0 \leq t < 2009.0 \\
 GPST &= UTC + 15s, 2009.0 \leq t
 \end{aligned}
 \tag{2.16}$$

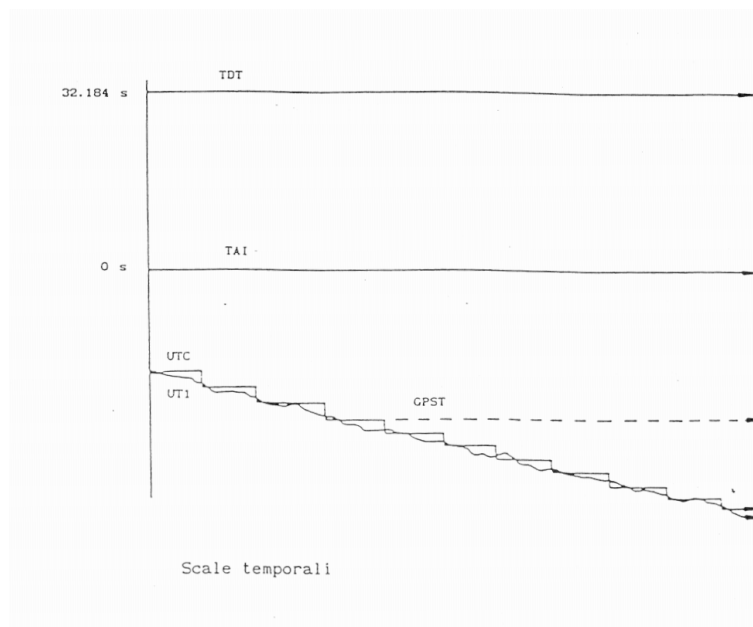


Figura 2.38. Offset fra scale dei tempi TAI, UTC e GPST.

Capitolo 3. Il sistema GPS

Il progetto GPS venne sviluppato negli anni '70 presso il DoD, ovvero il Department of Defence degli USA e nel 1978 fu lanciato il primo satellite; dal 1993 il GPS iniziò la piena operatività H24 in fase sperimentale, detta IOC (Initial Operational Capability); dal 1995 è ufficialmente dichiarata la FOC (Full Operational Capability). Il progetto del GPS nacque per la navigazione in tempo reale: all'epoca del progetto erano previste le massime accuratèzze potenziali, ovvero il Precise Positioning Service (PPS) con accuratèzze dell'ordine di 10 metri in planimetria, per le applicazioni militari; era anche previsto un servizio civile, lo Standard Positioning Service (SPS) con accuratèzze dell'ordine di 100 metri in planimetria. Negli ultimi 40 anni le strategie degli Stati Uniti d'America sono cambiate, anche in funzione degli sviluppi del mercato, e oggi le accuratèzze inizialmente previste per il PPS sono garantite a tutti: gli aspetti della navigazione in modalità di posizionamento assoluto saranno discusse nel Cap. 5. Parallelamente allo sviluppo delle tecniche di posizionamento assoluto per la navigazione si è avuto anche lo sviluppo delle tecniche di posizionamento relativo per la navigazione, le applicazioni topografiche e quelle geodetiche; si può ora dire che anche queste sono ormai quasi completamente sviscerate, e la relativa discussione impegnerà i Capp. 6 e 7. Il presente e il prossimo capitolo sono finalizzati all'introduzione del sistema e alla discussione dei suoi segnali e delle sue equazioni di osservazione.

3.1 Le orbite e le effemeridi dei satelliti

Nel seguito ci occuperemo solo dell'orbita definitiva di un satellite GPS e non delle complesse procedure per il suo lancio; inoltre si considera l'approssimazione puntiforme, ovvero non si considerano le rotazioni del satellite ma solo il moto del suo centro di massa. In generale in un sistema inerziale l'accelerazione $\ddot{\mathbf{x}}_p$ di un corpo puntiforme P di massa m_p è proporzionale alla somma delle forze agenti sul corpo stesso, $\mathbf{f}_i(\mathbf{x}_p, \dot{\mathbf{x}}_p, t)$: queste possono dipendere dalla posizione del corpo, $\mathbf{x}_p$, dalla sua velocità, $\dot{\mathbf{x}}_p$, e dal tempo t ; si ha cioè

$$\ddot{\mathbf{x}}_p(t) = \frac{1}{m_p} \sum_i \mathbf{f}_i(\mathbf{x}_p, \dot{\mathbf{x}}_p, t)$$

nel caso di satelliti artificiali terrestri, come il GPS, la forza dominante è quella centrale gravitazionale terrestre, la cui espressione è data dalla

$$\mathbf{f}_G = -GM_T m_p \frac{(\mathbf{x}_P - \mathbf{x}_G)}{\|\mathbf{x}_P - \mathbf{x}_G\|^3}$$

ove G è la costante universale di gravitazione, M_T è la massa della Terra, $\mathbf{x}_G$ è la posizione del geocentro.

Le altre forze agenti sul satellite, in ordine di significatività, sono la componente non centrale della forza gravitazionale terrestre, dovuta alle disomogeneità del pianeta, le perturbazioni temporali alla forza gravitazionale terrestre, dovute alle maree terrestri e oceaniche, l'attrazione gravitazionale di altri corpi del sistema solare, in particolare della Luna e il Sole, la pressione di radiazione solare, dovuta alla radiazione solare sui pannelli e infine l'attrito atmosferico, che per il GPS è dovuto a particelle libere nello spazio. In particolare la forza gravitazionale centrale terrestre supera di tre ordini di grandezza le altre forze: si può dunque scrivere

$$\ddot{\mathbf{x}}_P = -GM_T \frac{(\mathbf{x}_P - \mathbf{x}_G)}{\|\mathbf{x}_P - \mathbf{x}_G\|^3} + \delta\ddot{\mathbf{x}}_P \quad (3.1)$$

ove $\delta\ddot{\mathbf{x}}_P$ indica l'effetto delle forze perturbative agenti sul satellite; naturalmente l'effetto di $\delta\ddot{\mathbf{x}}_P$ sulla posizione del satellite dopo un certo arco di tempo è ottenuta mediante integrazione: in Tab. 3.1 sono riportati alcuni valori di esempio.

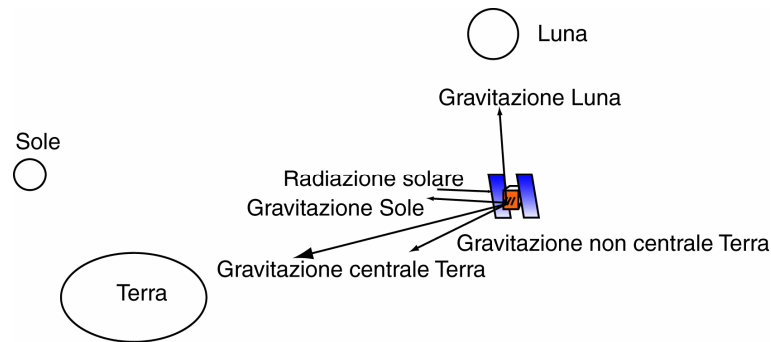


Figura 3.1. Rappresentazione grafica, non proporzionale, delle forze agenti su un satellite.

3.2.1 L'orbita Kepleriana

Si consideri solo la forza dominante sul satellite GPS, ovvero la forza gravitazionale terrestre centrale; in un sistema di riferimento inerziale, ovvero fisso rispetto alle stelle fisse, i satelliti GPS si muovono lungo orbite perfettamente ellittiche e predicibili nel tempo, in accordo alle leggi di Keplero. Si ricorda che le leggi di Keplero vennero formulate per descrivere il moto dei pianeti solari intorno al Sole: si prestano però

anche alla descrizione del moto dei satelliti artificiali terrestri (e della Luna) intorno alla Terra.

K1. L'orbita di un satellite ha forma ellittica e il centro di massa della Terra occupa uno dei due fuochi; si definiscono i seguenti punti particolari dell'orbita: apogeo, ovvero il punto dell'orbita più lontano dalla Terra, perigeo, ovvero il punto dell'orbita più vicino alla Terra. L'apogeo e il perigeo occupano posizioni stazionarie in un sistema di riferimento inerziale.



Figura 3.2. L'orbita di un satellite artificiale terrestre:
il satellite spazza aree uguali in tempi uguali.

K2. Il vettore posizione, ovvero il vettore che congiunge il geocentro al satellite, spazza aree uguali in tempi uguali: in altri termini il satellite si muove più rapidamente al perigeo e più lentamente all'apogeo (Fig. 3.2).

K3. Fra periodo di rivoluzione del satellite T e semiasse maggiore a dell'orbita vale una relazione precisa:

$$T^2 / a^3 = 4\pi^2 / GM$$

In un sistema inerziale un'orbita Kepleriana è stazionaria e predicibile all'infinito; in particolare può essere descritta mediante sei parametri che sono

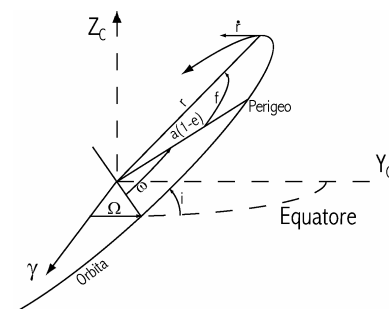


Figura 3.3. I parametri dell'orbita Kepleriana rispetto a un sistema inerziale

i : inclinazione dell'orbita, ovvero l'angolo fra il piano orbitale e il piano equatoriale di riferimento.

Ω : ascensione retta del nodo ascendente, ovvero l'angolo sul piano equatoriale di riferimento fra la direzione dell'equinozio di riferimento γ e l'intersezione fra piano equatoriale e piano orbitale;

ω : argomento del perigeo, ovvero l'angolo sul piano orbitale fra l'intersezione del piano orbitale con il piano equatoriale di riferimento e la direzione del perigeo dell'orbita;

T_0 : epoca di passaggio per il perigeo;

a : semiasse maggiore dell'orbita.

e : eccentricità dell'orbita.

Si indica con $f(t)$ (detto anomalia vera) l'angolo istantaneo nel fuoco dell'orbita fra il perigeo e la posizione del satellite. f è funzione del tempo e, conoscendo il momento T_0 di passaggio dal perigeo e i parametri a e e dell'orbita è possibile calcolarlo per qualunque epoca; quindi è possibile calcolare la posizione del satellite $[X_{Orb}, Y_{Orb}]$ nel piano orbitale. Non si riportano qui le formule, reperibili in qualunque testo di calcolo orbitale: si riportano invece in Eq. 3.2 le loro derivazioni applicate per il calcolo delle orbite GPS.

Dunque, utilizzando gli angoli i , Ω , ω per effettuare una rotazione fra sistemi di riferimento, si possono trasformare le coordinate dal piano orbitale al sistema di riferimento inerziale celeste; infine si utilizzano le trasformazioni tabulate e pubblicate per trasformare queste ultime al sistema di riferimento terrestre (ad esempio IGS).

3.2.2 Gli effetti delle altre forze

Le altre forze agenti sul satellite provocano perturbazioni dell'orbita reale rispetto all'orbita Kepleriana che, per brevi archi di tempo, possono essere descritte mediante un insieme limitato di parametri: in particolare l'effetto dovuto alla forma ellissoidica della Terra fa sì che l'ellissi kepleriana non mantenga una geometria costante rispetto alle stelle fisse, e che gli angoli Ω, ω, i varino linearmente nel tempo. La modellizzazione finale dell'orbita prende il nome di modellizzazione quasi Kepleriana e richiede 16 parametri numerici anziché i sei del caso ideale.

Dati i 16 parametri è possibile, mediante gli algoritmi descritti al termine del Par. 3.3.1, calcolare la posizione di un satellite GPS per qualunque epoca nell'intervallo di validità dei suddetti parametri: la modellizzazione quasi Kepleriana descrive con accuratezza di circa un metro l'orbita di un satellite per un arco di tempo di circa un'ora.

Tipo di forza	Effetto sull'orbita in 24h (Km)
Forza gravitazionale centrale terrestre	350000.0
Forza gravitazionale non centrale terrestre	10.0
Forza gravitazionale lunare	3.0
Forza gravitazionale solare	0.8
Pressione di radiazione solare	0.2

Tabella 3.1. Effetti delle forze agenti sul satellite GPS integrati su 24 ore

3.3 Le effemeridi del GPS

Le effemeridi di un satellite sono l'insieme di parametri e algoritmi che permettono di calcolarne la posizione ad ogni epoca in un sistema di riferimento assegnato: nel caso dei satelliti GPS si hanno le effemeridi trasmesse e le effemeridi a posteriori.

3.3.1 Le effemeridi trasmesse e la relativa procedura di calcolo

I satelliti GPS trasmettono in tempo reale le proprie effemeridi mediante la parametrizzazione quasi Kepleriana e i parametri vengono aggiornati ogni ora; queste sono le effemeridi trasmesse, che vengono utilizzate per la navigazione in tempo reale. Naturalmente il satellite non è in grado di predire quotidianamente la propria orbita e calcolare i parametri che la descrivono numericamente: tale compito è affidato alla rete statunitense di stazioni permanenti per il controllo del sistema, detta Control Segment. La rete di controllo si componeva inizialmente di cinque stazioni, circa equidistanziate lungo l'equatore terrestre, ma dal 2005 alla rete originaria sono state aggiunte ulteriori sei stazioni (Fig. 3.4); il calcolo delle orbite è implementato nel seguente processo:

1. le osservazioni di tutte le stazioni della rete a tutti i satelliti vengono inviate alla stazione Master (sita in Colorado Spring) che, per ciascun satellite:
2. interpola a posteriori l'orbita dei sette giorni immediatamente precedenti;
3. utilizza l'orbita interpolata per estrapolarne l'evoluzione nelle 24 ore immediatamente successive;
4. si stimano le 24 modellazioni quasi Kepleriane orarie che meglio si adattano alle 24 ore estrapolate,
5. i parametri delle modellazioni vengono trasmessi al satellite,
6. il quale durante la giornata successiva li trasmetterà agli utenti.



Figura 3.4. La rete di controllo del NIMA.
In rosso le stazioni originarie, in rosa quelle aggiunte nel 2005.
Dal sito <http://www.kowoma.de>.

Le effemeridi trasmesse dai satelliti sono anche dette effemeridi predette e permettono il calcolo della posizione dei satelliti con accuratezza di circa 1 m; esse sono espresse nel

sistema di riferimento WGS84, che sostanzialmente è una realizzazione di ITRS materializzata dalla rete di controllo del GPS. La differenza fra WGS84 e ITRF è inferiore al decimetro, ovvero molto inferiore all'accuratezza delle effemeridi trasmesse, quindi, da un punto di vista operativo, i due sistemi di riferimento sono indistinguibili: tali aspetti sono discussi nel Par. 5.6 per il posizionamento assoluto e Par. 6.13 per quello relativo.

Per completezza si riportano qui le formule necessarie per calcolare la posizione di un satellite dalle effemeridi trasmesse. I parametri comunicati tramite ogni pacchetto sono: M_0 : anomalia media all'epoca di riferimento; Δn : differenza media di moto rispetto al moto di riferimento; e : eccentricità dell'orbita; $\sqrt{a}$: radice quadrata del semiasse maggiore; Ω_0 : longitudine del nodo ascendente all'inizio della settimana GPS; i_0 : inclinazione dell'orbita all'epoca di riferimento; ω : argomento del perigeo; $\dot{\Omega}$: derivata nel tempo dell'ascensione retta; $IDOT$: derivata nel tempo dell'inclinazione; C_{uc} , C_{us} : coseno e seno della correzione all'argomento della latitudine; C_{rc} , C_{rs} : coseno e seno della correzione al raggio dell'orbita; C_{ic} , C_{is} : coseno e seno della correzione all'inclinazione dell'orbita; t_{oe} : epoca di riferimento per le effemeridi, in secondi della settimana GPS; $IODE$: epoca di aggiornamento delle effemeridi. Le costanti da utilizzarsi nei seguenti calcoli sono quella gravitazionale terrestre, $\mu = 3.986005 \cdot 10^{14} m^3 / s^2$, e la velocità angolare di rotazione terrestre, $\dot{\Omega}_E = 7.2921151467 \cdot 10^{-5} rad / s$. I calcoli avvengono nel seguente ordine:

$a = (\sqrt{a})^2$:	semiasse maggiore dell'orbita
$n_0 = \sqrt{\frac{\mu}{a^3}}$:	moto di riferimento
$\Delta t = t - t_{oe}$:	differenza fra tempo corrente e epoca di riferimento
$n = n_0 + \Delta n$:	moto medio
$M(t) = M_0 - n\Delta t$:	anomalia media

Si trova quindi il valore di $E(t)$, ovvero dell'anomalia dell'eccentricità, risolvendo iterativamente l'equazione $M(t) = E(t) - e \sin E(t)$; quindi si prosegue con

$f(t) = \tan^{-1} \left(\frac{\sqrt{1-e^2} \sin E(t)}{\cos E(t) - e} \right)$:	anomalia vera
$\phi(t) = f(t) + \omega$:	argomento della latitudine

$$\begin{aligned} \delta u(t) &= C_{us} \sin 2\phi(t) + C_{uc} \cos 2\phi(t) \\ \delta r(t) &= C_{rc} \cos 2\phi(t) + C_{rs} \sin 2\phi(t) : \text{termini di correzione} \\ \delta i(t) &= C_{ic} \cos 2\phi(t) + C_{is} \sin 2\phi(t) \\ u(t) &= \phi(t) + \delta u(t) : \text{argomento della latitudine corretto} \\ r(t) &= a \cdot (1 - e \cos E(t)) + \delta r(t) : \text{raggio corretto} \\ i(t) &= i_0 + \delta i(t) + IDOT \cdot \Delta t : \text{inclinazione corretta} \\ \begin{aligned} x(t) &= r(t) \cos u(t) \\ y(t) &= r(t) \sin u(t) \end{aligned} : \text{posizione del satellite nel piano orbitale} \\ \Omega(t) &= \Omega_0 + (\dot{\Omega} - \dot{\Omega}_e) \Delta t - \dot{\Omega}_e t_{oe} \quad \text{longitudine del nodo ascendente corretta} \\ X(t) &= x(t) \cos \Omega(t) - y(t) \cos i(t) \sin \Omega(t) \\ Y(t) &= x(t) \sin \Omega(t) + y(t) \cos i(t) \cos \Omega(t) \quad \text{posizione del satellite in WGS84} \\ Z(t) &= y(t) \sin i(t) \end{aligned}$$

Equazione 3.2. Il calcolo della posizione di un satellite GPS dalle effemeridi trasmesse.

3.3.2 Le effemeridi calcolate a posteriori

In rete sono disponibili anche le effemeridi dei satelliti GPS calcolate a posteriori, che presentano maggiori accuratezze di quelle trasmesse; naturalmente le effemeridi a posteriori non sono utilizzabili per navigazione ma hanno ampia applicazione nella postelaborazione di precisione dei dati. Innanzitutto ricordiamo le effemeridi precise dell'esercito statunitense, disponibili presso il sito della NGA, ma molto più importanti in ambito internazionale sono le effemeridi precise calcolate e distribuite da IGS; queste vengono distribuite nel formato SP3, ovvero file ASCII (Fig. 3.5) contenenti le posizioni di tutti i satelliti GPS tabulate ad intervalli regolari di tempo (attualmente di 15 minuti).

Per calcolare la posizione di un satellite a un'epoca diversa da quelle tabulate si ricorre tipicamente a un'interpolazione Lagrangiana, condotta indipendentemente sulle tre coordinate; sia x la generica coordinata tabulata di un satellite per un certo numero n di epoche t : $x(t_1), x(t_2), \dots, x(t_n)$; il polinomio interpolante di grado $n-1$ può essere calcolato mediante la

$$P(t) = \sum_{i=1}^n \alpha_i(t) x(t_i) \quad (3.3)$$

ove

$$\alpha_i(t) = \frac{\prod_{j=1, n; j \neq i} (t - t_j)}{\prod_{j=1, n; j \neq i} (t_i - t_j)}$$

per calcolare la posizione all'epoca t si utilizza una finestra di valori tabulati centrata su t : tipicamente si adotta $n = 10$.

```

... HEADER OMESSO...
* 2007  9 30  0  0  0.00000000
PG01  2667.489628  14512.617647  22299.283328  155.952112
PG02 -12204.536387 -23228.393961  4878.362107  140.117083
...
PG32 -2811.094639  22575.212168  14287.586567  99999.999999
* 2007  9 30  0 15  0.00000000
PG01  173.660695  14683.334461  22339.970226  155.952820
...
* 2007  9 30 23 45  0.00000000
...
EOF

```

Figura 3.5. Esempio di file SP3. Il file inizia con alcune righe di intestazione, qui non riportate; poi, per ogni epoca vengono riportati anno, mese, giorno, ora, minuto, secondo; per ogni satellite sono riportate le tre coordinate cartesiane geocentriche (in Km) e l'offset di orologio (in μs)...; così si ripete per tutte le epoche tabulate fino all'ultima epoca della giornata.

Le effemeridi finali IGS hanno accuratezza dell'ordine del centimetro ma sono distribuite con una latenza di 14 giorni; l'elevata latenza pone evidenti problemi a chi desidera utilizzare effemeridi molto accurate ma abbia urgenza per elaborare i propri dati: perciò IGS rilascia e distribuisce anche effemeridi a bassa latenza, ovviamente caratterizzate da accuratezze inferiori al prodotto finale: le effemeridi ultrarapide sono pubblicate in tempo reale e hanno accuratezza di 10 cm, le effemeridi rapide sono pubblicate con latenza di tre giorni e hanno accuratezza migliore di 5 cm.

3.4 Il satellite GPS

Ogni satellite è caratterizzato dal suo codice SVN (Space Vehicle Number) e dal blocco di appartenenza:

- blocco I: prima generazione di satelliti, lanciati dal 1978 al 1985 che hanno permesso le prime sperimentazioni ma sono ormai completamente dismessi;
- blocchi II e IIA, SVN dal 13 al 40: lanciati dal 1989 al 1997 hanno segnato l'avvio della piena operatività del sistema e alcuni di essi sono ancora in orbita;
- blocchi IIR e IIR-M, SVN dal 41: lanciati dal 1997 a oggi, hanno apportato significativi miglioramenti tecnologici rispetto alla generazione precedente; in

particolare quelli più recenti del blocco IIR-M implementano il nuovo codice civile L2C (vedi Par. 3.6): ne sono oggi operativi sei.

- blocco IIF: il loro lancio era inizialmente previsto a partire dal 2008, ma è stato ritardato; implementeranno la terza frequenza L5.

Si riportano schematicamente le caratteristiche e le dotazioni dei satelliti del blocco IIR-M (Fig. 3.6): i satelliti hanno dimensioni di circa $2m \times 2m \times 1.5m$, massa di circa $2 \cdot 10^3 kg$, sono forniti di pannelli solari per il proprio approvvigionamento e di retrorazzi per le periodiche manovre di correzione dell'orbita: hanno vita media prevista di dieci anni. Al fine di generare i segnali di tempo, i satelliti dispongono di tre oscillatori al rubidio; inoltre possono, seppur con degradazione delle accuratèzze, calcolare e predire le proprie orbite senza dipendere dalla rete di controllo.



Figura 3.6. Raffigurazione pittorica di un satellite del Blocco IIR.
Dal sito <http://gps.losangeles.af.mil/index.html>

3.5 La costellazione GPS

Il sistema è stato progettato in modo tale da garantire globalmente H24 la visibilità di almeno quattro satelliti: questo è infatti il requisito per la piena operatività; in particolare essa è garantita da una costellazione nominale di 24 satelliti, ripartiti a gruppi di quattro su sei piani orbitali; l'inclinazione media di ciascun piano orbitale rispetto al piano equatoriale è di 55° . Le orbite sono quasi circolari, ovvero l'eccentricità è quasi nulla, e hanno raggio di circa 26000 km: il periodo è quindi di 12 ore siderali e ciascun satellite si muove nello spazio a circa 4 km/s.

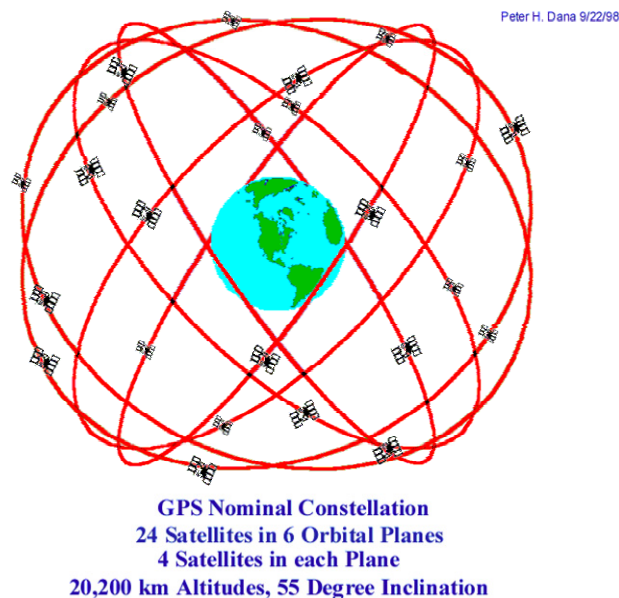


Figura 3.7. Le orbite dei satelliti GPS.

Dal sito http://www.colorado.edu/geography/gcraft/notes/gps/gps_f.html.

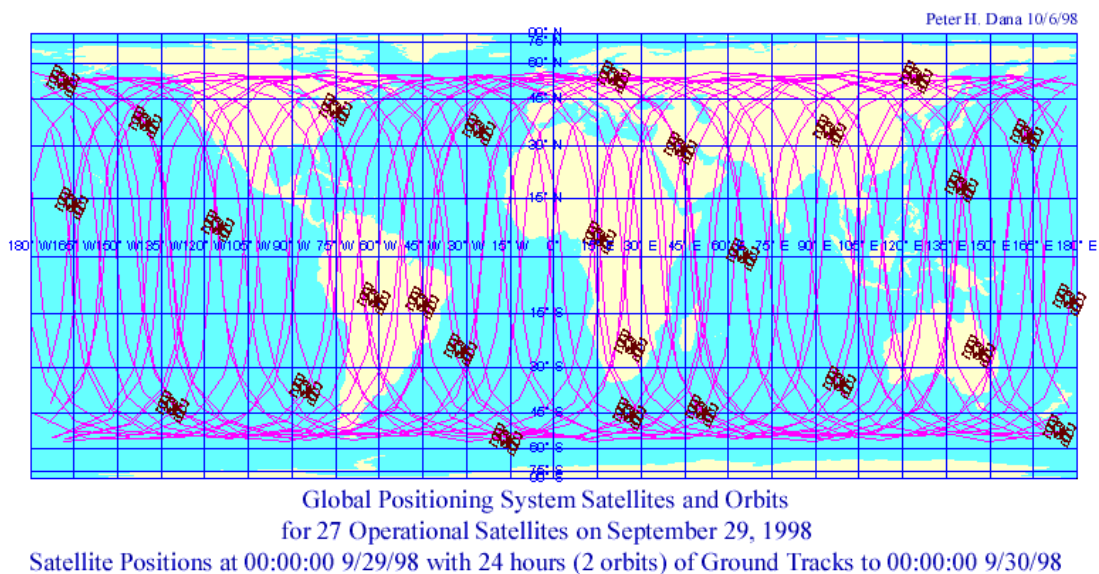


Figura 3.8. Traccia delle orbite dei satelliti GPS al suolo.

Dal sito http://www.colorado.edu/geography/gcraft/notes/gps/gps_f.html.

L'elevazione di un satellite S rispetto a un ricevitore R , indicata con η_R^S , è definita come l'angolo fra il piano orizzontale nel punto del ricevitore e la direzione del satellite; eccettuate applicazioni particolari, si utilizzano solo segnali provenienti da satelliti con elevazione superiore a 10° , al di sotto della quale vi è un significativo

effetto dei disturbi atmosferici; tipicamente ciascun satellite rimane sopra i 10° di elevazione per 1-6 ore.

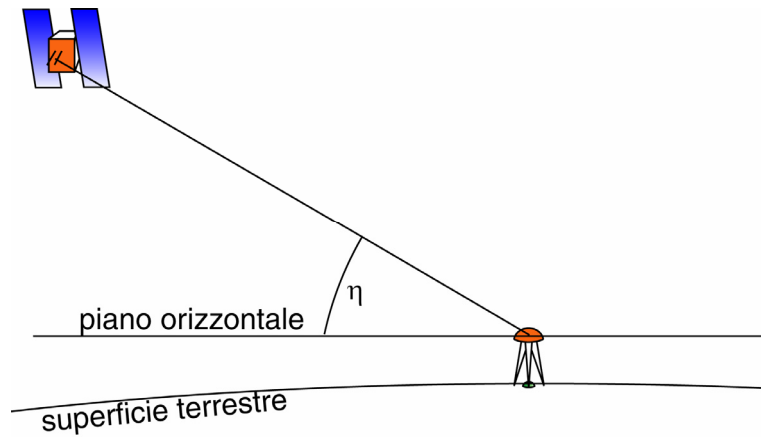


Figura 3.9. Elevazione di un satellite rispetto a un ricevitore.

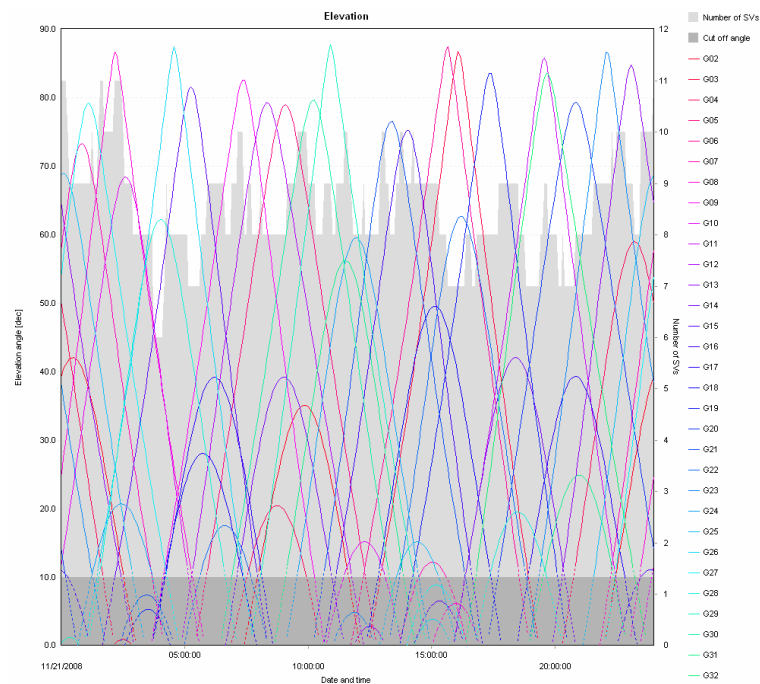


Figura 3.10. Elevazione e numero di satelliti a Como, per il 21/11/2009.

Si è detto che l'operatività globale H24 è garantita con quattro satelliti in vista, ovvero 24 satelliti in orbita; attualmente sono in orbita 31 satelliti, che garantiscono in effetti ridondanza al sistema: in Fig. 3.11 è mostrata la visibilità in Como.

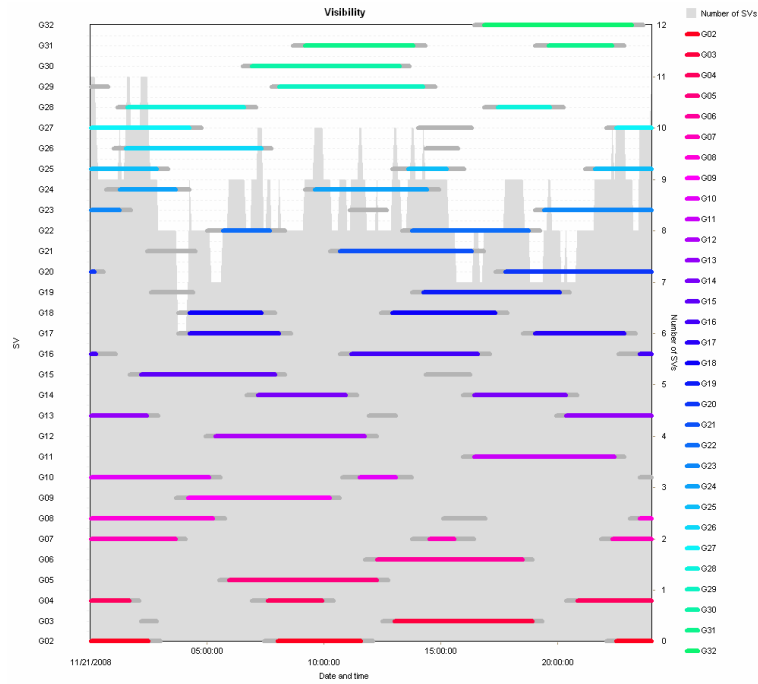


Figura 3.11. Visibilità (cut off a 10°) e numero di satelliti in Como, per il 21/11/2009.

3.6 Il segnale GPS

Gli oscillatori a bordo dei satelliti producono un segnale di frequenza nominale f_0 pari a 10.23 MHz; tale segnale è caratterizzato da elevatissima stabilità, $\Delta f / f = 10^{-13} - 10^{-14}$, ma non è comunque perfettamente costante nel tempo: nel Cap. 4 si vedranno gli effetti delle fluttuazioni di frequenza sul segnale di tempo emesso da un oscillatore atomico.

A partire da f_0 vengono generate due sinusoidi (o portanti), dette L1 e L2 e diversi codici binari: C/A (Coarse Acquisition Code), L2C (L2 Civilian), P(Y) (EncrYpted Precise Code) e M (Military). Viene infine emesso il messaggio di navigazione D (Navigation Data).

3.6.1 Le portanti

Un fenomeno sinusoidale che si ripete nel tempo con esatta ciclicità è descritto dalla seguente equazione:

$$A(t) = A_0 \sin(\omega t + \varphi_0) \quad (3.4)$$

dove A_0 è l'ampiezza del segnale, ω è la pulsazione (rad/s), φ_0 è la fase iniziale (rad); l'espressione $\omega t + \varphi_0$ è detta fase di $A(t)$. Il periodo di $A(t)$ è il tempo necessario per il completamento di un intero ciclo ed è dato dalla

$$T = \frac{2\pi}{\omega}$$

infatti

$$\begin{aligned} A(t+T) &= A_0 \sin\left(\omega\left(t + \frac{2\pi}{\omega}\right) + \varphi_0\right) = \\ &= A_0 \sin(\omega t + 2\pi + \varphi_0) = A(t) \end{aligned}$$

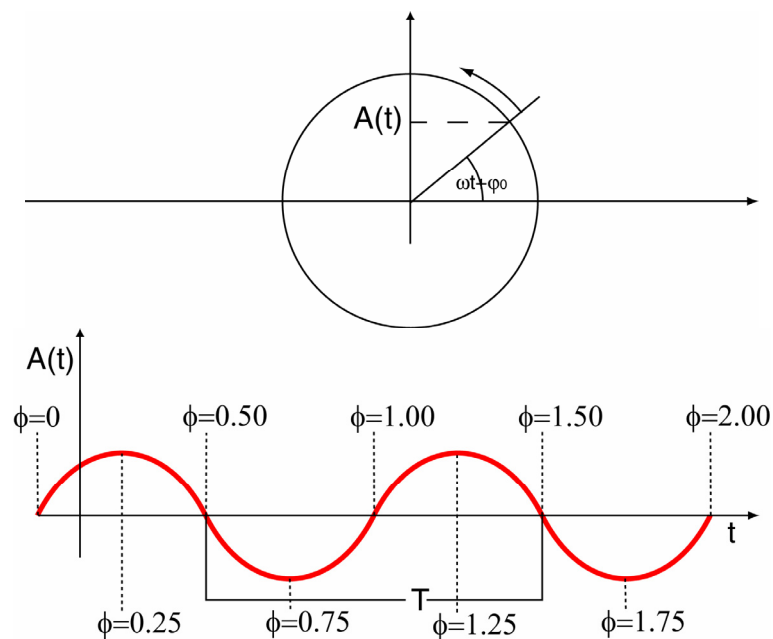


Figura 3.12. Un fenomeno ciclico, la sua ampiezza e la sua fase nel tempo.

La fase viene usualmente misurata in radianti, ma nelle nostre applicazioni è più utile utilizzare le frazioni o percentuali di ciclo:

$$\phi(t) = \frac{\varphi(t)}{2\pi}$$

Si introduce la frequenza, ovvero la derivata prima nel tempo della fase espressa in frazioni di ciclo:

$$f = \frac{d\phi}{dt} = \frac{\omega}{2\pi} \quad (3.5)$$

la (3.4) diviene

$$A(t) = A_0 \sin(2\pi f t + \varphi_0)$$

Per un fenomeno ideale, a frequenza costante f_0 , si ha la seguente

$$\int_0^T f dt = f_0 \int_0^T dt = f_0 T = \int_0^1 d\phi = 1$$

ovvero

$$f_0 = \frac{1}{T} \quad (3.6)$$

Si vedrà nel prossimo capitolo il caso degli oscillatori reali, la cui frequenza fluttua nel tempo. Pensiamo ora a un fenomeno oscillatorio che si propaghi sia nel tempo sia nello spazio, come ad esempio le onde nel mare; la sua legge di propagazione lungo la traiettoria x è data da

$$A(x, t) = A_0 \sin(2\pi f(t - \frac{x}{c}) + \varphi_0) = A_0 \sin(2\pi(\frac{t}{T} - \frac{x}{\lambda}) + \varphi_0) \quad (3.7)$$

rispetto a prima compaiono c , che rappresenta la velocità di propagazione del segnale lungo la traiettoria, e λ , la lunghezza d'onda ovvero la distanza nello spazio, alla medesima epoca, fra due successive ripetizioni del segnale; vale la seguente relazione

$$\lambda = cT = c / f$$

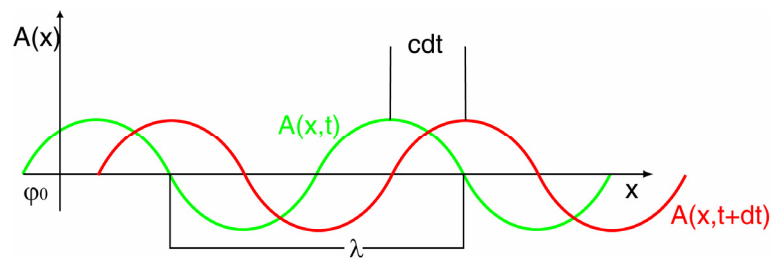


Figura 3.13. Segnale sinusoidale che si propaga nello spazio.

Le caratteristiche fondamentali delle due portanti generate dal satellite GPS sono riportate nella seguente tabella.

Nome	Frequenza	Lunghezza d'onda
L1	$154f_0 = 1572.42\text{MHz}$	19.029 cm
L2	$120f_0 = 1227.60\text{MHz}$	24.421 cm

Tabella 3.2. Caratteristiche delle portanti del GPS: le lunghezze d'onda sono approssimate al millesimo di centimetro: in un'implementazione numerica il loro calcolo deve essere effettuato mediante la formula esatta

$$\lambda_i = c / f_i, \quad c = 299792458\text{m/s} \quad (3.8)$$

3.6.2 I codici binari

Un codice binario è una successione di impulsi con valori pari a +1 e -1: la sequenza di trasmissione degli impulsi, secondo un'opportuna chiave di lettura, costituisce il contenuto del segnale.

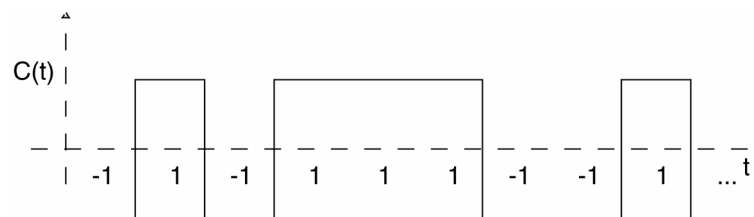


Figura 3.14. Un esempio di codice binario composto dalla sequenza (-1,1,-1,1,1,1,-1,-1,1) ovvero 9 impulsi.

Gli impulsi hanno tutti durata uguale: il periodo del codice è il tempo necessario per la trasmissione dell'intero codice ed è dato dal prodotto del numero di impulsi che costituiscono il codice per la durata di un impulso. La frequenza è convenzionalmente definita come l'inverso della durata di un singolo impulso; la lunghezza d'onda del codice è convenzionalmente definita come la lunghezza nello spazio di un singolo impulso. I codici binari trasmessi dai satelliti GPS vengono detti Pseudo Random (PR) perché, apparentemente, i loro impulsi seguono un ordine quasi casuale; in effetti hanno un contenuto informativo ben preciso. I tre codici PR hanno le seguenti caratteristiche:

Codice C/A: è composto da 1023 impulsi, ciascuno di durata pari a 1/1023 ms, ovvero con lunghezza nello spazio di 293.0 m, per una durata totale di 1 ms; è trasmesso da tutti i satelliti attualmente in orbita, è diverso per ogni satellite ed è pubblico.

Codice L2C: è composto da 10230 impulsi, ciascuno di durata pari a 1/1023 ms, ovvero con lunghezza nello spazio di 293.0 m, per una durata totale di 10 ms; è trasmesso solo dai satelliti del Blocco IIR-M ed è pubblico.

Nome	f (MHz)	λ (m)	Numero Impulsi	T
C/A	$0.1f_0 = 1.023$	293.0	1023	1 ms
L2C	$0.1f_0 = 1.023$	293.0	10230	10 ms
P(Y)	$f_0 = 10.230$	29.3	$3.2703264 \cdot 10^{16}$	37 settimane
M	$f_0 = 10.230$	29.3	?	?

Tabella 3.3. Caratteristiche dei codici binari del GPS

Codice P(Y): è il codice originariamente pensato per le applicazioni militari. Ogni impulso dura $1/10230$ ms, per una lunghezza di 29.3 m; il codice completo si compone di 3.2703264×10^{16} impulsi, dura 37 settimane e ogni satellite ne trasmette una differente porzione di durata di una settimana. Prima del 1994 il codice aveva nome di codice P, ed era pubblico; dall'agosto 1994, con la dichiarazione di piena operatività del sistema, il codice P è stato commutato nel codice Y, criptato e noto solo nelle memorie dei ricevitori in possesso del US Army; tecnicamente, il codice Y è ottenuto moltiplicando l'originario codice P, pubblico, con un codice W, segreto, caratterizzato da bassa frequenza.

Codice M. Oltre al codice L2C, i nuovi satelliti hanno introdotto anche un nuovo codice militare criptato, trasmesso su entrambe le portanti, con frequenza uguale a quella di P(Y); allo stato attuale non si sono trovate nella letteratura consultata informazioni univoche su tale codice.

Oltre ai codici pseudocasuali viene inviato anche il messaggio navigazionale D (Navigation Data), che serve per comunicare le effemeridi dei satelliti e altre informazioni. La frequenza di D è molto bassa, pari a 50 Hz, e il messaggio si compone di 25 blocchi di 30 secondi ($25 \times 30 \times 50$ bit) ciascuno, per una durata complessiva di 12.5 minuti; ogni blocco contiene le effemeridi e i parametri d'offset d'orologio del satellite che sta inviando il segnale, i parametri del modello di Klobuchar per il disturbo ionosferico e le informazioni orbitali approssimate (almanacchi) per gli altri satelliti della costellazione.

3.6.3 La modulazione dei segnali

Le portanti sono sinusoidi, mentre i codici binari sono una sequenza di impulsi ± 1 ; con grande semplificazione, modulare le portanti con i codici equivale a farne il prodotto: il segnale ottenuto riproduce la portante ma, in corrispondenza di una transizione di stato del codice, si ha un salto di 180° nella fase del segnale prodotto.

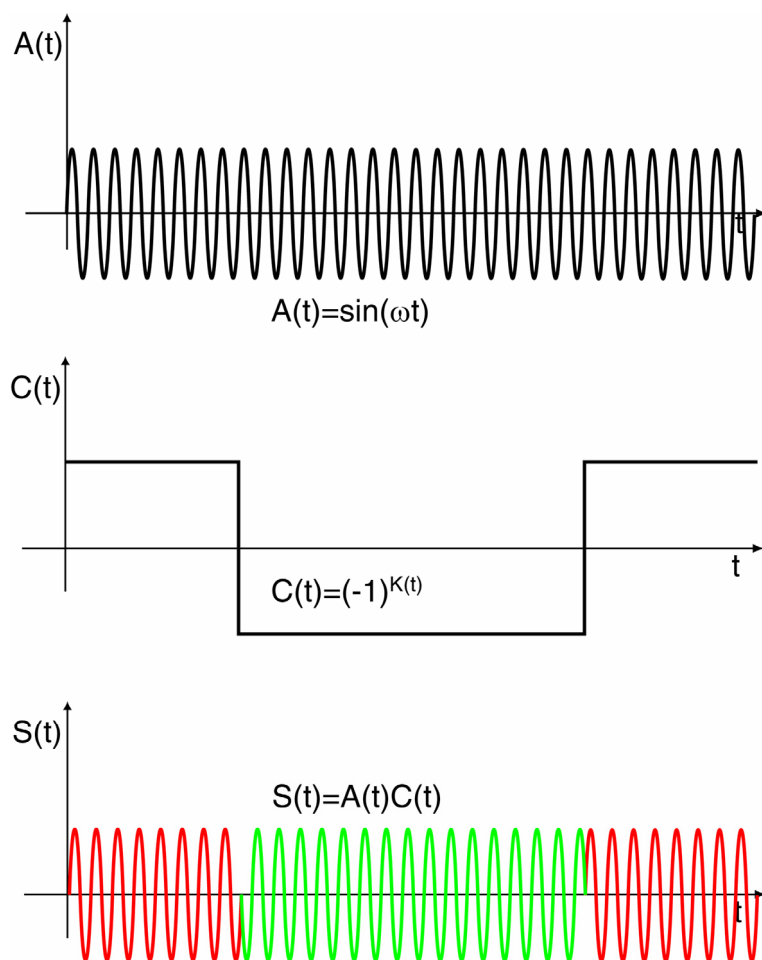


Figura 3.15. Modulazione della portante con il codice

Nota: nel disegno non sono rispettate
le scale di frequenza fra portante e codice binario

Il segnale finale emesso dai satelliti è così composto: sia L1 sia L2 vengono emesse in due repliche, con una differenza di fase di 90° ; la prima replica di L1 è modulata dal codice P(Y), la seconda dal codice C/A; la prima replica di L2 è modulata dal codice P(Y), la seconda dal codice L2C; tutti i segnali vengono infine modulati con il messaggio navigazionale. Nuovamente non sono disponibili informazioni univoche sul ruolo di M, che viene quindi omesso dalla presente formulazione del segnale finale.

$$S_{L_1}(t) = A_0 C(t) D(t) \sin(2\pi f_{L_1} + \varphi_{L_1}) + A_0 P(t) D(t) \cos(2\pi f_{L_1} + \varphi_{L_1}) \quad (3.9a)$$

$$S_{L_2}(t) = A_0 C(t) D(t) \sin(2\pi f_{L_2} + \varphi_{L_2}) + A_0 P(t) D(t) \cos(2\pi f_{L_2} + \varphi_{L_2}) \quad (3.9.b)$$

Le ragioni della grande complessità del segnale GPS sono varie. Innanzitutto solo grazie alla presenza delle due frequenze L1 e L2 è possibile rimuovere il disturbo ionosferico, che costituisce la più grande fonte di errore: a tale riguardo si veda il Par. 4.7.4. I codici C/A, L2C, P e M possono essere tutti utilizzati per le misure di codice: si veda il Cap. 4; la ragione di ben quattro codici finalizzati al medesimo scopo è nei seguenti fattori, per capire i quali si deve anche ricordare che il progetto originario è degli anni '60, ovvero di 40 anni fa:

- C/A: il codice è necessario per l'utilizzo pubblico del GPS; la sua brevità permette al ricevitore un rapido riconoscimento del satellite; inizialmente non si riteneva strategico garantire la rimozione del disturbo ionosferico agli utenti civili e, anzi, le linee di indirizzo iniziali prevedevano una limitazione delle accuratezze disponibili per l'utilizzo pubblico, SPS: perciò C/A era previsto su un'unica frequenza;
- P(Y): per le applicazioni militari un codice criptato, ovvero non utilizzabile ma soprattutto non contraffabile dai nemici, è strategico: per i primi anni di test si rese disponibile P ma era già prevista la sua commutazione a Y una volta raggiunta la piena operatività del sistema; la grande lunghezza del codice è necessaria per impedirne la decriptazione e la sua disponibilità su due frequenze permette la rimozione del ritardo ionosferico, in modo da garantire il cosiddetto Precise Positioning Service (PPS); in effetti gli algoritmi adottati per commutare P in Y sono parzialmente noti e quindi i ricevitori geodetici di moderna generazione, ovvero quelli di fascia di costo superiore, ricorrono a complessi algoritmi di ricostruzione del segnale che permettono di leggere e utilizzare anche il codice Y.
- L2C: la nascita di GLONASS, l'avvio del progetto GALILEO e l'inventiva del mercato hanno stimolato la crescita delle applicazioni di navigazione civile, che si sono scoperte ricchissime di ritorno economico; per non essere inadeguato rispetto alle richieste di maggiori accuratezze, il GPS si è dotato di un codice civile anche sulla seconda frequenza; attualmente L2C è disponibile solo per i 6 satelliti IIR-M già in orbita.
- il codice M sicuramente introduce importanti novità per le applicazioni strategiche per l'esercito statunitense, tali probabilmente da sostituire in prospettiva il codice Y, che forse verrà definitivamente reso pubblico; a tale riguardo, come detto, non si sono trovate molte informazioni.

Un'ultima nota è importante. Nei futuri satelliti del blocco IIF è prevista l'implementazione di una terza portante, che prenderà nome di L5, esattamente con le stesse caratteristiche dell'omonima combinazione che è attualmente possibile costruire su L1 e L2 (Par. 4.7.2); L5 avrà sicuramente grande applicazione nell'elaborazione delle osservazioni di fase, per quanto riguarda gli algoritmi di fissaggio delle ambiguità, e segnerà quindi un'evoluzione nell'utilizzo di precisione del GPS; il lancio dei primi satelliti IIF è attualmente previsto per l'autunno del 2009.

Capitolo 4.

Le osservazioni, la propagazione e i disturbi atmosferici

Nel presente capitolo si discute innanzitutto il comportamento degli oscillatori e degli orologi atomici; quindi vengono presentate le equazioni di osservazione di codice e fase del GPS, che costituiscono il mattone fondamentale del posizionamento, e vengono discussi gli effetti dell'atmosfera sulla propagazione del segnale. Da ultimo si presentano le possibili combinazioni lineari delle osservazioni, con attenzione a quelle più utilizzate per l'elaborazione dei dati; viene discusso il loro comportamento nel tempo e viene infine analizzato il filtro di Hatch, molto diffuso per il filtraggio delle osservazioni di codice.

4.1 La frequenza, la fase e il tempo di un oscillatore

Dato un oscillatore atomico, ad ogni epoca t la sua fase e la sua frequenza sono legate dalla

$$f(t) = \frac{d\phi(t)}{dt}, \quad \phi(t) = \int_{t_0}^t f(\tau) d\tau + \phi_0 \quad (4.1)$$

Sia i un orologio basato su un oscillatore atomico di frequenza nominale f_0 ; l'orologio calcola il tempo come differenza fra la fase corrente dell'oscillatore e la fase iniziale, ovvero all'accensione dell'oscillatore: al tempo t la stima di tempo $t_i(t)$ è

$$t_i(t) = \frac{\phi_i(t) - \phi_i(t_0)}{f_0} = \frac{\phi_i(t)}{f_0} - \frac{\phi_i(t_0)}{f_0}$$

considerando la precedente relazione si ottiene

$$t_i(t) = \frac{1}{f_0} \int_{t_0}^t f_i(\tau) d\tau \quad (4.2)$$

la frequenza di un oscillatore subisce fluttuazioni e derivate nel tempo: si può scrivere

$$f_i(t) = f_0 + \delta f_i(t)$$

ove $\delta f_i(t)$ rappresenta la fluttuazione rispetto alla frequenza nominale; si ottiene la

$$t_i(t) = \frac{1}{f_0} \int_{t_0}^t f_0 d\tau + \frac{1}{f_0} \int_{t_0}^t \delta f_i(\tau) d\tau$$

il primo integrale è uguale a t ; il secondo esprime l'errore accumulato dell'orologio, viene indicato con $dt_i(t)$ e prende nome di errore o offset dell'orologio i ; in definitiva si ha

$$t_i(t) = t + dt_i(t) \quad (4.3)$$

In assenza di reimpostazioni e correzioni, l'errore di un orologio assume valori significativi nel tempo. Consideriamo ora la fase dell'oscillatore dell'orologio all'epoca t , che è calcolata mediante la

$$\phi_i(t) = f_0 t_i(t) + \phi(t_0)$$

ovvero

$$\phi_i(t) = f_0[t + dt_i(t)] + \phi(t_0) = f_0 t + f_0 dt_i(t) + \phi(t_0) \quad (4.4)$$

Nel caso del GPS si hanno sostanzialmente due tipi di orologi, ovvero quelli dei satelliti e quelli dei ricevitori; i primi sono orologi atomici di altissima accuratezza e, su periodi di alcune ore, il loro errore ha un andamento ben descritto da un polinomio di secondo grado:

$$dt_i(t) = a_{0i} + a_{1i}(t - t_0) + a_{2i}(t - t_0)^2 + \delta_i(t) \quad (4.5)$$

lo scostamento rispetto all'andamento polinomiale, denotato nella precedente formula con il simbolo $\delta_i(t)$, è tipicamente inferiore a dieci nanosecondi ($10^{-8}s$) per periodi di alcune ore. Il segmento di controllo monitora e stima il comportamento degli orologi dei satelliti GPS: in particolare per ognuno di loro vengono quotidianamente predetti i parametri a_{0s} , a_{1s} , a_{2s} per 24 blocchi di un'ora ciascuno, per la giornata successiva; i parametri vengono inviati al satellite che li trasmette all'utenza attraverso il messaggio navigazionale, insieme alle effemeridi: l'accuratezza dei parametri d'orologio trasmessi dai satelliti è dell'ordine di 10 ns, corrispondenti a tre metri in lunghezza.

Come per le effemeridi, è possibile anche utilizzare gli offset stimati a posteriori: quelli IGS vengono pubblicati nei medesimi file SP3 delle effemeridi e hanno accuratezza migliore di 0.1 ns; in questo caso sono riportati direttamente i valori di offset tabulati a intervalli regolari, mentre i valori a qualunque altra epoca possono essere calcolati mediante l'interpolazione lagrangiana (3.3).

Gli orologi dei ricevitori hanno comportamento più erratico nel tempo, che non può banalmente essere descritto da un polinomio: nelle elaborazioni, l'errore d'orologio del ricevitore viene considerato incognito da stimare per ogni epoca di osservazione.

4.1.1 Il termine relativistico nell'offset d'orologio del satellite

L'offset d'orologio del satellite, $dt^S(t)$, è in effetti dato dalla somma di due offset distinti:

$$dt^S(t) = dt_0^S(t) + dt_{rel}^S(t)$$

il primo è l'offset in senso proprio, appena discusso, il secondo è l'effetto calcolato in base alla teoria della relatività ristretta: esso è dovuto all'eccentricità dell'orbita del satellite ed è calcolabile mediante la seguente espressione

$$dt_{rel}^S(t) = -2 \cdot \mathbf{x}^S(t) \cdot \dot{\mathbf{x}}^S(t) / c \quad (4.6)$$

ove $\mathbf{x}^S(t)$, $\dot{\mathbf{x}}^S(t)$ sono rispettivamente posizione e velocità del satellite, c è la velocità di propagazione della luce nel vuoto. L'effetto relativistico ha valori dell'ordine di 0-5 m e nell'elaborazione delle osservazioni deve essere modellizzato insieme all'offset propriamente detto; il suo calcolo necessita solo delle effemeridi, poiché dipende solo dalla posizione e dalla velocità del satellite e l'errore con cui lo si può calcolare è sostanzialmente trascurabile. Nel seguito, per mantenere semplicità di notazione, si scriverà un unico termine di offset d'orologio del satellite, $dt^S(t)$, considerandolo un termine calcolabile tramite le effemeridi: rimane inteso che il suo valore deve essere ottenuto dalla somma dei due termini appena discussi.

Si noti da ultimo che in senso proprio, l'effetto relativistico di un orologio in orbita dovrebbe comprendere anche un termine di relatività generale che è di un ordine di grandezza superiore a quello appena discusso; tale termine, che è assimilabile all'effetto di un'orbita perfettamente circolare, è costante nel tempo e viene corretto dall'oscillatore del satellite prima della generazione del segnale.

4.2 Le osservazioni di codice (o pseudorange, pseudo distanza)

Il ricevitore innanzitutto identifica il satellite mediante il riconoscimento del codice C/A. Quindi inizia il processo di misura, che può essere condotto indifferentemente sui codici pseudocasuali C/A, L2C e (per i ricevitori militari) P(Y) o M.

Il correlatore del ricevitore misura il ritardo del codice ricevuto rispetto a quello generato internamente; l'epoca di invio è quella locale dell'orologio del satellite, quella di ricezione è quella locale dell'orologio del ricevitore: quindi sono epoche affette dall'offset dei due orologi. Sia t l'epoca di osservazione nella scala dei tempi GPS, sia

$\Delta T_R^S(t)$ il ritardo osservato dal ricevitore R al satellite S : l'equazione di osservazione è data dalla

$$\Delta T_R^S(t) = t_R(t) - t^S(t - \tau_R^S)$$

ove $t_R(t)$ indica l'epoca di ricezione del segnale secondo l'orologio del ricevitore; $t^S(t - \tau_R^S)$ indica l'epoca di invio del segnale secondo l'orologio del satellite: $t - \tau_R^S$ è l'epoca di invio nella scala dei tempi GPS, τ_R^S è il tempo di tragitto del segnale dal satellite al ricevitore.

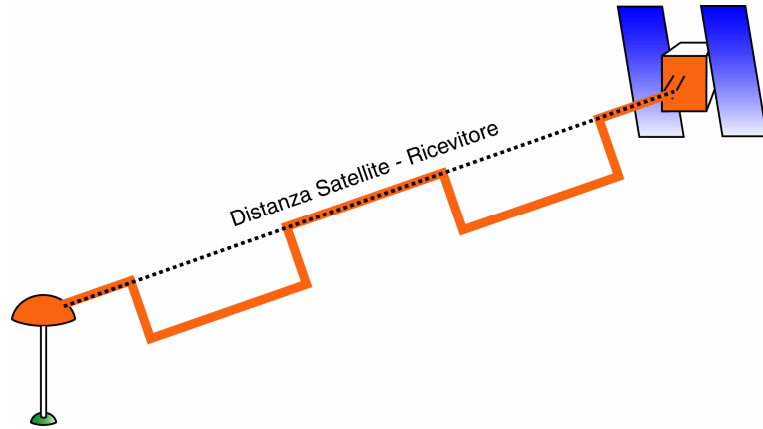


Figura 4.1. Una rappresentazione grafica delle osservazioni di codice.

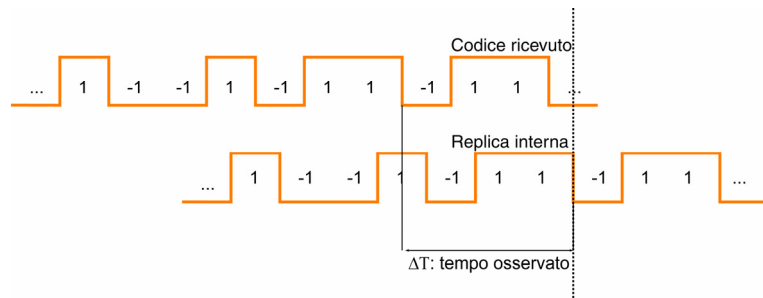


Figura 4.2. La correlazione fra codice ricevuto e codice generato internamente.

L'orologio del satellite non è allineato al tempo GPS ma si ha:

$$t^S(t - \tau_R^S) = t - \tau_R^S + dt^S(t - \tau_R^S)$$

il tempo di tragitto (circa 66 ms) è molto breve e si può ipotizzare che durante esso l'offset d'orologio del satellite non cambi: perciò si ha

$$dt^S(t - \tau_R^S) \cong dt^S(t)$$

Analogamente, per l'orologio del ricevitore,

$$t_R(t) = t + dt_R(t)$$

Abbiamo quindi

$$\Delta T_R^S(t) = t - t + \tau_R^S + dt_R(t) - dt^S(t) = \tau_R^S + dt_R(t) - dt^S(t)$$

L'osservazione di tempo viene moltiplicata per la velocità di propagazione del segnale nel vuoto e fornisce l'osservazione di codice (o pseudorange, pseudodistanza) P_R^S , la cui equazione finale di osservazione è

$$P_R^S(t) = c\Delta T_R^S(t) = c\tau_R^S + c(dt_R(t) - dt^S(t)) \quad (4.7)$$

4.2.1 L'ambiguità delle osservazioni C/A

Si discute subito un aspetto tecnicamente rilevante per l'utilizzo delle osservazioni di codice C/A nel posizionamento assoluto (Cap. 5), ovvero quello di tutti i ricevitori non abilitati ai codici militari: il codice C/A dura nel tempo $1ms$ ed è lungo nello spazio $1023 \times 300m = 306.9km$ mentre il tempo di transito del segnale è di circa $66ms$, essendo la distanza fra satellite e ricevitore maggiore di $20 \cdot 10^5 km$. Quindi, nel caso del codice C/A, la (4.7) è vera a meno del numero intero di ripetizioni dell'intero codice occorse fra la sua emissione e la sua ricezione, che sono appunto circa 66: in altri termini la misura è ambigua temporalmente di un multiplo intero di $1ms$, metricamente di un multiplo intero di $306.9km$. L'ambiguità viene risolta direttamente nel ricevitore, mediante una stima di tempo basata sulle osservazioni al messaggio navigazionale D: tale stima è molto grossolana ma comunque migliore di $1ms$ oppure $306.9km$, ovvero sufficiente allo scopo. La questione non è rilevante nell'elaborazione del dato, poiché gli stream in tempo reale e i file per la postelaborazione prodotti dai ricevitori riportano l'osservazione metrica già corretta per l'ambiguità di codice, ovvero nella forma (4.7).

4.3 Le osservazioni di fase

L'osservazione di codice è un'osservazione di differenza di stato fra il codice ricevuto dal satellite e una sua replica generata internamente al ricevitore; analogamente può essere effettuata un'osservazione di differenza di fase in cicli (Fig. 4.4) fra portante ricevuta dal satellite (ad es. L1) e una sinusoidale di uguale frequenza f generata da un oscillatore interno al ricevitore.

L'equazione di osservazione all'epoca t è data dalla

$$\phi_R^S(t) = \phi_R(t) - \phi_S(t - \tau_R^S)$$

ove $\phi_R(t)$ è la fase dell'oscillatore del ricevitore all'epoca di osservazione, $\phi_S(t - \tau_R^S)$ è la fase generata dall'oscillatore del satellite all'epoca di invio del segnale.

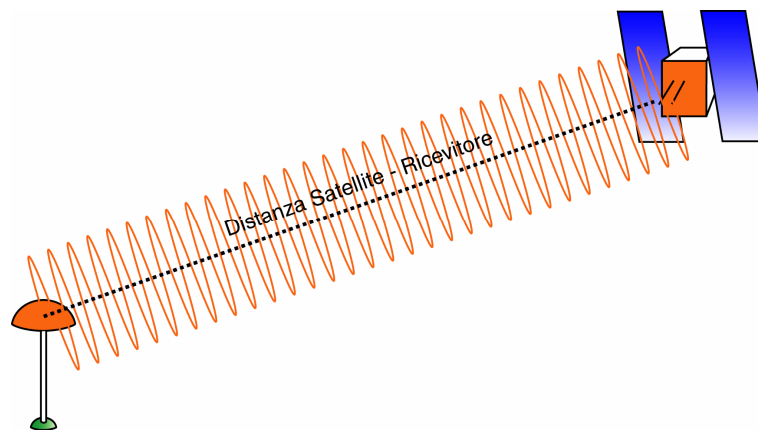


Figura 4.3. Una rappresentazione grafica delle osservazioni di fase.

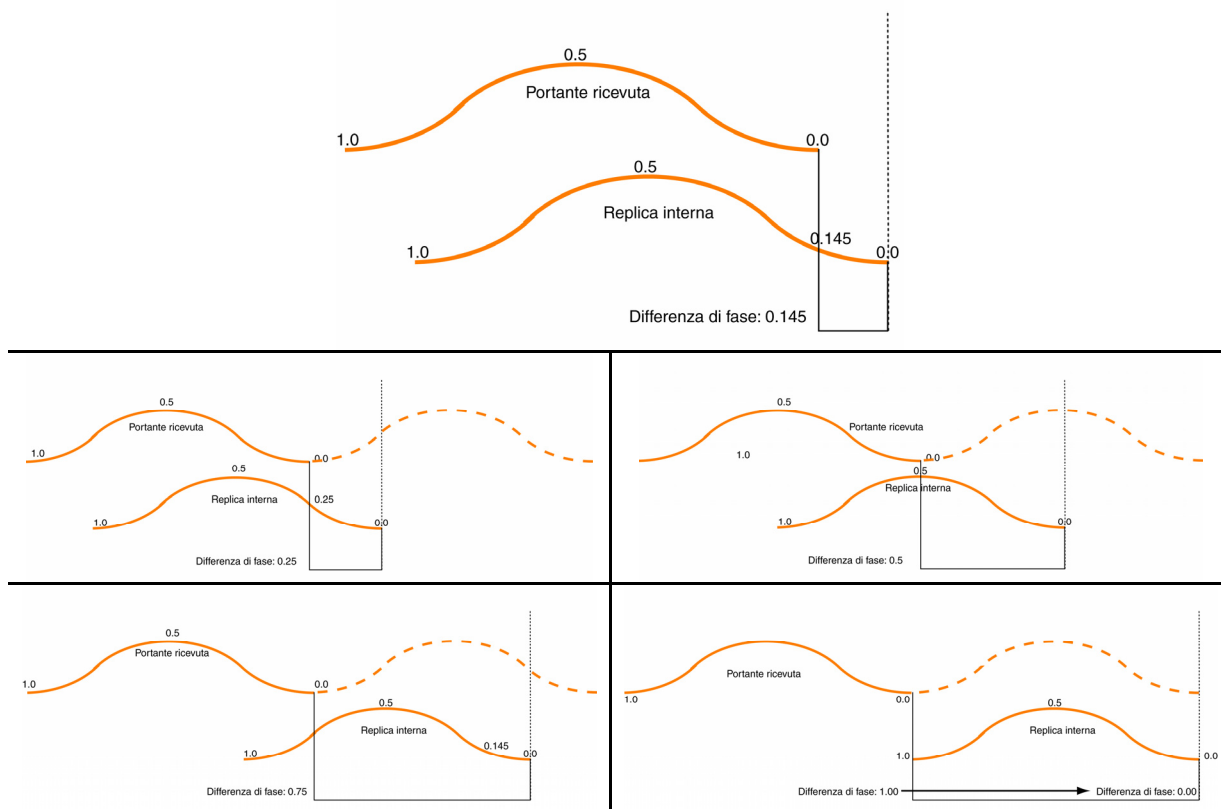


Figura 4.4. Vari esempi di misura di differenza di fase; nell'ultimo caso, viene introdotto il concetto di ambiguità intera.

Ricordando la (4.4) si ottiene

$$\begin{aligned}\phi_R^S(t) &= f_0 t - f_0(t - \tau_R^S) + f_0 dt_R(t) - f_0 dt^S(t - \tau_R^S) + \phi_R - \phi^S \\ &= f_0 \tau_R^S + f_0(dt_R(t) - dt^S(t)) + \phi_R - \phi^S\end{aligned}$$

In effetti il correlatore del ricevitore misura solo la parte frazionaria della differenza di fase e non il numero intero di cicli trascorsi dall'invio del segnale alla sua ricezione (Fig. 4.5); perciò nell'equazione di osservazione deve essere aggiunto un ultimo termine

$$\phi_R^S(t) = f_0 \tau_R^S + f_0(dt_R(t) - dt^S(t)) + \phi_R - \phi^S + N_R^S(t) \quad (4.8)$$

$N_R^S(t)$ è detta ambiguità intera e, da un punto di vista geometrico, rappresenta il numero intero di cicli di portante compresi fra il satellite e il ricevitore, non direttamente osservabile.

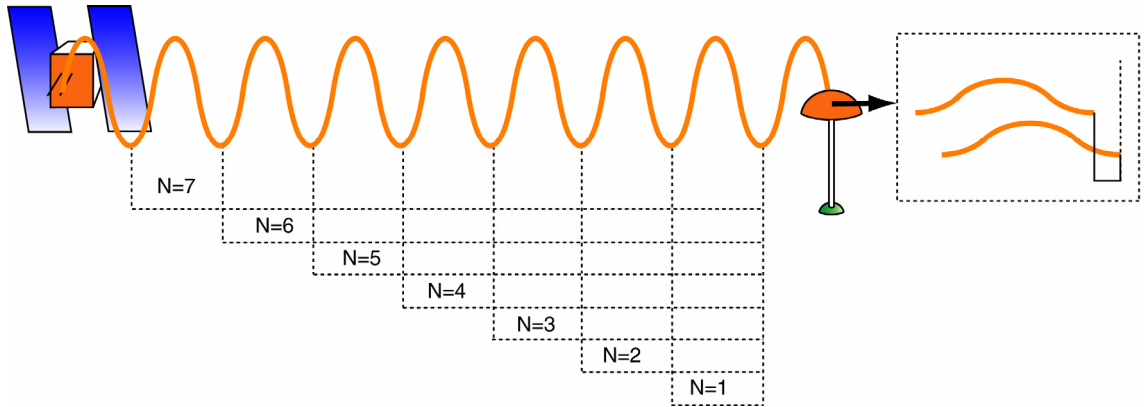


Figura 4.5. L'ambiguità iniziale: il ricevitore misura solo la parte fratta della portante ma non il numero intero di cicli dalla sua trasmissione.

L'equazione di osservazione in cicli viene moltiplicata per la lunghezza d'onda del segnale, in modo da ottenere l'osservazione in unità metriche

$$L_R^S(t) = \lambda \phi_R^S(t) = c \tau_R^S(t) + c(dt_R(t) - dt^S(t)) + \lambda(N_R^S(t) + \phi_R - \phi^S) \quad (4.9)$$

Il termine di ambiguità intera delle fasi è un multiplo di 19cm per L1, di 24cm per L2 e non esiste alcun algoritmo interno al ricevitore capace di risolverlo in fase di acquisizione del segnale: perciò, a differenza dell'ambiguità di C/A, essa rimane nelle osservazioni, esattamente come si è scritto nella (4.9) e deve essere opportunamente stimata nell'elaborazione dei dati.

4.4 L'errore di osservazione, i disturbi e gli errori di modellizzazione

L'errore di osservazione nei codici e nelle fasi, che indicheremo nel seguito con i simboli rispettivamente ν_p e ν_L , è dovuto al rumore elettronico del correlatore; storicamente, ovvero per i ricevitori di prima generazione si poneva tale errore uguale circa a 1% della lunghezza d'onda del segnale osservato: si avevano dunque 3 m per le osservazioni C/A, 30 cm per le osservazioni P e 2 mm per le fasi. Nei decenni la qualità dell'elettronica ha fatto significativi progressi e attualmente tali valori sono decisamente superati; in particolare si può dire che oggi i ricevitori hanno accuratezze dell'ordine del metro, con punte sino a 30 cm, anche su C/A. Le osservazioni di fase presentano accuratezza del millimetro.

Con il termine di disturbo intendiamo quei fenomeni che alterano il contenuto metrico delle osservazioni, come ad esempio i disturbi atmosferici: al fine di stimare posizioni il più accuratamente possibile è evidentemente necessario modellarne e correggerne l'effetto; inoltre si consideri che nelle osservazioni sono presenti termini che devono essere considerati noti, ovvero modellizzabili, come ad esempio la posizione dei satelliti. Tutti i modelli possono contenere significativi errori: ad esempio, le effemeridi trasmesse hanno errori dell'ordine del metro, il modello di Klobuchar per la correzione del disturbo ionosferico (Par. 4.5.1) ha errori che possono arrivare a qualche metro; chiameremo errori di modellizzazione tutti quelli che si commettono nel calcolo dei termini considerati noti e li denoteremo con il simbolo δ .

4.5 Il disturbo atmosferico

In assenza di atmosfera il segnale si muoverebbe con velocità costante c lungo tutto il tragitto; si avrebbe dunque

$$c \cdot \tau_R^S = \rho_R^S = \sqrt{(X_R - X^S)^2 + (Y_R - Y^S)^2 + (Z_R - Z^S)^2}$$

ove X, Y, Z sono le coordinate cartesiane di ricevitore e satellite all'epoca di osservazione. Però nell'atmosfera terrestre la velocità di propagazione del segnale subisce variazioni dovute allo stato fisico del mezzo attraversato, che causano il cosiddetto disturbo o ritardo atmosferico, discusso in questi paragrafi.

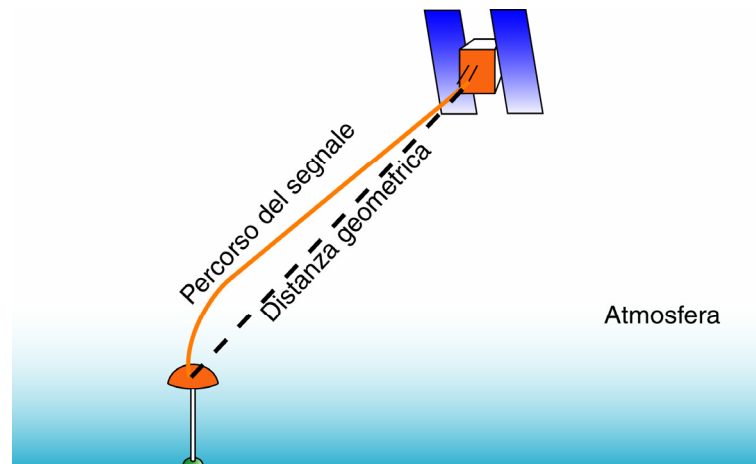


Figura 4.6. L'effetto di curvatura del segnale.

Innanzitutto si deve considerare che qualunque segnale elettromagnetico, nell'attraversare un mezzo, segue il percorso di minimo tempo (legge di Fermat) che non necessariamente coincide con la distanza geometrica. Il segnale GPS proveniente da un satellite al ricevitore ha seguito un percorso curvo, ovvero geometricamente più lungo della distanza: l'effetto, sintetizzato nella tabella, è ben modellizzabile e può essere calcolato e sottratto dall'osservazione originaria. Nel seguito si ometterà il termine di curvatura, dando per assodato che sia già stato sottratto dall'equazione di osservazione effettivamente registrata.

Angolo di elevazione (°)	Effetto di curvatura (mm)
5	75
10	19
20	5
45	1
90	0

Tabella 4.1. Effetto di curvatura sul segnale.

Il tempo di tragitto del segnale elettromagnetico dipende dalla sua velocità di propagazione; nello spazio esterno all'atmosfera questa è costante e uguale a c ma nell'atmosfera varia da punto a punto: si deve dunque porre $v = v(r)$ e si può scrivere

$$\tau_R^S = \int_{\rho_R^S} \frac{dr}{v(r)}$$

ove ρ_R^S è il percorso del segnale. Si ha quindi

$$c\tau_R^S = \int_{\rho_R^S} \frac{cdr}{v(r)} = \int_{\rho_R^S} n(r)dr = \int_{\rho_R^S} dr + \int_{\rho_R^S} (n(r)-1)dr = \rho_R^S + \Delta_R^S$$

ove $n(r) = c/v(r)$ è definito indice di rifrazione del mezzo nel punto r ; con Δ_R^S (disturbo atmosferico) si indica la differenza, in unità metriche

$$\Delta_R^S = c\tau_R^S - \rho_R^S = \int_{\rho_R^S} (n(r)-1)dr \quad (4.10)$$

In particolare due sono gli strati atmosferici che più interagiscono con il segnale GPS: la ionosfera e la troposfera.

4.5.1 Il disturbo ionosferico

Il disturbo ionosferico è dovuto agli strati di atmosfera fra 100 km e 1000 km di quota ove sono presenti ioni e elettroni liberi che interferiscono con la propagazione del segnale GPS: l'indice di rifrazione locale assume il seguente valore

$$n(r) = 1 \pm \frac{A \cdot N_E(r)}{f^2} \quad (4.11)$$

ove $A = 40.3m^3 s^{-2}$; f è la frequenza del segnale (in Hz); $N_E(r)$ è la densità di elettroni (in numero $\times m^3$) nel punto; il segno è positivo per i codici e negativo per le portanti: infatti, come per ogni mezzo dispersivo, la velocità di gruppo diminuisce ma la velocità di fase aumenta. Integrando lungo il percorso attraverso la ionosfera si ottiene, in unità metriche

$$I_R^S = \int_{Iono_R^S} (n(r)-1)dr = \pm \int_{Iono_R^S} A \frac{N_E(r)}{f^2} dr = \pm \frac{A}{f^2} \int_{Iono_R^S} N_E(r)dr = \pm A \frac{TEC_R^S}{f^2} \quad (4.12)$$

il simbolo TEC_R^S (Total Electron Content) indica la densità di elettroni liberi lungo il percorso del segnale (per unità di superficie, ovvero in numero $\times m^2$). Il contenuto di elettroni nella ionosfera varia significativamente in funzione dell'intensità di radiazione solare incidente sull'atmosfera che a sua volta dipende da dell'intensità di attività solare e ora del giorno: di conseguenza il disturbo ionosferico presenta grande variabilità (Tab. 4.2, Fig. 4.87).

Esiste un modello sintetico, detto di Klobuchar, per la descrizione matematica dello stato ionosferico, in funzione di latitudine, longitudine e epoca dell'osservazione; la funzione a sua volta dipende da otto parametri numerici $\alpha_i, i=0, \dots, 3$ e $\beta_i, i=0, \dots, 3$ che la rete di controllo del NIMA stima e invia quotidianamente ai satelliti i quali, a

loro volta, li trasmettono nel messaggio navigazionale: i ricevitori GPS utilizzano i parametri e il modello di Klobuchar per modellizzare il disturbo.

Elevazione (°)	Basso (m)	Medio (m)	Alto (m)
15	3.9	18.0	180
30	2.0	10.0	100
45	1.4	7.0	70
90	1.0	5.0	50

Tabella 4.2. Alcuni valori indicativi di disturbo ionosferico per diversi angoli di elevazione. Basso: periodo notturno. Medio: periodo diurno, normale attività solare. Alto: periodo diurno, intensa attività solare.

Il calcolo viene effettuato secondo il seguente schema; innanzitutto si ricorre al semicerchio come unità di misura degli angoli, ovvero $1SC = 180^\circ = \pi$: sono espressi in semicerchi la latitudine φ_R e la longitudine λ_R del ricevitore, l'azimuth α_R^S e l'elevazione η_R^S del satellite rispetto al ricevitore. Prima di tutto è necessario stimare l'angolo ψ esistente tra la posizione del ricevitore, il centro della Terra e il cosiddetto punto ionosferico, ovvero il punto medio della traiettoria del segnale attraverso la ionosfera:

$$\psi = \frac{0.0137}{a_K^p + 0.11} - 0.022$$

Si procede quindi alla stima delle coordinate del punto ionosferico: posto

$$\varphi' = \varphi_R + \psi \cdot \cos(\alpha_R^S)$$

si definiscono latitudine e longitudine del punto ionosferico

$$\varphi = \begin{cases} \varphi' & \Leftrightarrow |\varphi'| \leq 0.416 \\ 0.416 & \Leftrightarrow \varphi' > 0.416 \\ -0.416 & \Leftrightarrow \varphi' < -0.416 \end{cases},$$

$$\lambda = \lambda_R + \frac{\psi \cdot \sin(\alpha_R^S)}{\cos(\varphi)}$$

Si calcola poi la cosiddetta latitudine geomagnetica

$$\phi = \varphi + 0.064 \cdot \cos(\lambda - 1.617)$$

Inoltre viene definito l'argomento di tempo τ come:

$$\tau' = \lambda \cdot 43200 + t$$

$$\tau = \begin{cases} \tau' & \Leftrightarrow 0 \leq \tau' < 86400 \\ \tau' - 86400 & \Leftrightarrow \tau' \geq 86400 \\ \tau' + 86400 & \Leftrightarrow \tau' < 0 \end{cases}$$

ove t è l'epoca del giorno in secondi. Si calcola dunque:

$$x = \frac{2\pi \cdot (\tau - 50400)}{P}$$

ove P è dato dalla:

$$P' = \sum_{n=0}^3 \beta_n \cdot \phi^n$$

$$P = \begin{cases} P' & \Leftrightarrow P' \geq 72000 \\ 72000 & \Leftrightarrow P' < 72000 \end{cases}$$

Quindi si calcola A :

$$A' = \sum_{n=0}^3 \alpha_n \cdot \phi^n$$

$$A = \begin{cases} A' & \Leftrightarrow A' \geq 0 \\ 0 & \Leftrightarrow A' < 0 \end{cases}$$

e F

$$F = 1 + 16 \cdot (0.53 - a_R^S)^3$$

Infine si calcola il disturbo ionosferico, in metri, mediante la formula

$$I_R^S(t) = \begin{cases} c \cdot F \cdot \left[5 \cdot 10^{-9} + A \cdot \left(1 - \frac{x^2}{2} + \frac{x^4}{24} \right) \right] & |x| < 1.57 \\ c \cdot F \cdot 5 \cdot 10^{-9} & |x| > 1.57 \end{cases} \quad (4.13)$$

Il modello di Klobuchar presenta errori che possono superare il 5-10% del disturbo totale: l'utente deve quindi considerare che, in casi particolarmente sfavorevoli, può

rimanere un errore residuo superiore a 10-40 m nelle osservazioni. Si noti da ultimo che, per quanto riguarda i disturbi ionosferici sulle due portanti, che chiamiamo $I1, I2$, vale la seguente relazione:

$$I2_R^S(t) = A \frac{TEC_R^S(t)}{f_2^2} = \frac{f_1^2}{f_2^2} A \frac{TEC_R^S(t)}{f_2^2} = \frac{f_1^2}{f_2^2} A \frac{TEC_R^S(t)}{f_1^2} = \frac{f_1^2}{f_2^2} I1_R^S(t) \quad (4.14)$$

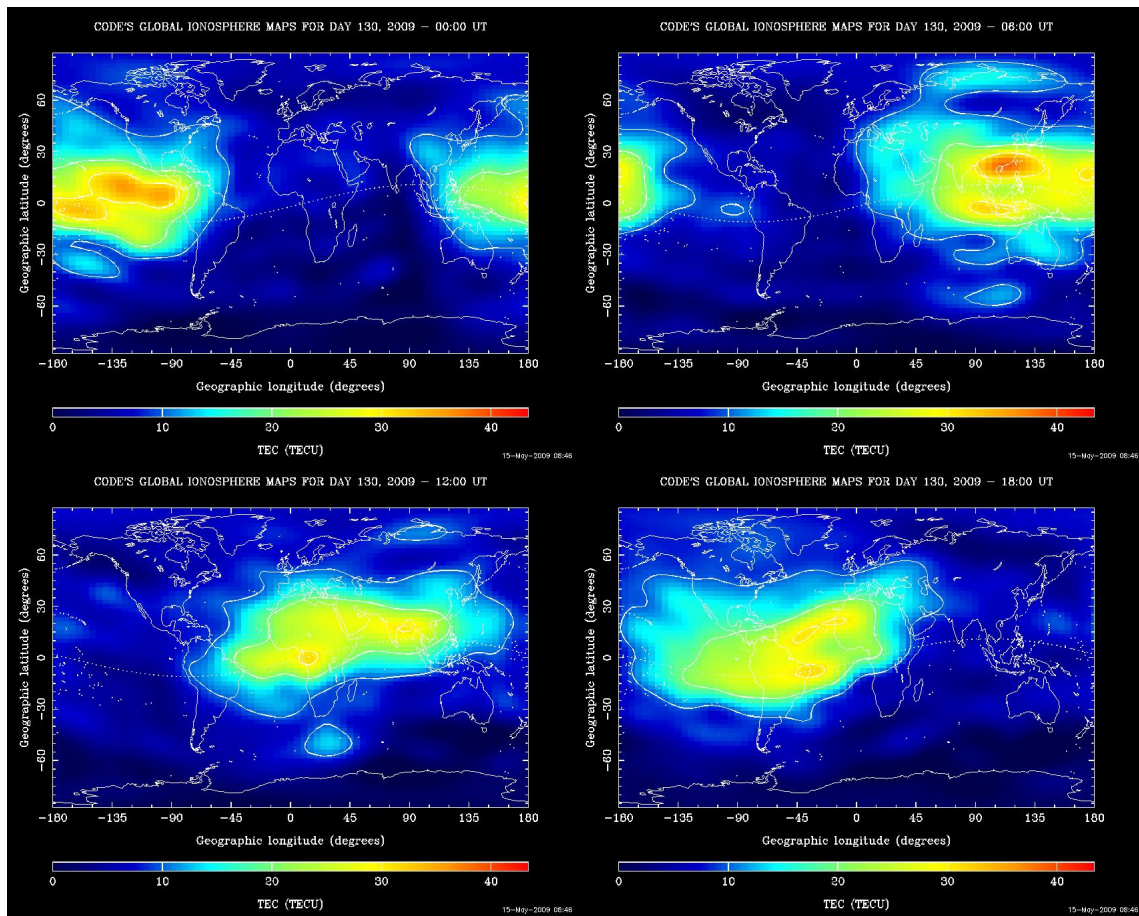


Figura 4.7. TEC nella direzione zenitale; mappe relative alle ore 00.00, 06.00, 12.00, 18.00 UTC del 10.5.2009: si notino le differenze fra notte (basso TEC) e pieno giorno (alto TEC). Dal sito del CODE dell'Astronomisches Institut Universität Bern.

4.5.2 Il disturbo troposferico (T)

Il disturbo troposferico è dovuto agli strati compresi fra il suolo e i 40 km di quota: è causato dall'aria e, nello strato dei primi 10 km, dalla presenza di vapor d'acqua. L'indice di rifrazione locale è dato dalla:

$$n(r) = 1 + k_1 \frac{P(r)}{T(r)} + k_2 \frac{e(r)}{T(r)} + k_3 \frac{e(r)}{T^2(r)}$$

ove $P(r)$ è la pressione (in mBar) $T(r)$ è la temperatura (in °K), e è la pressione parziale di vapor d'acqua (in mBar), le costanti sono rispettivamente

$$\begin{aligned} k_1 &= 77.624 \cdot 10^{-6} K \cdot mBar^{-1}, \\ k_2 &= -12.920 \cdot 10^{-6} K \cdot mBar^{-1}, \\ k_3 &= 37.19 \cdot 10^{-2} K^2 \cdot mBar^{-1} \end{aligned}$$

Integrando si ottiene

$$T_R^S = \int_{Tropo_R^S} (k_1 \frac{P(r)}{T(r)} + k_2 \frac{e(r)}{T(r)} + k_3 \frac{e(r)}{T^2(r)}) dr = \int_{Tropo_R^S} k_1 \frac{P(r)}{T(r)} dr + \int_{Tropo_R^S} (k_2 \frac{e(r)}{T(r)} + k_3 \frac{e(r)}{T^2(r)}) dr$$

Il primo integrale è detto componente secca o idrostatica del disturbo e contribuisce a circa il 90% del disturbo totale, il secondo è detto componente umida e contribuisce per il restante 10%. Esistono modelli standard per il calcolo del disturbo troposferico, che tipicamente sono basati su ipotesi di atmosfera in condizioni ideali: da tali ipotesi vengono ricavati profili verticali standard per pressione, temperatura e umidità; questi infine vengono integrati per calcolare il disturbo. Si riporta a titolo di esempio l'espressione finale del modello di Saastamoinen:

$$T_R^S = \frac{0.002277}{\sin \eta_R^S} \left[P_R + \left(\frac{1255}{T_R} + 0.05 \right) e_R - \tan^{-2} \eta_R^S \right] \quad (4.15)$$

ove η_R^S è l'elevazione del satellite rispetto al ricevitore, P_R , T_R , e_R sono la pressione, la temperatura e la pressione parziale di vapor d'acqua nel sito di ricezione: queste possono essere misurate con sensori meteorologici o, più semplicemente, propagate in quota dai valori standard al livello del mare: ad esempio, utilizzando le formule di Berg si ha

$$\begin{aligned} P_R &= P_0 (1 - 2.26 h_R)^{5.225}, & P_0 &= 1013.25 \text{ mBar} \\ T_R &= T_0 - 6.5 \cdot 10^{-3} h_R, & T_0 &= 291.15^\circ K \end{aligned} \quad (4.16)$$

$$e_R = \frac{H_R e^{c(T)}}{100}, \quad c(T) = -37.2465 + 0.213166^\circ T - 2.56908 \cdot 10^{-4} T^2,$$

H_R è l'umidità relativa: si ottiene dalla

$$H_R = H_0 e^{-0.0006396 h_R}, \quad H_0 = 50\%$$

I modelli troposferici standard non sono mai pienamente rappresentativi della reale condizione meteorologica al momento e nel luogo della misura: in particolare la componente umida del disturbo è quella più variabile nel tempo e nello spazio, e peggio modellizzabile. Si stima che il modello di Saastamoinen, in condizioni sfavorevoli, possa presentare errori residui fino al 10% del disturbo totale (Tab. 4.3, Fig. 4.8).

Elevazione (°)	Basso (m)	Medio (m)	Alto (m)
5	22.8	23.6	26.6
10	12.9	13.3	15.1
15	8.9	9.2	10.4
30	4.6	4.8	5.4
45	3.3	3.4	3.8
90	2.3	2.4	2.7

Tabella 4.3. Valori esemplificativi di disturbo troposferico per vari angoli di elevazione e diverse condizioni climatiche al livello del mare. Basso: clima secco e freddo. Medio: clima normale. Alto: clima caldo e umido.

Quota (°)	Totale(m)	Idrostatico (m)	Umido (m)
0	2.411	2.307	0.104
500	2.236	2.174	0.062
1000	2.084	2.047	0.037
2000	1.825	1.812	0.013
3000	1.603	1.599	0.004
4000	1.408	1.406	0.001
5000	1.233	1.233	0.000

Tabella 4.4. Valori esemplificativi di disturbo troposferico zenitale (modello di Saastamoinen) per varie quote in condizioni climatiche normali.

Il modello di Niell, pubblicato nel 1996, introduce nel calcolo della componente umida un termine di dipendenza dalla latitudine, che apporta qualche miglioramento rispetto agli altri modelli standard; il modello è però ancora scarsamente implementato nei programmi commerciali per l'elaborazione dei dati GPS.

Il disturbo troposferico relativo a un segnale proveniente da un satellite a 90° di elevazione è detto zenitale, e spesso viene indicato con l'acronimo inglese ZTD (Zenital Tropospheric Delay).

4.6 L'equazione di osservazione finale

Ricordando che $c\tau_R^S = \rho_R^S(t) + \Delta_R^S(t)$ e che $\Delta_R^S(t) = T_R^S(t) + I_R^S(t)$ si ottiene la

$$c\tau_R^S = \rho_R^S(t) + T_R^S(t) + I_R^S(t)$$

L'equazione di osservazione finale di codice diviene

$$\begin{aligned} P_R^S(t) &= c\tau_R^S + c(dt_R(t) - dt^S(t)) \\ &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + I_R^S(t) + T_R^S(t) \end{aligned} \quad (4.17)$$

in essa compaiono X^S, Y^S, Z^S , ovvero le coordinate del satellite: nelle elaborazioni di nostro interesse vengono considerate note dalle effemeridi, a meno dei relativi errori; X_R, Y_R, Z_R , ovvero le coordinate del ricevitore, sono considerate incognite da stimare; $dt_R(t)$, ovvero l'offset dell'orologio del ricevitore, è incognito, da stimare; $dt^S(t)$, l'offset dell'orologio del satellite, viene considerato noto dal messaggio navigazionale, a meno dei relativi errori; $T_R^S(t), I_R^S(t)$ vengono modellizzati e rimossi, a meno degli errori di modellizzazione. Quindi l'equazione contiene quattro incognite dipendenti dal ricevitore e non dal satellite.

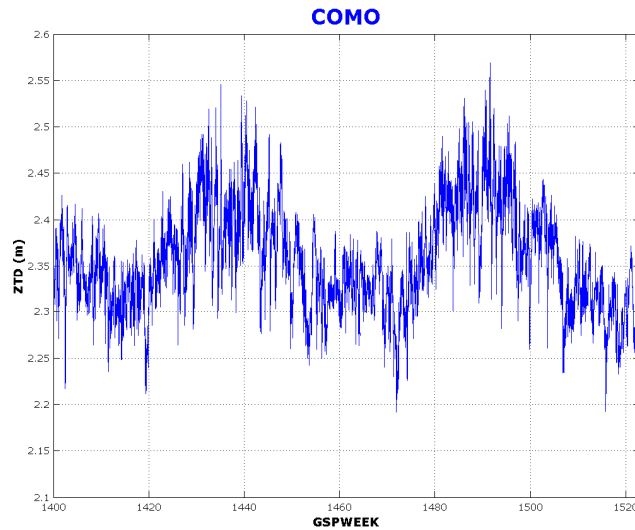


Figura 4.8. Disturbo troposferico zenitale stimato a posteriori per la stazione di Como nel periodo dalla settimana GPS 1400 (novembre 2006) alla 1525 (marzo 2009): si notino le ciclicità annuali.

L'equazione di osservazione finale di fase è data dalla

$$L_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) - I_R^S(t) + T_R^S(t) + \lambda(N_R^S(t) + \phi_R - \phi^S) \quad (4.18)$$

essa ricorda quella di codice, ma contiene l'ambiguità intera $N_R^S(t)$, che dipende dal satellite ed è incognita; contiene inoltre le fasi frazionarie iniziali ϕ_R e ϕ^S , costanti ma anch'esse incognite.

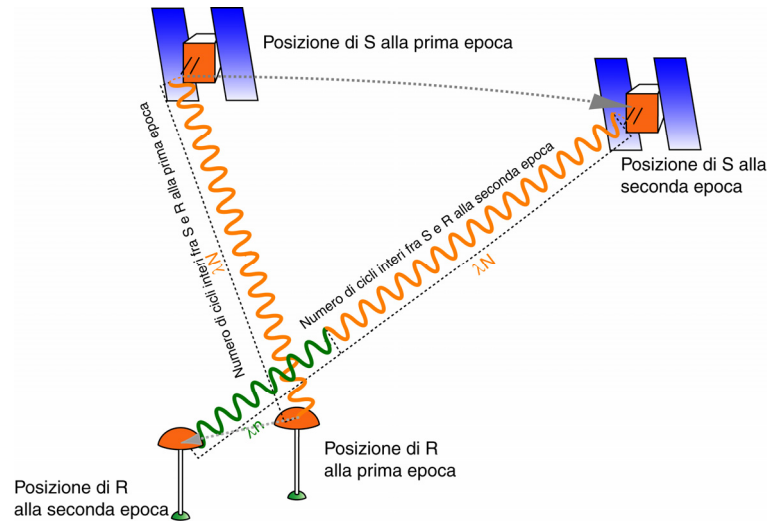


Figura 4.9. Evoluzione nel tempo dell'ambiguità intera.

Abbiamo detto che $N_R^S(t)$ rappresenta il numero intero di cicli compresi fra satellite e ricevitore all'epoca di osservazione; consideriamo due epoche consecutive di misura, t_1 e t_2 , e le relative osservazioni

$$L(t_1) = \rho(t_1) + c(dt_R(t_1) - dt^S(t_1)) - I(t_1) + T(t_1) + \lambda(N(t_1) + \phi_R - \phi^S)$$

$$L(t_2) = \rho(t_2) + c(dt_R(t_2) - dt^S(t_2)) - I(t_2) + T(t_2) + \lambda(N(t_2) + \phi_R - \phi^S)$$

L'ambiguità all'epoca t_2 è data da

$$N(t_2) = N(t_1) + n(t_1, t_2)$$

dove $n(t_1, t_2)$ è il numero intero di cicli trascorsi nell'intervallo di tempo $[t_1, t_2]$ corrispondente alla variazione della distanza fra satellite e ricevitore (Fig. 4.9): questa può arrivare a 1 km/s, ovvero in un secondo si possono avere anche

$$n(t, t+1s) \approx \Delta\rho / \lambda \approx 1000m / 0.2m = 5000 \text{ (cicli)}$$

Quindi per l'osservazione della seconda epoca si può scrivere

$$L(t_2) = \rho(t_2) + c(dt_R(t_2) - dt^S(t_2)) - I(t_2) + T(t_2) + \lambda(N(t_1) + \phi_R - \phi^S) + \lambda n(t_1, t_2)$$

A partire da t_1 , per ogni satellite osservato con continuità, il ricevitore conteggia e registra la variazione del numero intero di cicli di fase, ovvero $n(t_1, t_2)$: per qualunque epoca seguente la prima osservazione, l'ambiguità effettivamente incognita è solo quella relativa alla prima epoca: in altri termini, per periodi continui di osservazione,

l'ambiguità intera è incognita ma costante nel tempo e perciò viene anche detta ambiguità iniziale. Ciò è vero se e solo se non viene perso il contatto fra satellite e ricevitore, altrimenti si interrompe il conteggio di $n(t)$ e, per le epoche successive all'interruzione, si deve considerare una nuova ambiguità intera iniziale; le perdite di contatto fra satellite e ricevitore sono dette cycle slip (traducibile con salto di ciclo) e avvengono allorché:

- la linea di mira dal satellite al ricevitore incontra un'ostruzione fisica come muri, pali, automezzi,...;
- si verificano fenomeni anomali a livello ionosferico, che causano le cosiddette scintillazioni del segnale;
- si verificano problemi di elettronica nel ricevitore che fanno "saltare" il conteggio: con i ricevitori moderni questo fenomeno è ormai piuttosto raro.

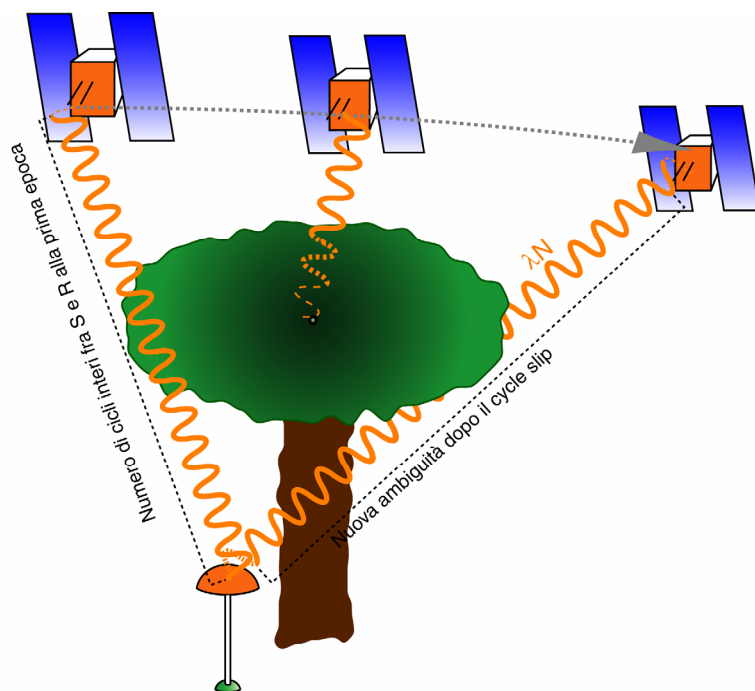


Figura 4.10. Esempio di cycle slip.

Le frequenze delle portanti sono dell'ordine di 1.5 GHz e quindi un'interruzione anche di pochi millisecondi implica il trascorrere di un numero di cicli dell'ordine di $1.5 \cdot 10^9 \cdot 10^{-3} = 1.5 \cdot 10^6$, ovvero del milione di cicli interi. Per capire intuitivamente il significato dell'ambiguità intera di fase e dei cycle slip possiamo pensare al seguente esempio: supponiamo di disporre di un orologio sul quale sia presente la lancetta dei secondi ma non quella dei minuti. Se guardiamo l'orologio non sapremo che minuto è ma avremo solo l'indicazione della parte frazionaria, quindi con un'ambiguità rispetto al numero intero di minuti. Se continuiamo a guardare l'orologio e conteggiamo il numero di passaggi della lancetta dei secondi dallo zero, possiamo però sapere esattamente quanti minuti sono trascorsi dalla prima osservazione: l'unica ambiguità è

quindi quella iniziale; se viceversa interrompiamo la lettura dell'orologio per un certo intervallo di tempo, quando ricominceremo a guardarlo avremo una nuova ambiguità perché non sapremo quanti minuti sono passati: questa nuova ambiguità è l'equivalente di un cycle slip.

Ci si può chiedere ora perché nel simbolismo si mantiene la distinzione fra l'ambiguità intera e le fasi iniziali frazionarie: ovvero, perché non si accorpi nell'equazione di osservazione i tre termini in uno unico, ad esempio ponendo:

$$\lambda(N_R^S(t) + \phi_R - \phi^S) = \eta_R^S(t) \quad (4.19)$$

la ragione per mantenere in generale la distinzione è che i tre termini hanno significato fisico ben diverso; inoltre l'ambiguità è un numero intero che rimane costante solo in assenza di cycle slip, mentre la fase del ricevitore e del satellite sono frazionarie ma rimangono costanti indipendentemente dalla continuità delle osservazioni.

Queste quantità vengono gestite in modo molto diverso nell'elaborazione dei dati per il posizionamento relativo e quindi è utile tenerle distinte nella formulazione; in altre occasioni, come nel posizionamento assoluto mediante osservazioni di fase, è utile accorparle.

4.7 Le osservazioni complessivamente disponibili da un ricevitore a un satellite a un'epoca e le loro combinazioni

Ad ogni epoca qualunque ricevitore può effettuare due osservazioni di codice ad ogni satellite osservato: C/A su L1 e L2C su L2; un ricevitore abilitato ai codici militari aggiunge le quattro osservazioni sui codici crittati, Y e M, entrambi sia su L1 sia su L2: si noti che la disponibilità dei codici militari rende inutili le osservazioni sui codici civili, più rumorose. Nel seguito, quando si discuteranno le osservazioni di codice in generale si userà il simbolo $P_R^S(t)$ come fatto sinora; quando sarà necessario distinguere le osservazioni di codice delle due frequenze si useranno i simboli $P1_R^S(t), P2_R^S(t)$. Ad ogni epoca un ricevitore può effettuare due osservazioni di fase allo stesso satellite, su L1 e su L2; nel seguito si userà il simbolo $L_R^S(t)$ per discutere in generale le osservazioni di fase; ove necessaria una distinzione, si useranno invece $L1_R^S(t), L2_R^S(t)$. In definitiva si hanno almeno quattro osservazioni utili ad ogni epoca

$$\begin{aligned} P1_R^S(t) &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) + I1_R^S(t) \\ P2_R^S(t) &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) + I2_R^S(t) \\ L1_R^S(t) &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) - I1_R^S(t) + \lambda(N1_R^S(t) + \phi1_R - \phi1^S) \\ L2_R^S(t) &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) - I2_R^S(t) + \lambda(N2_R^S(t) + \phi2_R - \phi2^S) \end{aligned} \quad (4.20)$$

Le ambiguità e le fasi iniziali sono diverse per le due portanti: perciò si sono utilizzati i simboli $N1, N2, \phi1, \phi2$; i disturbi ionosferici $I1, I2$ sono differenti ma fra loro vale la (4.14). La distanza, gli offset di orologio e il disturbo troposferico sono uguali per tutte le osservazioni e vengono talvolta accorpati in un termine cosiddetto non dispersivo

$$p_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) \quad (4.21)$$

ovvero

$$\begin{aligned} P_R^S(t) &= p_R^S(t) + I_R^S(t) \\ L_R^S(t) &= p_R^S(t) - I_R^S(t) + \lambda \eta_R^S(t) \end{aligned}$$

$$\text{ove } \eta_R^S(t) = N_R^S(t) + \phi_R - \phi^S$$

Può anche essere effettuato il raccoglimento vettoriale di tutte le osservazioni da un ricevitore a un satellite a un'epoca:

$$\mathbf{y}_R^S(t) = \begin{bmatrix} P1_R^S(t) \\ P2_R^S(t) \\ L1_R^S(t) \\ L2_R^S(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & f_1^2 / f_2^2 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -f_1^2 / f_2^2 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_R^S(t) \\ I1_R^S(t) \\ \eta1_R^S(t) \\ \eta2_R^S(t) \end{bmatrix} = \mathbf{A} \boldsymbol{\xi}_R^S(t) \quad (4.22)$$

la matrice a secondo membro della precedente equazione è di rango pieno e quindi il sistema è ben risolvibile rispetto alle quantità $p_R^S(t), I1_R^S(t), \eta1_R^S(t), \eta2_R^S(t)$; su tale aspetto torneremo nella discussione del posizionamento assoluto.

A partire dalle osservazioni di codice e di fase sulle due frequenze fatte da un ricevitore a un satellite a un'epoca si possono anche costruire combinazioni del tipo

$$O_R^S(t) = \alpha_1 P1_R^S(t) + \alpha_2 P2_R^S(t) + \beta_1 L1_R^S(t) + \beta_2 L2_R^S(t)$$

Nella storia del GPS sono state proposte diverse combinazioni per l'elaborazione dei dati GPS, ciascuna caratterizzata da propri vantaggi; all'atto pratico molte di esse hanno però anche dimostrato notevoli svantaggi che ne hanno fermato l'utilizzo: in questa sede si ritiene opportuno ricordare le combinazioni oggi piuttosto popolari e utilizzate dai programmi di elaborazione.

4.7.1 La combinazione Geometry Free

Iniziamo con le quattro combinazioni che coinvolgono le sole fasi. La prima è chiamata Geometry Free; viene denotata in genere con il simbolo $L4$, ed è data dalla

$$\beta_{41} = 1, \beta_{42} = -1 \quad (4.23)$$

si ha dunque

$$L4_R^S(t) = L1_R^S(t) - L2_R^S(t) = -I1_R^S(t) + I2_R^S(t) + \lambda_1(N1_R^S(t) + \phi1_R - \phi1^S) - \lambda_2(N2_R^S(t) + \phi2_R - \phi2^S)$$

ovvero

$$L4_R^S(t) = \left(\frac{f_1^2}{f_2^2} - 1\right)I1_R^S(t) + \lambda_1(N1_R^S(t) + \phi1_R - \phi1^S) - \lambda_2(N2_R^S(t) + \phi2_R - \phi2^S) \quad (4.24)$$

la $L4$ contiene solo il termine di disturbo ionosferico e una combinazione delle ambiguità, non contiene la geometria e da ciò deriva il suo nome: non può essere quindi utilizzata per la stima della posizione del ricevitore. Viene però utilizzata da alcuni programmi per stimare un modello locale di disturbo ionosferico, da sostituire a quello di Klobuchar nell'elaborazione finale delle osservazioni (Fig. 4.11).

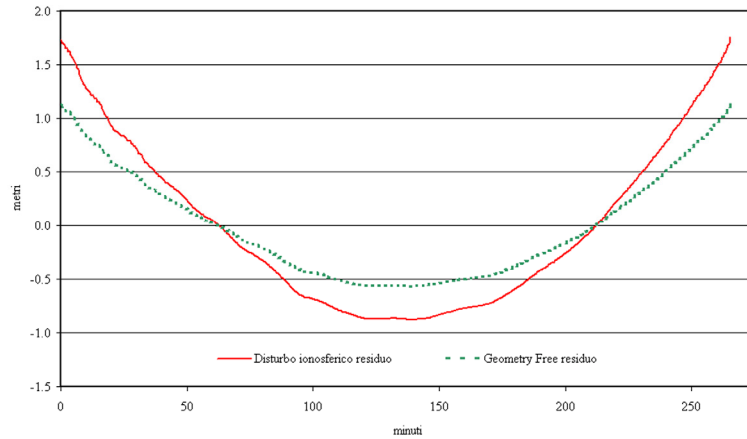


Figura 4.11. Serie temporale di Geometry Free residue ($\delta L4(t)$) di un satellite per un arco di osservazione di circa 4.5 ore; $\delta L4(t)$ viene calcolata mediante la $\delta L4(t) = L4(t) - E(L4)$, ove $E(L4)$ è la media su tutte le epoche, che contiene la media delle differenze di ambiguità e di ionosfera (si veda la 4.24); dalla $\delta L4(t)$ si può calcolare la ionosfera residua, ovvero $\delta I1(t) = f_2^2 / (f_1^2 - f_2^2) \delta L4(t)$, anch'essa rappresentata nel grafico.

4.7.2 La combinazione Wide Lane

La combinazione Wide Lane è denotata tipicamente con il simbolo L5, i suoi coefficienti sono dati da

$$\beta_{51} = \frac{f_1}{f_1 - f_2} = \frac{154f_0}{(154-120)f_0} = \frac{154}{34} \cong 4.53, \quad \beta_{52} = -\frac{f_2}{f_1 - f_2} = -\frac{120}{34} \cong -3.53 \quad (4.25)$$

si ha dunque

$$\begin{aligned} L5_R^S(t) &= \frac{f_1}{f_1 - f_2} L1_R^S(t) - \frac{f_2}{f_1 - f_2} L2_R^S(t) \\ &= \frac{f_1 - f_2}{f_1 - f_2} (\rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t)) \\ &\quad - \frac{f_1}{f_1 - f_2} I1_R^S(t) + \frac{f_2}{f_1 - f_2} I2_R^S(t) \\ &\quad + \frac{f_1 \lambda_1}{f_1 - f_2} (N1_R^S(t) + \phi1_R - \phi1^S) - \frac{f_2 \lambda_2}{f_1 - f_2} (N2_R^S(t) + \phi2_R - \phi2^S) \end{aligned}$$

ovvero

$$\begin{aligned} L5_R^S(t) &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) + \frac{f_1}{f_2} I1_R^S(t) \\ &\quad + \frac{c}{f_1 - f_2} (N1_R^S(t) - N2_R^S(t) + \phi1_R - \phi1^S - \phi2_R + \phi2^S) \end{aligned}$$

ponendo

$$N5_R^S(t) = N1_R^S(t) - N2_R^S(t),$$

$$\phi5_R = \phi1_R - \phi2_R,$$

$$\phi5^S = \phi1^S - \phi2^S,$$

$$\lambda_5 = \frac{c}{f_1 - f_2},$$

$$I5_R^S(t) = \frac{f_1}{f_2} I1_R^S(t)$$

si ottiene l'equazione di osservazione

$$L5_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) + I5_R^S(t) + \lambda_5(N5_R^S(t) + \phi 5_R - \phi 5^S) \quad (4.26)$$

Il disturbo ionosferico e il rumore di osservazione aumentano nella L5 rispetto a L1 o L2: si ha infatti $I5 = 1.3I1$ e $\nu_{L5} = 5.7\nu_{L1}$, inoltre il disturbo ionosferico cambia segno. Per contro alla L5 compete un'ambiguità iniziale con coefficiente moltiplicativo, ovvero lunghezza d'onda fittizia della combinazione, $\lambda_5 \cong 86cm$: con una lunghezza d'onda così ampia, essa presenta numerosi vantaggi nel fissaggio delle ambiguità, che verrà discusso nel Par. 6.9.5.

4.7.3 La combinazione Narrow lane

La combinazione speculare alla Wide Lane è chiamata Narrow Lane, che indicheremo con L6 ed è data dalla

$$\beta_{61} = \frac{f_1}{f_1 - f_2} \cong 4.53, \beta_{62} = \frac{f_2}{f_1 - f_2} \cong 3.53 \quad (4.27)$$

si ha dunque

$$\begin{aligned} L6_R^S(t) &= \frac{f_1}{f_1 - f_2} L1_R^S(t) + \frac{f_2}{f_1 - f_2} L2_R^S(t) \\ &= \frac{f_1 + f_2}{f_1 + f_2} (\rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t)) \\ &\quad - \frac{f_1}{f_1 + f_2} I1_R^S(t) - \frac{f_2}{f_1 + f_2} I2_R^S(t) \\ &\quad + \frac{f_1 \lambda_1}{f_1 + f_2} (N1_R^S(t) + \phi 1_R - \phi 1^S) + \frac{f_2 \lambda_2}{f_1 + f_2} (N2_R^S(t) + \phi 2_R - \phi 2^S) \end{aligned}$$

ovvero

$$\begin{aligned} L6_R^S(t) &= \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) - \frac{f_1}{f_2} I1_R^S(t) \\ &\quad + \frac{c}{f_1 + f_2} (N1_R^S(t) + N2_R^S(t) + \phi 1_R - \phi 1^S + \phi 2_R - \phi 2^S) \end{aligned}$$

ponendo

$$N6_R^S(t) = N1_R^S(t) + N2_R^S(t),$$

$$\phi 6_R = \phi 1_R + \phi 2_R,$$

$$\phi 4^S = \phi 1^S + \phi 2^S,$$

$$\lambda_6 = \frac{c}{f_1 + f_2},$$

$$I6_R^S(t) = \frac{f_1}{f_2} I1_R^S(t)$$

si ottiene l'equazione di osservazione

$$\begin{aligned} L6_R^S(t) = & \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) - I_{WL_R}^S(t) \\ & + \lambda_6(N6_R^S(t) + \phi 6_R - \phi 6^S) \end{aligned} \quad (4.28)$$

in sintesi, alla L6 compete $\lambda_6 \cong 10.7 \text{ cm}$. Si noti che la sua ambiguità ha la stessa parità di $N5$; infatti se sia $N1$ sia $N2$ sono pari, lo sono anche la loro somma e la loro differenza e lo stesso vale se entrambi sono dispari; se uno di loro è dispari e l'altro è pari sia la loro somma sia la loro differenza sono dispari: perciò L6 viene usata congiuntamente alla L5 nei processi di fissaggio delle ambiguità. Anche per la L6, si ha $I6 = 1.3I1$ e $\nu_{L6} = 5.7\nu_{L1}$.

4.7.4 La combinazione Ionospheric Free

L'ultima combinazione qui considerata è la cosiddetta Ionospheric Free, e viene denotata con il simbolo L3; i coefficienti sono:

$$\beta_{31} = \frac{f_1^2}{f_1^2 - f_2^2} = \frac{154^2}{154^2 - 120^2} \cong 2.55, \beta_{32} = -\frac{f_2^2}{f_1^2 - f_2^2} = -\frac{120^2}{154^2 - 120^2} \cong -1.55 \quad (4.29)$$

La L3 può essere calcolata sulle osservazioni di fase

$$L3_R^S(t) = \frac{f_1^2}{f_1^2 - f_2^2} L1_R^S(t) - \frac{f_2^2}{f_1^2 - f_2^2} L2_R^S(t)$$

con conti analoghi ai precedenti, si ottiene

$$\begin{aligned}
L3_R^S(t) = & \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) \\
& - \frac{f_1^2}{f_1^2 - f_2^2} I1_R^S(t) + \frac{f_2^2}{f_1^2 - f_2^2} I2_R^S(t) \\
& - \frac{f_1^2 \lambda_1}{f_1^2 - f_2^2} (N1_R^S(t) + \phi1_R - \phi1^S) + \frac{f_2^2 \lambda_2}{f_1^2 - f_2^2} (N2_R^S(t) + \phi2_R - \phi2^S) + \varepsilon_{IFR}^S(t)
\end{aligned}$$

ricordando la (4.14) si ottiene

$$\begin{aligned}
- \frac{f_1^2}{f_1^2 - f_2^2} I1_R^S(t) + \frac{f_2^2}{f_1^2 - f_2^2} I2_R^S(t) &= - \frac{f_1^2}{f_1^2 - f_2^2} I1_R^S(t) + \frac{f_2^2}{f_1^2 - f_2^2} \frac{f_1^2}{f_2^2} I1_R^S(t) \\
&= 0
\end{aligned}$$

ovvero la Ionospheric Free non contiene in alcun modo il disturbo ionosferico: da ciò, evidentemente, il suo nome. Per quanto riguarda il termine di ambiguità, esso può essere riscritto nel seguente modo:

$$\begin{aligned}
& \frac{f_1^2 \lambda_1}{f_1^2 - f_2^2} (N1_R^S(t) + \phi1_R - \phi1^S) - \frac{f_2^2 \lambda_2}{f_1^2 - f_2^2} (N2_R^S(t) + \phi2_R - \phi2^S) = \\
&= \frac{f_1 c}{f_1^2 - f_2^2} (N1_R^S(t) + \phi1_R - \phi1^S) - \frac{f_2 c}{f_1^2 - f_2^2} (N2_R^S(t) + \phi2_R - \phi2^S) = \\
&= \frac{c}{f_1^2 - f_2^2} (f_1 (N1_R^S(t) + \phi1_R - \phi1^S) - f_2 (N2_R^S(t) + \phi2_R - \phi2^S))
\end{aligned}$$

Ricordando che $f_1 = 154 f_0$, $f_2 = 120 f_0$, si ha

$$\begin{aligned}
& \frac{f_1 c}{f_1^2 - f_2^2} (N1_R^S(t) + \phi1_R - \phi1^S) - \frac{f_2 c}{f_1^2 - f_2^2} (N2_R^S(t) + \phi2_R - \phi2^S) = \\
&= \frac{c f_0}{f_1^2 - f_2^2} (154 (N1_R^S(t) + \phi1_R - \phi1^S) - 120 (N2_R^S(t) + \phi2_R - \phi2^S)) = \\
&= \frac{c f_0}{f_1^2 - f_2^2} (154 N1_R^S(t) - 120 N2_R^S(t) + 154 \phi1_R - 120 \phi2_R - 154 \phi1^S + 120 \phi2^S)
\end{aligned}$$

Ponendo

$$\lambda_3 = \frac{cf_0}{f_1^2 - f_2^2} = \frac{c}{(154^2 - 120^2)f_0}$$

$$N3_R^S(t) = 154N1_R^S(t) - 120N2_R^S(t)$$

$$\phi3_R = 154\phi1_R - 120\phi2_R$$

$$\phi3^S = 154\phi1^S + 120\phi2^S$$

si ottiene infine

$$L3_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) + \lambda_3(N3_R^S(t) + \phi3_R - \phi3^S) \quad (4.30)$$

La lunghezza d'onda associata alla L3 è $\lambda_3 \cong 0.77mm$, ovvero non è nulla come invece viene erroneamente indicato in alcuni testi; in ogni caso è più piccola del rumore di osservazione, $\nu_{L3} \cong 3\nu_{L1}$, e quindi, da un punto di vista pratico, non è separabile da esso. Per quanto riguarda i codici, l'unica combinazione effettivamente utilizzata nell'elaborazione dei dati GPS è la Ionospheric Free:

$$P3_R^S(t) = \frac{f_1^2}{f_1^2 - f_2^2} P1_R^S(t) - \frac{f_2^2}{f_1^2 - f_2^2} P2_R^S(t)$$

Si noti che le frequenze f_i indicate nei coefficienti sono nuovamente quelle delle portanti; l'equazione di osservazione finale è

$$P3_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) + T_R^S(t) \quad (4.31)$$

All'atto pratico oggi non è operabile la combinazione P3 mediante C/A e L2C, poiché solo sei satelliti hanno già implementato quest'ultimo; è viceversa già possibile e implementata la combinazione Ionospheric free nei ricevitori abilitati ai codici Y1 e Y2.

Nome	β_1	β_2	λ (cm)	I/I_1	ν/ν_1
L1	1	0	19	1	1.0
L2	0	1	24	1.6	1.0
L4	1	-1	∞	0.6	1.4
L5	4.53	-3.53	86	1.3	5.7
L6	4.53	3.53	11	1.3	5.7
L3	2.55	-1.55	0.77	0.0	3.0

Tabella 4.5. Elenco delle principali combinazioni di fase e delle loro caratteristiche. I coefficienti sono approssimati a due cifre decimali, le lunghezze d'onda alle prime due cifre significative. Ionosfera e rumore elettronico sono in rapporto a quelli su L1.

4.8 L'andamento nel tempo delle osservazioni da un ricevitore a un satellite

Si vogliono qui discutere le variazioni nel tempo delle quantità contenute nelle osservazioni di fase e codice; siano t_{k-1} e t_k due epoche consecutive: si adotta la convenzione $\Delta(\bullet)(t_k) = (\bullet)(t_k) - (\bullet)(t_{k-1})$ e si scrive

$$\begin{aligned} P_R^S(t_k) - P_R^S(t_{k-1}) &= \Delta P_R^S(t_k) \\ &= \Delta \rho_R^S(t_k) + c(\Delta dt_R(t_k) - \Delta dt^S(t_k)) + \Delta T_R^S(t_k) + \Delta I_R^S(t_k) \\ L_R^S(t_k) - L_R^S(t_{k-1}) &= \Delta L_R^S(t_k) \\ &= \Delta \rho_R^S(t_k) + c(\Delta dt_R(t_k) - \Delta dt^S(t_k)) + \Delta T_R^S(t_k) - \Delta I_R^S(t_k) + \lambda \Delta N_R^S(t_k) \end{aligned}$$

Nell'equazione di osservazione, il termine geometrico $\rho_R^S(t)$ presenta le maggiori variazioni nel tempo, poiché il moto relativo fra ricevitore e satellite può arrivare a 1 km/s: date le caratteristiche delle orbite dei satelliti, $\rho_R^S(t)$ ha un andamento molto liscio e semplicemente modellizzabile con polinomi. Considerando una distanza fra satellite e ricevitore di circa 20000 km, la velocità angolare ω del satellite rispetto al ricevitore è approssimativamente calcolabile mediante la

$$\omega \leq \frac{1 \text{ km/s}}{20000 \text{ km}} = 5 \cdot 10^{-5} \text{ rad/s} \cong 10''/\text{s} \quad (4.32)$$

quindi, perché l'orbita di un satellite spazzi un angolo di un grado rispetto ad un osservatore, sono necessari almeno sei minuti.

La traccia del segnale si muove attraverso la troposfera molto lentamente: se ad esempio consideriamo lo strato a 10 km di quota, il segnale si muove in esso a

$$v \leq 10 \text{ km} \cdot 5 \cdot 10^{-5} \text{ rad/s} \cong 0.5 \text{ m/s}$$

perciò il disturbo troposferico cambia lentamente nel tempo e in modo molto liscio: indicativamente le sue variazioni sull'arco del minuto sono inferiori al centimetro (Fig. 4.8). La ionosfera è soggetta a minore variabilità temporale e spaziale ma il relativo disturbo è più grande la traccia del segnale vi si muove con velocità dell'ordine di alcune decine di metri al secondo; alle nostre latitudini, nelle ore centrali della giornata, le variazioni di disturbo ionosferico su ogni segnale possono raggiungere alcuni centimetri al minuto (Fig. 4.11).

Tipicamente gli offset degli orologi dei satelliti presentano sull'arco del minuto variazioni inferiori al nanosecondo che, convertito in termini metrici, significa 30 cm; il loro comportamento, come già detto, è ben descritto da semplici polinomi (Fig. 4.12). Gli orologi dei ricevitori hanno comportamento più erratico di quelli dei satelliti, legato all'età del ricevitore e alla sua fascia di costo. Come già detto le fasi frazionarie iniziali sono costanti mentre le ambiguità intere lo rimangono finché non occorrono cycle slip:

in questo caso i salti possono andare da pochi a milioni di cicli, e il loro effetto metrico sull'osservazione è dato da $\lambda_i n$.

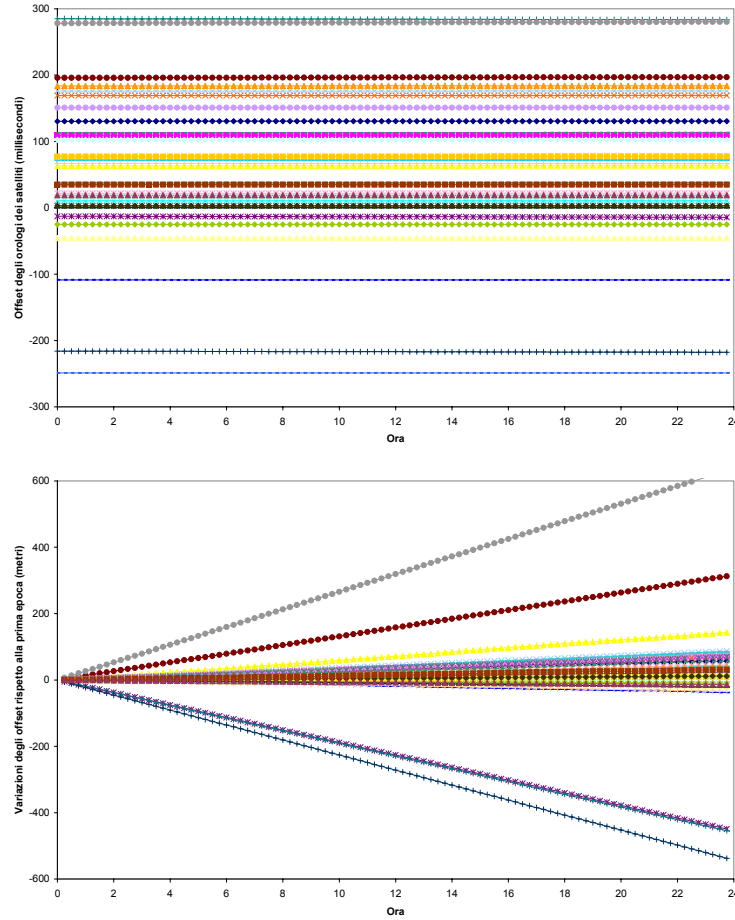


Figura 4.12. In alto, offset degli orologi dei satelliti, in millisecondi, per una giornata di esempio; in basso, variazione dei suddetti offset rispetto alla prima epoca, già convertita in metri. Si noti l'andamento molto liscio nel tempo.

4.9 Il filtro di Hatch

Le osservazioni di fase hanno piccolo rumore di osservazione ma presentano l'ambiguità intera e la fase frazionaria iniziali; le osservazioni di codice sono non ambigue ma anche molto rumorose. Ai fini della navigazione può essere costruita una combinazione ricorsiva sulla serie temporale di osservazioni di codice e fase che sfrutti le caratteristiche positive delle due osservabili, riducendone le problematiche; particolarmente popolare è il filtro di Hatch, che viene quindi descritto nel seguito.

L'osservazione filtrata all'epoca t , denominata nel seguito $PF_R^S(t)$, viene costruita da un codice e una fase relative alla stessa frequenza, mediante la seguente espressione:

$$PF_R^S(t) = w(t)P_R^S(t) + (1 - w(t))(PF_R^S(t-1) + L_R^S(t) - L_R^S(t-1)) \quad (4.33)$$

dove, alla prima epoca si inizializza il filtro mediante la

$$w(t_1) = 1$$

$$PF_R^S(t_1) = P_R^S(t_1)$$

per $1 < i \leq I$, $w(t_i)$ viene fatto linearmente decrescere:

$$w(t_i) = 1 - \frac{(i-1)}{I};$$

infine per $i > I$, $w(t_i)$ assume valore costante,

$$w(t_i) = w = 1/I.$$

Si consideri il filtro a regime, ovvero per un'epoca k dopo la I-esima; in tal caso l'espressione del filtro di Hatch è

$$PF_R^S(t_k) = wP_R^S(t_k) + (1 - w)(PF_R^S(t_{k-1}) + L_R^S(t_k) - L_R^S(t_{k-1}))$$

$$= \sum_{i=0}^I (1 - w)^i (wP_R^S(t_{k-i}) + (1 - w)\Delta\phi_R^S(t_{k-i})) + (1 - w)^{I-1} PF_R^S(t_{k-I}) \quad (4.34)$$

Per illustrare le caratteristiche del filtro di Hatch in modo semplice si considera il caso $I = 100$, che è quello comunemente adottato; si ipotizza l'assenza di cycle slip nelle osservazioni di fase fra l'epoca t_{k-1} e t_k ; si ha quindi

$$\Delta L_R^S(t_k) = L_R^S(t_k) - L_R^S(t_{k-1}) = \Delta\rho_R^S(t_k) + c(\Delta dt_R(t_k) - \Delta dt^S(t_k)) + \Delta T_R^S(t_k) - \Delta I_R^S(t_k)$$

Si può inoltre scrivere

$$PF_R^S(t_{k-1}) = \rho_R^S(t_{k-1}) + c(dt_R(t_{k-1}) - dt^S(t_{k-1})) + T_R^S(t_{k-1}) + I_R^S(t_{k-1}) + \delta_R^S(t_{k-1})$$

ove il termine $\delta_R^S(t_{k-1})$ esprime gli eventuali errori di modello del filtro all'epoca t_{k-1} . Si ha quindi

$$\begin{aligned}
PF_R^S(t_k) &= 0.01P_R^S(t_k) + 0.99(PF_R^S(t_{k-1}) + \Delta L_R^S(t_k)) \\
&= 0.01(\rho_R^S(t_k) + c(dt_R(t_k) - dt^S(t_k)) + T_R^S(t_k) + I_R^S(t_k)) \\
&\quad + 0.99(\rho_R^S(t_{k-1}) + c(dt_R(t_{k-1}) - dt^S(t_{k-1})) + T_R^S(t_{k-1}) + I_R^S(t_{k-1}) + \delta_R^S(t_{k-1}) + \Delta L_R^S(t_k))
\end{aligned}$$

In assenza di cycle slip vale la seguente

$$\begin{aligned}
&\rho_R^S(t_{k-1}) + c(dt_R(t_{k-1}) - dt^S(t_{k-1})) + T_R^S(t_{k-1}) + I_R^S(t_{k-1}) + \delta_R^S(t_{k-1}) + \Delta L_R^S(t_k) \\
&= \rho_R^S(t_k) + c(dt_R(t_k) - dt^S(t_k)) + T_R^S(t_k) + 2I_R^S(t_{k-1}) - I_R^S(t_k) + \delta_R^S(t_{k-1})
\end{aligned}$$

Si ottiene perciò

$$\begin{aligned}
PF_R^S(t_k) &= 0.01(\rho_R^S(t_k) + c(dt_R(t_k) - dt^S(t_k)) + T_R^S(t_k) + I_R^S(t_k)) \\
&\quad + 0.99(\rho_R^S(t_k) + c(dt_R(t_k) - dt^S(t_k)) + T_R^S(t_k) + 2I_R^S(t_{k-1}) - I_R^S(t_k) + \delta_R^S(t_{k-1})) \\
&= \rho_R^S(t_k) + c(dt_R(t_k) - dt^S(t_k)) + T_R^S(t_k) \\
&\quad + 1.98I_R^S(t_{k-1}) - 0.98I_R^S(t_k) + 0.99\delta_R^S(t_{k-1})
\end{aligned} \tag{4.35}$$

Vale la $1.98I_R^S(t_{k-1}) - 0.98I_R^S(t_k) = I_R^S(t_{k-1}) - 0.98\Delta I_R^S(t_k)$, ovvero nel filtro compare un termine ionosferico che è combinazione dei disturbi delle due ultime epoche; per la geometria, l'equazione di osservazione è identica a quella dei codici e non vi è deriva degli errori di modello nel tempo: quindi l'osservazione di codice filtrata può essere utilizzata per il posizionamento, senza il problema delle ambiguità e delle fasi iniziali. Il grande vantaggio è che il rumore di osservazione del codice viene abbattuto di un fattore I , ovvero, trascurando il rumore delle fasi, si ha

$$\nu_{PF} \simeq \frac{1}{I} \nu_P$$

Si consideri però che un cycle slip di n cicli causa un salto di $0.99\lambda n$ nell'osservazione; in tal senso è necessario applicare un algoritmo di identificazione di discontinuità prima di utilizzare le osservazioni filtrate per il posizionamento: allorché si identificano cycle slip, è necessario inizializzare ex novo il filtro. Si noti infine che il filtro di Hatch può essere anche applicato alle combinazioni ionofree di codice (4.31) e fase (4.30). In tal caso l'equazione di osservazione è data dalla

$$PF3_R^S(t_k) = \rho_R^S(t_k) + c(dt_R(t_k) - dt^S(t_k)) + T_R^S(t_k) + 0.99\delta_R^S(t_{k-1}) \tag{4.36}$$

e non contiene disturbo ionosferico, che quindi non richiede modellizzazione.

Capitolo 5.

Il posizionamento assoluto

Si intende con posizionamento assoluto di un ricevitore la stima della sua posizione a partire da:

- osservazioni ai satelliti, di codice e di fase,
- messaggio navigazionale dei satelliti.

Rigorosamente parlando, il termine di posizionamento assoluto è improprio: il ricevitore si posiziona rispetto ai satelliti, e quindi si tratta di un posizionamento relativo; il termine posizionamento relativo è però normalmente utilizzato per la tecnica discussa nel capitolo successivo e ci si attiene quindi alla terminologia ormai invalsa. Il posizionamento assoluto può essere effettuato in singola epoca, ovvero in rilievi cinematici e per la navigazione; in alternativa può essere effettuato mediante stazionamento prolungato, ovvero occupando lo stesso punto per più epoche. Per effettuare il posizionamento assoluto si ricorre in genere alla stima mediante minimi quadrati, che però richiede la linearizzazione delle equazioni di osservazione rispetto a valori approssimati delle incognite: il capitolo è quindi introdotto da un breve paragrafo tecnico su come sia possibile stimare tali coordinate approssimate; poi si prosegue e si dettaglia la discussione più rilevante, ovvero il metodo di stima mediante minimi quadrati e il relativo bilancio di errore. Viene quindi presentato l'aspetto del posizionamento assoluto mediante osservazioni di fase e il cosiddetto Precise Point Positioning: a tali aspetti non si dedica molto dettaglio poiché sono ancora oggi argomento di ricerca e non rappresentano la tipica prassi di elaborazione. Il capitolo è chiuso da alcuni dettagli relativi alla linearizzazione dell'equazione di osservazione.

5.1 La soluzione di Bancroft

Si utilizzano le sole osservazioni di codice su C/A; si considerano note le coordinate e l'offset d'orologio dei satelliti dal messaggio navigazionale e si pongono uguali a zero i disturbi troposferici, poiché non si conosce ancora la posizione del ricevitore, e quindi la sua quota e l'elevazione dei satelliti; per il disturbo ionosferico si può utilizzare un valore zenitale medio $\bar{I}_R^S(t)$, in modo da correggerne almeno una parte. Si consideri l'osservazione ridotta da un ricevitore a un satellite

$$\Pi_R^S(t) = P_R^S(t) + cdt_S(t) - \bar{I}_R^S(t) = \rho_R^S(t) + cdt_R(t)$$

omettiamo nella notazione la dipendenza dal tempo, poiché per definizione stiamo lavorando in singola epoca. Si può scrivere

$$\Pi_R^S - \theta_R = \rho_R^S$$

ove si è posto $cdt_R = \theta_R$; elevando al quadrato entrambi i membri dell'equazione si ha

$$\begin{aligned} \Pi_R^{S^2} + \theta_R - 2\Pi_R^S \theta_R &= (X_R - X^S)^2 + (Y_R - Y^S)^2 + (Z_R - Z^S)^2 \\ &= X_R^2 + Y_R^2 + Z_R^2 + X^{S^2} + Y^{S^2} + Z^{S^2} - 2X_R X^S - 2Y_R Y^S - 2Z_R Z^S \end{aligned}$$

con un diverso raggruppamento si scrive

$$(X^{S^2} + Y^{S^2} + Z^{S^2} - \Pi_R^{S^2}) - 2(X_R X^S + Y_R Y^S + Z_R Z^S - \Pi_R^S \theta_R) = -(X_R^2 + Y_R^2 + Z_R^2 - \theta_R^2) \quad (5.1)$$

Si definisce il prodotto di Lorenz in $\mathbb{R}^4$

$$\langle \mathbf{g}, \mathbf{h} \rangle = \mathbf{g}^T \mathbf{M} \mathbf{h}, \quad \mathbf{g}, \mathbf{h} \in \mathbb{R}^4, \quad \mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

con la proprietà fondamentale $\langle \mathbf{M} \mathbf{g}, \mathbf{M} \mathbf{h} \rangle = \langle \mathbf{g}, \mathbf{h} \rangle$, $\mathbf{M} \mathbf{M} = \mathbf{I}$. Nel seguito si pone

$$\xi = \begin{bmatrix} \mathbf{x}_R \\ \theta_R \end{bmatrix}, \quad \xi = \frac{1}{2} \langle \xi, \xi \rangle, \quad \mathbf{w}^S = \begin{bmatrix} \mathbf{x}^S \\ \Pi_R^S \end{bmatrix}$$

La (5.1) può quindi essere scritta come

$$\frac{1}{2} \langle \mathbf{w}^S, \mathbf{w}^S \rangle - \langle \mathbf{w}^S, \xi \rangle + \xi = 0 \quad (5.2)$$

Si considerino ora quattro satelliti 1,2,3,4, le cui linee di mira dal ricevitore non giacciono sullo stesso piano; il sistema che considera nel complesso le relative osservazioni può essere scritto come

$$\mathbf{a} - \mathbf{W} \mathbf{M} \xi + \xi \mathbf{i} = 0$$

ove $\mathbf{a}$ e $\mathbf{i}$ sono vettori 4×1 mentre $\mathbf{W}$ è una matrice 4×4 , rispettivamente dati dalle

$$a_i = \frac{1}{2} \langle \mathbf{w}^i, \mathbf{w}^i \rangle, \mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{W} = \begin{bmatrix} \mathbf{w}^{1T} \\ \mathbf{w}^{2T} \\ \mathbf{w}^{3T} \\ \mathbf{w}^{4T} \end{bmatrix} = \begin{bmatrix} X^1 & Y^1 & Z^1 & \Pi_R^1 \\ X^2 & Y^2 & Z^2 & \Pi_R^2 \\ X^3 & Y^3 & Z^3 & \Pi_R^3 \\ X^4 & Y^4 & Z^4 & \Pi_R^4 \end{bmatrix}$$

La non complanarità delle linee di mira ai satelliti garantisce il pieno rango di $\mathbf{W}$ e si può scrivere

$$\xi = \mathbf{M}\mathbf{W}^{-1}(\xi\mathbf{i} + \mathbf{a}) \quad (5.3)$$

Si sostituisce in (5.2) ogni occorrenza di ξ , e ricordando la proprietà del prodotto di Lorenz, si ottiene

$$\langle \mathbf{W}^{-1}\mathbf{i}, \mathbf{W}^{-1}\mathbf{i} \rangle \xi^2 + 2(\langle \mathbf{W}^{-1}\mathbf{i}, \mathbf{W}^{-1}\mathbf{a} \rangle - 1)\xi + \langle \mathbf{W}^{-1}\mathbf{a}, \mathbf{W}^{-1}\mathbf{a} \rangle = 0$$

questa è un'equazione di secondo grado in ξ , che quindi ammette due soluzioni le quali, una volta inserite nella (5.3), generano due soluzioni alternative per le coordinate del ricevitore: una di esse è quella corretta, ovvero geometricamente fondata, l'altra è artificialmente generata dalla quadratura iniziale delle equazioni di osservazione. Per identificare la soluzione corretta si hanno due vie; se solo quattro satelliti sono effettivamente in vista si considerano le due stime di quota: nella quasi totalità dei casi, una sola di queste assume valori ragionevoli, ed è quella relativa alla stima geometricamente fondata; esistono rarissime eccezioni, in cui entrambe le stime portano a valori plausibili delle coordinate e in tal caso per effettuare la scelta bisogna attendere qualche epoca: la soluzione non fondata si muove a velocità inverosimili. Se sono in vista più di quattro satelliti, come tipicamente avviene, si itera il calcolo selezionandone di volta in volta solo quattro; la soluzione geometricamente fondata è quella condivisa, o meglio la più simile, fra le varie coppie di soluzioni.

Nel caso solo quattro satelliti siano in vista la soluzione di Bancroft è l'unica calcolabile; in ogni caso è necessario iterarla perlomeno due volte, in modo da inserire alla seconda iterazione dei valori approssimati, calcolati nella posizione del ricevitore, per i disturbi troposferico e ionosferico. Nel caso di più di quattro satelliti in vista la soluzione di Bancroft non è ottimale poiché non viene fatta alcuna modellizzazione statistica degli errori; peraltro essa permette di ottenere i valori approssimati delle coordinate del ricevitore, con accuratezza dell'ordine del centinaio di metri, necessari per linearizzare le equazioni di osservazione e applicare i minimi quadrati, sia in singola epoca sia nel caso statico.

5.2 La soluzione ridondante in singola epoca mediante osservazioni di codice

Sono disponibili dei valori approssimati $(\tilde{X}_R, \tilde{Y}_R, \tilde{Z}_R)$ per le coordinate del ricevitore; si considerano noti, a meno dei relativi errori di modellizzazione, i disturbi ionosferico e troposferico $(\tilde{I}_R^S, \tilde{T}_R^S)$ e la posizione e l'offset d'orologio dei satelliti $(\tilde{\mathbf{x}}^S, d\tilde{t}^S)$. L'equazione di osservazione di codice a ciascun satellite viene linearizzata rispetto alle coordinate del ricevitore, ottenendo la:

$$\begin{aligned} P_{R0}^S(t) = & \frac{\tilde{X}_R - \tilde{X}_R^S(t)}{\tilde{\rho}_R^S(t)} (X_R(t) - \tilde{X}_R) + \frac{\tilde{Y}_R - \tilde{Y}_R^S(t)}{\tilde{\rho}_R^S(t)} (Y_R(t) - \tilde{Y}_R) \\ & + \frac{\tilde{Z}_R - \tilde{Z}_R^S(t)}{\tilde{\rho}_R^S(t)} (Z_R(t) - \tilde{Z}_R) + cdt_R(t) \\ & + (\tilde{\rho}_R^S(t) - cd\tilde{t}^S(t) + \tilde{T}_R^S(t) + \tilde{I}_R^S(t)) + \delta(\mathbf{x}^S, dt^S, I_R^S, T, t) + \nu_R^S(t) \end{aligned}$$

ove il termine $\delta(\mathbf{x}^S, dt^S, I_R^S, T_R^S)$ contiene l'effetto degli errori di modello. Con la notazione introdotta in App. 1 e indicando con

$$b_R^S(t) = \tilde{\rho}_R^S(t) - cd\tilde{t}^S(t) + \tilde{T}_R^S(t) + \tilde{I}_R^S(t), \quad \theta_R(t) = cdt_R(t)$$

si ha

$$P_{R0}^S(t) = b_R^S(t) + \tilde{\mathbf{e}}_R^S(t) \cdot \delta \mathbf{x}_R(t) + \theta_R(t) + \delta(\mathbf{x}^S, dt^S, I_R^S, T_R^S, t) + \nu_R^S(t)$$

Considerando I satelliti in vista all'epoca t , si può scrivere

$$\mathbf{y}_0(t) = \mathbf{P}_0(t) - \mathbf{b}(t) = \mathbf{A}(t)\boldsymbol{\xi}(t) + \mathbf{v}(t) \quad (5.4)$$

ove

$$\begin{aligned} \mathbf{P}_0(t) &= \begin{bmatrix} P_{R0}^1(t) \\ P_{R0}^2(t) \\ \dots \\ P_{R0}^I(t) \end{bmatrix}, \quad \mathbf{A}(t) = [\mathbf{E}_R(t) \quad \mathbf{i}], \quad \mathbf{E}_R(t) = \begin{bmatrix} \tilde{e}_{RX}^1(t) & \tilde{e}_{RY}^1(t) & \tilde{e}_{RZ}^1(t) \\ \tilde{e}_{RX}^2(t) & \tilde{e}_{RY}^2(t) & \tilde{e}_{RZ}^2(t) \\ \dots & \dots & \dots \\ \tilde{e}_{RX}^I(t) & \tilde{e}_{RY}^I(t) & \tilde{e}_{RZ}^I(t) \end{bmatrix}, \quad \mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix}, \\ \boldsymbol{\xi}(t) &= \begin{bmatrix} \delta \mathbf{x}_R(t) \\ \theta_R(t) \end{bmatrix}, \quad \mathbf{b}(t) = \begin{bmatrix} b_R^1(t) \\ b_R^2(t) \\ \dots \\ b_R^I(t) \end{bmatrix}, \quad \mathbf{v}(t) = \begin{bmatrix} \delta(X^1, \dots, T_R^1, t) + \nu_R^1 \\ \delta(X^2, \dots, T_R^2, t) + \nu_R^2 \\ \dots \\ \delta(X^I, \dots, T_R^I, t) + \nu_R^I \end{bmatrix} \end{aligned}$$

che è risolvibile rispetto a ξ mediante minimi quadrati, quando $m \geq 4$. Gli errori di modello (atmosfera, effemeridi e offset d'orologio del satellite) vengono considerati parte del bilancio di errore insieme al rumore elettronico e accorpati nel vettore di errore $\mathbf{v}(t)$; le osservazioni ai diversi satelliti vengono considerate reciprocamente incorrelate e di uguale precisione e quindi la loro matrice di covarianza viene posta uguale a $\mathbf{C}_{\varepsilon\varepsilon} = \sigma_0^2 \mathbf{I}$: la soluzione dei minimi quadrati è quella discussa nell'App. 1. La soluzione in singola epoca può essere calcolata in tempo reale, sfruttando effemeridi e offset d'orologio dei satelliti trasmessi mediante il messaggio navigazionale, oppure in postelaborazione utilizzando effemeridi e offset calcolati a posteriori.

5.3 Un'analisi degli errori di modello

La presenza di errori di modello causa un errore nell'equazione di osservazione: cerchiamo di analizzarne gli effetti sulla stima finale.

1. Definiamo $\delta \mathbf{x}^S(t)$ il vettore d'errore delle effemeridi utilizzate per il satellite S . Linearizzando l'equazione di osservazione rispetto a $\delta \mathbf{x}^S(t)$ si ottiene:

$$\rho_R^S(t) = \tilde{\rho}_R^S(t) + \tilde{\mathbf{e}}_R^S(t) \cdot \delta \mathbf{x}^S(t) \quad (5.5)$$

ovvero l'errore introdotto nel modello, $\tilde{\mathbf{e}}_R^S(t) \cdot \delta \mathbf{x}^S(t)$, è uguale alla proiezione dell'errore delle effemeridi nella direzione di mira dal ricevitore al satellite (Fig. 5.1).

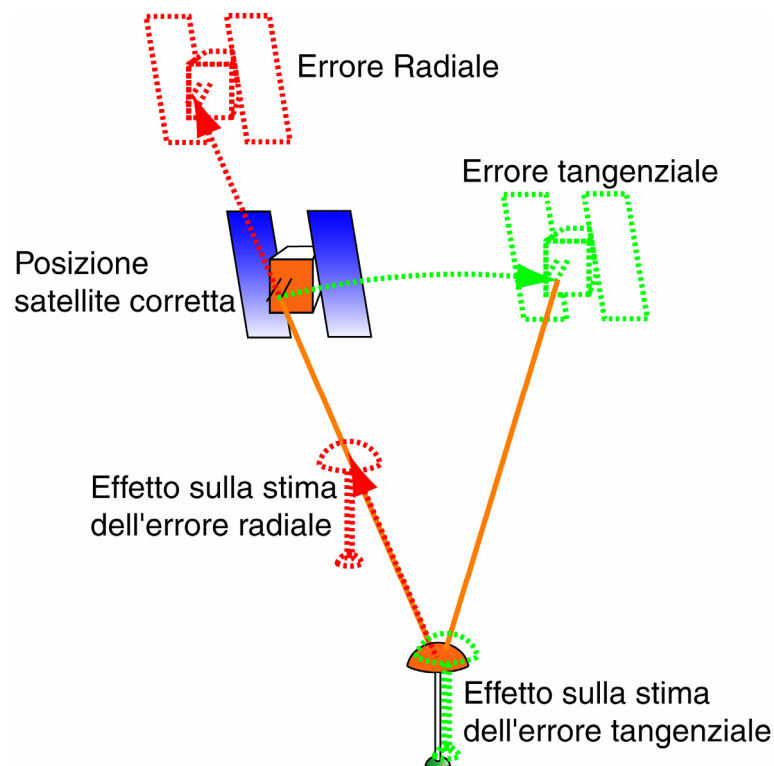


Figura 5.1. L'effetto di un errore radiale delle effemeridi sulla stima della posizione è maggiore dell'effetto di un errore tangenziale.

2. Si definisce $\delta t^S(t)$ l'errore nel valore utilizzato per modellizzare l'offset d'orologio del satellite S e si può scrivere

$$cdt^S(t) = cd\tilde{t}^S(t) + c\delta t^S(t) \quad (5.6)$$

Quindi l'errore causa un errore proporzionale nell'equazione di osservazione, in cui la costante di proporzionalità è la velocità di propagazione del segnale elettromagnetico nel vuoto: ad esempio

$$\delta t^S(t) = 5ns \Rightarrow c\delta t^S(t) = 5^{-9}s \cdot 3 \cdot 10^8 m/s = 1.5m$$

3. L'errore metrico nella modellizzazione dei disturbi atmosferici ($\delta I_R^S(t)$ e $\delta T_R^S(t)$) comporta un identico errore nell'equazione di osservazione. L'errore di modello troposferico è sostanzialmente simile fra i diversi satelliti in vista, i cui segnali hanno sostanzialmente percorso la medesima troposfera; l'errore di modello ionosferico non è necessariamente simile poiché i segnali provenienti da satelliti con differenti direzioni di mira hanno percorso strati ionosferici molto lontani.

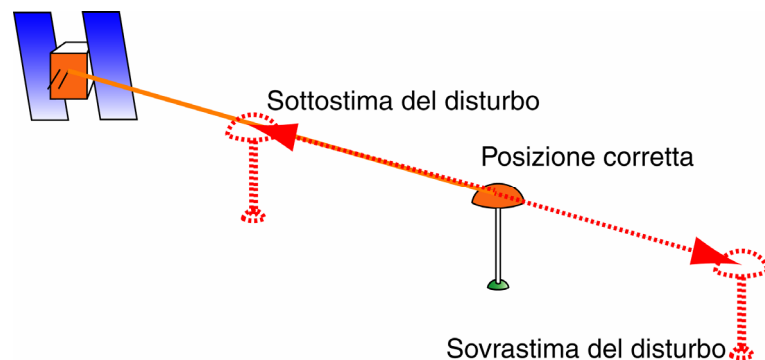


Figura 5.2. Gli effetti di una errata modellizzazione del disturbo atmosferico.

4. Il multipath è un fenomeno di interferenza del segnale con repliche dello stesso provenienti da direzioni spurie e può avvenire ad esempio per riflessioni dovute alla presenza di cartelloni metallici o pareti vetrate in prossimità dell'antenna: dipende dalle condizioni ambientali del punto di misura e dalla letteratura risulta che nelle peggiori delle ipotesi può causare errori di osservazione dell'ordine del metro per il codice, di alcuni centimetri per le fasi; in Fig. 5.3 è riportato un esempio di stima di multipath per la stazione permanente di Como.

5. Il rumore elettronico di misura dipende dal tipo di ricevitori e osservabili, ed è già stato discusso nei Parr. 4.4 e 4.9 per le varie osservabili.

Tipo errore	Modelli trasmessi	Modelli a posteriori	Note
Effemeridi	$\cong 1 \text{ m}$	$\cong 5 \text{ cm}$	
Offset d'orologio	$\cong 1 \text{ m}$	$\cong 5 \text{ cm}$	
Modello ionosferico	$\cong 0-20 \text{ m}$	$\cong 0-10 \text{ m}$	(2)(3)
Modello troposferico	$\leq 1 \text{ m}$		(1)
Multipath	$\cong 1 \text{ m}$		
Rumore C/A o L2C	$0.3-3.0 \text{ m}$		(4)
Rumore P(Y)	$0.1-0.3 \text{ m}$		(4)
Rumore PF	0.01 m		(5)

Tabella 5.1. Ordini di grandezza degli errori residui sul posizionamento di codice. Note: (1) simile per tutti i satelliti in vista; (2) molto variabile in funzione dello stato ionosferico; (3) i modelli di ionosfera calcolati a posteriori comunque non riescono a modellizzare le variabilità locali; (4) molto variabile in funzione della vetustà elettronica del ricevitore; (5): ci si riferisce all'osservabile costruita con il filtro di Hatch, come descritto al Par. 4.9.

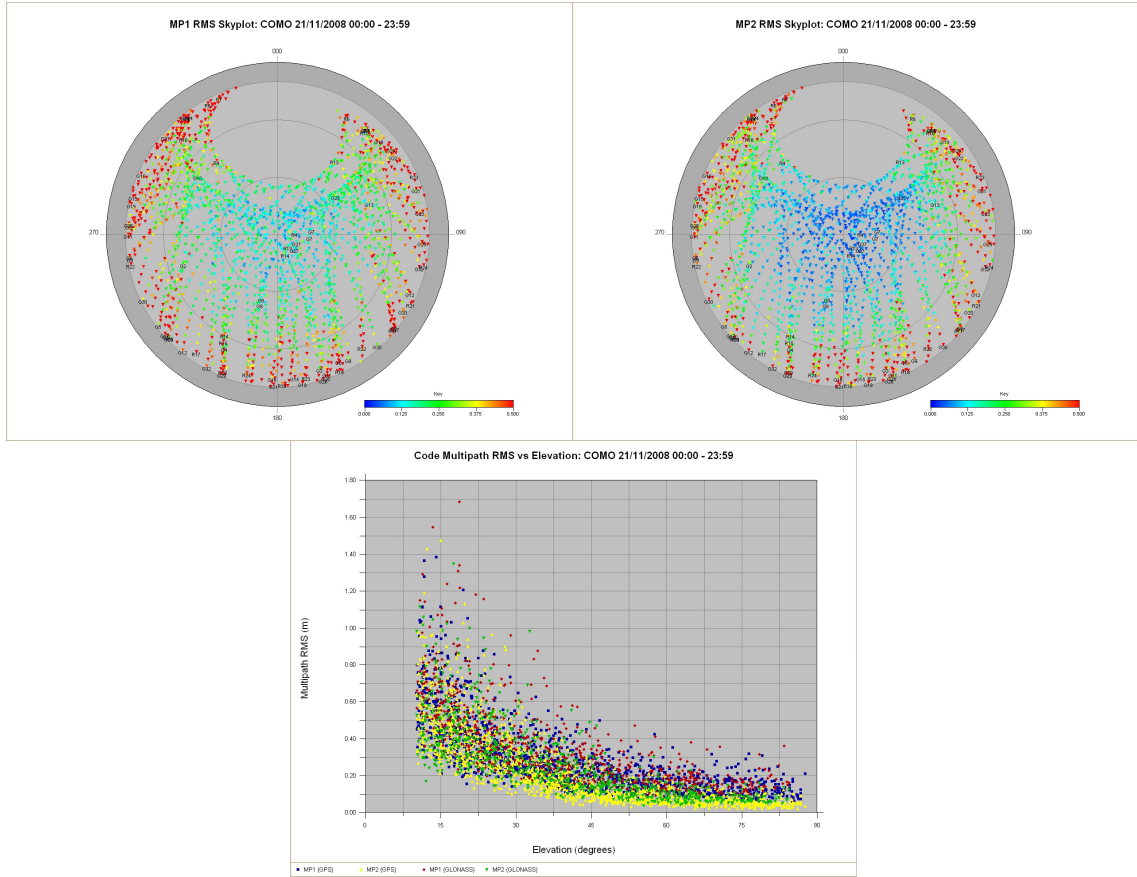


Figura 5.3. Effetto del multipath sui codici della stazione permanente di Como, per la giornata del 21/11/2008. Sopra a sinistra: multipath su P1, in funzione di elevazione e direzione azimutale, plottato sullo skyplot; Sopra a destra: multipath su P2, in funzione di elevazione e direzione azimutale, plottato sullo skyplot; sotto: multipath per tutti i segnali, in funzione dell'elevazione del satellite.

6. L'accuratezza nella stima della posizione di un ricevitore dipende dagli errori presenti nelle singole osservazioni, appena discussi, ma anche dalla geometria dei satelliti rispetto al ricevitore; infatti si ricorda che nella stima finale mediante minimi quadrati si ha

$$\mathbf{C}_{\hat{\mathbf{x}}} = \hat{\sigma}_0^2 (\mathbf{A}^T \mathbf{Q}^{-1} \mathbf{A})^{-1}$$

ove σ è la varianza delle osservazioni e $\mathbf{Q}$ è la loro matrice dei cofattori: in particolare nel caso del posizionamento assoluto si pone $\mathbf{Q} = \mathbf{I}$; si ha dunque

$$\mathbf{C}_{\hat{\mathbf{x}}} = \hat{\sigma}_0^2 (\mathbf{A}^T \mathbf{A})^{-1} = \hat{\sigma}_0^2 \mathbf{Q}_{xx}$$

$\mathbf{Q}_{xx}$ dipende dalla geometria dei satelliti, ed è della forma

$$\mathbf{Q}_{xx} = \begin{bmatrix} q_{XX} & q_{XY} & q_{XZ} & q_{Xt} \\ q_{XY} & q_{YY} & q_{YZ} & q_{Yt} \\ q_{XZ} & q_{YZ} & q_{ZZ} & q_{Zt} \\ q_{Xt} & q_{Yt} & q_{Zt} & q_{tt} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{XYZ} & \mathbf{q}_t \\ \mathbf{q}_t^T & q_{tt} \end{bmatrix}$$

Per esprimere le precisioni relative in un sistema di coordinate intuitive si ruota la sottomatrice dei cofattori della geometria dal sistema cartesiano geocentrico alle coordinate locali, in modo da separare la planimetria dalla quota (2.14):

$$\mathbf{Q}_{\lambda\phi h} = \begin{bmatrix} q_{\lambda\lambda} & q_{\lambda\phi} & q_{\lambda h} \\ q_{\lambda\phi} & q_{\phi\phi} & q_{\phi h} \\ q_{\lambda h} & q_{\phi h} & q_{hh} \end{bmatrix} = \mathbf{R} \mathbf{Q}_{XYZ} \mathbf{R}^T ;$$

si definiscono quindi gli indicatori Dilution of Precision (DOP), che vengono derivati dagli elementi diagonali della matrice $\mathbf{Q}_{XYZ}$ ruotata, e forniscono un indice geometrico dell'accuratezza ottenibile nella stima delle coordinate. In particolare si hanno:

1. HDOP, Horizontal DOP: indice geometrico di accuratezza per il posizionamento planimetrico; somma dei due elementi diagonali $q_{\phi\phi} + q_{\lambda\lambda}$, ovvero relativi alle coordinate planimetriche;
2. VDOP, Vertical DOP: indice geometrico di accuratezza per la stima della quota; elemento diagonale q_{hh} , relativo alla quota;
3. PDOP, Position DOP: somma di HDOP e VDOP;
4. TDOP, Time DOP: indice di accuratezza per la stima dell'offset d'orologio del satellite, riportato in termini metrici; elemento diagonale q_{tt} relativo alla stima dell'offset d'orologio del ricevitore;
5. GDOP, Global DOP: indice globale di accuratezza; somma di PDOP e TDOP, ovvero traccia completa di $\mathbf{Q}_{xx}$.

Con una buona copertura di satelliti, ovvero basso PDOP (Figg. 5.4, 5.5), in tempo reale in singola frequenza si può ottenere un'accuratezza dell'ordine di 10 m; in postelaborazione si ottengono accuratèzze lievemente migliori per la possibilità di filtrare i risultati. I ricevitori che possono effettuare letture di codice su entrambe le frequenze costruiscono la combinazione di codici Ionospheric free e nel bilancio d'errore non compare più il disturbo ionosferico, che attualmente è il fattore limitante del posizionamento: in questo caso in tempo reale si possono avere accuratèzze sino a

circa 1-2 metri. Infine l'elaborazione delle osservazioni di codice lisciate con la fase permette l'attenuazione del rumore elettronico: in questo caso si può arrivare a accuratze di singola epoca migliori del metro. Una sintesi dei risultati è riportata in Tab. 5.2. Da ultimo si noti che in linea di massima si ha una miglior determinazione planimetrica rispetto a quella altimetrica, poiché gli errori nelle modellizzazioni atmosferiche influiscono di più sulla stima della quota.

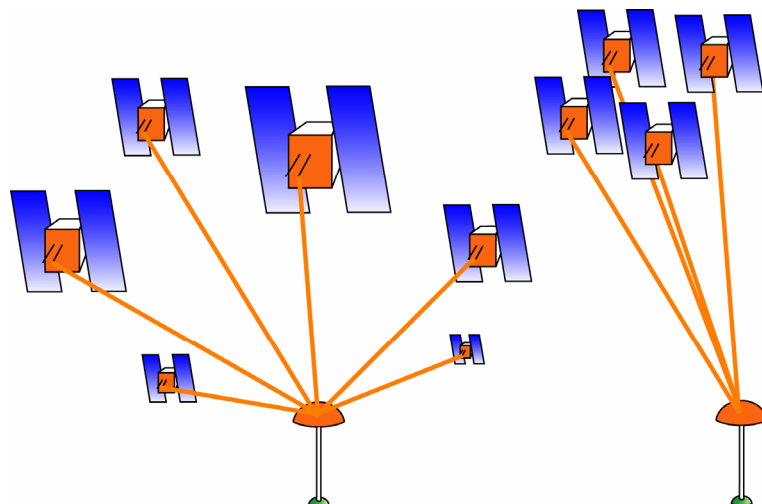


Figura 5.4. Esempi di buona (a sinistra) e cattiva (a destra) configurazioni satellitari

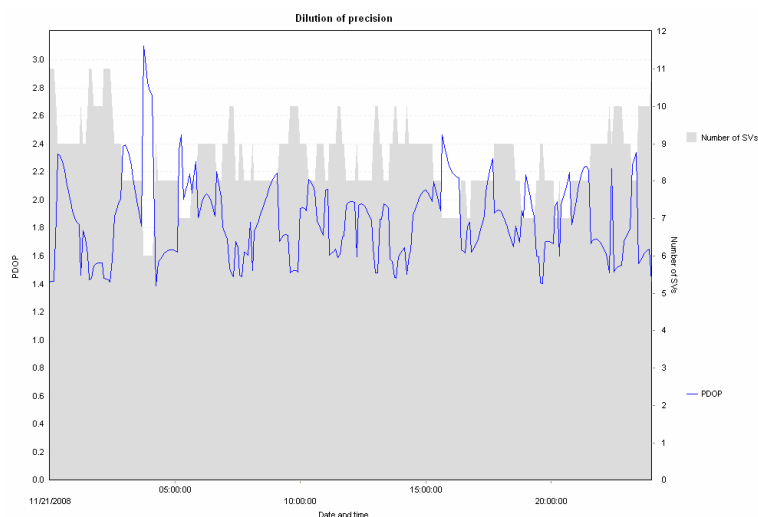


Figura 5.5. Esempio di andamento di PDOP per Como: giornata del 21/11/2008

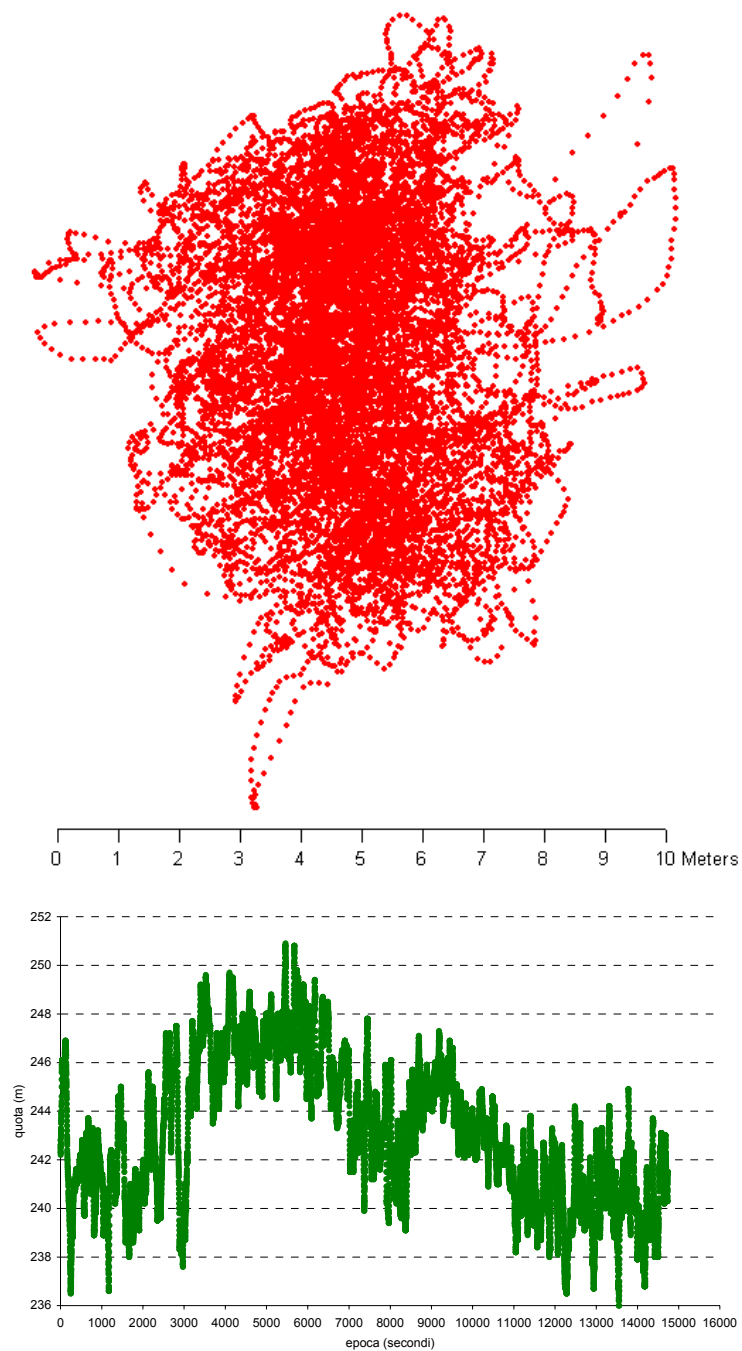


Figura 5.6. Esempio di risultati di posizionamento in tempo reale: in alto la dispersione planimetrica delle coordinate Est, Nord; in basso il comportamento della stima della quota nel tempo. Ricevitore eBonTek, 4 ore di rilievo, 1 osservazione al secondo.

5.4 La soluzione statica mediante osservazioni di codice

Con l'eccezione del multipath, che ha andamento periodico con periodo di circa 20 minuti, gli errori di modello prima discussi variano lentamente nel tempo: perciò, per mediarli sarebbero necessarie sessioni statiche molto lunghe; nella pratica, raramente si effettua il posizionamento statico assoluto mediante osservazioni di codice poiché, per migliorare le accuratezze è conveniente passare all'approccio del posizionamento relativo; per completezza ne viene comunque presentato il formalismo.

Si sia effettuato uno stazionamento statico per T epoche sul medesimo punto: per semplificare l'esempio si ipotizza che gli I satelliti osservati siano comuni a tutte le epoche. Siano $\mathbf{P}_0(t_k)$, $\mathbf{b}(t_k)$ e $\mathbf{v}(t_k)$ rispettivamente i vettori contenenti le osservazioni di codice, il termine noto e gli errori per una singola epoca; siano $\mathbf{E}_R(t_k)$ le matrici di geometria per ogni epoca, in accordo a quanto esposto nella (5.4): si può scrivere il sistema

$$\mathbf{y}_0 = \mathbf{P}_0 - \mathbf{b} = \mathbf{A}\boldsymbol{\xi} + \mathbf{v} \quad (5.7)$$

$$\mathbf{P}_0 = \begin{bmatrix} \mathbf{P}_0(t_1) \\ \mathbf{P}_0(t_2) \\ \dots \\ \mathbf{P}_0(t_T) \end{bmatrix}, \quad \mathbf{A}_{T \times (3+T)} = \begin{bmatrix} \mathbf{E}_R(t_1) & \mathbf{i} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{E}_R(t_2) & \mathbf{0} & \mathbf{i} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{E}_R(t_T) & \mathbf{0} & \mathbf{0} & \dots & \mathbf{i} \end{bmatrix}, \quad \boldsymbol{\xi} = \begin{bmatrix} \delta \mathbf{x}_R \\ \theta_R(t_1) \\ \theta_R(t_2) \\ \dots \\ \theta_R(t_T) \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}(t_1) \\ \mathbf{b}(t_2) \\ \dots \\ \mathbf{b}(t_T) \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}(t_1) \\ \mathbf{v}(t_2) \\ \dots \\ \mathbf{v}(t_T) \end{bmatrix}$$

che può essere risolto rispetto alle coordinate e agli offset d'orologio del ricevitore. Si noti che nel posizionamento statico non è necessario che siano in vista almeno quattro satelliti per ogni epoca; è però necessario che il numero di osservazioni ($[T]$) superi il numero di incognite ($[3+T]$). Ragionevolmente, con un'ora di stazionamento, si possono ottenere accuratezze migliori del metro.

5.5 Il posizionamento assoluto mediante osservazioni di fase

Con ricevitori di codice e fase è possibile combinare le osservabili: la soluzione adottata per da quasi tutti i ricevitori è quella del filtraggio dei codici mediante le fasi, in accordo al filtro di Hatch descritto nel Par. 4.9 e richiamato nel Par. 5.3. Si vuole però ora descrivere il posizionamento assoluto mediante le sole osservazioni di fase, che potenzialmente garantisce le massime accuratezze. Ricordiamo l'equazione di osservazione di fase, per una delle due portanti o per qualunque loro combinazione:

$$L_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) - I_R^S(t) + T_R^S(t) + \lambda(N_R^S(t) + \phi_R - \phi^S)$$

accorpiamo l'ambiguità intera e le fasi frazionarie come nella (4.19); si ha quindi

$$L_R^S(t) = \rho_R^S(t) + c(dt_R(t) - dt^S(t)) - I_R^S(t) + T_R^S(t) + \eta_R^S(t)$$

Rispetto all'equazione di codice compare un termine di ambiguità incognita per ogni satellite: quindi è necessario aumentare il numero di epoche di osservazione per risolvere il sistema. Nel seguito si adotteranno le ipotesi semplificate di costanza nel tempo della costellazione dei satelliti visibili e di assenza di cycle slip: per ogni satellite si può dunque scrivere

$$\eta_R^S(t) = \eta_R^S(t_0) = \eta_R^S$$

si pone inoltre

$$b_R^S(t) = \tilde{\rho}_R^S(t) - c\tilde{dt}^S(t) - \tilde{I}_R^S(t) + \tilde{T}_R^S(t)$$

L'equazione di osservazione diviene:

$$L_{RO}^S(t) = b_R^S(t) + \tilde{\mathbf{e}}_R^S(t) \cdot \delta \mathbf{x}_R + \theta_R(t) + \eta_R^S + \delta(\mathbf{x}^S, dt^S, I_R^S, T_R^S) + \nu_R^S(t)$$

Per una singola epoca in singola frequenza con I satelliti si scrive

$$\mathbf{L}_0(t) = \begin{bmatrix} \mathbf{E}_R(t) & \mathbf{i} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \delta \mathbf{x}_R(t) \\ \theta_R(t) \\ \boldsymbol{\eta} \end{bmatrix} + \mathbf{b}(t) + \mathbf{v}(t), \quad (5.8)$$

ove $\mathbf{I}$ è la matrice identità $[I \times I]$, $\mathbf{E}_R$, $\mathbf{b}$, $\mathbf{i}$ e $\mathbf{v}$ hanno il medesimo significato discusso nella (5.4), ricordando però che in questo caso la ionosfera ha segno negativo; inoltre compare il termine

$$\boldsymbol{\eta}_{I \times 1} = \begin{bmatrix} \eta_R^1 \\ \eta_R^2 \\ \dots \\ \eta_R^I \end{bmatrix}$$

Per più epoche si può scrivere quindi

$$\mathbf{y}_0 = \mathbf{L}_0 - \mathbf{b} = \begin{bmatrix} \mathbf{L}_0(t_1) \\ \mathbf{L}_0(t_2) \\ \dots \\ \mathbf{L}_0(t_T) \end{bmatrix} = \mathbf{A}\boldsymbol{\xi} + \mathbf{v} \quad (5.9)$$

ove

$$\mathbf{A}_{Tl \times (3+T+I)} = \begin{bmatrix} \mathbf{E}_R(t_1) & \mathbf{i} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} \\ \mathbf{E}_R(t_2) & \mathbf{0} & \mathbf{i} & \dots & \mathbf{0} & \mathbf{I} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{E}_R(t_T) & \mathbf{0} & \mathbf{0} & \dots & \mathbf{i} & \mathbf{I} \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \delta \mathbf{x}_R \\ \theta_R(t_1) \\ \theta_R(t_2) \\ \dots \\ \theta_R(t_T) \\ \boldsymbol{\eta} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} \mathbf{b}(t_1) \\ \mathbf{b}(t_2) \\ \dots \\ \mathbf{b}(t_T) \end{bmatrix}, \mathbf{v} = \begin{bmatrix} \mathbf{v}(t_1) \\ \mathbf{v}(t_2) \\ \dots \\ \mathbf{v}(t_T) \end{bmatrix}$$

Si noti che con cinque satelliti in vista bastano due epoche per avere ridondanza poiché si hanno dieci osservazioni e dieci incognite, ovvero tre coordinate, due offset d'orologio del ricevitore e cinque ambiguità iniziali: in effetti il sistema non è comunque banalmente risolvibile, per i motivi che discuteremo al termine del paragrafo.

In ogni caso il posizionamento assoluto mediante fasi in singola frequenza non risulta conveniente poiché richiede algoritmi molto più complessi rispetto all'elaborazione dei codici e non abbatta la principale fonte di errore, ovvero quello ionosferico (Tab. 5.1); disponendo invece di osservazioni su entrambe le frequenze si può procedere in due vie. Innanzitutto si può scrivere un sistema che incorpori i disturbi ionosferici nel vettore dei parametri incogniti anziché nel termine noto, ovvero per ogni epoca si pone

$$\mathbf{L}_0(t) = \begin{bmatrix} \mathbf{L}_{10}(t) \\ \mathbf{L}_{20}(t) \end{bmatrix}_{2I} = \begin{bmatrix} \mathbf{E}_R(t)_{I \times 3} & \mathbf{i}_I & \mathbf{I}_{I \times I} & \mathbf{I}_{I \times I} & \mathbf{0} \\ \mathbf{E}_R(t)_{I \times 3} & \mathbf{i}_I & k_2 \mathbf{I}_{I \times I} & \mathbf{0} & \mathbf{I}_{I \times I} \end{bmatrix} \begin{bmatrix} \delta \mathbf{x}_R(t) \\ \theta_R(t) \\ \mathbf{u}_R(t) \\ \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{b}_1(t) \\ \mathbf{b}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{bmatrix} \quad (5.10)$$

ove gli indici 1 e 2 caratterizzano la prima e la seconda frequenza,

$$k_2 = \frac{f_1^2}{f_2^2}, \mathbf{u}_R(t) = \begin{bmatrix} I\mathbf{1}_R^1(t) \\ I\mathbf{1}_R^2(t) \\ \dots \\ I\mathbf{1}_R^I(t) \end{bmatrix}, I\mathbf{1}_R^i(t) = \frac{A \cdot TEC_R^i(t)}{f_1^2}$$

Per più epoche si costruisce un sistema del tipo

$$\mathbf{y}_0 = \mathbf{L}_0 - \mathbf{b} = \begin{bmatrix} \mathbf{L}_o(t_1) \\ \mathbf{L}_o(t_2) \\ \dots \\ \mathbf{L}_o(t_T) \end{bmatrix} = \mathbf{A}\boldsymbol{\xi} + \mathbf{v} \quad (5.11)$$

$2TI$

ove in questo caso il vettore dei parametri incogniti è dato da

$$\boldsymbol{\xi}_{3+T+IT+2I} = \begin{bmatrix} \delta \mathbf{x}_R \\ \boldsymbol{\theta}_R \\ \mathbf{v}_R \\ \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{bmatrix}, \boldsymbol{\theta}_R = \begin{bmatrix} \theta_R(t_1) \\ \theta_R(t_2) \\ \dots \\ \theta_R(t_T) \end{bmatrix}, \mathbf{v} = \begin{bmatrix} \mathbf{v}_1(t_1) \\ \mathbf{v}_2(t_1) \\ \mathbf{v}_1(t_2) \\ \dots \\ \mathbf{v}_2(t_T) \end{bmatrix}, \mathbf{v}_R = \begin{bmatrix} \mathbf{v}_R(t_1) \\ \mathbf{v}_R(t_2) \\ \dots \\ \mathbf{v}_R(t_T) \end{bmatrix}$$

IT

Il sistema, di cui per sintesi non si esplicita la matrice disegno, prevede la stima di $3+T(1+I)+2I$ incognite a partire da $2TI$ osservazioni: con cinque satelliti e quattro epoche si hanno 40 osservazioni in 37 incognite; con un minuto di osservazione all'intervallo di un secondo, si hanno già $60 \times 10 = 600$ osservazioni, e $3+60+300+10=373$ incognite. La ridondanza è inferiore al caso (5.9) ma con questo formalismo non è più necessario modellizzare il disturbo ionosferico. In alternativa è possibile costruire la combinazione Ionospheric Free e poi risolvere il relativo sistema, ovvero identico a quello della (5.9) senza però il modello ionosferico nel termine noto. In effetti i sistemi (5.9) o (5.10) non sono risolvibili semplicemente quando si hanno più osservazioni che incognite, perché si possono dimostrare le seguenti proprietà, derivate dalla (4.32):

1. con poche epoche ravvicinate di osservazione, la geometria dei satelliti rispetto all'osservatore non cambia;
2. in tale condizione le incognite geometriche non sono separabili dalle incognite di ambiguità e i sistemi presentano deficienza di rango anche in condizioni di ridondanza numerica;
3. si ha un significativo cambio di configurazione satellitare fra la prima e l'ultima epoca di osservazione solo dopo un'ora circa di osservazione;
4. in tale condizione i sistemi da risolvere sono finalmente ben condizionati numericamente ed è quindi possibile stimare posizione e ambiguità (e ionosfera, nel caso della 5.11).

Il rumore delle osservazioni di fase è dell'ordine del millimetro e con la doppia frequenza si elimina l'errore di modellizzazione ionosferica. Quindi un posizionamento assoluto statico mediante elaborazione delle fasi può fornire risultati molto più accurati rispetto a quello basato sui codici: sperimentalmente, con stazionamenti dell'ordine di una giornata, postelaborazione delle osservazioni e adozione delle effemeridi precise di

IGS si sono ottenute accuratissime di alcuni centimetri. Però tali risultati sono ancora relegati ad un ambito di ricerca universitaria e frutto di complesse elaborazioni: per le applicazioni di alta accuratezza le tecniche di posizionamento relativo risultano oggi molto più convenienti, perché richiedono tempi di occupazione inferiori e algoritmi di elaborazione più semplici.

Si riporta a conclusione della discussione una tabella riassuntiva delle accuratissime ottenibili nel posizionamento assoluto. Si nota solo che, al fine di migliorarne le stime, il GPS può anche essere integrato da altri sensori o informazioni esterne: ad esempio in navigazione si possono utilizzare sensori inerziali per filtrare il rumore dalle stime di traiettoria. Oppure, nel caso di navigazione di automezzi, possono essere utilizzate informazioni cartografiche ausiliarie che vincolino la stima della quota a un modello digitale della superficie o la posizione orizzontale alle linee definite da un grafo stradale. Per i motivi descritti nell'introduzione, tali aspetti non vengono qui discussi.

Tipo di posizionamento	Osservabili	Requisiti e note	Ordine di accuratezza
Cinematico in tempo reale	Codici in singola frequenza	Ricevitori a basso costo	dieci metri
	Codici in doppia frequenza	Attualmente solo ricevitori abilitati ai codici militari, di elevato costo	circa un metro
	Codici in doppia frequenza lisciati con le fasi (filtro di Hatch)		meglio del metro
Statico in postelaborazione	Fasi in doppia frequenza	Ricevitori geodetici, metodi ancora oggetto di sperimentazione universitaria	alcuni centimetri

Tabella 5.2. Riepilogo delle accuratissime ottenibili nelle varie modalità del posizionamento assoluto.

5.6 Il sistema di riferimento in cui viene fornita la stima

Il sistema di riferimento in cui si effettua il posizionamento assoluto è lo stesso in cui si sono date le coordinate dei satelliti, ovvero quello delle effemeridi; se si sono utilizzate le effemeridi trasmesse, le stime sono in WGS84; se le elaborazioni dei dati sono state svolte a posteriori con le effemeridi a posteriori di IGS, le stime sono in IGS05. Peraltro le differenze di sistema di riferimento sono dell'ordine del decimetro: nel contesto del posizionamento assoluto mediante codici sono trascurabili perché sono inferiori alle

accuratezze di stima del posizionamento assoluto; nel caso di posizionamento postelaborato mediante fasi si usano tipicamente le effemeridi IGS.

5.7 Osservazioni sulla linearizzazione della distanza fra satellite e ricevitore

L'elaborazione delle osservazioni GPS prevede la linearizzazione delle equazioni di osservazione rispetto alle coordinate del ricevitore; l'espressione sinora utilizzata è adatta e corretta per una discussione generale, ma non considera alcuni problemi che invece vanno implementati in un programma di elaborazione: in questo paragrafo si vuole per completezza tecnica riportare lo sviluppo numerico effettivamente da implementarsi. Si parte dalla notazione generale

$$\rho_R^S(t) = \tilde{\rho}_R^S(t) + \tilde{\mathbf{e}}_R^S(t) \cdot \delta \mathbf{x}_R(t)$$

Le coordinate del satellite vengono calcolate mediante le effemeridi, per le quali si usa il simbolo $\mathbf{x}_{EPH}^S(t)$. Si devono considerare due aspetti: si consideri che

$$\rho_R^S(t) = \|\mathbf{x}_R(t) - \mathbf{x}^S(t - \tau)\|$$

ovvero la distanza osservata è quella fra il ricevitore all'epoca t e il satellite all'epoca $t - \tau$: durante il tempo di tragitto ($\tau \approx 66ms$), il satellite si è mosso di circa $1000m/s \cdot 0.066s = 66m$; inoltre, durante il tragitto del segnale il pianeta compie una rotazione pari a $\tau \dot{\Omega}_E$, ove $\dot{\Omega}_E$ rappresenta la velocità angolare di rotazione terrestre: infine si ricorda che la vera epoca di ricezione t è incognita poiché si conosce solo l'epoca t_R registrata dall'orologio del ricevitore, e fra le due vale la $t = t_R + dt_R$; nel calcolo si ha dunque

$$\mathbf{x}^S(t - \tau_R^S) = \mathbf{x}^S(t_R + dt_R(t) - \tau_R^S) = \mathbf{x}_{EPH}^S(t_R) + \dot{\mathbf{x}}_{EPH}^S(t_R) \cdot (dt_R(t) - \tau_R^S) - \tau_R^S [\boldsymbol{\omega} \times] \mathbf{x}_{EPH}^S(t_R)$$

ove $\mathbf{x}_{EPH}^S$ indica le coordinate del satellite calcolate mediante le effemeridi, indifferente se predette o precise, $\boldsymbol{\omega} = \begin{bmatrix} 0 & 0 & \dot{\Omega}_E \end{bmatrix}^T$ è il vettore di velocità angolare terrestre.

Nella soluzione di Bancroft si conosce solo t_R , ovvero l'epoca di ricezione; si pone $dt_R \approx 0$ mentre per il tempo di tragitto si utilizza il classico valore approssimato $\tilde{\tau}_R^S \approx 20000km/c = 66ms$: le approssimazioni indotte da tali scelte influiscono significativamente sul bilancio d'errore della stima e, se si desidera una maggiore accuratezza, si deve ricorrere a un processo di stima iterativa.

Nella soluzione ai minimi quadrati si suppone di disporre di valori approssimati per le quantità incognite. Partendo dalla

$$\begin{aligned}\rho_R^S(t) &= \left\| \mathbf{x}_R(t_R + dt_R(t)) - \mathbf{x}^S(t_R + dt_R(t) - \tau_R^S) \right\| \\ &= \left\| \tilde{\mathbf{x}}_R(t_R + dt_R(t)) + \delta \mathbf{x}_R(t_R + dt_R(t)) - \mathbf{x}_{EPH}^S(t_R) - \dot{\mathbf{x}}_{EPH}^S(t_R) \cdot (dt_R(t) - \tau_R^S) + \tau_R^S [\boldsymbol{\omega} \times] \mathbf{x}_{EPH}^S(t_R) \right\|\end{aligned}$$

ricordando che

$$\tau_R^S = \frac{\rho_R^S(t)}{c}$$

si deve quindi linearizzare la distanza rispetto a τ_R^S e dt_R , oltre che rispetto alla correzione alle coordinate del ricevitore; trascurando i termini di entità non significativa (ovvero inferiori a $5 \times 10^{-4} m$) l'equazione finale risulta

$$\rho_R^S(t) = k_{1R}^S(t_R) \tilde{\rho}_R^S(t_R) + \tilde{\mathbf{e}}_R^S(t_R) \cdot \delta \mathbf{x}_R(t_R) + k_{2R}^S(t_R) dt_R(t_R) \quad (5.12)$$

ove

$$\begin{aligned}k_{1R}^S(t_R) &= 1 - \alpha_{1R}^S(t_R) / c, \\ k_{2R}^S(t_R) &= -\alpha_{2R}^S(t_R) + \alpha_{1R}^S(t_R) \alpha_{2R}^S(t_R) / c, \\ \alpha_{1R}^S(t_R) &= \tilde{\mathbf{e}}_R^S(t_R) \cdot (\dot{\mathbf{x}}_{EPH}^S(t_R) + [\boldsymbol{\omega} \times] \mathbf{x}_{EPH}^S(t_R)); \\ \alpha_{2R}^S(t_R) &= \tilde{\mathbf{e}}_R^S(t_R) \cdot \dot{\mathbf{x}}_{EPH}^S(t_R) \\ \tilde{\rho}_R^S(t) &= \left\| \tilde{\mathbf{x}}_R(t_R) - \mathbf{x}_{EPH}^S(t_R) \right\| \\ \tilde{\mathbf{e}}_R^S(t_R) &= (\mathbf{x}_{EPH}^S(t_R) - \tilde{\mathbf{x}}_R(t_R)) / \tilde{\rho}_R^S(t_R)\end{aligned}$$

Ciò comporta due principali necessità nell'implementazione numerica della stima rispetto alla scrittura generale: il termine noto geometrico viene moltiplicato per $k_{1R}^S(t_R)$ e l'incognita $dt_R(t_R)$ per $c + k_{2R}^S(t_R)$: ne segue che il coefficiente di $dt_R(t_R)$ dipende dal satellite.

Capitolo 6.

Il posizionamento relativo

La dizione posizionamento relativo indica la stima di posizioni di punti incogniti rispetto a un punto di coordinate note, detto anche di riferimento. Nell'ambito del GPS, il posizionamento relativo viene effettuato utilizzando particolari combinazioni di osservazioni, dette differenze doppie, che coinvolgono due ricevitori: il primo di questi è detto ricevitore indietro e occupa il punto di riferimento, il secondo è detto ricevitore avanti e occupa il punto incognito; il vettore tridimensionale che congiunge i due ricevitori è detto base; nel seguito si supporrà sempre che il ricevitore di riferimento sia fermo, mediante stazionamento su caposaldo o perché monumentato stabilmente. Il ricevitore incognito può a sua volta effettuare uno stazionamento statico ma anche un rilievo cinematico, battendo consecutivamente più punti.

La costruzione delle differenze doppie attraverso quelle singole e la discussione del loro contenuto introducono il capitolo, con particolare attenzione agli svantaggi e ai vantaggi che esse presentano rispetto alle osservazioni non differenziate; segue una parte più tecnica, necessaria però per chiarire alcuni aspetti su come si possono costruire le differenze doppie da un insieme di osservazioni di partenza e su come si propaghi la covarianza. La parte centrale del capitolo è dedicata all'elaborazione delle differenze doppie; si noti che, a seconda delle applicazioni e del tipo di ricevitori disponibili, sono possibili e sono stati implementati negli ultimi decenni approcci di elaborazione completamente differenti: fra i molti casi possibili, si è scelto di discutere nel dettaglio due esempi all'estremo opposto. Il primo è quello delle doppie differenze di codice in singola epoca. Il secondo è quello di un'elaborazione statica mediante ricevitore di codice e fase a doppia frequenza, come viene in genere condotta dai programmi di elaborazione per basi sino alla lunghezza di circa 50 km; per i motivi già detti nell'introduzione, non si fornisce una discussione dell'elaborazione di reti con basi di lunghezza superiore. Il capitolo si chiude con la descrizione molto sintetica dei passi di elaborazione per i rilievi di tipo cinematico e per il cosiddetto tempo reale; sono inoltre fornite alcune tabelle sulle accuratezze tipicamente ottenibili.

6.1 Le differenze singole

Consideriamo due ricevitori R_1 e R_2 che abbiano effettuato osservazioni ad un satellite S alla medesima epoca t : la differenza singola è la differenza fra le osservazioni del ricevitore avanti e di quello indietro. Sui codici si ottiene

$$\begin{aligned}
P_{R1,R2}^S(t) &= P_{R1}^S(t) - P_{R2}^S(t) \\
&= \rho_{R1}^S(t) - \rho_{R2}^S(t) + c(dt_{R1}(t) - dt_{R2}(t)) - c(dt^S(t) - dt^S(t)) \\
&\quad + T_{R1}^S(t) - T_{R2}^S(t) + I_{R1}^S(t) - I_{R2}^S(t) \\
&= \rho_{R1}^S(t) - \rho_{R2}^S(t) + T_{R1}^S(t) - T_{R2}^S(t) + c(dt_{R1}(t) - dt_{R2}(t)) + I_{R1}^S(t) - I_{R2}^S(t)
\end{aligned} \tag{6.1}$$

Sulle fasi

$$\begin{aligned}
L_{R1,R2}^S(t) &= L_{R1}^S(t) - L_{R2}^S(t) \\
&= \rho_{R1}^S(t) - \rho_{R2}^S(t) + c(dt_{R1}(t) - dt_{R2}(t)) - c(dt^S(t) - dt^S(t)) \\
&\quad + T_{R1}^S(t) - T_{R2}^S(t) - I_{R1}^S(t) + I_{R2}^S(t) + \lambda(N_{R1}^S(t) - N_{R2}^S(t) + \phi_{R1} - \phi_{R2} - \phi^S + \phi^S) \tag{6.2} \\
&= \rho_{R1}^S(t) - \rho_{R2}^S(t) + c(dt_{R1}(t) - dt_{R2}(t)) + T_{R1}^S(t) - T_{R2}^S(t) - I_{R1}^S(t) + I_{R2}^S(t) \\
&\quad + \lambda(N_{R1}^S(t) - N_{R2}^S(t) + \phi_{R1} - \phi_{R2})
\end{aligned}$$

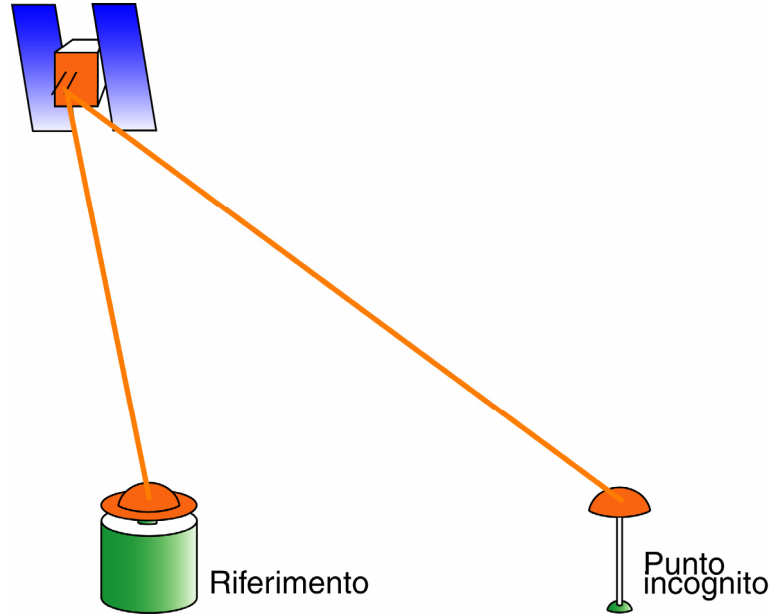


Figura 6.1. Le differenze singole.

La differenza singola costituisce una nuova osservazione costruita da due originali, ove la perdita di numerosità è compensata da alcuni vantaggi: nella differenza singola non compare più l'errore d'orologio del satellite, che contribuiva al bilancio di errore; inoltre la differenziazione riduce sensibilmente gli effetti del disturbo troposferico e degli errori delle effemeridi e attenua anche l'errore ionosferico (Par. 6.4); infine, nel caso delle fasi, scompare la fase frazionaria iniziale del satellite.

6.2 Le differenze doppie (DD)

Le differenze doppie sono la differenza di due differenze singole contemporanee riferite alla medesima coppia di ricevitori e a due satelliti distinti S_1 e S_2 . Sui codici si ottiene

$$\begin{aligned}
 P_{R1,R2}^{S1,S2}(t) &= P_{R1,R2}^{S1}(t) - P_{R1,R2}^{S2}(t) \\
 &= \rho_{R1}^{S1}(t) - \rho_{R2}^{S1}(t) - \rho_{R1}^{S2}(t) + \rho_{R2}^{S2}(t) + c(dt_{R1}(t) - dt_{R2}(t)) - c(dt_{R1}(t) - dt_{R2}(t)) \\
 &\quad + T_{R1}^{S1}(t) - T_{R2}^{S1}(t) - T_{R1}^{S2}(t) + T_{R2}^{S2}(t) + I_{R1}^{S1}(t) - I_{R2}^{S1}(t) - I_{R1}^{S2}(t) + I_{R2}^{S2}(t) \\
 &= \rho_{R1}^{S1}(t) - \rho_{R2}^{S1}(t) - \rho_{R1}^{S2}(t) + \rho_{R2}^{S2}(t) \\
 &\quad + T_{R1}^{S1}(t) - T_{R2}^{S1}(t) - T_{R1}^{S2}(t) + T_{R2}^{S2}(t) + I_{R1}^{S1}(t) - I_{R2}^{S1}(t) - I_{R1}^{S2}(t) + I_{R2}^{S2}(t)
 \end{aligned} \tag{6.3}$$

Sulle fasi, con uno sviluppo analogo:

$$\begin{aligned}
 L_{R1,R2}^{S1,S2}(t) &= L_{R1,R2}^{S1}(t) - L_{R1,R2}^{S2}(t) \\
 &= \rho_{R1}^{S1}(t) - \rho_{R2}^{S1}(t) - \rho_{R1}^{S2}(t) + \rho_{R2}^{S2}(t) + c(dt_{R1}(t) - dt_{R2}(t)) - c(dt_{R1}(t) - dt_{R2}(t)) \\
 &\quad + T_{R1}^{S1}(t) - T_{R2}^{S1}(t) - T_{R1}^{S2}(t) + T_{R2}^{S2}(t) - I_{R1}^{S1}(t) + I_{R2}^{S1}(t) + I_{R1}^{S2}(t) - I_{R2}^{S2}(t) \\
 &\quad + \lambda(N_{R1}^{S1}(t) - N_{R2}^{S1}(t) - N_{R1}^{S2}(t) + N_{R2}^{S2}(t) + \phi_{R1} - \phi_{R2} - \phi_{R1} + \phi_{R2}) \\
 &= \rho_{R1}^{S1}(t) - \rho_{R2}^{S1}(t) - \rho_{R1}^{S2}(t) + \rho_{R2}^{S2}(t) \\
 &\quad + T_{R1}^{S1}(t) - T_{R2}^{S1}(t) - T_{R1}^{S2}(t) + T_{R2}^{S2}(t) - I_{R1}^{S1}(t) + I_{R2}^{S1}(t) + I_{R1}^{S2}(t) - I_{R2}^{S2}(t) \\
 &\quad + \lambda(N_{R1}^{S1}(t) - N_{R2}^{S1}(t) - N_{R1}^{S2}(t) + N_{R2}^{S2}(t))
 \end{aligned} \tag{6.4}$$

Nel seguito si adotterà la semplificazione di notazione

$$(\bullet)_{Ri,Rj}^{Sk,Sl} = (\bullet)_{i,j}^{k,l}$$

essendo implicito che gli indici in apice riguardano i satelliti mentre quelli in pedice sono relativi ai ricevitori. L'utilizzo delle differenze singole o doppie non permette il posizionamento assoluto ma solo quello relativo (Par. 6.3); inoltre il rumore elettronico di osservazione si propaga, aumentando e correlandosi, nella differenziazione delle osservazioni (Par. 6.6.1). Per contro il vantaggio offerto dalle differenze doppie è tale che negli ultimi 30 anni esse sono divenute l'osservabile fondamentale per il posizionamento relativo: innanzitutto sono eliminate le incognite di orologio dei ricevitori, ma non è questo il punto principale poiché per tale scopo vengono anche sacrificate osservazioni; il beneficio maggiore è la completa eliminazione delle fasi frazionarie iniziali: infatti ora rimane solo il termine intero di ambiguità, al quale è possibile applicare i cosiddetti algoritmi di stima e fissaggio, che verranno discussi nei Parr. 6.9.4 e 6.9.5; solo il successo del fissaggio delle ambiguità permette di ottenere

accuratezze dell'ordine di alcuni centimetri nell'elaborazione dei rilievi statico rapidi e cinemati, come verranno discussi al termine del capitolo.

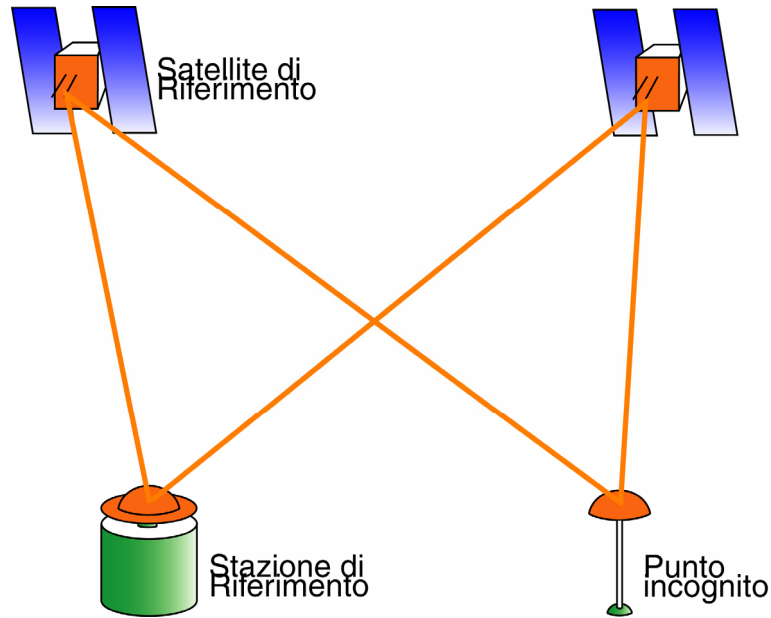


Figura 6.2. Le differenze doppie.

6.3 Il contenuto geometrico delle differenze singole e doppie

Le differenze doppie possono essere usate solo nel posizionamento relativo, ovvero nella stima della base fra due ricevitori, e non nella stima simultanea delle coordinate di entrambi. Per capire tale aspetto, si consideri il contenuto geometrico della generica osservazione da un ricevitore a un satellite, $\rho_R^S(t)$, e lo si linearizzi rispetto alle coordinate approssimate del ricevitore.

$$\rho_R^S = \tilde{\rho}_R^S + \tilde{\mathbf{e}}_R^S \cdot \delta \mathbf{x}_R$$

ove, con la consueta notazione, $\tilde{\rho}_R^S$ è la distanza approssimata fra satellite e ricevitore, $\tilde{\mathbf{e}}_R^S$ è il relativo versore approssimato, $\delta \mathbf{x}_R$ è la correzione alla posizione approssimata del ricevitore. Si considerino due equazioni relative al medesimo satellite ma a due ricevitori diversi

$$\begin{aligned}\rho_{R1}^S &= \tilde{\rho}_{R1}^S + \tilde{\mathbf{e}}_{R1}^S \cdot \delta \mathbf{x}_{R1} \\ \rho_{R2}^S &= \tilde{\rho}_{R2}^S + \tilde{\mathbf{e}}_{R2}^S \cdot \delta \mathbf{x}_{R2}\end{aligned}$$

e se ne faccia la differenza singola:

$$\rho_{1,2}^S = \tilde{\rho}_{1,2}^S + \tilde{\mathbf{e}}_1^S \cdot \delta \mathbf{x}_1 - \tilde{\mathbf{e}}_2^S \cdot \delta \mathbf{x}_2$$

La precedente può essere scritta come

$$\rho_{1,2}^S = \tilde{\rho}_{1,2}^S + \tilde{\mathbf{e}}_1^S \cdot \delta \mathbf{x}_1 - \tilde{\mathbf{e}}_2^S \cdot \delta \mathbf{x}_2 + \tilde{\mathbf{e}}_2^S \cdot \delta \mathbf{x}_1 - \tilde{\mathbf{e}}_2^S \cdot \delta \mathbf{x}_1$$

ovvero

$$\rho_{1,2}^S = \tilde{\rho}_{1,2}^S - \tilde{\mathbf{e}}_2^S \cdot (\delta \mathbf{x}_2 - \delta \mathbf{x}_1) + (\tilde{\mathbf{e}}_1^S - \tilde{\mathbf{e}}_2^S) \cdot \delta \mathbf{x}_1$$

Consideriamo ora il caso in cui la distanza fra i due ricevitori sia inferiore a 100 km, che è il tipico caso delle reti locali o regionali. In questo caso le distanze da R_1 e R_2 al satellite (in orbita a 20000 km di quota) sono sicuramente molto simili e si può approssimare

$$\tilde{\mathbf{e}}_1^S - \tilde{\mathbf{e}}_2^S = \frac{\tilde{\mathbf{x}}_1 - \mathbf{x}^S}{\tilde{\rho}_1^S} - \frac{\tilde{\mathbf{x}}_2 - \mathbf{x}^S}{\tilde{\rho}_2^S} \cong \frac{\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_2}{\bar{\rho}^S} = -\frac{\Delta \tilde{\mathbf{x}}_{1,2}}{\bar{\rho}^S}$$

ove si è posto $\bar{\rho}^S = \frac{1}{2}(\tilde{\rho}_1^S + \tilde{\rho}_2^S)$; $\Delta \mathbf{x}_{1,2} = \mathbf{x}_2 - \mathbf{x}_1$ è la base fra i due ricevitori.

L'equazione di osservazione di differenza singola si approssima quindi nel seguente modo:

$$\rho_{1,2}^S \cong \tilde{\rho}_{1,2}^S - \tilde{\mathbf{e}}_2^S \cdot \delta(\mathbf{x}_2 - \mathbf{x}_1) - \frac{\Delta \tilde{\mathbf{x}}_{1,2}}{\bar{\rho}^S} \cdot \delta \mathbf{x}_1$$

ove $\delta(\mathbf{x}_2 - \mathbf{x}_1) = \delta \Delta \mathbf{x}_{1,2}$ è la correzione da apportare alle componenti approssimate della base fra i due ricevitori. Si consideri il vettore che moltiplica $\delta \mathbf{x}_1$: esso ha componenti dell'ordine di $100\text{km}/20000\text{km} \cong 5 \times 10^{-3}$, ovvero molto più piccole del versore che moltiplica $\delta \Delta \mathbf{x}_{1,2}$, che ha componenti dell'ordine di $1/\sqrt{3} = 0.57$. In sostanza un sistema che preveda insieme la stima delle coordinate di entrambi i ricevitori risulta mal condizionato, prossimo alla deficienza di rango; il ragionamento è diverso per basi più lunghe che però non vengono considerate nel presente volume. Per la differenza doppia ai satelliti S_1, S_2 si ha la seguente equazione di osservazione

$$\rho_{1,2}^{1,2} \cong \tilde{\rho}_{1,2}^{1,2} + (\tilde{\mathbf{e}}_2^2 - \tilde{\mathbf{e}}_2^1) \cdot \delta(\mathbf{x}_2 - \mathbf{x}_1) + \Delta \tilde{\mathbf{x}}_{1,2} \left(\frac{1}{\bar{\rho}^2} - \frac{1}{\bar{\rho}^1} \right) \cdot \delta \mathbf{x}_1$$

anch'essa presenta il medesimo problema di malcondizionamento: nell'elaborazione si deve dunque supporre che le coordinate del punto indietro siano già state determinate e si può operare in due modi alternativi.

Il primo modo è quello di linearizzare le equazioni di osservazione rispetto alle coordinate di entrambi i ricevitori, ovvero costruendo un sistema del tipo

$$\mathbf{d} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \end{bmatrix} \begin{bmatrix} \delta \mathbf{x}_1 \\ \delta \mathbf{x}_2 \end{bmatrix}$$

ove $\mathbf{d}$ indica il vettore di differenze doppie di distanza; per rimuovere il malcondizionamento appena discusso si inserisce nel sistema di equazioni anche l'osservazione delle coordinate a priori del punto indietro, ottenendo il sistema finale

$$\mathbf{d}_0 = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \end{bmatrix} \begin{bmatrix} \delta \mathbf{x}_1 \\ \delta \mathbf{x}_2 \end{bmatrix} + \mathbf{v}, \mathbf{C}_{dd} = \sigma_0^2 \mathbf{Q}_{dd}$$

$$0 = \delta \mathbf{x}_1 + \mathbf{v}_1, \mathbf{C}_1 = \sigma_{01}^2 \mathbf{Q}_1$$

ovvero

$$\begin{bmatrix} \mathbf{d}_0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{I} & 0 \end{bmatrix} \begin{bmatrix} \delta \mathbf{x}_1 \\ \delta \mathbf{x}_2 \end{bmatrix} + \mathbf{v}, \mathbf{C} = \sigma_0^2 \begin{bmatrix} \mathbf{Q}_{dd} & 0 \\ 0 & \frac{\sigma_{01}^2}{\sigma_0^2} \mathbf{Q}_1 \end{bmatrix}$$

che risulta ben risolvibile. In tale approccio le coordinate del punto di riferimento sono considerate osservazioni, caratterizzate da una propria varianza.

L'alternativa più semplice è quella di linearizzare le equazioni di osservazione solo rispetto alle correzioni alla base, ovvero si costruisce un sistema del tipo:

$$\mathbf{d}_0 = \mathbf{A}_2 \delta \Delta \mathbf{x}_{1,2} + \mathbf{v}_d, \mathbf{C}_{dd} = \sigma_0^2 \mathbf{Q}_{dd}$$

che non presenta problemi di mal condizionamento; ciò equivale a considerare le coordinate a priori del punto indietro come costanti note senza errore: si vedrà nel Par. 6.4 la propagazione di un errore in tali costanti sulla stima della base.

Da un punto di vista rigoroso il primo approccio è quello corretto, poiché permette di costruire il sistema più completo, con l'attribuzione di varianze corrette a tutte le informazioni. Per contro, nell'elaborazione in singola base il secondo approccio è concettualmente più semplice e fornisce risultati identici al primo, inoltre è quello tipicamente implementato nei programmi commerciali. Poiché in tale volume si discuterà sempre la soluzione di singola base con attenzione alle scelte dei programmi commerciali si adotta il secondo approccio per tutte le discussioni seguenti: le

coordinate del ricevitore indietro sono quindi note a priori e vengono vincolate deterministicamente; si adotta la seguente linearizzazione per la differenza doppia di distanze e si risolve il sistema rispetto alla base fra i due ricevitori:

$$\rho_{1,2}^{1,2} = \tilde{\rho}_{1,2}^{1,2} + (\tilde{\mathbf{e}}_2^2 - \tilde{\mathbf{e}}_2^1) \cdot \delta \Delta \mathbf{x}_{1,2} \quad (6.5)$$

6.4 Discussione dei termini presenti nella differenza doppia

Ci si chiede innanzitutto quale sia l'effetto di un errore nelle coordinate del ricevitore di riferimento e nelle effemeridi dei satelliti: mediante linearizzazione analoga a quella del paragrafo precedente, si trova che un loro errore δ introduce un errore residuo il cui ordine di grandezza è

$$\delta \frac{\Delta \mathbf{x}_{1,2}}{\bar{\rho}^S}$$

Ciò vale per un errore sia nelle coordinate del punto indietro sia nelle effemeridi dei satelliti. A titolo di esempio consideriamo due basi, una di 10 km e una di 100 km; supponiamo inoltre due casi distinti: dapprima un errore dell'ordine di 1 m, quindi un errore di 5 cm. Nella stima della prima base si ha un effetto di 0.5 mm per $\delta = 1m$ e di 0.025 mm per $\delta = 5cm$; per la seconda base gli errori salgono rispettivamente a 5 mm e 0.25 mm. I due esempi di errore sopra riportati sono tipici rispettivamente delle effemeridi trasmesse e di quelle precise; per le coordinate del punto di riferimento sono rappresentativi di un posizionamento assoluto prolungato e di una monografia di rete statica, come ad esempio di un punto IGM95; una stazione permanente ha stime di coordinate tipicamente caratterizzate da accuratezze migliori del centimetro.

Negli approcci di elaborazione più semplici si utilizzano i modelli standard per la stima e la rimozione dei disturbi atmosferici. Evidentemente il disturbo troposferico di differenza doppia è tanto maggiore quanto più differiscono le condizioni meteorologiche nei due siti di osservazione ed è ragionevole attendersi di avere condizioni meteorologiche tanto più differenti quanto più essi sono lontani: perciò, empiricamente, ci si attende che l'errore residuo troposferico aumenti linearmente con la distanza fra i ricevitori, con effetto inferiore a due parti per milione (2 ppm) della lunghezza della base: quindi, per una base di 10 km, l'errore residuo può arrivare a 2 cm, per una base di 100 km può arrivare a 20 cm. Anche il disturbo ionosferico si attenua nelle differenze doppie, ma rimane il termine di errore più grande: può arrivare anche a 8 ppm della base, ovvero 8 cm su 10 km, 80 cm su 100 km.

Si può dire che, a seconda delle condizioni di rilievo, in singola epoca, l'errore residuo di differenza doppia può variare da 0.5 ppm a 10 ppm della lunghezza della base; nel caso di ricevitori a doppia frequenza è possibile costruire le osservabili Ionospheric Free e da queste costruire le differenze doppie, che non subiscono quindi l'effetto

ionosferico: in questo caso l'errore residuo finale rimane tipicamente al di sotto dei 2 ppm. Alcuni valori di esempio per i diversi casi sono riportati in Tab. 6.1.

Lunghezza base	Buone condizioni	Condizioni normali	Medioci condizioni
1 Km	0.1 cm	0.2 cm	1.0 cm
10 Km	0.5 cm	2.0 cm	10.0 cm
50 Km	2.5 cm	10.0 cm	50.0 cm
100 Km	5.0 cm	20.0 cm	100.0 cm

Tabella 6.1. Errore residuo di differenza doppia, singola epoca, in funzione delle condizioni di osservazione. Buone condizioni: troposfera vicina alla condizione standard, basso stato di attività ionosferico; condizioni normali: quanto normalmente ci si può attendere; mediocri condizioni: ionosfera in attività anomala, troposfera molto differente fra i siti.

6.5 L'offset d'antenna

In tutte le discussioni, $\mathbf{x}_R(t)$ rappresenta il punto di ricezione del segnale GPS nell'antenna del ricevitore, ovvero il centro di fase dell'antenna, che ovviamente non è il punto fisico del quale si vuole determinare la posizione: questo può essere ad esempio il centro di presa di una fotocamera su aeromobile per un rilievo cinematico fotogrammetrico, oppure un caposaldo nel caso di rilievo di una rete topografica/geodetica.

Evidentemente l'offset dell'antenna non è specifico del posizionamento relativo ma, visti gli ordini di grandezza delle quantità coinvolte, esso può contribuire significativamente al bilancio d'errore nel posizionamento relativo e quindi se ne fa discussione solo in questo momento. Se denotiamo con il simbolo $\mathbf{x}_P(t)$ la posizione del punto P da stimare, si può scrivere

$$\mathbf{x}_R(t) = \mathbf{x}_P(t) + \Delta\mathbf{x}_{P,R}(t)$$

$\Delta\mathbf{x}_{P,R}(t)$ è l'offset d'antenna, ovvero il vettore fra P e il centro di fase dell'antenna, che viene usualmente espresso come somma di tre termini distinti

$$\Delta\mathbf{x}_{P,R}(t) = \Delta\mathbf{x}_{P,ARP_R}(t) + \Delta\mathbf{x}_{ARP_R,\bar{R}}(t) + \delta\mathbf{x}_{\bar{R},R}(\eta_R^S(t), \alpha_R^S(t))$$

Il primo è l'offset di stazionamento, ovvero il vettore fra P e il punto di riferimento dell'antenna (ARP, Antenna Reference Point), che in generale coincide con il centro geometrico della base dell'antenna: nel caso più comune dello stazionamento verticale su punti topografici, $\Delta\mathbf{x}_{P,ARP}(t)$ è costante per tutta la sessione e dato dalla

$$\Delta \mathbf{x}_{P,ARP}(t) = \Delta h_{P,ARP} \mathbf{z}_P$$

ove $\Delta h_{P,ARP}$ è la distanza verticale fra P e la base dell'antenna, $\mathbf{z}_P$ è il versore verticale nel punto. L'offset di stazionamento deve essere misurato e correttamente registrato dall'utente.

$\Delta \mathbf{x}_{ARP,\bar{R}}(t)$ è il vettore fra ARP e il centro di fase nominale o medio $\bar{R}$ dell'antenna ed è caratteristico dell'antenna; il centro di fase medio è lungo l'asse baricentrico dell'antenna e la distanza fra ARP e $\bar{R}$ può arrivare a due decimetri, viene comunicata dal costruttore dell'antenna o è pubblicata nei medesimi file dei PCV, discussi qui di seguito. Anche in questo caso, per uno stazionamento verticale si ha $\Delta \mathbf{x}_{ARP,\bar{R}}(t) = \Delta h_{ARP,\bar{R}} \mathbf{z}_P$, mentre nel caso più generale esso può essere calcolato conoscendo gli angoli istantanei di assetto dell'antenna.

$\delta \mathbf{x}_{\bar{R},R}(\eta_R^S(t), \alpha_R^S(t))$ è il vettore fra il centro di fase nominale dell'antenna e l'effettivo punto di acquisizione del segnale, che dipende dalla posizione del satellite rispetto al ricevitore, ovvero dalla sua elevazione $\eta_R^S(t)$ e dal suo azimut $\alpha_R^S(t)$.

$\delta \mathbf{x}_{\bar{R},R}(\eta_R^S(t), \alpha_R^S(t))$ viene detto variazione del centro di fase (PCV, Phase Center Variation), è un vettore le cui componenti cambiano nel tempo in funzione del moto relativo del satellite rispetto al ricevitore, con modulo che va da diversi millimetri sino a uno o due centimetri. I PCV sono diversi da antenna a antenna e vengono resi noti mediante tabelle di valori per intervalli regolari di elevazione e azimut: il loro studio e le loro stime sono iniziati una decina di anni fa e sono ancora argomento in evoluzione. Si consideri ora l'elaborazione delle differenze doppie: una volta noti i vari termini che compongono l'offset d'antenna, la stima delle basi fra i punti è ottenuta mediante tre semplici passaggi.

1. Ogni osservazione singolarmente presa dai due ricevitori viene ridotta dal centro di fase effettivo al centro di fase nominale mediante la

$$O_{\bar{R}}^S(t) = O_R^S(t) - \tilde{\mathbf{e}}_{\bar{R}}^S(t) \cdot \delta \mathbf{x}_{\bar{R},R}(\eta_R^S(t), \alpha_R^S(t))$$

2. le osservazioni ridotte vengono utilizzate per stimare la base fra i centri di fase nominali, secondo il processo di stima illustrato nei paragrafi successivi;

3. la stima della base viene corretta per i due offset fra i centri di fase nominali e i due punti

$$\Delta \hat{\mathbf{x}}_{P1,P2}(t) = \Delta \hat{\mathbf{x}}_{\bar{R}1,\bar{R}2}(t) - (\Delta \mathbf{x}_{P2,ARP2}(t) + \Delta \mathbf{x}_{ARP2,\bar{R}2}(t)) + \Delta \mathbf{x}_{P1,ARP1}(t) + \Delta \mathbf{x}_{ARP1,\bar{R}1}(t)$$

L'ultima correzione avviene epoca per epoca nel caso di un rilievo cinematico, è invece una sola per tutta la sessione nel caso di rilievo statico.

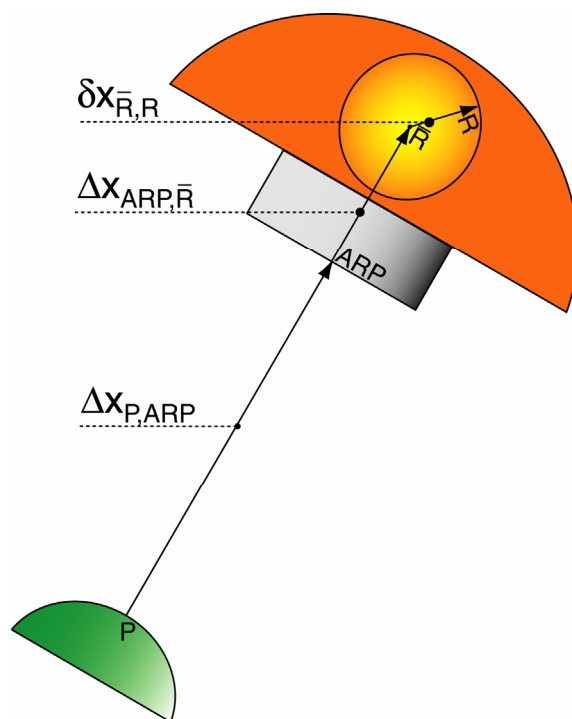


Figura 6.3. Le componenti dell'offset d'antenna.

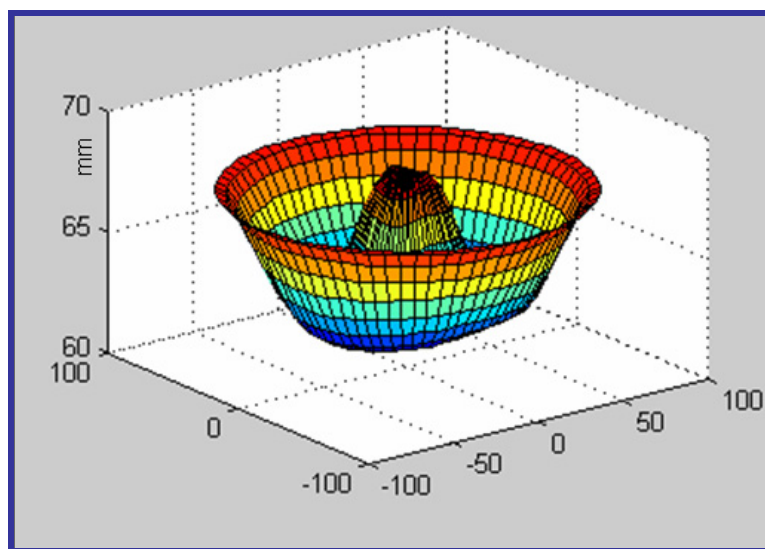


Figura 6.4. Un esempio di comportamento d'antenna: la stazione permanente di Como. In funzione di azimut ed elevazione è graficata la componente altimetrica del vettore fra ARP e centro di fase istantaneo.

Da un punto di vista operativo si rimarca che, per ottenere accuratèzze migliori del metro è necessario misurare e inserire correttamente perlomeno l'offset di stazionamento, mentre per ottenere accuratèzze migliori del decimetro è anche

necessario inserire i dati dell'antenna, al fine di applicare il corretto offset fra ARP e centro di fase medio. Per le applicazioni di altissima precisione è fondamentale anche l'utilizzo dei PCV, possibilmente in una versione aggiornata: quelli stimati e distribuiti da IGS (attualmente nel file IGS_05.atx) sono i più accurati e vengono periodicamente aggiornati; non tutti i programmi di elaborazione sono però in grado di leggere il file di IGS e, in questo caso, è necessario avvalersi dei PCV cablati all'interno dei programmi stessi, che sono però spesso obsoleti e meno accurati.

6.6 La costruzione del vettore di differenze doppie relative a un'epoca di osservazione

Si consideri la configurazione con due ricevitori e I satelliti per una singola epoca: è possibile costruire m differenze singole indipendenti per ciascuna osservabile O , di codice e fase ($P1, P2, C1, C2, L1, L2$); si ha quindi

$$O_{1,2}^1(t), O_{1,2}^2(t), \dots, O_{1,2}^I(t)$$

Per costruire le differenze doppie dalle differenze singole, si può procedere utilizzando un unico satellite di riferimento, ovvero secondo lo schema: S_1 con S_2 , poi S_1 con S_3 , fino a S_1 con S_I

$$O_{1,2}^{1,2}(t), O_{1,2}^{1,3}(t), \dots, O_{1,2}^{1,I}(t)$$

oppure a cascata, ovvero secondo lo schema: S_1 con S_2 , poi S_2 con S_3 , fino a S_{I-1} con S_I

$$O_{1,2}^{1,2}(t), O_{1,2}^{2,3}(t), O_{1,2}^{3,4}(t), \dots, O_{1,2}^{I-1,I}(t)$$

Nelle discussioni seguenti ci si atterrà all'approccio più tipicamente implementato, che è quello con un unico satellite di riferimento; tipicamente si sceglie come riferimento il satellite con elevazione maggiore perché si suppone che abbia osservazioni meno disturbate e rumorose. In ogni caso lo schema di costruzione non è rilevante dal punto di vista metodologico: è però importante ricordare che, per ogni epoca, per ogni osservabile, in presenza di due ricevitori e I satelliti è possibile costruire solo $I-1$ differenze doppie indipendenti (Fig. 6.5).

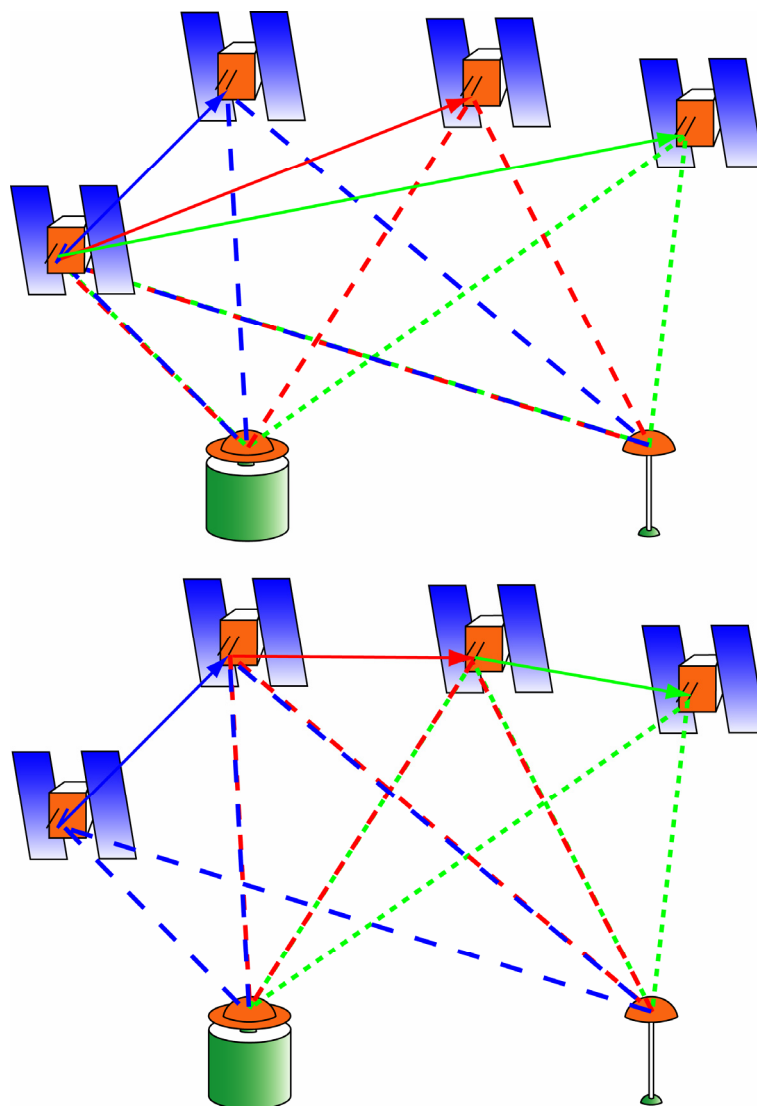


Figura 6.5. Lo schema di costruzione delle differenze doppie nell'esempio di quattro satelliti. In alto con un unico satellite di riferimento, in basso a cascata.

6.6.1 La propagazione del rumore di misura: la matrice di covarianza delle osservazioni di differenza doppia

Siano $O_1^i(t), O_2^i(t)$, con $1 \leq i < I$, le osservazioni contemporanee da 2 ricevitori a I satelliti, per un'epoca e su un'osservabile O : l'ipotesi classica è quella di considerare le osservazioni scorrelate e di uguale precisione, ovvero

$$C_{00} = \sigma_0^2 I$$

con $\sigma_0 \cong 0.1-1.0m$ per le osservazioni di codice e $\sigma_0 \cong 1mm$ per le osservazioni di fase. Si vuole calcolare la propagazione della matrice di covarianza dalle osservazioni non differenziate alle differenze doppie: si inizi con la costruzione delle differenze singole

$$s_{1,2}^s = O_1^s - O_2^s$$

ovvero

$$\begin{bmatrix} s_{1,2}^1 \\ s_{1,2}^2 \\ \dots \\ s_{1,2}^I \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & -1 \end{bmatrix} \begin{bmatrix} O_1^1 \\ O_1^2 \\ \dots \\ O_1^I \\ O_2^1 \\ O_2^2 \\ \dots \\ O_2^I \end{bmatrix}$$

$$\mathbf{s} = \mathbf{S} \begin{bmatrix} \mathbf{O}_1 \\ \mathbf{O}_2 \end{bmatrix}$$

Per la legge di propagazione della covarianza si ha

$$\mathbf{C}_{ss} = \mathbf{S} \mathbf{C}_{00} \mathbf{S}^T$$

ovvero

$$\mathbf{C}_{ss} = \sigma_0^2 \begin{bmatrix} 2 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 2 \end{bmatrix} = 2\mathbf{C}_{00} \quad (6.6)$$

Le differenze singole rimangono scorrelate e sono caratterizzate da varianza doppia rispetto alle osservazioni iniziali. Si costruiscano ora le differenze doppie

$$\begin{bmatrix} d_{1,2}^{1,2} \\ d_{1,2}^{1,3} \\ \dots \\ d_{1,2}^{1,I} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & -1 \end{bmatrix} \begin{bmatrix} s_{1,2}^1 \\ s_{1,2}^2 \\ s_{1,2}^3 \\ \dots \\ s_{1,2}^I \end{bmatrix}$$

$$\mathbf{d} = \mathbf{D}\mathbf{s}$$

Si ha

$$\mathbf{C}_{dd} = \mathbf{D}\mathbf{C}_{ss}\mathbf{D}^T$$

ovvero

$$\mathbf{C}_{dd} = \sigma_0^2 \begin{bmatrix} 4 & 2 & \dots & 2 \\ 2 & 4 & \dots & 2 \\ \dots & \dots & \dots & \dots \\ 2 & 2 & \dots & 4 \end{bmatrix} = 4\sigma_0^2 \begin{bmatrix} 1 & 0.5 & \dots & 0.5 \\ 0.5 & 1 & \dots & 0.5 \\ \dots & \dots & \dots & \dots \\ 0.5 & 0.5 & \dots & 1 \end{bmatrix} \quad (6.7)$$

Le differenze doppie sono reciprocamente correlate e caratterizzate da varianza quadrupla rispetto alle osservazioni originali. Un approccio più rigoroso nell'elaborazione geodetica suggerisce di pesare le osservazioni non differenziate con una funzione del loro angolo di elevazione; le attuali linee guida internazionali suggeriscono la scelta

$$\sigma_R^{S^2}(t) = \frac{\sigma_0^2}{\sin^2(\eta_R^S(t))} \quad (6.8)$$

In tal caso la matrice di covarianza delle differenze singole mantiene la struttura diagonale ma non è l'identità; la propagazione alle differenze doppie non è analiticamente sintetizzabile come fatto nella (6.7), ma rimangono valide le considerazioni generali che ne discendono: le differenze doppie sono correlate epoca per epoca e sono caratterizzate da varianza circa quadrupla rispetto alle osservazioni iniziali.

6.7 Introduzione all'elaborazione delle differenze doppie

Si consideri una coppia di ricevitori in grado di effettuare osservazioni a tutti i segnali: si hanno quindi quattro osservazioni di differenza doppia per ogni coppia di satelliti e per ogni epoca

$$\begin{aligned}
P1_{1,2}^{1,i}(t) &= \rho_{1,2}^{1,i}(t) + T_{1,2}^{1,i}(t) + I1_{1,2}^{1,i}(t) \\
P2_{1,2}^{1,i}(t) &= \rho_{1,2}^{1,i}(t) + T_{1,2}^{1,i}(t) + I2_{1,2}^{1,i}(t) \\
L1_{1,2}^{1,i}(t) &= \rho_{1,2}^{1,i}(t) + T_{1,2}^{1,i}(t) - I1_{1,2}^{1,i}(t) + \lambda N1_{1,2}^{1,i}(t) \\
L2_{1,2}^{1,i}(t) &= \rho_{1,2}^{1,i}(t) + T_{1,2}^{1,i}(t) - I2_{1,2}^{1,i}(t) + \lambda N2_{1,2}^{1,i}(t)
\end{aligned}$$

I termini geometrici, ovvero la posizione dei satelliti e dei ricevitori, sono uguali sui codici e sulle fasi, per tutte le frequenze e lo stesso vale per il disturbo troposferico. Il disturbo ionosferico è uguale in modulo per i codici e le fasi della stessa frequenza, ma ha segno opposto fra codici e fasi; è invece diverso per le due frequenze ma i due valori sono legati dalla (4.14); le ambiguità iniziali $N1_{1,2}^{1,i}(t)$ e $N2_{1,2}^{1,i}(t)$ sono differenti e non esiste alcuna relazione fra di esse. Ripetendo l'accorpamento del termine non dispersivo visto nel Par. 4.7, $p_{1,2}^{1,i}(t) = \rho_{1,2}^{1,i}(t) + T_{1,2}^{1,i}(t)$, si ottiene

$$\mathbf{y}_{1,2}^{1,i}(t) = \begin{bmatrix} P1_{1,2}^{1,i}(t) \\ P2_{1,2}^{1,i}(t) \\ L1_{1,2}^{1,i}(t) \\ L2_{1,2}^{1,i}(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & f_1^2 / f_2^2 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -f_1^2 / f_2^2 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{1,2}^{1,i}(t) \\ I1_{1,2}^{1,i}(t) \\ N1_{1,2}^{1,i}(t) \\ N2_{1,2}^{1,i}(t) \end{bmatrix} = \mathbf{A} \boldsymbol{\xi}_{1,2}^{1,i}(t) \quad (6.9)$$

ovvero un sistema di quattro equazioni in quattro incognite $d_{1,2}^{1,i}(t), N1_{1,2}^{1,i}(t), N2_{1,2}^{1,i}(t), I1_{1,2}^{1,i}(t)$, qui raccolte nel vettore $\boldsymbol{\xi}_{1,2}^{1,i}(t)$; il sistema può essere risolto epoca per epoca poiché la matrice disegno $\mathbf{A}$ ha rango pieno: considerando la disponibilità di I satelliti, ovvero di $I-1$ differenze doppie, si ottengono le stime $\hat{p}_{1,2}^{1,2}(t), \hat{p}_{1,2}^{1,3}(t), \dots, \hat{p}_{1,2}^{1,I}(t)$, che possono a loro volta essere utilizzate per stimare la base.

In effetti tale approccio non è quello più tipicamente implementato; in generale non esiste un modo univoco di processare le differenze doppie, ma sono state implementate numerose varianti di approcci di elaborazione, ottimizzate in funzione dei seguenti aspetti:

- tipologia di ricevitori utilizzati nella misura, che possono essere di solo codice in singola frequenza, di codice su entrambe le frequenze, di codice e fase in singola frequenza, di codice e fase in doppia frequenza;
- distanza fra i ricevitori, che può andare dal centinaio di metri (basi cortissime) a diverse centinaia di chilometri (basi continentali);
- durata dell'occupazione, ovvero tempo di stazionamento sui punti, che può andare da una singola epoca (rilievo cinematografico) all'occupazione statica continua (stazioni permanenti).

Inoltre si deve considerare che oggi sono disponibili molti programmi di elaborazione dei dati GPS, che si differenziano significativamente per le scelte adottate. Alcuni programmi, tipicamente quelli sviluppati da enti di ricerca, implementano algoritmi chiaramente ed esaustivamente descritti nella documentazione ma la maggior parte,

ovvero quelli sviluppati a scopo commerciale, si limita a fornire documentazione di utilizzo, senza una descrizione dettagliata degli algoritmi implementati. Perciò, senza pretendere alcuna completezza, si mostrano con dettaglio due possibili approcci per l'elaborazione di due casi all'estremo opposto:

1. la soluzione di solo codice in singola epoca, in quanto rappresentativa dell'approccio più semplice al posizionamento relativo, implementato nei ricevitori più economici;
2. la stima statica di una base mediante elaborazione delle fasi su due frequenze, in quanto rappresentativa dell'approccio più accurato.

Verranno quindi illustrate al termine del capitolo le caratteristiche degli altri approcci, le relative problematiche e accuratezze.

6.8 La soluzione di solo codice in singola epoca

Si considera il caso più semplice, ovvero quello dei tradizionali ricevitori civili che effettuano osservazioni solo su C/A: si consideri la disponibilità di I satelliti per un'epoca. Innanzitutto si calcolano le differenze doppie di disturbo troposferico e ionosferico mediante i modelli standard, $\tilde{T}_{1,2}^{1,i}(t)$, $\tilde{I}_{1,2}^{1,i}(t)$; per ciascuna osservazione si scrive l'equazione

$$P_{01,2}^{1,i}(t) = \tilde{\epsilon}_2^{1,i}(t) \cdot \delta \Delta \mathbf{x}_{1,2}(t) + b_{1,2}^{1,i}(t) + \nu_{1,2}^{1,i}(t) + \delta_{1,2}^{1,i}(t)$$

$$b_{1,2}^{1,i}(t) = \tilde{\rho}_{1,2}^{1,i}(t) + \tilde{T}_{1,2}^{1,i}(t) + \tilde{I}_{1,2}^{1,i}(t)$$

ove $\delta_{1,2}^{1,i}(t)$ esprime l'errore residuo di modellizzazione. Considerando tutte le differenze doppie per un'epoca, e accorpare errori di modello e rumore di osservazione si ha un sistema da risolvere del tipo

$$\mathbf{y}_0(t) = \mathbf{P}_0(t) - \mathbf{b}(t) = \mathbf{E}(t) \delta \Delta \mathbf{x}_{R1R2}(t) + \mathbf{v}(t)$$

ove

$$\mathbf{P}_0(t) = \begin{bmatrix} P_{1,2}^{1,2}(t) \\ P_{1,2}^{1,3}(t) \\ \dots \\ P_{1,2}^{1,I}(t) \end{bmatrix}_{(I-1) \times 1}, \quad \mathbf{E}(t) = \begin{bmatrix} \tilde{\epsilon}_{2_x}^{2,1}(t) & \tilde{\epsilon}_{2_x}^{2,1}(t) & \tilde{\epsilon}_{2_x}^{2,1}(t) \\ \tilde{\epsilon}_{2_x}^{3,1}(t) & \tilde{\epsilon}_{2_y}^{3,1}(t) & \tilde{\epsilon}_{2_z}^{3,1}(t) \\ \dots & \dots & \dots \\ \tilde{\epsilon}_{2_x}^{I,1}(t) & \tilde{\epsilon}_{2_y}^{I,1}(t) & \tilde{\epsilon}_{2_z}^{I,1}(t) \end{bmatrix}_{(I-1) \times 3}, \quad \mathbf{b}(t) = \begin{bmatrix} \tilde{\rho}_{1,2}^{1,2}(t) + \tilde{T}_{1,2}^{1,2}(t) + \tilde{I}_{1,2}^{1,2}(t) \\ \tilde{\rho}_{1,2}^{1,3}(t) + \tilde{T}_{1,2}^{1,3}(t) + \tilde{I}_{1,2}^{1,3}(t) \\ \dots \\ \tilde{\rho}_{1,2}^{1,I}(t) + \tilde{T}_{1,2}^{1,I}(t) + \tilde{I}_{1,2}^{1,I}(t) \end{bmatrix}_{(I-1) \times 1}$$

Con almeno quattro satelliti in vista per ogni epoca si possono costruire abbastanza differenze doppie da risolvere rispetto alle tre componenti della base: in tal senso le differenze doppie di codice si prestano a soluzioni di navigazione. La matrice di covarianza del rumore è stata discussa nel Par. 6.6.1, mentre l'errore di modello può arrivare a 10 ppm in mediocri condizioni di osservazione: pertanto una coppia di

ricevitori di codice in singola frequenza può determinare la base con accuratezze migliori del metro solo se si tratta di ricevitori di buona qualità a distanza massima di alcune decine di chilometri; altrimenti le accuratezze sono peggiori del metro.

Nel caso di ricevitori di codice a doppia frequenza si può elaborare la Ionospheric Free sui codici e l'errore di modello scende a 2 ppm: con buoni ricevitori l'accuratezza del posizionamento può rimanere migliore del metro anche a distanze dell'ordine di 100 km. L'elaborazione delle differenze doppie di codice può essere effettuata anche sulle osservazioni di codice preliminarmente filtrate con le fasi, come descritto nel Par. 4.9: l'effetto del rumore elettronico diminuisce sensibilmente e quindi si migliorano ulteriormente le accuratezze. Alcune sintesi delle accuratezze sono riportate nella tabella finale del capitolo (Tab. 6.5).

6.9 L'elaborazione di un rilievo statico in doppia frequenza

Si considerino due ricevitori che abbiano effettuato un rilievo statico, ovvero che abbiano occupato la stessa posizione per un periodo prolungato; si considerano ricevitori geodetici, ovvero in grado di acquisire osservazioni di codice e fase su entrambe le frequenze; infine, si considera il caso di postelaborazione dei dati. I passi dell'elaborazione e le relative finalità sono qui sintetizzate:

1. la preelaborazione dei dati: controllo di qualità generale, correzione degli offset d'orologio, prima stima della base;
2. la rimozione dei cycle slip: identificazione dei cycle slip nelle differenze doppie di fase L1 e L2 e, se possibile, loro correzione;
3. la stima della ionosfera: per basi molto lunghe, stima di modelli locali di ionosfera da utilizzarsi ai passi successivi;
4. la stima delle ambiguità: stima della base ma soprattutto stima reale delle differenze doppie di ambiguità iniziali e della relativa matrice di covarianza;
5. il fissaggio delle ambiguità iniziali: se possibile, le ambiguità iniziali vengono vincolate ai loro valori interi;
6. la stima finale: stima della base, possibilmente ad ambiguità fissate, con l'obiettivo di minimizzare l'effetto residuo atmosferico.

6.9.1 La preelaborazione dei dati

Il primo passo dell'elaborazione è sui ricevitori singolarmente presi: per ogni epoca del rilievo, vengono elaborate le osservazioni di codice non differenziate, in accordo a quanto visto per il posizionamento assoluto: questo serve per un primo controllo di qualità dei dati, ma soprattutto per la stima e la correzione degli offset d'orologio di entrambi i ricevitori. Nel Par. 5.7 si è visto che quando si linearizzano le equazioni di osservazione compare un termine contenente l'offset d'orologio del ricevitore, del tipo $k_R^S(t)dt_R$; quindi nella linearizzazione delle differenze doppie di distanza rimane un termine di orologio dato dalla

$$\theta_{1,2}^{1,i}(t) = (k_1^1(t) - k_1^i(t))dt_1(t) - (k_2^1(t) - k_2^i(t))dt_2(t)$$

Il termine $\theta_{1,2}^{1,i}(t)$ andrebbe considerato incognito e da stimare; considerando però i massimi valori che i coefficienti $k_R^S(t)$ possono assumere, si vede che quando $dt_R(t) \leq 1\mu s$, si ha $\theta_{1,2}^{1,i}(t) < 0.5mm$ ovvero il termine è trascurabile. Gli offset d'orologio del ricevitore possono arrivare anche a diversi millisecondi e quindi la precedente condizione non è mai soddisfatta. Però un posizionamento assoluto in singola epoca permette sicuramente la stima, e quindi la correzione, dei $dt_R(t)$ con accuratezze migliori del μs per ogni epoca: per tale scopo vengono elaborati i codici non differenziati, gli offset d'orologio vengono corretti e nelle seguenti elaborazioni si potrà trascurare $\theta_{1,2}^{1,i}(t)$ in tutte le differenze doppie.

Al fine di avere una stima migliore della base si elaborano le differenze doppie di codice fra ricevitore indietro e ricevitore avanti: la stima risultante della base è tipicamente caratterizzata da accuratezza dell'ordine di alcuni decimetri.

6.9.2 L'identificazione e la rimozione dei cycle slip

Nell'elaborazione di una sessione è fondamentale identificare tutti i cycle slip perché questi generano nuove ambiguità iniziali, trascurando le quali si introducono errori di modello pari al numero di cicli persi, moltiplicato per la lunghezza d'onda della portante: ricordando che $\lambda \cong 20cm$, ad esempio un cycle slip di un milione di cicli comporta $10^6 \cdot 20cm = 2 \cdot 10^5 m = 200km$. La presenza di molti cycle slip implica un significativo aumento del numero di ambiguità da stimare, ovvero una diminuzione della ridondanza nella successiva soluzione del sistema; tenendo conto che i cycle slip sono incogniti ma per definizione interi si possono costruire algoritmi per la loro identificazione, la loro stima e, in caso di successo, la loro rimozione dai dati: si propone qui un semplice esempio di approccio. Si consideri una sessione statica di due ricevitori; si consideri la serie temporale di differenze doppie residue di fase riferite alla medesima coppia di satelliti, ovvero

$$\delta L_{1,2}^{1,i}(t_k) = L_{1,2}^{1,i}(t_k) - \tilde{\rho}_{1,2}^{1,i}(t_k) - \tilde{T}_{1,2}^{1,i}(t_k) - \tilde{I}_{1,2}^{1,i}(t_k)$$

si ha

$$\delta L_{1,2}^{1,i}(t_1) = \tilde{\mathbf{e}}_2^{1,1}(t_1) \cdot \delta \Delta \mathbf{x}_{R1R2} + \lambda N_{1,2}^{1,i}(t_1) + \delta_{1,2}^{1,i}(t_1)$$

$$\delta L_{1,2}^{1,i}(t_2) = \mathbf{e}_2^{1,1}(t_2) \cdot \delta \Delta \mathbf{x}_{R1R2} + \lambda N_{1,2}^{1,i}(t_2) + \delta_{1,2}^{1,i}(t_2)$$

...

$$\delta L_{1,2}^{1,i}(t_T) = \tilde{\mathbf{e}}_2^{1,1}(t_T) \cdot \delta \Delta \mathbf{x}_{R1R2} + \lambda N_{1,2}^{1,i}(t_T) + \delta_{1,2}^{1,i}(t_T)$$

Si calcoli ora la differenza fra differenze doppie residue relative a epoche consecutive: si ha

$$\begin{aligned}\Delta\delta L_{1,2}^{1,i}(t_k) &= \delta L_{1,2}^{1,i}(t_k) - \delta L_{1,2}^{1,i}(t_{k-1}) \\ &= (\tilde{\mathbf{e}}_2^{i,1}(t_k) - \tilde{\mathbf{e}}_2^{i,1}(t_{k-1})) \cdot \delta \Delta \mathbf{x}_{R1,R2} \\ &\quad + \lambda(N_{1,2}^{1,i}(t_k) - N_{1,2}^{1,i}(t_{k-1})) + \delta_{1,2}^{1,i}(t_k) - \delta_{1,2}^{1,i}(t_{k-1})\end{aligned}\tag{6.10}$$

Supponendo che le componenti approssimate della base siano sufficientemente accurate (meglio di un metro, come sono state ottenute al passo precedente) il primo termine a destra dell'uguaglianza è pressoché nullo: infatti il versore dal ricevitore al satellite cambia molto lentamente nel tempo e quindi la sua differenza fra epoche successive è trascurabile (le componenti hanno valori inferiori a 10^{-4}); analogamente le differenze nel tempo degli errori di modello sono inferiori al centimetro, almeno per epoche vicine. Si ha quindi

$$\Delta\delta L_{1,2}^{1,i}(t_k) \cong \lambda(N_{1,2}^{1,i}(t_k) - N_{1,2}^{1,i}(t_{k-1}))$$

In assenza di cycle slip, $N_{1,2}^{1,i}(t_k) = N_{1,2}^{1,i}(t_{k-1})$ e quindi

$$\Delta\delta L_{1,2}^{1,i}(t_k) \cong 0$$

In presenza di cycle slip $\Delta\delta L_{1,2}^{1,i}(t_k) \neq 0$; in particolare

$$N_{1,2}^{1,i}(t_k) - N_{1,2}^{1,i}(t_{k-1}) \cong \frac{\Delta\delta L_{1,2}^{1,i}(t_k)}{\lambda}$$

Da un punto di vista numerico l'algoritmo viene implementato nel seguente modo: si definisce un livello di soglia per le discontinuità, ad esempio $\delta = 5cm$. Ciascuna serie di differenze doppie relativa alla coppia di satelliti $S1, Si$ viene analizzata separatamente; per ogni serie si parte dalla seconda epoca, ovvero t_2 ;

1. si calcola $\Delta\delta L_{1,2}^{1,i}(t_2)$;

2a. se $|\Delta\delta L_{1,2}^{1,i}(t_2)| \leq \delta$ si ipotizza l'assenza di cycle slip e si passa al punto 4;

2b. in caso contrario, si ipotizza la presenza di un cycle slip e si calcola

$$\Delta N_{1,2}^{1,i}(t_2) = \frac{\Delta\delta L_{1,2}^{1,i}(t_2)}{\lambda}; \text{ sia ora } N \text{ l'intero più vicino a } \Delta N_{1,2}^{1,i}(t_2); \text{ si hanno due casi:}$$

3a. se $|N - \Delta N_{1,2}^{1,i}(t_2)| \leq \delta$ si ritiene che N sia il valore esatto del cycle slip, che viene quindi utilizzato per correggere (riparare) tutte le osservazioni da t_2 all'ultima epoca;

3b. in caso contrario si ritiene di aver identificato un cycle slip ma non si è in grado di stimare il suo valore intero; si deve quindi introdurre una nuova ambiguità iniziale incognita per t_2 .

4. Il processo riprende dalla coppia costituita dalla seconda e dalla terza epoca e prosegue sequenzialmente sino all'ultima epoca.

5. Terminata l'analisi di $L1_{1,2}^{1,i}(t)$, si effettua l'analisi di $L2_{1,2}^{1,i}(t)$.

Note. La grandezza $\Delta \delta L_{1,2}^{1,i}(t_k)$, calcolata in accordo alla (6.10) può risentire della propagazione del rumore di osservazione di $\delta L_{1,2}^{1,i}(t_{k-1})$ e $\delta L_{1,2}^{1,i}(t_k)$ che, per epoche in prossimità di un cycle slip, può essere significativo e inficiare la stima: una variante più raffinata prevede l'utilizzo di polinomi interpolanti, stimati a partire dalle epoche già validate, per predire $\delta L_{1,2}^{1,i}(t_{k-1})$, confrontarlo con il valore all'epoca successiva e così identificare eventuali cycle slip e stimarne il valore intero. Inoltre, per basi lunghe, ovvero superiori ai 50 km, il metodo appena descritto può risentire dell'effetto residuo ionosferico, qui considerato trascurabile: ulteriori varianti, che non discutiamo qui, prevedono l'analisi congiunta di L1 e L2 per rimuovere, almeno parzialmente, tale fonte di errore.

6.9.3 La stima della ionosfera

Per basi di lunghezza superiore a circa 10 km, i passi di stima e di fissaggio delle ambiguità possono risentire dell'errore residuo ionosferico. E' possibile utilizzare le osservazioni dei ricevitori per calcolare un modello locale di ionosfera che risulta più accurato rispetto al modello standard. Il processo di stima non viene qui discusso; si evidenziano solo alcuni punti:

- il modello locale viene tipicamente calcolato utilizzando l'osservabile L4, che appunto dipende solo dalla ionosfera e dalle ambiguità iniziali;
- il modello locale è comunque un modello approssimato: risulta d'ausilio nella stima delle ambiguità, ma per la stima finale della base è meglio ricorrere alla osservabile Ionospheric Free.

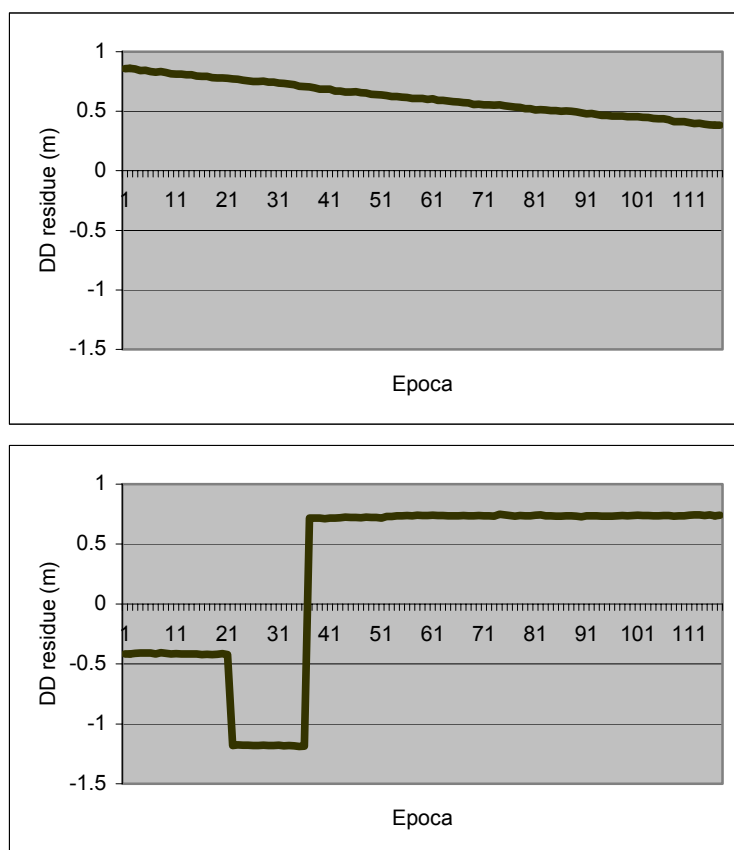


Figura 6.6. Esempi di serie temporali di differenze doppie residue sulla frequenza L1.
Sopra: senza cycle slip; sotto: con due cycle slip,
in corrispondenza delle epoche 21 e 36.

6.9.4 La stima delle ambiguità

Per semplicità di esposizione si discute il caso di costellazione costante, con I satelliti per tutta la durata di una sessione, composta di T epoche, e due frequenze, ovvero $2T(I-1)$ osservazioni di differenza doppia di fase; si pone inoltre il caso ideale di assenza di cycle slip. Si vuole risolvere il sistema rispetto alla base e alle ambiguità iniziali di L1 e L2: si hanno quindi 3 componenti della base e $2(I-1)$ ambiguità iniziali. Il sistema è il seguente:

$$\mathbf{y}_0 = \mathbf{L}_0 - \mathbf{b} = \mathbf{A}\boldsymbol{\xi} + \mathbf{v} \quad (6.11)$$

ove

$$\begin{aligned}
\mathbf{L}_{2T(I-1)} &= \begin{bmatrix} \mathbf{L}(t_1) \\ \mathbf{L}(t_2) \\ \dots \\ \mathbf{L}(t_T) \end{bmatrix}, \quad \mathbf{L}(t) = \begin{bmatrix} \mathbf{L1}(t) \\ \mathbf{L2}(t) \end{bmatrix}, \quad \mathbf{Li}(t) = \begin{bmatrix} Li_{1,2}^{1,2}(t) \\ Li_{1,2}^{1,3}(t) \\ \dots \\ Li_{1,2}^{1,I}(t) \end{bmatrix} \\
\mathbf{b}_{2T(I-1)} &= \begin{bmatrix} \mathbf{b}(t_1) \\ \mathbf{b}(t_2) \\ \dots \\ \mathbf{b}(t_T) \end{bmatrix}, \quad \mathbf{b}(t) = \begin{bmatrix} \mathbf{b1}(t) \\ \mathbf{b2}(t) \end{bmatrix}, \quad \mathbf{bi}(t) = \begin{bmatrix} \tilde{\rho}_{1,2}^{1,2}(t) + \tilde{T}_{1,2}^{1,2}(t) - \tilde{li}_{1,2}^{1,2}(t) \\ \tilde{\rho}_{1,2}^{1,3}(t) + \tilde{T}_{1,2}^{1,3}(t) - \tilde{li}_{1,2}^{1,3}(t) \\ \dots \\ \tilde{\rho}_{1,2}^{1,I}(t) + \tilde{T}_{1,2}^{1,I}(t) - \tilde{li}_{1,2}^{1,I}(t) \end{bmatrix} \\
\mathbf{A}_{\begin{smallmatrix} 2T(I-1) \\ \times \\ 3+2(I-1) \end{smallmatrix}} &= \begin{bmatrix} \mathbf{A}(t_1) \\ \mathbf{A}(t_2) \\ \dots \\ \mathbf{A}(t_T) \end{bmatrix}, \quad \mathbf{A}(t) = \begin{bmatrix} \mathbf{E}(t) & \lambda_1 \mathbf{I} & \mathbf{0} \\ \mathbf{E}(t) & \mathbf{0} & \lambda_2 \mathbf{I} \end{bmatrix}, \quad \mathbf{E}(t) = \begin{bmatrix} \tilde{e}_{2_x}^{2,1}(t) & \tilde{e}_{2_x}^{2,1}(t) & \tilde{e}_{2_x}^{2,1}(t) \\ \tilde{e}_{2_x}^{3,1}(t) & \tilde{e}_{2_y}^{3,1}(t) & \tilde{e}_{2_z}^{3,1}(t) \\ \dots & \dots & \dots \\ \tilde{e}_{2_x}^{I,1}(t) & \tilde{e}_{2_y}^{I,1}(t) & \tilde{e}_{2_z}^{I,1}(t) \end{bmatrix}_{(I-1) \times 3} \\
\boldsymbol{\xi}_{3+2(I-1)} &= \begin{bmatrix} \delta \Delta \mathbf{x}_{1,2} \\ \mathbf{N} \end{bmatrix}, \quad \mathbf{N}_{2(I-1)} = \begin{bmatrix} \mathbf{N1} \\ \mathbf{N2} \end{bmatrix}, \quad \mathbf{Ni}_{I-1} = \begin{bmatrix} Ni_{1,2}^{1,2} \\ Ni_{1,2}^{1,3} \\ \dots \\ Ni_{1,2}^{1,I} \end{bmatrix}
\end{aligned}$$

Il sistema presenta ridondanza con l'accumularsi delle epoche: ad esempio nel caso di quattro satelliti già con due epoche si hanno dodici osservazioni da risolversi rispetto a nove incognite; con un'ora di osservazione e intervallo di osservazione a quindici secondi, si hanno 240 epoche, ovvero 1440 osservazioni nelle stesse incognite. Costruita la matrice di covarianza delle differenze doppie (Par. 6.6.1), il sistema viene risolto mediante minimi quadrati, ottenendo le stime di $\hat{\boldsymbol{\xi}}$, $\mathbf{C}_{\hat{\boldsymbol{\xi}}\hat{\boldsymbol{\xi}}}$ ovvero di $\delta \Delta \mathbf{x}_{1,2}$, di $\mathbf{N}$ e delle relative covarianze: tali stime prendono il nome di stime float.

Molti programmi, anziché elaborare L1 e L2, elaborano le combinazioni L5 e L6: in tal caso il vettore delle osservazioni diviene

$$\mathbf{L}_{2T(I-1)} = \begin{bmatrix} \mathbf{L}(t_1) \\ \mathbf{L}(t_2) \\ \dots \\ \mathbf{L}(t_T) \end{bmatrix}, \quad \mathbf{L}(t) = \begin{bmatrix} \mathbf{L5}(t) \\ \mathbf{L6}(t) \end{bmatrix} = \begin{bmatrix} \beta_{1_5} \mathbf{L1}(t) + \beta_{2_5} \mathbf{L2}(t) \\ \beta_{1_6} \mathbf{L1}(t) + \beta_{2_6} \mathbf{L2}(t) \end{bmatrix}$$

Rispetto a prima cambiano i seguenti elementi del modello deterministico

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{E}(t) & \lambda_3 \mathbf{I} & \mathbf{0} \\ \mathbf{E}(t) & \mathbf{0} & \lambda_6 \mathbf{I} \end{bmatrix}, \mathbf{b}(t) = \begin{bmatrix} \mathbf{b5}(t) \\ \mathbf{b6}(t) \end{bmatrix}$$

$\begin{matrix} 2(I-1) \\ \times \\ 3+2(I-1) \end{matrix}$
 $\begin{matrix} 2(I-1) \end{matrix}$

dove gli elementi del termine noto cambiano solo per il contributo di disturbo ionosferico. Il vettore delle incognite contiene le componenti della base e le ambiguità Wide e Narrow Lane, ovvero

$$\xi_{3+2(I-1)} = \begin{bmatrix} \delta \Delta \mathbf{x}_{1,2} \\ \mathbf{N} \end{bmatrix}, \mathbf{N}_{2(I-1)} = \begin{bmatrix} \mathbf{N5} \\ \mathbf{N6} \end{bmatrix}$$

Tale approccio, rispetto all'elaborazione di L1 e L2, rivela un fondamentale vantaggio nel passaggio successivo, ovvero il fissaggio delle ambiguità: L5 ha lunghezza d'onda di circa 86 cm: risulta quindi più facile discriminare l'intero più probabile a partire dalla stima reale; L6 ha in effetti lunghezza d'onda di circa 11 cm, però la sua ambiguità deve avere la stessa parità di L5: perciò il suo discreto intero è circa uguale a quelli di L1 o L2.

Alcuni programmi elaborano altre combinazioni ancora di fasi e codici, qui non trattate ma questo particolare aspetto non è essenziale, poiché in definitiva lo scopo di tutti gli approcci, più o meno ottimizzati, è quello di stimare le ambiguità iniziali di differenza doppia.

Prima di proseguire nella discussione, è necessario ricordare che l'esempio è chiaramente irrealistico poiché in un rilievo statico possono sempre esservi cambi di configurazione e cycle slip; nei casi reali quindi, le incognite di ambiguità sono di più e la matrice disegno ha una struttura più complessa: in ogni caso questo è un problema puramente numerico di corretta costruzione del sistema risolutivo, che deve essere affrontato nelle implementazioni ma non inficia il senso di un'esposizione generale. Si vuole però ricordare un aspetto fondamentale relativo al caso di una configurazione satellitare variabile, con cambi del satellite di riferimento per la costruzione delle differenze doppie: tali cambi generano nuove ambiguità, che però possono essere funzionalmente dipendenti dalle precedenti. Per chiarire il concetto si cita il seguente semplicissimo esempio: siano $S1, S2, S3$ gli indici di tre satelliti, con uno schema di occupazione temporale come esemplificato in Fig. 6.7: si ipotizza che nella prima parte della sessione sia stato utilizzato il satellite $S1$ come riferimento, nella seconda parte sia stato utilizzato $S2$. Si hanno quindi quattro ambiguità iniziali: due relative al primo periodo, $N_{1,2}^{1,2}, N_{1,2}^{1,3}$, due relative al secondo periodo, $N_{1,2}^{2,1}, N_{1,2}^{2,3}$; però, in assenza di cycle slip, si hanno anche due relazioni di vincolo

$$N_{1,2}^{1,2} = N_1^1 - N_1^2 - N_2^1 + N_2^2 = -N_{1,2}^{2,1}$$

$$\begin{aligned}
 N_{1,2}^{2,3} &= N_1^2 - N_1^3 - N_2^2 + N_2^3 + N_1^1 - N_1^1 - N_2^1 + N_2^1 \\
 &= -N_{1,2}^{1,2} + N_{1,2}^{1,3}
 \end{aligned}$$

ovvero le incognite indipendenti sono solo due. Individuare le dipendenze e formalizzarle nella costruzione del sistema mediante le opportune relazioni di vincolo evita l'introduzione di incognite superflue, che ridurrebbero la ridondanza.

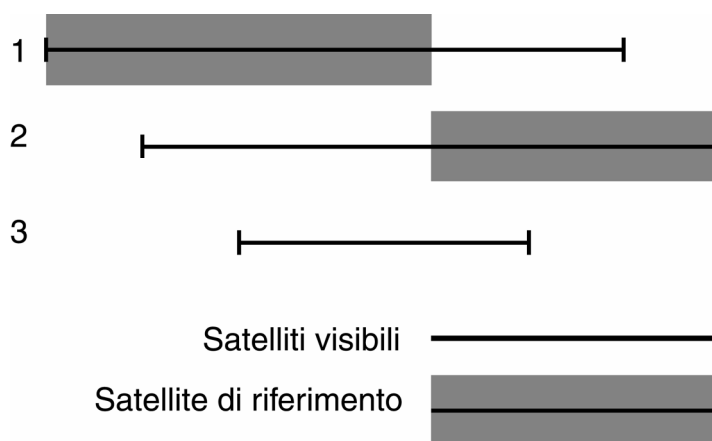


Figura 6.7. Dipendenza funzionale delle ambiguità al cambio del satellite di riferimento.

6.9.5 Il fissaggio delle ambiguità

Senza dimostrarla, si riporta una caratteristica fondamentale della stima float: essa risulta tipicamente ben condizionata nel caso di una significativa variabilità nella geometria relativa fra ricevitori e satelliti, come può ad esempio avvenire con una sessione di almeno un'ora; nel caso di sessioni più corte, sino al limite di pochi minuti, la geometria relativa non cambia sostanzialmente e, anche in condizioni di ridondanza fra osservazioni e incognite, la stima float può essere malcondizionata: in particolare le componenti della base possono essere caratterizzate da incertezze anche del decimetro, a prescindere dall'entità degli errori di modello. In tal senso è fondamentale il cosiddetto fissaggio delle ambiguità, il cui principio è il seguente; le stime float delle ambiguità di differenza doppia sono valori reali: ci si chiede se è possibile utilizzare l'informazione a priori sulla loro natura intera per vincolare le stime reali a opportuni numeri interi, in modo da eliminarle dall'insieme dei parametri incogniti. In altri termini: sono le stime abbastanza vicine a numeri interi, in modo da poter ragionevolmente affermare che quegli interi sono le vere ambiguità iniziali? Si presentano nel seguito due approcci di fissaggio: il primo di essi è il più semplice fra tutti, il secondo è il più rigoroso. A livello di notazione, normalmente un'ambiguità di differenza doppia viene contrassegnata dagli indici dei ricevitori, dei satelliti e della frequenza cui si riferisce, ovvero viene indicata con la notazione $N_{1,2}^{1,i}$; nel presente paragrafo tale distinzione non è essenziale e quindi ci limiteremo a indicare il vettore di

tutte le ambiguità iniziali con $\mathbf{N}$, le singole ambiguità con un loro indice d'ordine progressivo, ovvero N_i . Si dispone dunque di $\hat{\mathbf{N}}$ e della stima della relativa matrice di covarianze:

$$\hat{\mathbf{N}} = \begin{bmatrix} \hat{N}_1 \\ \hat{N}_2 \\ \dots \\ \hat{N}_p \end{bmatrix}, \quad \mathbf{C}_{NN} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2p} \\ \dots & \dots & \dots & \dots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{bmatrix}$$

Il primo approccio di fissaggio analizza le ambiguità singolarmente prese (Fig. 6.8); per ognuna di esse si considerano la stima reale e la relativa accuratezza di stima, $\hat{N}_i$, σ_i e si adotta l'ipotesi di distribuzione normale delle stime: si ha quindi il seguente intervallo di confidenza per il valore vero dell'ambiguità

$$P(\hat{N}_i - 3\sigma_i \leq N_i \leq \hat{N}_i + 3\sigma_i) = 0.997 \quad (6.12)$$

Se nell'intervallo di confidenza ricade uno e un solo intero $\bar{N}$ si può, con probabilità 99.7%, ipotizzare che $N_i = \bar{N}$: l'ambiguità d'ora innanzi viene fissata, ovvero considerata nota e uguale a $\bar{N}$. Se nell'intervallo non ricade alcun intero, a questo livello di elaborazione non è possibile fissare l'ambiguità, se viceversa ricadono più interi non è possibile discriminare fra i candidati ammissibili; in entrambi i casi si rinuncia al fissaggio o si deve adottare un ulteriore criterio per ottenerlo: tale situazione può occorrere per scarsa accuratezza nelle osservazioni o per la presenza di errori residui grossolani nelle equazioni di osservazione. Il processo viene condotto iterativamente per tutte le ambiguità nel vettore $\hat{\mathbf{N}}$.

Soprattutto in sessioni corte, può accadere che le stime delle ambiguità siano fortemente correlate e che quindi le loro regioni di confidenza siano iperellissoidi, centrati sulle stime reali e molto elongati e inclinati rispetto agli assi: i candidati interi più ovvi, ovvero quelli singolarmente più vicini alle stime reali, non sempre compongono il vettore intero più probabile (si veda l'esempio in Fig. 6.9); l'approccio basato sulle ambiguità singolarmente prese ignora completamente le correlazioni fra le stime e quindi, in questi casi, può causare errori di fissaggio. Si introduce perciò un approccio più rigoroso che considera le ambiguità complessivamente prese, analizzando il vettore $\hat{\mathbf{N}}$ nel suo insieme e congiuntamente alla sua matrice di covarianza $\mathbf{C}_{NN}$.

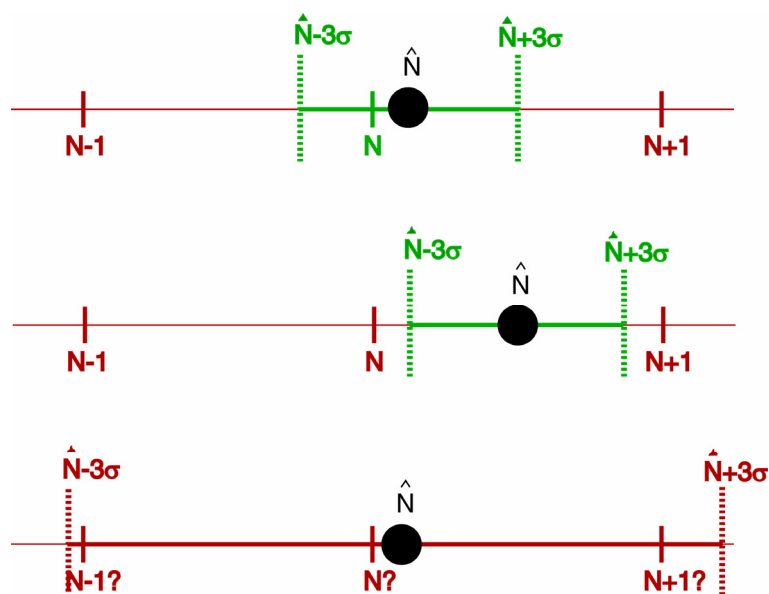


Figura 6.8. I diversi casi nel fissaggio delle ambiguità singolarmente prese. Sopra: l'intervallo di confidenza contiene un unico intero: l'ambiguità viene fissata. In mezzo e sotto: l'intervallo di confidenza non contiene interi, oppure ne contiene più d'uno: l'ambiguità non può essere fissata.

Satellite di riferimento	Secondo Satellite	Ambiguità stimata	Accuratezza di stima	Ambiguità fissata
17	3	-6.194	+0.132	-6
17	6	+5.132	+0.065	+5
17	15	+0.955	+0.028	+1
17	22	-2.125	+0.080	-2
17	23	+8.270	+0.088	-----
17	23	-41.787	+0.081	-42
17	25	+5.871	+0.094	+6

Tabella 6.2. Stima del vettore delle ambiguità di differenze doppie per una sessione su una sola frequenza e relativi fissaggi. Nota: al satellite 23 corrispondono due ambiguità poiché è interessato da un cycle slip non risolto; tutte le ambiguità sono state fissate con successo, eccettuata $N_{1,2}^{17,23}$ che rimane float, poiché nel relativo intervallo di confidenza non sono presenti interi.

Si indica con $\mathbf{N}$ il vettore delle ambiguità vere: in accordo alla teoria della stima, si ha la seguente

$$(\hat{\mathbf{N}} - \mathbf{N})^T \mathbf{C}_{NN}^{-1} (\hat{\mathbf{N}} - \mathbf{N}) \sim F_{p,k} \quad (6.13)$$

ove p è il numero di ambiguità, k sono i gradi di libertà della stima float, $F_{p,k}$ indica la distribuzione di Fisher a p, k gradi di libertà. Per fissare si procede nel seguente modo:

1. si sceglie un livello di significatività α : tipicamente $\alpha = 5\%$;
2. si calcola il valore limite per la $F_{p,k}$: $F_{p,k}(\alpha)$;
3. ciò definisce la regione di confidenza

$$(\hat{\mathbf{N}} - \mathbf{N})^T \mathbf{C}_{NN}^{-1} (\hat{\mathbf{N}} - \mathbf{N}) \leq F_{p,k}(\alpha) \quad (6.14)$$

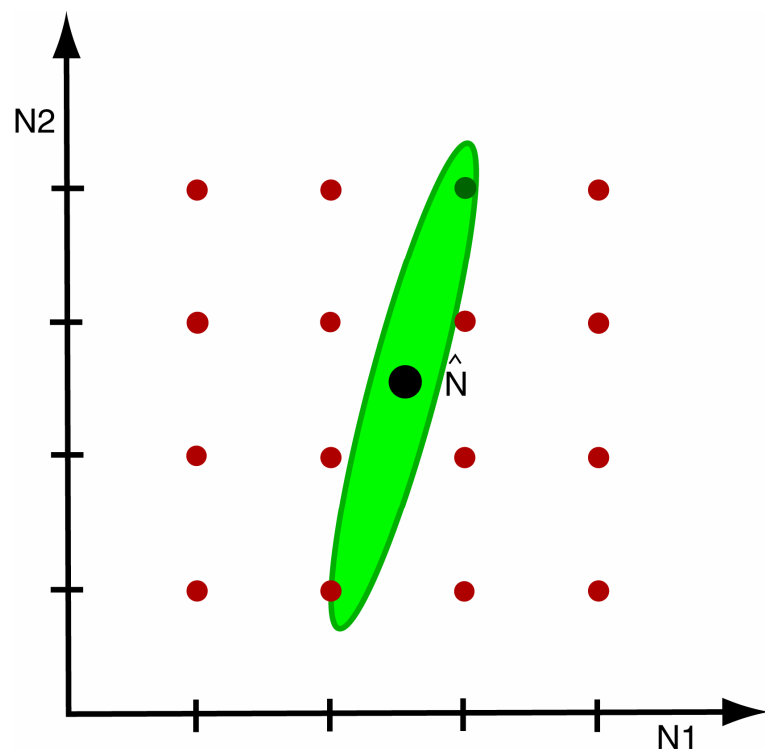


Figura 6.9. Esempio di stima per un vettore di due ambiguità iniziali N_1 e N_2 e di relativa regione di confidenza: gli interi più vicini alle stime reali non sono i candidati più probabili.

Si hanno quindi tre casi:

- 4a. nella regione di confidenza cade uno e un solo vettore $\bar{\mathbf{N}}$ di interi;
- 4b. nella regione di confidenza non è presente alcun vettore di interi;
- 4c. nella regione di confidenza sono presenti più vettori interi.

Nel primo caso le ambiguità vengono fissate a $\bar{\mathbf{N}}$; nel secondo e nel terzo caso si rinuncia al fissaggio o si deve adottare un ulteriore criterio per ottenerlo.

L'approccio ad ambiguità singolarmente prese è troppo semplificato ma permette l'analisi delle ambiguità una per una e risulta molto efficiente da un punto di vista numerico; l'approccio globale è il più rigoroso ma impone una scelta su tutto il vettore delle ambiguità, risultando molto restrittivo e non ottimizzato numericamente. Perciò i programmi di elaborazione implementano approcci di fissaggio che, da un punto di vista del rigore e dell'efficienza, si collocano in una via intermedia fra i due esempi qui illustrati. In ogni caso, al termine del fissaggio delle ambiguità si hanno due casi: quello peggiore in cui nessuna ambiguità è stata fissata con successo e quello tipico, in cui alcune o tutte le ambiguità sono state fissate con successo; in entrambi i casi si compie un ultimo passo di elaborazione, necessario per stimare finalmente la base, con il fine della massima riduzione dei disturbi atmosferici. Nel caso si siano elaborate le osservazioni L5 e L6 e quindi fissate le ambiguità $\bar{N}5$ e $\bar{N}6$, si devono ricavare le ambiguità fissate $\bar{N}1$ e $\bar{N}2$, che servono all'ultimo passo: la trasformazione è data da:

$$\begin{bmatrix} \bar{N}1 \\ \bar{N}2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \bar{N}5 \\ \bar{N}6 \end{bmatrix}$$

Si noti che la divisione per due non compromette in alcun modo l'interezza del risultato: $N5$ e $N6$ hanno la stessa parità, quindi sia la loro somma sia la loro differenza sono sempre pari.

6.10 La stima finale della base

Se nessuna ambiguità è stata fissata si parla di stima finale float, se tutte le ambiguità sono state fissate si parla di stima finale completamente fixed, mentre è corretto riferirsi ai casi intermedi con il termine di stima parzialmente fixed. Nel seguito si ipotizzerà una stima finale completamente fixed, per mantenere semplicità espositiva e concentrare il discorso sulla geometria e i disturbi atmosferici. La stima finale viene effettuata utilizzando le differenze doppie di fase e cercando di minimizzare l'effetto residuo dei disturbi atmosferici; per eliminare il disturbo ionosferico si elabora l'osservabile Ionospheric Free e quindi l'equazione di osservazione delle differenze doppie è

$$\begin{aligned} L3_{1,2}^{1,i}(t) &= \rho_{1,2}^{1,i}(t) + T_{1,2}^{1,i}(t) + \beta_{3_1} \lambda_1 \bar{N}1_{1,2}^{1,i}(t) + \beta_{3_2} \lambda_2 \bar{N}2_{1,2}^{1,i}(t) \\ &= \tilde{\rho}_{1,2}^{1,i}(t) + \tilde{T}_{1,2}^{1,i}(t) + \beta_{3_1} \lambda_1 \bar{N}1_{1,2}^{1,i}(t) + \beta_{3_2} \lambda_2 \bar{N}2_{1,2}^{1,i}(t) \\ &\quad + \mathbf{e}_2^{i,1}(t) \cdot \delta \Delta \mathbf{x}_{1,2} + \delta_{1,2}^{1,i}(t) \\ &= \tilde{b}_{1,2}^{1,i}(t) + \mathbf{e}_2^{i,1}(t) \cdot \delta \Delta \mathbf{x}_{1,2} + \delta_{1,2}^{1,i}(t) \end{aligned} \tag{6.15}$$

ove nell'errore di modello non compare il termine di ionosfera.

6.10.1 La stima finale per le sessioni brevi

Nel caso di sessioni con durata inferiore a due ore, si modella il disturbo troposferico mediante i modelli standard e si inserisce il relativo errore residuo nel bilancio d'errore complessivo della stima. Si risolve quindi il sistema

$$\mathbf{y}_0 = \mathbf{L}_0 - \mathbf{b} = \mathbf{A}\boldsymbol{\xi} + \mathbf{v} \quad (6.16)$$

ove

$$\begin{aligned} \mathbf{L}_{T(I-1)} &= \begin{bmatrix} \mathbf{L}(t_1) \\ \mathbf{L}(t_2) \\ \dots \\ \mathbf{L}(t_T) \end{bmatrix}, \mathbf{L}(t) = \begin{bmatrix} L3_{1,2}^{1,2}(t) \\ L3_{1,2}^{1,3}(t) \\ \dots \\ L3_{1,2}^{1,I}(t) \end{bmatrix}, \boldsymbol{\xi} = \delta\Delta\mathbf{x}_{1,2} \\ \mathbf{A}_{T(I-1) \times 3} &= \begin{bmatrix} \mathbf{E}(t_1) \\ \mathbf{E}(t_2) \\ \dots \\ \mathbf{E}(t_T) \end{bmatrix}, \mathbf{E}(t) = \begin{bmatrix} \tilde{e}_{2_x}^{2,1}(t) & \tilde{e}_{2_x}^{2,1}(t) & \tilde{e}_{2_x}^{2,1}(t) \\ \tilde{e}_{2_x}^{3,1}(t) & \tilde{e}_{2_y}^{3,1}(t) & \tilde{e}_{2_z}^{3,1}(t) \\ \dots & \dots & \dots \\ \tilde{e}_{2_x}^{I,1}(t) & \tilde{e}_{2_y}^{I,1}(t) & \tilde{e}_{2_z}^{I,1}(t) \end{bmatrix}, \mathbf{b}_{T(I-1)} = \begin{bmatrix} \mathbf{b}(t_1) \\ \mathbf{b}(t_2) \\ \dots \\ \mathbf{b}(t_T) \end{bmatrix}, \mathbf{b}(t) = \begin{bmatrix} \tilde{b}_{1,2}^{1,2}(t) \\ \tilde{b}_{1,2}^{1,3}(t) \\ \dots \\ \tilde{b}_{1,2}^{1,I}(t) \end{bmatrix} \end{aligned}$$

6.10.2 La stima finale per le sessioni lunghe

Per sessioni di durata superiore ad alcune ore è possibile un approccio di stima del disturbo residuo troposferico; si riparta dalla scrittura del disturbo troposferico non differenziato come somma di un modello e di un errore di modello, ovvero

$$T_R^S(t) = \tilde{T}_R^S(t) + \delta T_R^S(t)$$

si pone

$$\delta T_R^S(t) = \delta T_R(t) \cdot F(\eta_R^S)$$

ovvero l'errore di modello viene a sua volta parametrizzato come il prodotto fra un disturbo residuo zenitale, dipendente solo dal tempo, e una mapping function, dipendente solo dall'elevazione del satellite: in pratica si pone un'ipotesi di isotropia locale dell'errore. Normalmente la dipendenza dal tempo del disturbo residuo zenitale viene espressa mediante una funzione lineare a tratti, con intervallo di campionamento fisso, ovvero

$$\delta T_R(t) = \begin{cases} \delta T_{R,1} + \dot{T}_{R,1} \cdot (t - t_0) & t_1 \leq t < t_2 \\ \delta T_{R,2} + \dot{T}_{R,2} \cdot (t - t_1) & t_2 \leq t < t_3 \\ \dots & \dots \\ \delta T_{R,K} + \dot{T}_{R,K} \cdot (t - t_K) & t_K \leq t < t_{K+1} \end{cases} \quad (6.17)$$

ove $t_k = t_1 + (k-1)\Delta t$ e K è il numero totale di intervalli; normalmente, nel caso di sessioni lunghe, si effettuano stazionamenti multipli dell'ora e si adottano intervalli di campionamento di un'ora. Apparentemente la funzione lineare a tratti dipende da $2K$ parametri indipendenti, ovvero da tutti i $\delta T_{R,k}$ e i $\dot{T}_{R,k}$; se però si pone la ovvia condizione di continuità nei punti di raccordo

$$\delta T_{R,k} + \dot{T}_{R,k} \Delta t = \delta T_{R,k+1}$$

si hanno solo $K+1$ parametri indipendenti da stimare; ad esempio si possono scegliere i $\delta T_{R,1}, \delta T_{R,2}, \dots, \delta T_{R,K}, \delta T_{R,K+1}$: le derivate nel tempo di ogni intervallo sono dati dalla

$$\dot{T}_{R,k} = (\delta T_{R,k+1} - \delta T_{R,k}) / \Delta t$$

e quindi la relazione funzionale diviene

$$\begin{aligned} \delta T_R^S(t) &= (\delta T_{R,k} + (\delta T_{R,k+1} - \delta T_{R,k})(t - t_k) / \Delta t) \cdot F(\eta_R^S) \\ &= [\delta T_{R,k}(1 - (t - t_k) / \Delta t) + \delta T_{R,k+1}(t - t_k) / \Delta t] \cdot F(\eta_R^S) \end{aligned} \quad (6.18)$$

Andando alle differenze doppie e linearizzandole rispetto alle componenti della base si ha quindi

$$L3_{1,2}^{1,i}(t) = \tilde{b}_{1,2}^{1,i}(t) + \mathbf{e}_2^{i,1}(t) \cdot \delta \Delta \mathbf{X}_{1,2} + F_1^{1,i}(t) \delta T_1(t) - F_2^{1,i}(t) \delta T_2(t) + \delta_{1,2}^{1,i}(t)$$

ove si è adottato l'accorpamento

$$F_R^{1,i}(t) = F(\eta_R^1(t)) - F(\eta_R^i(t))$$

Si danno ora due casi, ovvero quello delle basi corte (di lunghezza inferiore ai 30-40 km) e quello delle basi più lunghe. Nel caso delle basi corte, gli angoli di elevazione con cui i due siti vedono i satelliti sono pressoché uguali: si ha quindi

$$F_1^{1,i}(t) \cong F_2^{1,i}(t) \cong \bar{F}^{1,i}(t)$$

ovvero

$$\begin{aligned} L3_{1,2}^{1,i}(t) &= \mathbf{e}_2^{i,1}(t) \cdot \delta \Delta \mathbf{x}_{1,2} + \bar{F}^{1,i}(t)(\delta T_1(t) - \delta T_2(t)) + b_{1,2}^{1,i}(t) \\ &= \mathbf{e}_2^{i,1}(t) \cdot \delta \Delta \mathbf{x}_{1,2} + \bar{F}^{1,i}(t)\Delta T_{1,2}(t) + b_{1,2}^{1,i}(t) \end{aligned} \quad (6.19)$$

ove si è adottata la notazione

$$\Delta T_{1,2}(t) = \delta T_1(t) - \delta T_2(t)$$

Si noti che per ogni intervallo di campionamento k , si ha

$$\Delta T_{1,2}(t) = \delta T_{1,k} - \delta T_{2,k} + (\dot{T}_{1,k} - \dot{T}_{2,k})(t - t_k)$$

In pratica le osservazioni di differenza doppia non permettono di stimare i disturbi residui zenitali di ogni sito ma solo la loro differenza. Nel caso di basi più lunghe, i coefficienti $F_1^{1,i}(t)$ e $F_2^{1,i}(t)$ differiscono in modo sufficiente da permettere la stima congiunta dei disturbi residui su entrambi i siti. In generale è quindi da risolversi un sistema del tipo

$$\mathbf{y}_0 = \mathbf{L}_0 - \mathbf{b} = \mathbf{A}\xi + \mathbf{v} \quad (6.20)$$

ove $\mathbf{L}_0$ e $\mathbf{b}$ sono identici a quelli del caso delle sessioni brevi; nel caso delle basi corte, la matrice disegno e il vettore dei parametri incogniti sono rispettivamente dati da:

$$\underset{\substack{T(I-1) \\ \times \\ 3+K+1}}{\mathbf{A}} = \begin{bmatrix} \mathbf{A}(t_1) \\ \mathbf{A}(t_2) \\ \dots \\ \mathbf{A}(t_T) \end{bmatrix}, \quad \underset{\substack{(I-1) \\ \times \\ 3+K+1}}{\mathbf{A}(t)} = \begin{bmatrix} \mathbf{E}(t) & \bar{\mathbf{F}}(t) \end{bmatrix}, \quad \underset{3+K+1}{\xi} = \begin{bmatrix} \delta \Delta \mathbf{x}_{1,2} \\ \Delta \mathbf{T}_{1,2} \end{bmatrix}, \quad \Delta \mathbf{T}_{1,2} = \begin{bmatrix} \delta T_{1,1} - \delta T_{2,1} \\ \delta T_{1,2} - \delta T_{2,2} \\ \dots \\ \delta T_{1,K+1} - \delta T_{2,K+1} \end{bmatrix}$$

$\mathbf{E}(t)$ contiene le combinazioni di versori già viste; $\bar{\mathbf{F}}(t)$ contiene, nelle opportune posizioni di riga e colonna, le relazioni fra le osservazioni di differenza doppia all'epoca t e le relative differenze doppie di disturbo residuo zenitale; nel caso delle basi lunghe, la matrice disegno e il vettore dei parametri incogniti sono rispettivamente

$$\underset{\substack{T(I-1) \\ \times \\ 3+2(K+1)}}{\mathbf{A}} = \underset{\substack{(I-1) \\ \times \\ 3+2(K+1)}}{\begin{bmatrix} \mathbf{A}(t_1) \\ \mathbf{A}(t_2) \\ \dots \\ \mathbf{A}(t_T) \end{bmatrix}}, \quad \underset{\substack{(I-1) \\ \times \\ 3+2(K+1)}}{\mathbf{A}(t)} = \underset{\substack{(I-1) \times (K+1) \\ (I-1) \times (K+1)}}{\begin{bmatrix} \mathbf{E}(t) & \mathbf{F}_1(t) & -\mathbf{F}_2(t) \end{bmatrix}}, \quad \underset{3+2(K+1)}{\boldsymbol{\xi}} = \underset{\substack{\delta \Delta \mathbf{x}_{1,2} \\ \mathbf{T}_1 \\ \mathbf{T}_2}}{\begin{bmatrix} \delta \Delta \mathbf{x}_{1,2} \\ \mathbf{T}_1 \\ \mathbf{T}_2 \end{bmatrix}}, \quad \underset{\substack{\delta T_{R,1} \\ \delta T_{R,2} \\ \dots \\ \delta T_{R,K+1}}}{\mathbf{T}_R} = \begin{bmatrix} \delta T_{R,1} \\ \delta T_{R,2} \\ \dots \\ \delta T_{R,K+1} \end{bmatrix},$$

in questo caso $\mathbf{F}_R(t)$ contiene, nelle opportune posizioni di riga e colonna, le relazioni fra osservazioni di differenza doppia all'epoca t e relative incognite di disturbo residuo zenitale. Si noti che in questi sistemi il vettore degli errori di modello non contiene più l'errore residuo troposferico, poiché questo è stato incluso fra i parametri incogniti, mentre rimane un errore di secondo ordine, dovuto alla semplificazione nella modellizzazione degli errori troposferici residui.

6.10.3 Le accuratze tipicamente ottenibili

In funzione della lunghezza della base e del tempo di rilievo, nella Tab. 6.3 sono indicate alcune accuratze tipicamente conseguibili mediante postelaborazione di un rilievo statico, effettuato da ricevitori di fase a doppia frequenza.

In sostanza, a pari tempo di rilievo, l'accuratezza dipende dalla lunghezza della base e non è possibile stimare basi molto lunghe con sessioni brevi; per contro, con lunghi tempi di rilievo, sono comunque ottenibili accuratze migliori del centimetro anche per basi di scala continentale. Si notano le seguenti problematiche, raccolte per tipologia di sessione e di base.

Distanza/Tempo	1 km	10 km	20 km	100 km	1000 km
10 minuti	1.5 cm	2.5 cm	5 cm	ns	ns
1 ora	1.0 cm	1.5 cm	2 cm	5 cm	ns
24 ore	0.3 cm	0.5 cm	< 1 cm	1 cm	1.5 cm
1 settimana	0.1 cm	0.1 cm	0.3 cm	0.3 cm	<1 cm

Tabella 6.3. Accuratze ottenibili nel posizionamento statico relativo mediante osservazioni di fase in doppia frequenza; in funzione di distanza fra i ricevitori e di durata del rilievo. ns: non stimabile. Le accuratze indicate per le basi di 100 e 1000 km richiedono l'utilizzo di programmi scientifici, i quali implementano algoritmi significativamente più complessi di quelli qui discussi.

1. Sessioni statico rapide, $T \leq 15 \text{ min}$, per basi corte, $l \leq 10 \text{ km}$: la possibilità di ottenere le massime accuratze è strettamente legata al completo fissaggio delle ambiguità, poiché il sistema cosiddetto float è debolmente condizionato e la sua parziale instabilità numerica risulta in una mediocre stima della base. Per contro, in basi così corte, i disturbi atmosferici sono tipicamente trascurabili e non inficiano il buon esito del fissaggio; l'eventuale insuccesso di quest'ultimo è quindi dovuto a scarsa qualità delle osservazioni per uno dei due siti o a una pessima configurazione satellitare.

2. Sessioni statico rapide, $T \leq 30 \text{ min}$, per basi di media lunghezza, $10 \text{ km} < l \leq 50 \text{ km}$. Qui i disturbi atmosferici possono assumere valori significativi e, in condizioni di mediocre configurazione satellitare, è possibile il fallimento del fissaggio: in tal caso le soluzioni float sono decisamente mal condizionate, con accuratezze che possono degradare fino al decimetro.
3. Sessioni statico rapide, $T \leq 30 \text{ min}$, basi più lunghe. I disturbi atmosferici assumono valori decisamente significativi e in ogni caso, con così poche epoche, non ci si può attendere il fissaggio delle ambiguità e quindi una stima di accuratezza accettabile.
4. Sessioni statiche, $30 \text{ min} < T \leq 6 \text{ h}$ per basi sino alla media lunghezza, $l \leq 50 \text{ km}$. In tal caso l'effetto dei disturbi atmosferici inizia a mediarsi sulle epoche e ci si attende il fissaggio delle ambiguità: l'eventuale insuccesso può denotare qualche problema nelle osservazioni ma la geometria della stima float è ben condizionata e non impedisce accuratezze accettabili.
5. Sessioni statiche, $30 \text{ min} < T \leq 6 \text{ h}$ per basi più lunghe, $50 \text{ km} < l$. In tal caso si possono riproporre le stesse problematiche delle sessioni statico rapide per le basi con $10 \text{ km} < l$.
6. Sessioni statiche prolungate, $6 \text{ h} < T$. Per basi sino alla media lunghezza, $l \leq 50 \text{ km}$, tali sessioni garantiscono già accuratezze migliori del centimetro; per le basi con $l > 10 \text{ km}$ è necessaria la stima del disturbo residuo zenitale, perlomeno relativo. Anche per le basi più lunghe, con $l > 50 \text{ km}$, tali sessioni forniscono risultati di sicura accuratezza.
7. Serie di acquisizioni continue, $24 \text{ h} < T$. Le elaborazioni forniscono accuratezze potenzialmente migliori del centimetro anche nel caso di basi molto lunghe, ovvero $l \geq 100 \text{ km}$; per l'elaborazione di queste basi è però necessaria l'adozione degli standard di elaborazione scientificamente riconosciuti a livello internazionale, significativamente più complessi di quelli esposti in questo libro.

Si nota da ultimo che per basi molto corte, ovvero con $l \leq 5 \text{ km}$, l'errore ionosferico residuo sulle differenze doppie in singola epoca è tipicamente inferiore al centimetro. In tali condizioni è opportuno utilizzare le sole osservazioni di fase su L1, senza ricorrere alla combinazione Ionospheric Free che è più rumorosa; per contro, la stima e il fissaggio delle ambiguità risultano più robusti se effettuati congiuntamente su entrambe le frequenze piuttosto che sulla sola L1 e ciò anche per la maggior ridondanza del sistema. Perciò, tipicamente, per basi molto corte viene implementato il seguente approccio: 1) preelaborazione, stima e fissaggio delle ambiguità su entrambe le frequenze; 2) stima finale fixed utilizzando la sola osservabile L1.

6.11 L'elaborazione di un rilievo cinematico

Il termine cinematico caratterizza un rilievo in cui il ricevitore incognito abbia occupato una posizione diversa ad ogni epoca di osservazione: il fine ultimo di un'elaborazione cinematica è quindi quello di stimare la sua traiettoria, ad ambiguità fissate e quanto più possibile libera dagli errori residui atmosferici. In questo paragrafo si vuole dare solo

una sintetica descrizione dell'elaborazione cinematica delle osservazioni di fase, senza entrare in dettagli tecnici; ciò anche perché l'elaborazione cinematica è tipicamente implementata in programmi commerciali, i cui algoritmi non sono completamente pubblici. Le differenze doppie disponibili ad ogni epoca t per una coppia di ricevitori a doppia frequenza vengono accorpate nel vettore $\mathbf{y}(t)$ già introdotto

$$\mathbf{y}_{4(I-1)}(t) = \begin{bmatrix} \mathbf{y}_{1,2}^{1,2}(t) \\ \mathbf{y}_{1,2}^{1,3}(t) \\ \dots \\ \mathbf{y}_{1,2}^{1,I}(t) \end{bmatrix}, \quad (6.21)$$

$$\mathbf{y}_{1,2}^{1,i}(t) = \begin{bmatrix} P1_{1,2}^{1,i}(t) \\ P2_{1,2}^{1,i}(t) \\ L1_{1,2}^{1,i}(t) \\ L2_{1,2}^{1,i}(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & f_1^2 / f_2^2 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -f_1^2 / f_2^2 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_{1,2}^{1,i}(t) \\ I1_{1,2}^{1,i}(t) \\ N1_{1,2}^{1,i}(t) \\ N2_{1,2}^{1,i}(t) \end{bmatrix} = \mathbf{A} \boldsymbol{\xi}_{1,2}^{1,i}(t)$$

Con almeno quattro satelliti è possibile risolvere il sistema in singola epoca rispetto alla base: tale approccio permette la stima delle ambiguità e non risente dell'errore ionosferico, che viene stimato come incognita, però risente del rumore di osservazione dei codici, che può impedire il fissaggio delle ambiguità. Nell'elaborazione cinematica si ricorre a un approccio basato sui seguenti passaggi:

1. le osservazioni di differenza doppia relative a codici e fasi, per tutti i satelliti e tutte le epoche di un rilievo vengono impilate nel vettore delle osservazioni:

$$\mathbf{y}_{4T(I-1)} = \begin{bmatrix} \mathbf{y}(t_1) \\ \mathbf{y}(t_2) \\ \dots \\ \mathbf{y}(t_T) \end{bmatrix}$$

2. il vettore viene utilizzato per:

- ottenere una prima stima della traiettoria,
- identificare e rimuovere gli eventuali cycle slip,
- stimare e fissare le ambiguità intere iniziali.

3. Una volta terminata la preelaborazione, supponendo di aver ottenuto il completo fissaggio delle ambiguità, le differenze doppie di fase Ionospheric Free vengono utilizzate per la stima finale della traiettoria, epoca per epoca secondo il modello

$$\begin{aligned}
\mathbf{L}(t) &= \begin{bmatrix} L3_{1,2}^{1,2}(t) \\ L3_{1,2}^{1,3}(t) \\ \dots \\ L3_{1,2}^{1,I}(t) \end{bmatrix} = \mathbf{E}(t) \delta \Delta \mathbf{x}_{1,2}(t) + \mathbf{b}(t) + \boldsymbol{\varepsilon}(t), \\
\mathbf{E}(t) &= \begin{bmatrix} \tilde{e}_{2_x}^{2,1}(t) & \tilde{e}_{2_x}^{2,1}(t) & \tilde{e}_{2_x}^{2,1}(t) \\ \tilde{e}_{2_x}^{3,1}(t) & \tilde{e}_{2_y}^{3,1}(t) & \tilde{e}_{2_z}^{3,1}(t) \\ \dots & \dots & \dots \\ \tilde{e}_{2_x}^{I,1}(t) & \tilde{e}_{2_y}^{I,1}(t) & \tilde{e}_{2_z}^{I,1}(t) \end{bmatrix}, \quad \mathbf{b}(t) = \begin{bmatrix} \tilde{\rho}_{1,2}^{1,2}(t) + \tilde{T}_{1,2}^{1,2}(t) + \lambda_3 \bar{N}3_{1,2}^{1,2} \\ \tilde{\rho}_{1,2}^{1,3}(t) + \tilde{T}_{1,2}^{1,3}(t) + \lambda_3 \bar{N}3_{1,2}^{1,3} \\ \dots \\ \tilde{\rho}_{1,2}^{1,I}(t) + \tilde{T}_{1,2}^{1,I}(t) + \lambda_3 \bar{N}3_{1,2}^{1,I} \end{bmatrix}
\end{aligned} \tag{6.22}$$

Si noti che, se la base è molto corta, nella stima finale si può utilizzare L1 anziché L3. Il processo appare molto simile a quello relativo all'elaborazione statica e in effetti il flusso logico lo è; però, da un punto di vista formale, numerico e tecnico, vi sono fondamentali differenze.

1. Nel caso cinematico la base fra i due ricevitori cambia di epoca in epoca, e quindi la stima iniziale fornita dai codici ha minore accuratezza rispetto al caso statico: sono necessari algoritmi più complessi per l'identificazione e la stima dei cycle slip rispetto all'esempio discusso nel Par. 6.9.
2. La stima dei cycle slip è più difficile e si è spesso costretti, in corrispondenza di essi, a settare nuove ambiguità iniziali: il sistema che ne risulta è significativamente meno ridondante rispetto al caso statico, e questo comporta una maggiore difficoltà nel fissaggio delle ambiguità.
3. Il sistema finale è poco ridondante e quindi poco robusto rispetto a eventuali errori grossolani nelle osservazioni: l'accuratezza e l'affidabilità della soluzione dipendono significativamente dalla qualità delle osservazioni.

In particolare si consideri il caso di insuccesso parziale o totale nel fissaggio delle ambiguità: da un punto di vista strettamente numerico il sistema float, sia L3 sia L1, costruito su tutte le epoche è comunque risolvibile quando vi sono più osservazioni che incognite: per ogni epoca le osservazioni sono $I-1$, le incognite sono le $I-1$ ambiguità iniziali e le tre componenti della base per ogni epoca; si deve quindi avere

$$(I-1)T \geq I-1+3T$$

tale condizione non è mai soddisfatta se sono in vista solo quattro satelliti, ma può rapidamente essere soddisfatta con più satelliti; ad esempio, nel caso di cinque satelliti, quattro epoche sono sufficienti; con dieci satelliti due epoche garantiscono già la ridondanza. In realtà la soluzione ad ambiguità float nel caso cinematico accentua il malcondizionamento del caso statico rapido e risulta debolmente condizionata: anche in assenza di altri problemi, la sua accuratezza non è migliore del decimetro: per ottenere accuratezze dell'ordine che vanno da meglio del decimetro ad alcuni centimetri, a

seconda della qualità delle osservazioni, è necessario il successo del fissaggio delle ambiguità e quindi la soluzione fixed.

6.12 Il controllo dei risultati di una postelaborazione di base

Ogni elaborazione fornisce anche le stime di varianza della base; si considerino però gli errori di modello, come gli errori residui ionosferico e troposferico, l'eventuale cattiva modellizzazione dell'offset d'antenna, il multipath che, soprattutto per rilievi molto brevi, costituiscono sistematismi: questi non sono evidenziati nei residui, ovvero nella stima della varianza a posteriori e quindi le stime di deviazione standard sono tipicamente irrealistiche. Nella consapevolezza di tale problema, alcuni programmi moltiplicano le stime di varianza per coefficienti correttivi empiricamente determinati e cablati nel programma stesso: tale scelta è sostanzialmente arbitraria perché non tiene conto delle reali condizioni del rilievo. L'unico modo per stimare rigorosamente le accuratezze è quella di effettuare rilievi secondo schemi ridondanti e quindi analizzare la ripetibilità e compensare rigorosamente le basi, come verrà discusso nel Cap. 8.

Ogni programma fornisce poi uno o più indicatori di qualità che assumono spesso nomi esotici, ma tipicamente sono derivati da indici statistici rigorosi, come ad esempio il test del χ^2 ; una discreta esperienza pregressa permette all'utente la valutazione di tali indici e quindi la validazione dei risultati. Le cause di eventuali problemi possono essere diverse: ambiguità iniziali fissate male, cycle slip non identificati o risolti male, interferenze e disturbi anomali per qualche satellite. In presenza di problemi significativi, si può sperare che essi siano imputabili alla scarsa qualità delle osservazioni solo di un particolare satellite o solo di un particolare periodo della sessione e le possibili vie per verificare ciò sono:

1. escludere dall'elaborazione i satelliti iterativamente uno a uno, controllando di volta in volta i risultati;
2. suddividere la sessione in sottosessioni, elaborare in modo indipendente le sottosessioni e controllare i risultati di ogni sottosessione.

Nel caso di un rilievo statico rapido o cinematico, tali approcci abbattano la ridondanza, quindi le soluzioni parziali risultano deboli e non sempre si riesce a isolare i problemi; inoltre vi è il caso più sfortunato di problemi distribuiti omogeneamente sulla sessione e sui satelliti: si è quindi costretti a ripetere il rilievo e rielaborare i dati. La più efficace analisi di qualità viene fornita da schemi di rilievo ridondanti e dalle successive analisi di ripetibilità e compensazioni di rete.

6.13 Il rilievo in tempo reale mediante osservazioni di fase

Il rilievo statico rapido e quello cinematico vennero proposti come evoluzione dallo statico tradizionale negli anni '80; le problematiche relative all'elaborazione dei dati – in particolare il fissaggio delle ambiguità – vennero sviscerate nel decennio successivo e quindi implementate nei programmi di postelaborazione. Parallelamente venne

introdotta l'idea del rilievo statico rapido o cinematico in tempo reale, di cui si semplifica il processo.

1. Il ricevitore di riferimento e quello mobile acquisiscono i dati dalla costellazione GPS;
2. i dati di entrambi, ad ogni epoca e appena acquisiti, vengono inviati a una stazione di elaborazione.
3. La stazione di elaborazione elabora i dati contemporaneamente alla loro disponibilità, senza attendere il termine della sessione e utilizzando algoritmi ottimizzati per non accumulare latenza.
4. Le prime epoche vengono elaborate per arrivare al fissaggio delle ambiguità iniziali: il processo, che risulta più veloce se il ricevitore mobile rimane fermo fino al successo nel fissaggio, prende nome di inizializzazione del rilievo.
5. Una volta inizializzato il rilievo le differenze doppie di fase vengono utilizzate per stimare le posizioni del ricevitore mobile in (quasi) tempo reale, epoca per epoca: se il ricevitore mobile è fermo su un punto, si può parlare di rilievo statico rapido in tempo reale, se il ricevitore mobile è effettivamente in moto, si parla di rilievo cinematico.
6. Parallelamente alla stima delle coordinate viene effettuato un ciclo di controllo per identificare eventuali cycle slip: in caso di tale evenienza, è necessaria una nuova inizializzazione delle nuove ambiguità iniziali.

Non si vuole qui discutere tutti i problemi del tempo reale, ma solo ricordarne alcuni aspetti essenziali. Tipicamente la stazione di elaborazione dei dati coincide fisicamente con il ricevitore mobile, che riceve i dati della stazione di riferimento, effettua l'elaborazione, visualizza all'operatore e memorizza internamente la propria posizione e la traiettoria. La trasmissione fra i due ricevitori può avvenire attraverso canali radio o di telefonia cellulare: il primo caso è tipico quando entrambi i ricevitori sono operati dal medesimo ente per un rilievo su scala molto locale, come ad esempio nei lavori di cantiere; il secondo è quello tipico delle stazioni di riferimento installate per distribuire dati a più utenti, non necessariamente coordinati, come ad esempio le stazioni permanenti dei servizi di posizionamento.

L'elaborazione dei dati e la stima delle incognite può avvenire mediante un filtro di Kalman o mediante minimi quadrati sequenziali, come sono sinteticamente descritti nel Par. A1.3. Il rilievo in tempo reale esaspera le problematiche già evidenziate per la postelaborazione del cinematico e dello statico rapido, sostanzialmente per tre motivi specifici; innanzitutto la comunicazione fra i due ricevitori può subire eventuali interferenze, che possono causare la perdita, totale o parziale, di blocchi di epoche; inoltre vi è necessità di ottimizzare numericamente i passi dell'elaborazione dei dati, in modo da renderli paralleli al flusso di acquisizione degli stessi; infine gli algoritmi di stima devono procedere in modo sequenziale, ovvero impilando man mano il vettore delle osservazioni: in sostanza si ha un aumento del rischio di non identificare correttamente i cycle slip, di insuccesso nel fissaggio delle ambiguità e della conseguente scarsa qualità delle soluzioni.

Si forniscono in Tab. 6.4 alcune valutazioni empiriche, ovvero basate su numerose sperimentazioni condotte negli ultimi anni; per un'interdistanza fra ricevitore di riferimento e ricevitore mobile inferiore a 10 km, il fissaggio ha tipicamente successo e la soluzione finale garantisce accuratezze compatibili con i tipici requisiti topocartografici, ovvero dell'ordine di alcuni centimetri in planimetria; per interdistanze fra 10 km e 20 km, il fissaggio ha successo con una discreta configurazione di satelliti: in caso di successo le accuratezze sono simili alle precedenti; è difficile prevedere l'esito di un rilievo a interdistanze superiori.

Il tempo reale presenta un notevole vantaggio rispetto alla postelaborazione, che ne ha dettato la grande diffusione per applicazioni topocartografiche: quando si pianificano rilievi finalizzati alla postelaborazione dei dati è necessario prevedere tempi di stazionamento conservativi, ovvero superiori a quanto strettamente necessario, per essere sicuri del successo dell'elaborazione; viceversa, nel tempo reale la qualità dei dati e il tempo effettivamente necessario per l'inizializzazione vengono implicitamente controllati durante il rilievo stesso: perciò si minimizza sia il tempo di rilievo sia il rischio di dover ritornare sul punto.

Lunghezza base	Esito del fissaggio	Tempo di inizializzazione	Accuratezza fixed
0-10 km	Sostanzialmente garantito	1-2 min	2-4 cm
10-15 km	Con configurazione di satelliti normale, garantito	1-3 min	3-5 cm
15-20 km	Con buona configurazione di satelliti, garantito	1-4 min	3-6 cm
>20 km	Non predicibile		≤ 10 cm

Tabella 6.4. Tipici esiti di un rilievo in tempo reale. Le probabilità di esito del fissaggio e le seguenti indicazioni di accuratezza sono basate sull'ipotesi di buona qualità delle osservazioni e della comunicazione fra i due ricevitori; nel caso contrario, difficilmente si ottiene il fissaggio.

6.14 Il sistema di riferimento per il posizionamento relativo

Per stimare la posizione del ricevitore incognito si vincola la posizione della stazione di riferimento: si discute ora la scelta del sistema di riferimento in cui questa viene vincolata, ovvero la scelta del riferimento a terra. Le effemeridi dei satelliti sono in IGS, o in WGS84 che gli è sostanzialmente uguale, e queste costituiscono il sistema di riferimento spaziale: se anche il sistema di riferimento a terra è in IGS vi è piena coerenza, la base è stimata in IGS e quindi lo sono anche le coordinate del ricevitore incognito. Se invece il sistema di riferimento a terra è diverso da IGS si ha incongruenza fra vincoli spaziali e terrestri e quindi la stima della base ne risulta distorta, in ragione della propagazione discussa nel Par. 6.4. Si considerino dunque i

sistemi di riferimento diversi da IGS ma comunque utili nelle operazioni di rilievo GPS, iniziando da quello continentale ETRS89 e le sue varie realizzazioni, come IGM95; nel peggiore dei casi le loro differenze rispetto a IGS sono entro alcuni decimetri, diciamo ad esempio per l'Italia quattro decimetri: la propagazione rigorosa fornisce perciò

$$\delta_{\Delta x} \cong 4 \cdot 10^{-1} m / (2 \cdot 10^7 m) \Delta x = 2 \cdot 10^{-8} \Delta x$$

l'uso di coordinate ETRS89 per la stazione di riferimento provoca distorsioni trascurabili alla stima delle basi, inferiori al millimetro per basi sino al centinaio di chilometri: in tal senso si può dire che la stima della base è sostanzialmente invariante rispetto alla scelta del sistema di riferimento terrestre e quindi la stazione incognita viene stimata nel sistema di riferimento di quella vincolata.

Si consideri ora il caso dei sistemi di riferimento planoaltimetrici, come Roma40, o catastali, che possono differire anche di centinaia di metri dalle realizzazioni di IGS o ETRS; il loro utilizzo per vincolare la stazione di riferimento di una elaborazione comporta distorsioni inaccettabili, ad esempio di 10 cm per basi di 10 km: i risultati ottenuti da tale prassi sono quindi inutilizzabili. La prassi corretta per ottenere risultati in tali sistemi di riferimento è quella di stimare le coordinate nei sistemi di riferimento propri del GPS, ovvero IGS o ETRS89, e quindi di convertirle ai sistemi di interesse, mediante le opportune trasformazioni.

Tipo rilievo	Osservabili	Requisiti e note	Ordine di accuratezza
Posizionamento cinematico in tempo reale	Codici in singola frequenza	Interdistanza fra stazione di riferimento e mobile indicativamente entro 100 km: Par. 6.8	Uno o due metri
	Codici in doppia frequenza		Meglio del metro
	Codici in doppia frequenza lisciati con le fasi (filtro di Hatch)		Alcuni decimetri
	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 20 km oppure presenza di un servizio di posizionamento: Par 6.13	Con successo del fissaggio delle ambiguità: alcuni centimetri/ meglio del decimetro
Posizionamento cinematico mediante postelaborazione	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 20 km oppure presenza di un servizio di posizionamento: Parr. 6.11 e 6.13	Con successo del fissaggio delle ambiguità: alcuni centimetri
Posizionamento statico rapido (pochi minuti) in tempo reale			
Posizionamento statico rapido mediante postelaborazione	Fasi in singola frequenza	Interdistanza fra stazione di riferimento e mobile entro 5 km: Par. 6.10 (6.10.3)	Con successo del fissaggio delle ambiguità: alcuni centimetri
	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 20 km: Par. 6.10 (6.10.3)	
Posizionamento statico di un'ora mediante postelaborazione	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile entro 100 km Par. 6.10 (6.10.3)	Alcuni centimetri/ meglio del decimetro
		Interdistanza fra stazione di riferimento e mobile entro 10 km Par. 6.10 (6.10.3)	Dell'ordine del centimetro
Posizionamento statico con occupazione permanente mediante postelaborazione	Fasi in doppia frequenza	Interdistanza fra stazione di riferimento e mobile anche superiore a 100 km Par. 6.10 (6.10.3)	Meglio del centimetro

Tabella 6.5. Sintesi dei risultati ottenibili dal posizionamento relativo; vengono richiamati i paragrafi in cui i singoli approcci vengono discussi con maggior dettaglio.

Capitolo 7.

I servizi di posizionamento

Come scritto nell'introduzione, i problemi fondamentali del rilievo delle basi mediante GPS sono di costo e di complessità: tali problemi sono particolarmente limitanti nelle applicazioni topografiche alla scala locale, ove gli strumenti topografici tradizionali pongono meno problemi di costo e complessità. In particolare è emersa nel precedente capitolo l'idea che per tali applicazioni sia utile ricorrere alla tecnica del rilievo in tempo reale, il quale però presuppone piccola interdistanza fra stazione di riferimento e stazione mobile. Per ovviare ai problemi e per fornire supporto al rilievo in tempo reale è nata e si è sviluppata in molte nazioni l'idea di reti di stazioni permanenti GNSS per servizi di posizionamento. Il presente capitolo si apre con una discussione generale delle reti di stazioni permanenti e con una loro categorizzazione in funzione delle applicazioni cui si rivolgono. Le reti di stazioni permanenti per il monitoraggio geodinamico e locale, che pure erano già ampiamente sviluppate al termine degli anni '90, non sono strutturalmente disegnate per fornire servizi di posizionamento in tempo reale, che quindi si sono sviluppate solo in seguito: tale aspetto è chiarito all'inizio del capitolo. La descrizione dei principi di funzionamento di un servizio di posizionamento e dei prodotti che esso eroga occupano la parte principale del capitolo; questo si chiude poi con alcune indicazioni relative ai sistemi di riferimento distribuiti da un servizio di posizionamento e alcuni suggerimenti operativi per l'utenza.



Figura 7.1. Una stazione permanente: l'esempio della stazione permanente di Como, inserita nella rete europea EPN e nel servizio di posizionamento di Regione Lombardia.

7.1 Le stazioni permanenti

Una stazione permanente GNSS si compone di un ricevitore e della relativa antenna, in acquisizione 24 ore su 24, monumentata stabilmente; il ricevitore può essere controllato localmente da un PC o connesso remotamente a un centro di controllo. Una rete di stazioni permanenti (nel seguito rete permanente) si compone dei seguenti elementi minimi:

- un insieme di stazioni permanenti con distribuzione omogenea sulla porzione di territorio interessata;
- uno o più centri di controllo che:
 - gestiscono le stazioni permanenti e ne garantiscono la manutenzione;
 - ne assimilano i dati;
 - monitorano la qualità degli stessi e li elaborano;
 - stimano le coordinate delle stazioni permanenti;
 - eventualmente, distribuiscono dati e prodotti all'utenza.

Gli scopi di una rete permanente possono essere molteplici, e li elenchiamo in ordine di gerarchia spaziale e storica:

- il controllo globale del sistema di riferimento e della geodinamica,
- il controllo continentale del sistema di riferimento e della geodinamica,
- il controllo dalla scala nazionale a quella strettamente locale, dalla geodinamica al controllo di frane, strutture,
- i servizi di posizionamento.

Nella trattazione dei sistemi di riferimento abbiamo già citato le reti IGS e EPN come esempio rispettivamente di rete del primo e del secondo tipo. Non si discutono qui le reti permanenti per il controllo locale, perché non sono rilevanti alla presente esposizione, ma se ne citano alcuni esempi, rimandando alla relativa bibliografia; la rete RING, monumentata dall'Istituto Nazionale di Geofisica e Vulcanologia (INGV) è un esempio di rete per il controllo geodinamico alla scala della nazione; la rete Alps-GPSQuakeNet, progetto capitanato dall'Università di Trieste, Dipartimento di Scienze della Terra, si pone lo scopo del monitoraggio alla scala dell'arco alpino.

Le reti di stazioni permanenti per servizi di posizionamento – nel seguito denominate servizi di posizionamento - hanno lo scopo di favorire l'utilizzo dei metodi GNSS per le applicazioni di rilevamento, controllo del territorio e posizionamento in generale; per fare ciò devono permettere agli utenti l'utilizzo del GNSS nella modalità del posizionamento relativo, garantendone i vantaggi, ovvero alte accuratezze e bassi tempi, ma eliminandone gli svantaggi, ovvero i già citati fattori di costo e complessità. A un primo livello quindi un servizio di posizionamento deve perlomeno distribuire coordinate e dati delle proprie stazioni, che fungano da riferimento per l'utenza e deve effettuare gli opportuni controlli di qualità sui dati distribuiti; inoltre i servizi di posizionamento dovrebbero poter certificare l'accuratezza, almeno potenziale, dei risultati della propria utenza e devono svolgere attività di supporto e formazione: quindi

un servizio di posizionamento si compone di una rete di stazioni permanenti ma anche, necessariamente, di un centro qualificato di controllo, di calcolo e di supporto all'utenza.

7.2 I servizi di rete per il tempo reale

Perché un servizio di posizionamento svolga un ruolo veramente esaustivo verso l'utenza, la distribuzione delle sue stazioni deve garantire accuratezze topocartografiche in tempi ragionevoli di rilievo e in particolare dovrebbe essere garantita la possibilità del rilievo in tempo reale: questo, ricordando le tabelle e le discussioni del paragrafo precedente, implicherebbe una geometria delle stazioni a maglia regolare, con interdistanza reciproca di circa 20 km, in modo da avere una distanza massima di 15 km di ogni punto del territorio dalla stazione più vicina. Tale distribuzione significa una stazione ogni 400 km²: per una regione di 24000 km² come la Lombardia, sarebbero necessarie 60 stazioni, ovvero un costo e una complessità di gestione difficilmente giustificabili rispetto al potenziale mercato. Negli anni '90 si è iniziato lo studio degli approcci di rete per la distribuzione di servizi aggiunti rispetto alla semplice distribuzione di dati e coordinate: tali approcci hanno appunto lo scopo di permettere un diradamento delle stazioni senza pregiudizio per l'affidabilità e accuratezza dal lato dell'utente.

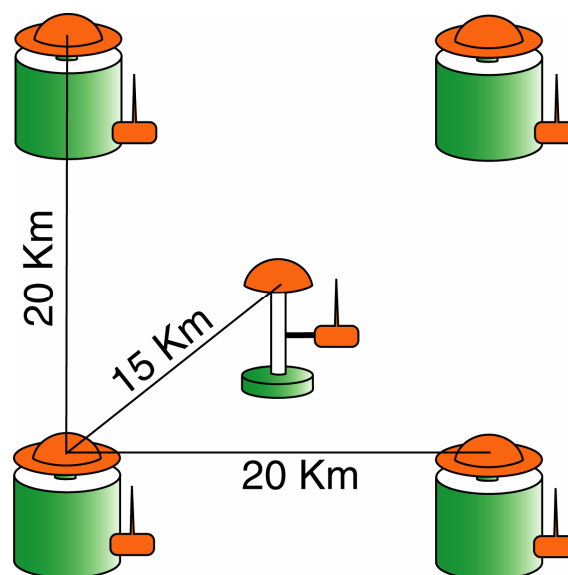


Figura 7.2. Interdistanza delle stazioni permanenti per garantire il tempo reale mediante posizionamento da singola stazione.

7.3 L'approccio di rete alla modellizzazione degli errori

Si consideri una rete permanente le cui coordinate siano state stimate mediante inquadramento nella rete IGS e che quindi possano essere considerate note con pochi

millimetri di incertezza. La rete può ribaltare il concetto geodetico classico: le stazioni acquisiscono le osservazioni e le inviano al centro di controllo; il centro di controllo considera note le coordinate delle stazioni, le vincola ed elabora le loro osservazioni per:

1. stimare gli errori e i disturbi sulle singole stazioni,
2. modellizzare gli errori e i disturbi nel tempo e nello spazio,
3. distribuire all'utenza i dati e i modelli, in modo che questa possa correggere i propri dati.

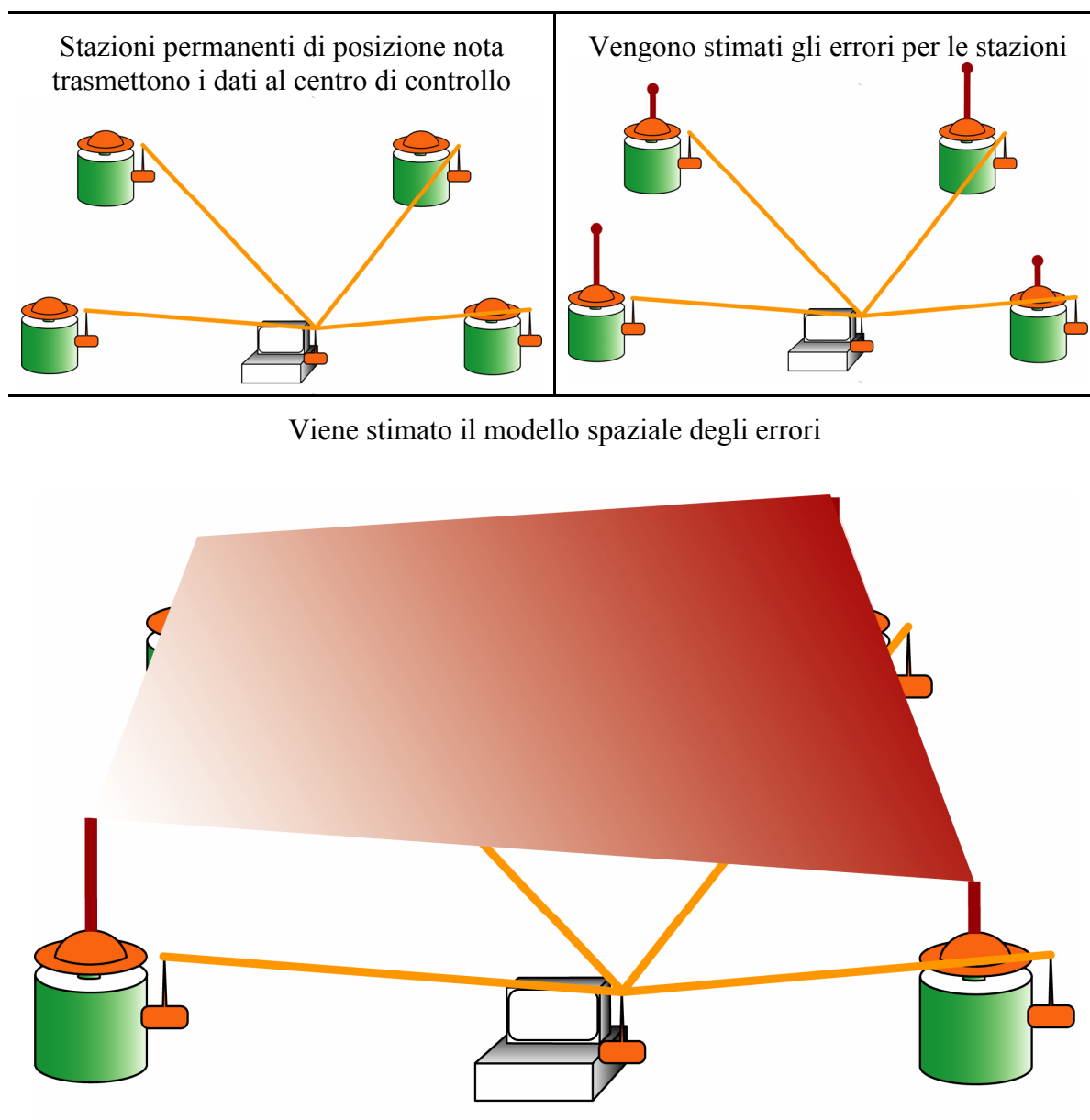


Figura 7.3. Approccio di rete alla modellizzazione spaziale degli errori.

Esistono diversi approcci per il calcolo dei modelli di errore e disturbo, che sono molto complessi e non completamente pubblicati, poiché costituiscono il cuore di programmi distribuiti commercialmente; nel seguito, senza volontà di completezza ma con lo scopo di semplificazione, viene schematizzato un semplice schema di flusso di calcolo. Ricordiamo che l'errore residuo di differenza doppia è dato dalla differenza di errori di modello

$$\delta_{1,2}^{i,j} = f(\delta(\tilde{\mathbf{x}}^i), \delta(\tilde{\mathbf{x}}^j), \delta(\tilde{I}), \delta(\tilde{T})) = \delta_1^i - \delta_2^i - (\delta_1^j - \delta_2^j)$$

ciascun termine d'errore può essere scritto come

$$\delta_R^S(t) = \delta_{\rho_R^S}(t) + \delta_{\tilde{T}_R^S}(t) + \delta_{\tilde{I}_R^S}(t)$$

per l'effetto dell'errore delle effemeridi si può inoltre scrivere

$$\delta_{\rho_R^S}(t) = \mathbf{e}_R^S(t) \cdot \boldsymbol{\delta}^S(t)$$

ove: $\delta_{\tilde{T}_R^S}(t)$ è l'errore nel modello troposferico, $\delta_{\tilde{I}_R^S}(t)$ è l'errore nel modello ionosferico, $\boldsymbol{\delta}^S(t)$ è l'errore nelle effemeridi del satellite. Nel caso del calcolo di rete dei disturbi e degli errori, non si utilizzano modelli a priori per troposfera e ionosfera, ovvero si ha $\delta_{\tilde{T}_R^S}(t) = I_R^S(t)$, $\delta_{\tilde{I}_R^S}(t) = T_R^S(t)$; per il disturbo ionosferico si ricorda inoltre che vale la (4.14), ovvero che si può scrivere

$$I1_R^S(t) = \frac{1}{f_1^2} I0_R^S(t), \quad I2_R^S(t) = \frac{1}{f_2^2} I0_R^S(t) \quad (7.1)$$

La rete si pone l'obiettivo di modellizzare i diversi errori da singola stazione a singolo satellite mediante funzioni parametriche e stimare quindi i parametri dei modelli.

7.3.1 La riduzione delle incognite di stazione

Data l'equazione di osservazione di fase per una stazione permanente:

$$L_{SP}^S(t) = \rho_{SP}^S + cdt_{SP}(t) - cdt^S(t) - I_{SP}^S(t) + T_{SP}^S(t) + \lambda(N_{SP}^S + \phi_{SP} - \phi^S)$$

il termine di distanza può essere calcolato nelle coordinate note della stazione e nelle coordinate dalle effemeridi trasmesse dal satellite:

$$\rho_{SP}^S(t) = \tilde{\rho}_{SP}^S(t) + \mathbf{e}_{SP}^S(t) \cdot \boldsymbol{\delta}^S(t) = \tilde{\rho}_{SP}^S(t) + \delta_{\rho_{SP}^S}(t)$$

Per stimare i disturbi atmosferici e l'errore nelle effemeridi è necessario innanzitutto, per ogni stazione della rete:

1. stimare e fissare le ambiguità intere di entrambe le frequenze per tutti i satelliti:

$$N_{SP}^S \Rightarrow \hat{N}_{SP}^S \Rightarrow \bar{N}_{SP}^S$$

2. modellizzare e stimare gli offset d'orologio e le fasi frazionarie di entrambe le frequenze, per tutti i satelliti:

$$dt_{SP}(t) \rightarrow \hat{dt}_{SP}(t), dt^S(t) \rightarrow \hat{dt}^S(t),$$

$$\phi_{SP} - \phi^S \Rightarrow \hat{\phi}_{SP}^S$$

4. rimuovere le ambiguità fissate e le quantità modellizzate da tutte le osservazioni di fase:

$$\begin{aligned} \delta L_{SP}^S(t) &= L_{SP}^S(t) - \tilde{\rho}_{SP}^S(t) - cd\hat{t}_{SP}(t) - cd\hat{t}^S(t) - \lambda(\bar{N}_{SP}^S + \hat{\phi}_{SP}^S) \\ &= \delta_{\rho_{SP}^S}(t) - I_{SP}^S(t) + T_{SP}^S(t) \end{aligned} \quad (7.2)$$

Tale processo prende nome di inizializzazione della rete, è evidentemente particolarmente complesso e non viene qui discusso: si cita solo il fatto che tipicamente l'inizializzazione della rete richiede ai relativi programmi un tempo di calcolo di alcuni minuti.

7.3.2 La modellizzazione e la stima degli errori e dei disturbi

E' ora necessario definire modelli d'area e di tempo per i diversi tipi di disturbo: ad esempio un semplice modello bilineare nelle coordinate planimetriche e polinomiale nel tempo è dato dalla:

$$\begin{aligned} \delta^S(t) &= \sum_{k=0,n} \mathbf{a}_k^S t^k \\ T_{SP}^S(t) &= T^S(\varphi_{SP}, \lambda_{SP}, t) \\ &= (b_0(t) + b_\varphi(t)\varphi_{SP} + b_\lambda(t)\lambda_{SP} + b_{\varphi\lambda}(t)\varphi_{SP}\lambda_{SP})F_T(\eta_{\varphi,\lambda}^S) \\ I_{SP}^S(t) &= I^S(\varphi_{SP}, \lambda_{SP}, t) \\ &= (c_0(t) + c_\varphi(t)\varphi_{SP} + c_\lambda(t)\lambda_{SP} + c_{\varphi\lambda}(t)\varphi_{SP}\lambda_{SP})F_I(\eta_{\varphi,\lambda}^S) \end{aligned} \quad (7.3)$$

ove

$$b(t) = \sum_{k=0,n} b_k t^k, \quad c(t) = \sum_{k=0,n} c_k t^k$$

$F_T(\eta)$ e $F_I(\eta)$ sono le mapping function per troposfera e ionosfera rispettivamente; per la prima si può ricorrere all'approssimazione piana

$$F_T(\eta_{\varphi,\lambda}^S) = \frac{1}{\sin \eta_{\varphi,\lambda}^S}$$

per la seconda sono invece necessari modelli più complessi che tengano conto della geometria sferica dello strato ionosferico, qui non discussi.

Il modello precedente prevede la stima dei parametri

$$\mathbf{a} = \begin{bmatrix} a_0^1 \\ \dots \\ a_n^1 \\ a_0^2 \\ \dots \\ a_n^m \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_{0_0} \\ \dots \\ b_{\varphi\lambda_0} \\ b_{0_1} \\ \dots \\ b_{\varphi\lambda_n} \end{bmatrix}, \mathbf{c} = \begin{bmatrix} c_{0_0} \\ \dots \\ c_{\varphi\lambda_0} \\ c_{0_1} \\ \dots \\ c_{\varphi\lambda_n} \end{bmatrix}$$

Innanzitutto, per stimare il vettore $\mathbf{c}$ si può ricorrere all'osservabile Geometric Free: ricordando infatti la (4.14) si avrà, per ogni stazione, ogni satellite e ogni epoca:

$$\delta L 4_{SP}^S(t) = \left(\frac{f_2^2 - f_1^2}{f_2^2 f_1^2} \right) I 0_{SP}^S(t)$$

ovvero osservazioni che contengono solo il disturbo ionosferico; una volta stimati i relativi parametri, si possono usare L1 e L2 per stimare i parametri relativi a troposfera e effemeridi. La stima dei parametri incogniti può essere effettuata ad esempio mediante minimi quadrati sequenziali (Par. A1.3) oppure mediante filtri di Kalman: anche tale processo non viene qui discusso, ma si ricorda solo che la stima dei parametri del modello inizia al termine dell'inizializzazione della rete e prosegue in (quasi) tempo reale per tutto il periodo di erogazione di servizi all'utenza.

7.3.3 La distribuzione dei modelli

Si dispone di modelli stimati

$$\hat{\delta}^S(t), \hat{T}^S(\varphi, \lambda, t), \hat{I}_0^S(\varphi, \lambda, t)$$

che possono essere trasmessi all'utenza attraverso diversi approcci: nel seguito si descrive dettagliatamente il metodo Virtual Reference Station, per la sua chiarezza

concettuale e perché è stato il primo implementato in ordine storico. All'accensione, il ricevitore dell'utente comunica al centro di controllo le proprie coordinate approssimate, ad esempio quelle fornite dal posizionamento assoluto sulle prime epoche: tali coordinate sono quelle della cosiddetta stazione virtuale VRS, $\mathbf{x}_{VRS}$. Presso il centro di calcolo, per ogni satellite, data la posizione della VRS, vengono calcolati i termini di disturbo:

$$\delta_{\hat{\rho}_{VRS}^S}, \hat{I}_0^S(\mathbf{x}_{VRS}), \hat{T}^S(\mathbf{x}_{VRS})$$

Si utilizzano i dati della stazione permanente più vicina all'utente; questi vengono ricombinati per costruire le osservazioni di una stazione nella posizione della VRS:

$$\begin{aligned} L_{VRS}^S(t) &= L_{SP}^S(t) - \tilde{\rho}_{SP}^S(t) + \tilde{\rho}_{VRS}^S(t) + \hat{\delta}_{\hat{\rho}_{VRS}^S}^S(t) - \hat{I}_{VRS}^S(t) + \hat{T}_{VRS}^S(t) - (\hat{\delta}_{\hat{\rho}_{SP}^S}^S - \hat{I}_{SP}^S + \hat{T}_{SP}^S) \\ &= \tilde{\rho}_{VRS}^S(t) + \hat{\delta}_{\hat{\rho}_{VRS}^S}^S(t) - \hat{I}_{VRS}^S(t) + \hat{T}_{VRS}^S(t) + cdt_{SP}(t) - cdt^S(t) + \lambda(N_{SP}^S + \phi_{SP}^S) \\ &= \hat{\rho}_{VRS}^S(t) - \hat{I}_{VRS}^S(t) + \hat{T}_{VRS}^S(t) + cdt_{SP}(t) - cdt^S(t) + \lambda(N_{SP}^S + \phi_{SP}^S) \end{aligned} \quad (7.4)$$

Il centro distribuisce coordinate e dati della VRS, per tutti i satelliti e per entrambe le frequenze; l'utente costruisce le differenze doppie rispetto alla VRS:

$$L_{VRS,R}^{1,i}(t) = \rho_{VRS,R}^{1,i}(t) - I_{VRS,R}^{1,i}(t) + T_{VRS,R}^{1,i}(t) + \lambda N_{VRS,R}^{1,i}(t) \quad (7.5)$$

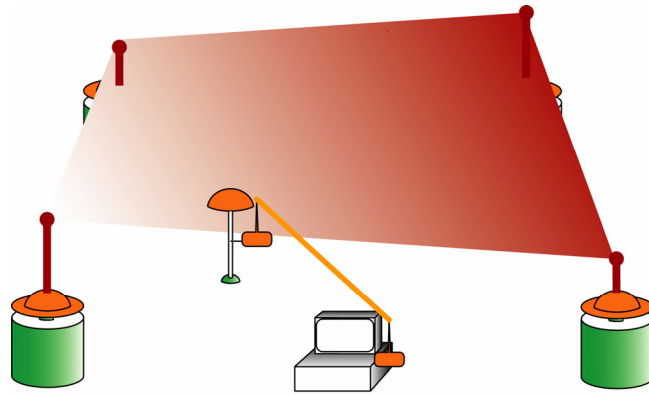
la VRS viene vincolata come stazione di riferimento e viene stimata la base fra essa e il ricevitore dell'utente, come discusso nel Cap. 6.

In alternativa alla VRS, negli anni passati veniva proposto l'approccio FKP (Flächenkorrekturparameter: parametri di correzione piana): il centro trasmette all'utente le coordinate e i dati di una stazione, tipicamente la più vicina, e i parametri delle funzioni bilineari che modellizzano gli errori. Il ricevitore dell'utente

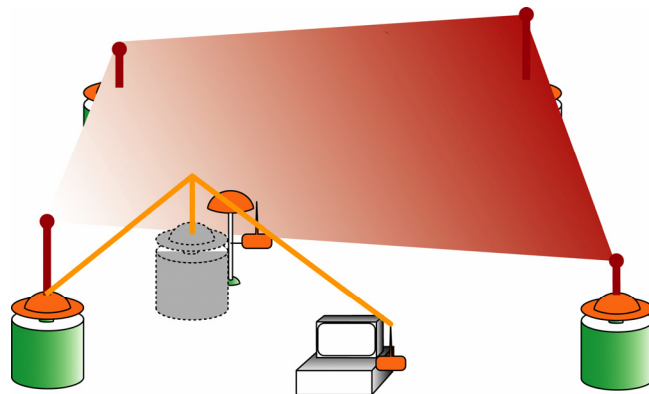
1. costruisce le differenze doppie rispetto alla stazione,
2. calcola gli errori e i disturbi con i parametri ricevuti,
3. corregge le differenze doppie per gli errori e i disturbi,
4. stima la propria posizione.

Il metodo FKP, rispetto al metodo VRS, ha il vantaggio di trasferire la parte del calcolo dal centro di controllo ai singoli utenti: in tal modo si evitano sovraccarichi del centro e possibili rallentamenti del flusso.

1. Il ricevitore mobile comunica la propria posizione approssimata alla rete



La rete genera una VRS nella posizione approssimata del ricevitore mobile



Il ricevitore mobile effettua posizionamento relativo rispetto alla VRS

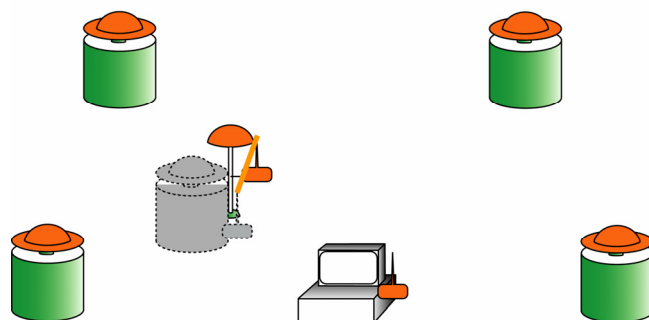


Figura 7.4. I passi dell'elaborazione VRS.

In effetti le tendenze internazionali stanno abbandonando il metodo FKP in favore di una sua ulteriore evoluzione detta metodo MAC (Master Auxiliary Concept): in tal caso il centro distribuisce all'utente:

1. le coordinate, le osservazioni e le stime di disturbi ed errori per una singola stazione permanente,
2. le coordinate e le differenze singole delle stime degli errori e dei disturbi per le altre stazioni della cella che contiene l'utente.

Il ricevitore dell'utente:

1. costruisce un proprio modello spaziale degli errori a partire dalle informazioni relative alle singole stazioni permanenti,
2. costruisce le differenze doppie con la stazione più vicina,
3. corregge le osservazioni mediante il modello degli errori appena calcolato,
4. stima la propria posizione.

L'approccio MAC sposta completamente l'onere del calcolo dei modelli di errore sul ricevitore dell'utente, poiché il centro di controllo si limita a stimare e trasmettere gli errori sulle singole stazioni permanenti; inoltre è l'approccio che distribuisce la maggior informazione possibile, non ancora filtrata e mediata a livello centrale: questo permette ai ricevitori degli utenti una maggior elasticità a livello locale nel controllo di qualità, nella soluzione di eventuali problemi e nell'elaborazione delle osservazioni.

7.4 Il sistema di riferimento per un servizio di posizionamento

Una rete di stazioni permanenti, distribuendo dati, prodotti e servizi, distribuisce anche le coordinate delle stazioni, ovvero materializza e distribuisce un sistema di riferimento. Per garantire la massima accuratezza, le coordinate delle stazioni devono essere stimate nel sistema di riferimento proprio del GNSS, che naturalmente è quello IGS; l'inquadramento della rete deve avvenire mediante postelaborazione quotidiana, seguendo le linee guida internazionali per l'elaborazione di reti permanenti; con tale prassi si hanno le serie quotidiane delle coordinate di tutte le stazioni (Fig. 7.5) che a loro volta possono essere interpolate con i consueti modelli lineari; ciò, oltre alla stima di massima accuratezza del modello dinamico di ogni stazione, garantisce il controllo continuo della qualità dei dati e la verifica che non vi siano stati spostamenti imprevisti, dovuti a cedimenti o altre anomalie.

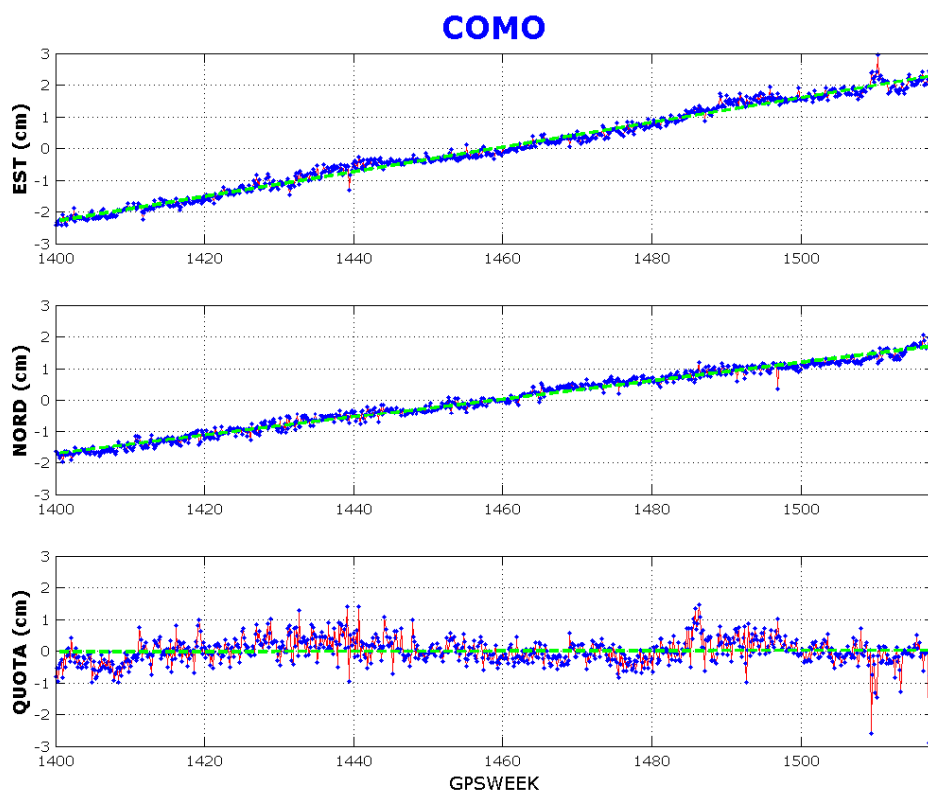


Figura 7.5. Le serie di coordinate della stazione di Como. Sono riportate le stime quotidiane e la loro interpolazione lineare a partire dalla settimana GPS 1400; i valori graficati sono rispetto alla media.

Tale processo fornisce il modello dinamico delle stazioni nel sistema di riferimento IGS corrente, ora IGS05; naturalmente l'utilizzo del modello dinamico completo costituisce la scelta naturale per l'inquadramento di sottoreti di alta accuratezza, effettuate mediante postelaborazione dei dati e questa prassi è già quella usuale per chi inquadri una propria sottorete locale nella rete IGS o in quella EPN.

Nel caso delle coordinate distribuite per le applicazioni in tempo reale, finalizzate al rilievo topografico, si pongono invece due problemi che richiedono un compromesso fra massima accuratezza e massima semplicità d'uso di un servizio di posizionamento. una stazione italiana si sposta, tipicamente di 1-3cm/y, ovvero di una frazione di millimetro al giorno: perciò, evidentemente, non è necessario aggiornare continuamente i valori delle coordinate distribuite in tempo reale, mentre è sufficiente effettuarne l'aggiornamento solo quando le coordinate siano cambiate significativamente rispetto all'ultimo valore pubblicato, ad esempio di 2-3 cm, ovvero al più una volta ogni sei mesi. A ciò si aggiunge però il fatto che l'utenza del tempo reale tipicamente richiede coordinate nel sistema di riferimento cartografico nazionale; per soddisfare tale requisito è possibile agire in due vie distinte, che definiamo A e B.

A.1. Il servizio di posizionamento stima e distribuisce le coordinate delle stazioni e i prodotti nell'IGS corrente;

A.2. inoltre stima e pubblica la trasformazione al sistema di riferimento cartografico nazionale;

A.3. il ricevitore dell'utente stima coordinate nell'IGS corrente e quindi le trasforma al sistema di riferimento cartografico nazionale.

B.1. Il servizio di posizionamento stima le coordinate delle stazioni nell'IGS corrente, ma anche nel sistema di riferimento cartografico nazionale;

B.2. distribuisce le coordinate IGS per la postelaborazione dei dati, quelle nel sistema di riferimento cartografico per il tempo reale;

B.3. il ricevitore dell'utente del tempo reale stima direttamente coordinate nel sistema di riferimento cartografico nazionale.

Come già detto nel Cap. 2, ETRF89-IGM95 era deformato e presentava anche errori sparsi; tecnicamente, seguire il percorso B comportava l'introduzione di significative approssimazioni nel calcolo di rete e quindi un potenziale degrado di accuratezza nei prodotti per il tempo reale: pertanto, finché è stato in vigore ETRF89-IGM95, l'unico approccio conservativo in Italia era quello contrassegnato con la A.

Oggi il sistema di riferimento italiano è ETRF2000-RDN, che allo stato attuale ha accuratezze molto superiori; il percorso più lineare, da un punto di vista del rigore formale, è ancora quello A; è però ragionevole attendersi che anche l'approccio B non ponga reali problemi nella gestione della rete e nella qualità dei prodotti distribuiti, a fronte di una effettiva semplificazione dal lato dell'utenza. Su ciò, naturalmente, sarà necessario un futuro lavoro di investigazione a livello nazionale.

7.5 Riepilogo finale dei prodotti tipicamente distribuiti da un servizio di posizionamento

Un servizio di posizionamento dovrebbe al minimo distribuire, in ordine di complessità:

1. l'aggiornamento periodico delle coordinate delle stazioni,
2. i dati delle stazioni permanenti per applicazioni di postelaborazione;
3. i prodotti di rete per il tempo reale;
4. la trasformazione fra il sistema di riferimento IGS corrente e il sistema di riferimento cartografico;
5. per applicazioni molto specialistiche, il modello dinamico delle coordinate.

Inoltre un servizio di posizionamento può svolgere anche altri servizi, come ad esempio la progettazione di rilievi, il controllo di qualità e la postelaborazione dei dati degli utenti, ma questi in effetti si configurano come elementi aggiuntivi; è viceversa fondamentale che il servizio implementi un sito web continuamente aggiornato che indichi:

- le principali caratteristiche della rete e dei prodotti distribuiti,
- chiare informazioni sull'utilizzo del servizio e di ogni prodotto,
- accesso alle osservazioni delle stazioni per la postelaborazione,
- preavvisi di manutenzione,
- notizie in tempo reale sulla funzionalità della rete,
- indicatori statistici sul funzionamento attuale della rete e sulla qualità dei dati,
- le serie storiche di stima delle coordinate delle stazioni,
- le trasformazioni fra sistema di riferimento IGS e il sistema di riferimento cartografico.

Sul sito debbono trovarsi anche programmi di utilità, che ad esempio permettano la trasformazione di sistema di riferimento, documenti che spieghino i settaggi dei ricevitori degli utenti e materiale didattico, scientifico o illustrativo. Sul sito deve trovare posto, ovviamente, l'interfaccia per l'accesso e l'eventuale tariffazione di tutti i servizi offerti.

7.5.1 Il formato di distribuzione dei dati per la postelaborazione

Il formato internazionalmente riconosciuto è il RINEX (Receiver INdependent EXchange format, ora arrivato alla versione 3.1) a cui è opportuno applicare le opportune compattazioni e compressioni, sempre codificate a livello internazionale; non si descrive qui il formato RINEX e i relativi standard di compressione, poiché sono ampiamente documentati, ad esempio nel sito di IGS: si ricorda solo che il RINEX è autoconsistente, poiché un file RINEX riporta al suo interno tutte le informazioni necessarie per il suo utilizzo: naturalmente l'ente gestore deve garantire la correttezza di tutti i metadati inseriti nel RINEX; gli standard internazionali di compattazione (Hatanaka) e compressione (Unix) sono quelli che minimizzano le dimensioni del file risultante. Non esiste un protocollo predeterminato per la distribuzione dei file RINEX: alcuni siti pubblicano i dati mediante protocollo ftp, che permette agli utenti lo scaricamento automatico, altri lavorano mediante interfaccia web, con selezione manuale; altri ancora distribuiscono i dati su richiesta email, inoltrandoli a indirizzi dedicati.

7.5.2 Il formato di distribuzione dei dati e dei prodotti per il tempo reale

Il formato internazionalmente normato per la trasmissione dei dati GPS in tempo reale è il formato RTCM (Radio Technical Commission for Maritime Services), ora arrivato alla versione 3.0. I dati per il tempo reale sono caratterizzati non solo dal formato ma anche dal cosiddetto protocollo di comunicazione fra la stazione di riferimento o il centro di controllo di una rete e il ricevitore mobile; ricordiamo qui la modalità di diffusione NTRIP (Networked Transport of RTCM via Internet Protocol), che all'inizio era semplicemente oggetto di studio e di sperimentazione in ambito EUREF ma che ora costituisce lo standard de facto anche su reti proprietarie di minor estensione, come i servizi di posizionamento in Italia. Il protocollo NTRIP prevede la connessione

dell'utente mediante modem GPRS a un punto d'accesso Internet e quindi il dialogo con il centro di controllo mediante un protocollo HTTP: il meccanismo di accesso, una volta impostato il modem, è del tutto trasparente all'utente. Rispetto ad altri possibili protocolli che non utilizzino la telefonia mobile, NTRIP presenta diversi vantaggi, sia dal lato rete sia dal lato utente: in primo luogo permette il dialogo bidirezionale fra centro di controllo e utente e offre la possibilità di autenticazione degli utenti; dal lato dell'utente fornisce una completa e chiara informazione, aggiornata in tempo reale, dei diversi prodotti erogati dalla rete, fra i quali è possibile scegliere mediante semplici interfacce, tipicamente implementate nei ricevitori mediante menu a scorrimento.

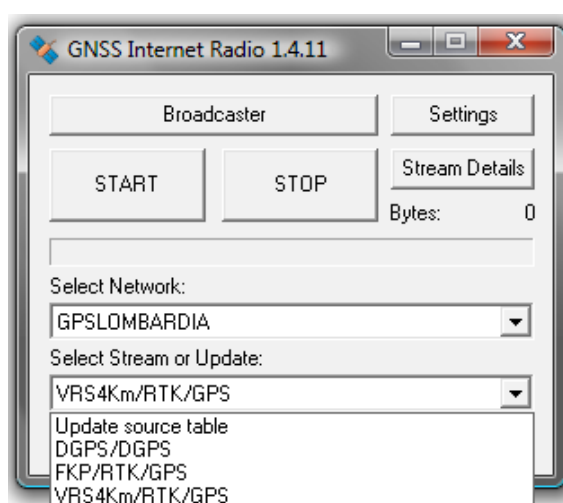


Figura 7.6. Un esempio di menu a tendina per la scelta dei diversi servizi erogati da una rete secondo il protocollo NTRIP.

7.6 I servizi di posizionamento in Italia e la sperimentazione del 2005

L'implementazione di servizi di posizionamento avviene alla scala locale, poiché non è previsto nelle loro linee di sviluppo che la rete globale IGS e quella continentale EPN assolvano tali compiti. In alcune nazioni dell'Europa i servizi di posizionamento si sono sviluppati alla scala nazionale, coordinati dalla rispettiva autorità cartografica nazionale; questo è l'esempio della Svizzera con la rete AGNES e della Germania con la rete SAPOS: vista la sua estensione, quest'ultima viene effettivamente gestita per sottoreti indipendenti al livello tecnico, ma vi è un coordinamento a livello nazionale. In Italia i primi servizi di posizionamento si sono sviluppati autonomamente alla scala delle regioni, poiché i relativi enti hanno relativa autonomia finanziaria, autorità normativa cartografica e hanno dimostrato interesse su tali tematiche: in particolare il Servizio di Regione Lombardia (GPSLombardia) è stato il primo operativo, ma ormai circa altre dieci regioni hanno un proprio servizio e quasi tutte le rimanenti sono in fase di progettazione o realizzazione; non si fornisce un elenco di dettaglio perché la situazione è in rapida evoluzione e si suggerisce agli interessati di informarsi direttamente presso i settori cartografici della propria Regione; inoltre le principali ditte

rivenditrici di strumentazione GPS per il rilievo topocartografico - Geotop, Leica Italia e Trimble Italia - stanno sviluppando propri progetti di reti alla scala nazionale, con l'obiettivo finale di creare dei servizi di posizionamento: anche in questo caso, per ulteriori informazioni, si suggerisce di contattare direttamente le ditte. Nel 2005 venne condotta un'ampia sperimentazione, condotta secondo un protocollo condiviso da più gruppi di ricerca indipendenti sulle reti che erano allora attive: Lombardia, Piemonte, Emilia Romagna, Umbria e Lazio.

Il rilievo venne condotto su più di 250 punti IGM95 o delle reti di raffittimento, omogeneamente distribuiti nelle aree coperte dalle rispettive reti e di cui si considerarono note, da monografia, le coordinate. Il protocollo di rilievo fu tale da simulare il comportamento del tipico utente dei servizi in tempo reale; in particolare si pose attenzione ai tempi di inizializzazione dei rilievi, alle motivazioni di eventuali insuccessi nell'inizializzazione stessa e al confronto metrico fra risultati del tempo reale e valori monografati. Le maggiori criticità furono dovute a problemi di connessione dati fra ricevitore utente e centro di controllo: è da notarsi che all'epoca delle sperimentazione erano utilizzati protocolli piuttosto instabili, cui appunto è subentrato NTRIP, molto più robusto. Pochissimi furono i punti in cui tutte le prove fallirono per cause dipendenti dalla tecnica GPS, e tipicamente si trattò di punti con significative ostruzioni verso il cielo, ovvero mai più di quattro satelliti in vista. Per il resto tutte le misure diedero esiti molto soddisfacenti: inizializzazioni tipicamente entro i tre minuti dall'accensione strumentale e accuratezze entro i due centimetri in planimetria, i quattro in quota: a titolo di esempio si riportano in Tab. 7.1 e in Fig. 7.7 alcuni estratti dei risultati ottenuti in Regione Lombardia. In tal senso si dimostrò sperimentalmente che i servizi di posizionamento costituiscono allo stato attuale un efficace supporto al rilievo GPS/GNSS per applicazioni topocartografiche.

Complessivo		Num	420	
	Est (cm)	Nord (cm)	2D (cm)	h (cm)
Media	1.3	1.7	3.2	0.6
Deviazione Standard	2.1	1.9	1.6	4.4
Minimo	-4.0	-5.3	0.4	-12.6
Massimo	6.1	8.7	8.9	14.8

Tabella 7.1. Statistiche di riepilogo dei confronti fra stime in tempo reale e monografie IGM nell'esperimento in Lombardia.

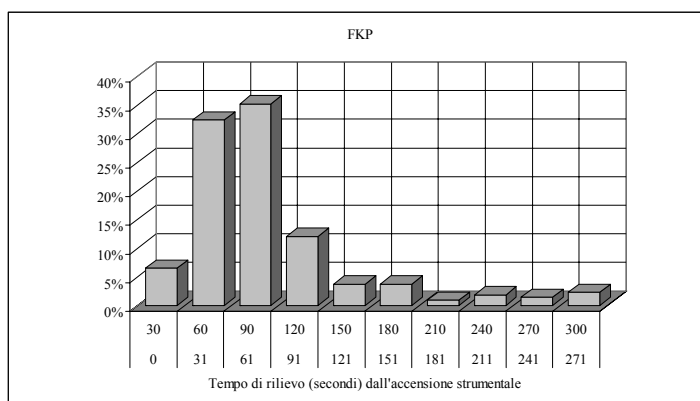


Figura 7.7. Istogramma di riepilogo sui tempi di rilievo ottenuti in modalità FKP nell'esperimento in Lombardia.

7.7 Alcune indicazioni all'utilizzo dei prodotti in tempo reale

Come nella postelaborazione, il tempo reale fornisce la stima della base fra la stazione di riferimento – eventualmente una stazione virtuale – e il punto incognito; dalle coordinate della stazione di riferimento e dalla stima della base viene quindi calcolata la posizione del punto, che è il risultato visualizzato all'utente e memorizzato. In associazione alla stima delle coordinate vengono memorizzati anche alcuni indici di qualità della soluzione; a differenza di quanto avviene nel caso di dati acquisiti per la postelaborazione, si perdono i dati grezzi e tipicamente non viene mantenuta memoria dei risultati parziali ottenuti nei diversi passi dell'elaborazione: ciò pone dei limiti nel controllo a posteriori dell'effettiva qualità dei risultati. Si consideri che il tempo reale è finalizzato a requisiti di accuratezza dell'ordine di alcuni centimetri e che tali accuratezze sono normalmente perseguibili, perlomeno in planimetria, con una soluzione fixed in discrete condizioni di osservazione; viceversa, in caso di significativi problemi nell'elaborazione, non si inizializza il rilievo, rimanendo quindi su una soluzione float, oppure si fissano le ambiguità a valori interi sbagliati. Innanzitutto si ricorda che la probabilità di successo dell'inizializzazione diminuisce in condizioni di cattiva geometria di osservazione, ovvero pochi satelliti visibili oppure racchiusi in un piccolo angolo solido: questa è la possibile condizione di punti in prossimità di edificato molto alto oppure in montagna impervia o bosco; in tali casi può essere necessario prevedere occupazioni finalizzate alla memorizzazione e successiva postelaborazione dei dati. Per gli altri casi, al fine di prevenire e identificare problemi di elaborazione si suggerisce la seguente prassi operativa:

1. nel caso di rilievo rispetto a una singola stazione di riferimento, mantenere la distanza da essa entro i 15 km: in tal modo si mantengono entro limiti accettabili i disturbi atmosferici;
2. pianificare i rilievi in finestre temporali ove siano presenti almeno cinque o meglio sei satelliti, condizione che, allo stato attuale, è garantita quasi sempre, almeno per

siti con buone condizioni di osservazione: in tal modo si alzano le probabilità di successo dell'inizializzazione;

3. pianificare almeno due ripetizioni del rilievo di ogni punto, se possibile in momenti lontani della giornata: in tal modo si ottengono risultati relativi a condizioni ambientali e configurazioni satellitari differenti;
4. in fase di rilievo, accettare e memorizzare solo le soluzioni fixed;
5. in caso vi siano difficoltà nell'inizializzazione, spegnere il ricevitore, eventualmente attendere qualche minuto e quindi riprovare il rilievo;
6. in caso di ulteriori difficoltà, in caso di rilievo appoggiato a servizi di rete, provare a utilizzare un prodotto di rete differente rispetto a quello del/dei tentativi precedenti;
7. in caso di permanenza delle difficoltà, abbandonare il punto e riprovare in un momento della giornata differente, per i motivi già detti al punto 3;
8. in fase di rilievo effettuare il confronto immediato dei diversi risultati ottenuti per ciascun punto: se questi differiscono di pochi centimetri, si consideri l'esito soddisfacente e si associ al punto la media dei risultati;
9. in caso contrario, ovvero di differenze significative fra i risultati, ripetere ulteriori rilievi, in modo da isolare e scartare quello affetto da problemi.

Durante un rilievo in tempo reale in genere si memorizzano alcune informazioni salienti (perlomeno il nome dei punti e l'altezza d'antenna) in associazione alla stima delle coordinate; è però anche opportuno redigere un foglio di campagna più esaustivo di quanto tipicamente memorizzabile tramite l'interfaccia utente del ricevitore, di cui si propone un esempio in Tab. 7.2: tale foglio potrà servire nell'archiviazione dei risultati, ove emergesse qualche dubbio.

Nome Operatore
 Tipo Ricevitore
 Tipo Antenna
 Metodo misura h ant.
 Data
 Nome Lavoro

Num.	Nome Punto	Prodotto di rete	H ant. nel ARP	Ora	PDOP	Num. Sat.	Durata	Note

Tab. 7.2. Esempio di foglio di campagna per rilievo statico rapido in tempo reale.

Capitolo 8.

Il rilievo e l'inquadramento di reti

L'inquadramento di una rete locale consiste nella stima delle coordinate dei suoi punti rispetto ai punti di una rete di ordine superiore scelta come riferimento. Nel presente capitolo viene trattato l'inquadramento di reti rilevate mediante GPS statico postelaborato, che implica diversi livelli di lavoro:

1. la progettazione del rilievo, ovvero la definizione di quali basi andranno rilevate e lo schema con cui rilevarle;
2. il rilievo,
3. la postelaborazione dei dati, per la stima delle basi,
4. la compensazione delle basi e l'inquadramento finale della rete.

La progettazione deve ottimizzare il bilancio fra spese e complessità del rilievo e accuratezza e affidabilità dei risultati ottenuti dalla compensazione finale; la progettazione è fondamentale, poiché da essa dipende l'esito del rilievo, e ciò porta ad alcune indicazioni metodologiche e tecniche che verranno discusse nel prossimo paragrafo. Per la gestione del rilievo sono possibili alcune indicazioni pratiche, dettate sostanzialmente dall'esperienza, che verranno semplicemente elencate; la postelaborazione dei dati è stata discussa nel Cap. 6, e non viene qui riproposta. La compensazione delle basi postelaborate e il successivo inquadramento finale richiedono una conoscenza essenziale del metodo dei minimi quadrati, riassunti nell'App. 1; qui si forniscono alcune linee guida operative su come condurre il relativo processo. Nel capitolo si utilizzeranno ricorrentemente alcuni simboli, che vengono sotto elencati:

- N_P : numero di punti di una rete,
 N_{PR} : numero di punti di riferimento inclusi nella rete,
 N_{PI} : numero di punti incogniti nella rete,
 N_S : numero di sessioni di rilievo per una rete,
 N_B : numero di basi rilevate in una rete.

8.1 La progettazione del rilievo

La postelaborazione delle differenze doppie fornisce stime di basi; innanzitutto si ricorda che in una sessione S_i in cui si siano utilizzati $N_R(i)$ ricevitori possono essere stimate solo $N_R(i) - 1$ basi indipendenti: il grafo con cui queste vengono connesse è indifferente e si possono vedere alcuni esempi in Fig. 8.1. E' però fondamentale che non vengano elaborati anelli chiusi, cioè che il grafo sia aperto: infatti essendo le differenze doppie le osservazioni di base per una rete, la presenza di un anello chiuso comporta

l'uso di osservazioni completamente correlate fra loro. N_S sessioni forniscono la stima di un numero di basi dato dalla

$$N_B = \sum_{i=1}^{N_S} (N_R(i) - 1) \quad (8.1)$$

Nel caso di disponibilità di un numero costante di ricevitori, la precedente si semplifica nella

$$N_B = N_S (N_R - 1)$$

Per determinare le coordinate di tutti i punti di una rete è evidentemente necessario rilevare perlomeno un grafo di basi indipendenti e che il grafo sia emanato da almeno un punto di coordinate note a priori, ovvero di riferimento. Si consideri ora una rete composta di N_P punti, con $N_P = N_{PR} + N_{PI}$: il numero di basi indipendenti della rete è pari a $N_P - 1$ e per costruirle si può ad esempio, emanare a stella da un punto di riferimento oppure sequenzialmente lungo un grafo connesso, come negli esempi illustrati in Fig. 8.1.

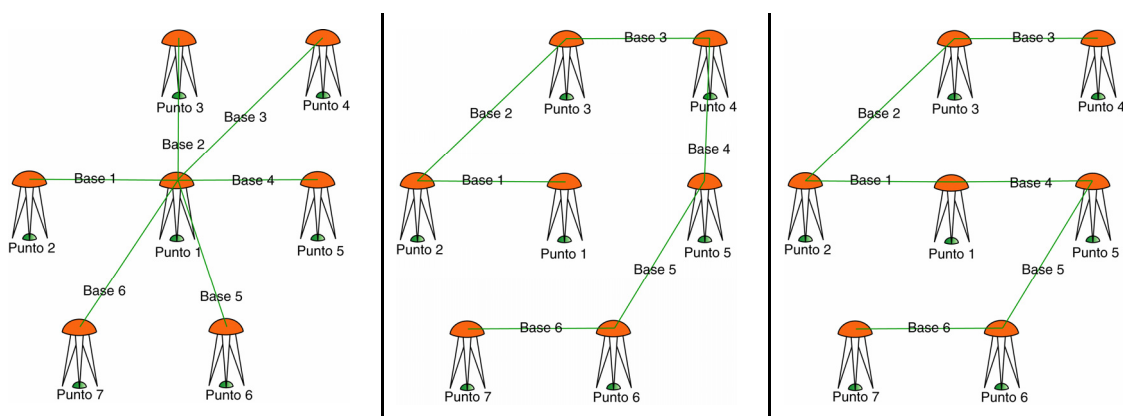


Figura 8.1. Esempio di schemi di grafo di rete. A sinistra grafo a stella, al centro e a destra esempi di grafi sequenziali: il numero di basi indipendenti è sempre uguale.

Lo schema di rilievo minimale ha ridondanza nulla, quindi non permette alcun controllo reciproco sulle stime delle basi e alcuna valutazione di accuratezza delle stime finali dei punti: è necessario prevedere schemi a ridondanza superiore. Per ottenere tale scopo è utile percorrere contemporaneamente due vie:

1. inserire più punti di riferimento nella propria rete,
2. stazionare ciascun punto della rete più volte.

Innanzitutto si ricorda che i punti di riferimento possono essere caposaldi di una rete statica, come IGM95, da stazionare con propri ricevitori durante il rilievo della rete locale; in questo caso è sempre possibile che le loro coordinate monografate presentino

significativi errori, come discusso nel Cap. 2: includerne più di uno aumenta i costi del rilievo poiché richiede più stazionamenti, ma aumenta la ridondanza del sistema ed è l'unico modo per avere un controllo reciproco delle coordinate ritenute note. In alternativa i punti di riferimento possono essere stazioni permanenti: se queste sono correttamente gestite e continuamente monitorate il rischio che le loro coordinate pubblicate presentino significativi errori è pressoché nullo, ma in ogni caso l'utilizzo di più stazioni di riferimento è utile, poiché aumenta la ridondanza del sistema senza aumentare i costi del rilievo: infatti è sufficiente acquisire i dati delle stazioni stesse senza necessità di stazionarvi. In linea di massima, la prassi usuale è quella di introdurre almeno tre punti di riferimento nel rilievo di reti locali; si parla in tal caso di inquadramento della rete locale nella rete di ordine superiore.

E' sempre opportuno prevedere per ciascun punto almeno due occupazioni in sessioni distinte, poiché così si realizza uno schema che permette perlomeno l'identificazione di eventuali errori grossolani di stazionamento; naturalmente, al crescere dei requisiti di accuratezza devono corrispondere schemi di rilievo via via sempre più ridondanti, sino a prevedere la ripetizione multipla di ogni base.

La ridondanza del rilievo è proporzionale alla differenza fra numero di basi rilevate e numero di punti incogniti: si veda a tale proposito il Par. 8.7 e la 8.17, 8.18. Nella progettazione del rilievo di una rete locale si devono dapprima scegliere i punti di inquadramento; quindi si stabilisce la ridondanza desiderata. A questo punto si devono considerare i seguenti aspetti: numero di ricevitori disponibili, posizione dei punti e lunghezza delle basi che compongono la rete, qualità di ricezione sui singoli punti, tempo e difficoltà degli spostamenti fra i punti della rete, qualità di stazionamento nei singoli punti. In funzione di tali aspetti, si deve operativamente decidere:

1. quali basi misurare, ovvero lo schema di rilievo della rete, con l'obiettivo di distribuire ridondanza, accuratezza e affidabilità il più possibile omogeneamente sui punti della rete;
2. i tempi di rilievo per le basi scelte, con l'obiettivo di soddisfare i requisiti di accuratezza previsti per i risultati finali;
3. la successione in cui rilevare le basi, con l'obiettivo principale di minimizzare i tempi, i costi e le complessità degli spostamenti fra i vari punti;
4. il cronoprogramma del rilievo, con l'obiettivo logistico di garantire la contemporaneità di rilievo per tutte le coppie di ricevitori che dovranno costituire le basi da postelaborare.

Si mostrano ora alcuni esempi.

8.1.1 Lo schema a un solo ricevitore, o a ricevitori indipendenti

Si supponga di operare in un'area servita da un servizio di posizionamento, ovvero che siano disponibili stazioni permanenti omogeneamente distribuite e a distanza ragionevole: in tal caso è possibile operare con un unico ricevitore e, per rilevare tutti i punti di una rete, è necessario (Fig. 8.2) effettuare un numero di sessioni dato da

$$N_S = N_{PI}$$

Ogni rilievo di un punto P_i della rete garantisce la stima di tante basi quante sono le stazioni permanenti disponibili per il punto, $N_B(P_i) = N_{PR}(P_i)$, ovvero per ogni punto è garantita ridondanza

$$r(P_i) = N_{PR}(P_i) - 1 \quad (8.2)$$

Si noti che in generale le stazioni permanenti utilizzate possono variare da punto a punto, in funzione della loro vicinanza; la ridondanza finale di un ciclo di rilievo è data dalla

$$r = \sum_{i=1}^{N_p} N_B(P_i) - N_{PI} \quad (8.3)$$

Se vi sono almeno due stazioni permanenti un singolo ciclo garantisce ridondanza ma, per avere più occupazioni per ogni punto, è necessario ripetere più cicli. L'approccio appena descritto è l'unico possibile quando si disponga di un solo ricevitore; disponendo di più ricevitori questi possono muoversi in modo indipendente, semplicemente per determinare le basi dalle stazioni a ciascun punto; inoltre è applicabile anche nel caso di inquadramento in una rete statica: i ricevitori necessari stazionano i punti di riferimento, mentre i restanti battono i punti incogniti. Però tipicamente lo schema a ricevitori indipendenti richiede il rilievo di basi piuttosto lunghe, ovvero dai punti di riferimento ai punti della rete locale singolarmente presi, quando le basi della rete locale sono molto più corte, e ciò implica un prolungamento dei tempi di stazionamento: per tale motivo, quando si disponga di più ricevitori, sono preferibili altri approcci, descritti nei paragrafi successivi. In ogni caso, con N_R ricevitori, per completare un ciclo di rilievo è necessario un numero di sessioni pari almeno a

$$N_S = \text{ceil}\left(\frac{N_{PI}}{N_R}\right) \quad (8.4)$$

$\text{ceil}()$ indica l'arrotondamento del rapporto all'intero superiore; in particolare nell'ultima sessione è necessario occupare gli ultimi $N_{PI} - (N_S - 1)N_R$ punti, mentre i rimanenti ricevitori possono iniziare un nuovo ciclo di occupazione degli altri punti.

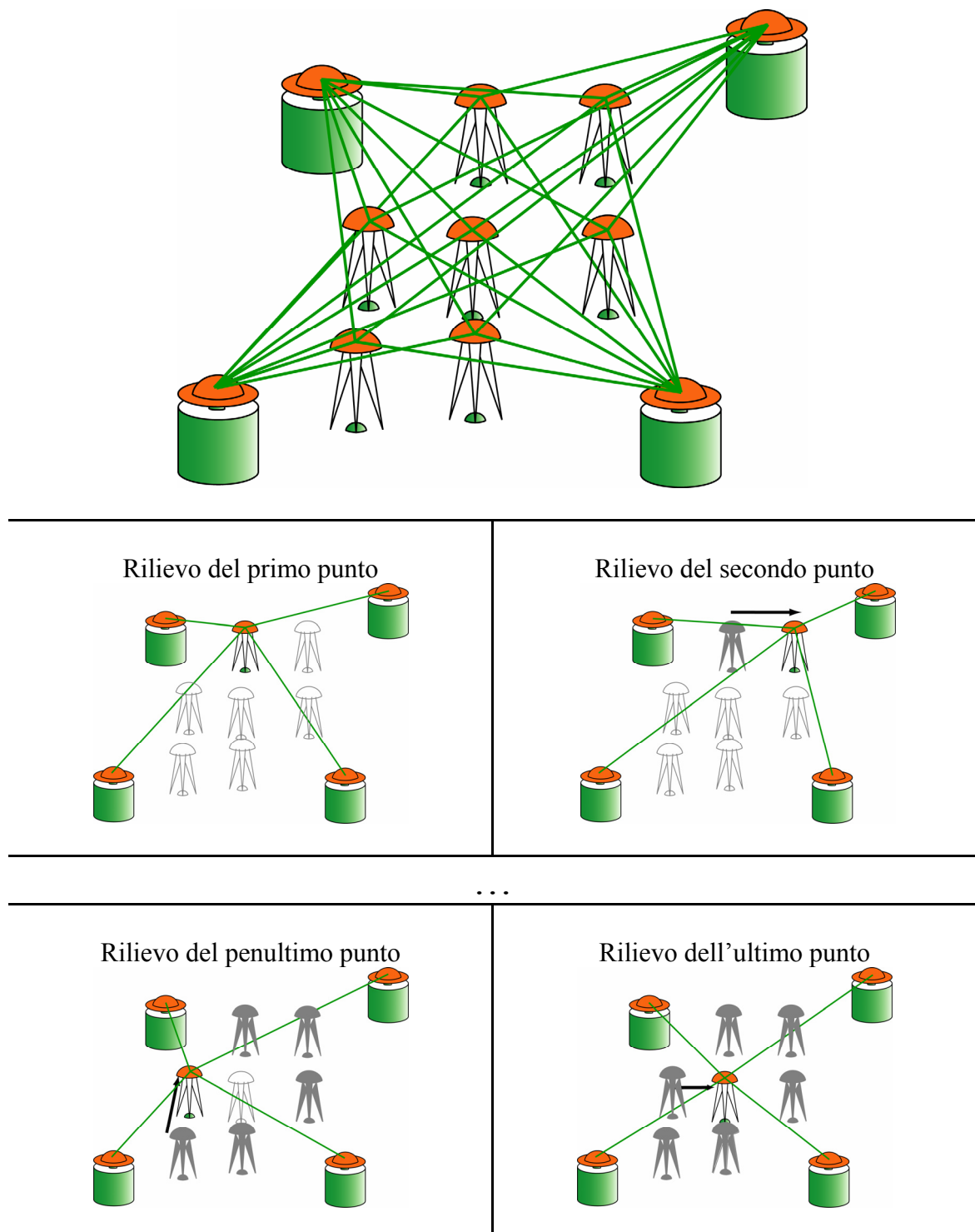


Figura 8.2. La stima di una rete locale con un solo ricevitore in presenza di 4 stazioni permanenti. Si ha buona ridondanza per ogni punto, 4-1, ma non si hanno rioccupazioni, che sono ottenibili solo con ulteriori cicli.

8.1.2 Lo schema a più ricevitori, a stella gerarchica

La via concettualmente più semplice per rilevare una rete locale con più ricevitori in modo coordinato è lo schema gerarchico a stella, di cui è dato un esempio in Fig. 8.3. Innanzitutto si definisce un punto di riferimento locale della rete, che chiamiamo interno, ovvero quello che garantisce la miglior visibilità al cielo fra quelli collocati in posizione più o meno baricentrica; il punto interno viene rilevato rispetto alle stazioni di inquadramento, quindi vengono rilevate le basi dal punto interno agli altri punti della rete locale.

Tale schema presenta i maggiori vantaggi quando le stazioni di inquadramento sono costituite da stazioni permanenti, ovvero non richiedono stazionamento con propri ricevitori: in tal caso sul punto interno si staziona continuamente con un ricevitore, mentre gli altri ricevitori girano e stazionano sugli altri punti della rete locale. Si supponga di disporre di N_{PR} stazioni permanenti, o di N_{PR} ricevitori che stazionano su altrettanti punti di riferimento, e di N_R ricevitori per stazionare i punti incogniti; ogni sessione fornisce la stima di N_{PR} basi al punto interno e di $N_R - 1$ basi da questo agli altri punti; per occupare almeno una volta tutti i punti della rete è necessario un ciclo di N_S sessioni, con

$$N_S = \text{ceil}\left(\frac{N_{PI} - 1}{N_R - 1}\right) \quad (8.5)$$

per l'arrotondamento $\text{ceil}()$ valgono le considerazioni già viste: in particolare nell'ultima sessione rimangono da rilevare $N_{PI} - 1 - (N_S - 1)(N_R - 1)$ punti.

Un singolo ciclo garantisce il rilievo di $N_B = N_{PI} - 1 + N_{PR}$ basi, con ridondanza

$$r = N_B - N_{PI} = N_{PI} - 1 + N_{PR} - N_P = N_{PR} - 1 \quad (8.6)$$

La ridondanza non è omogeneamente distribuita poiché un ciclo fornisce ridondanza solo per il punto interno e nessuna rioccupazione; per avere ridondanza e rioccupazioni per tutti i punti, è necessario effettuare più cicli e, per avere più occupazioni indipendenti anche sul punto interno, basta inizializzare lo stazionamento ad ogni nuovo ciclo. In definitiva, con N_C cicli, si garantisce ridondanza finale pari a

$$r = N_C(N_{PI} - 1 + N_{PR}) - N_{PI} \quad (8.7)$$

Lo schema gerarchico a stella apparentemente è sbilanciato verso il punto interno, sul quale si effettuano stazionamenti lunghi quanto un intero ciclo; in effetti lo sbilanciamento è tanto meno marcato quanti più sono i ricevitori disponibili: ad esempio su una rete di 6 punti con 2 ricevitori un ciclo richiede $N_{PI} - 1$ sessioni sugli altri punti,

ma con N_{PI} ricevitori un ciclo coincide con una sessione. Il maggior tempo di occupazione per il punto interno può peraltro essere un vantaggio quando la distanza dai punti di riferimento alla rete locale sia superiore all'interdistanza media fra i punti di quest'ultima: in tal caso lo schema lascia il giusto tempo al rilievo delle basi più lunghe e minimizza il tempo per rilevare le basi interne; inoltre minimizza anche la logistica, poiché non richiede la sincronizzazione dei ricevitori che si muovono fra i punti.

8.1.3 Altri schemi a più ricevitori

Altri schemi possono prevedere la concatenazione delle basi, con spostamenti sincronizzati di tutti i ricevitori ove la ridondanza del rilievo può essere garantita rilevando più volte le basi: un esempio in tal senso è dato dallo schema a stella illustrato in Fig. 8.5; più tradizionalmente si ricorre a schemi che prevedono la chiusura di poligoni, dal semplice anello chiuso sino alla triangolazione di Delaunay (Fig. 8.6) della rete. Si noti che per avere ridondanza omogeneamente distribuita con uno schema che preveda la ripetizione di un grafo di basi indipendenti è necessario rilevare ogni base del grafo almeno due volte, ovvero effettuare almeno $2(N_p - 1)$ rilievi; mediante uno schema a poligoni chiusi è già sufficiente rilevare gli N_p lati del perimetro della rete, ovvero realizzare il più semplice schema ad anello di Fig. 8.4: in tal senso lo schema a anello offre ridondanza al minimo costo, ma la sua ridondanza è comunque minimale. Definiamo il rapporto $R = N_B / (N_p - 1)$ fra basi rilevate e basi indipendenti; è corretto affermare che un rapporto $1 < R < 2$ è ottenibile solo mediante schemi a poligoni chiusi, mentre nel caso in cui si ponga l'obiettivo di avere $R \geq 2$, è indifferente operare mediante ripetizione delle basi o chiusura dei poligoni. La costruzione di poligoni chiusi è un requisito per la compensazione di osservazioni della topografia classica, poiché introduce varietà nelle osservazioni compensate; nel caso del GPS la stima delle basi è già frutto della varietà geometrica delle osservazioni, collegata al moto relativo dei satelliti rispetto ai ricevitori; in questo caso è importante massimizzare la ridondanza generale mantenendo omogeneità delle ridondanze e affidabilità locali, ma è irrilevante che tali obiettivi siano ottenuti mediante ripetizione delle basi piuttosto che mediante schemi a poligoni chiusi. Nelle figure 8.5 e 8.6 sono dati due esempi di schemi di rilievo per una rete, rispettivamente a ripetizione delle basi e a poligoni chiusi.

8.2 Il rilievo della rete

La progettazione del rilievo porta alla scrittura di una tabella di marcia per tutti i ricevitori, di cui si fornisce un esempio in Tab. 8.2. Si ricordi che l'elaborazione della base fra due ricevitori richiede la piena contemporaneità delle osservazioni: perciò, nella stesura della tabella di marcia si deve adottare un approccio conservativo sui tempi di spostamento. Si discute ora un semplice elenco di regole operative da seguire nel rilievo: ci si riferisce in questo caso al tipico caso di una rete in territorio servito da viabilità, alla scala locale ma con punti non intervisibili, ovvero alla scala dei 5-40 km.

Si noti che alcuni dei seguenti suggerimenti sfiorano l'ovvietà, ma si vuole comunque fornire un elenco completo.

1. Ogni ricevitore viene gestito da un'unità, che tipicamente si compone di uno o due operatori: non è necessario che gli operatori sappiano programmare il ricevitore, poiché tale operazione può essere stata fatta prima, ma è necessario che siano capaci di effettuare correttamente gli stazionamenti, di accendere il ricevitore e gestirne il buon funzionamento durante la sessione di osservazione, di compilare correttamente il foglio di campagna.
 2. Ogni ricevitore deve essere dotato di alimentazione e memoria sufficienti per coprire una intera giornata di misura; nel caso di campagna su più giornate, deve essere possibile per ogni unità effettuare la ricarica delle batterie durante la notte e lo scaricamento e il controllo dei dati alla sera.
 3. Ogni unità dispone di chiare indicazioni su come raggiungere i punti di sua competenza, ovvero monografie ed eventuali indicazioni stradali; a tal fine, e anche per verificare l'integrità e l'accessibilità dei punti, è utile effettuare una perlustrazione della rete prima della campagna vera e propria.
 4. Ogni unità è dotata di un telefono cellulare; è necessario che sia chiaramente indicato il nome di un coordinatore della campagna, che sarà sempre reperibile telefonicamente;
 5. se le unità si compongono di un solo operatore, si dovrebbero rendere disponibili uno o due operatori non direttamente coinvolti nella campagna ma pronti a sostituire o ad aiutare quelli sul campo; è utile anche disporre di cavi e batterie di scorta.
 6. Ogni unità deve strettamente attenersi alla propria tabella di marcia: ciò è particolarmente strategico per le campagne ove sia previsto il moto sincronizzato fra i ricevitori; qualunque problema deve essere segnalato al coordinatore, che deciderà e comunicherà alle unità quali cambiamenti apportare al programma iniziale.
 7. Ogni unità dovrebbe poter gestire autonomamente i propri spostamenti fra le diverse sessioni e disporre quindi di un automezzo che, nel caso di stazionamenti lunghi in località isolate, può fornire anche un riparo dalle intemperie.
 8. Nello stazionamento dei punti si deve misurare e registrare l'altezza d'antenna rispetto al caposaldo: è opportuno ripetere la misura sia prima sia dopo lo stazionamento del punto e verificare la coerenza dei due valori. E' inoltre utile effettuare alcune fotografie dello stazionamento, al fine di mantenerne memoria visiva: quattro fotografie prese dai quattro punti nadirali possono essere adeguate.
 9. Nel caso di stazionamento in luogo protetto, ovvero non accessibile a estranei, non è necessaria la presenza continua di un operatore; nel caso di stazionamento in luogo non protetto almeno un operatore deve presidiare lo strumento con continuità.
 10. Nel caso di campagne su più giorni è opportuno che ogni sera si verifichino speditivamente la completezza e la qualità dei dati acquisiti durante la giornata e si raccolgano, controllandoli, i fogli di campagna redatti dalle unità.
- Naturalmente il rilievo di reti molto locali, con eventuale intervisibilità dei punti risulta molto più semplice e ha minori requisiti logistici; non si trattano casi molto specifici, come potrebbe essere quello di un controllo geodinamico in ambiente alpino, poiché questi ovviamente vanno analizzati singolarmente.

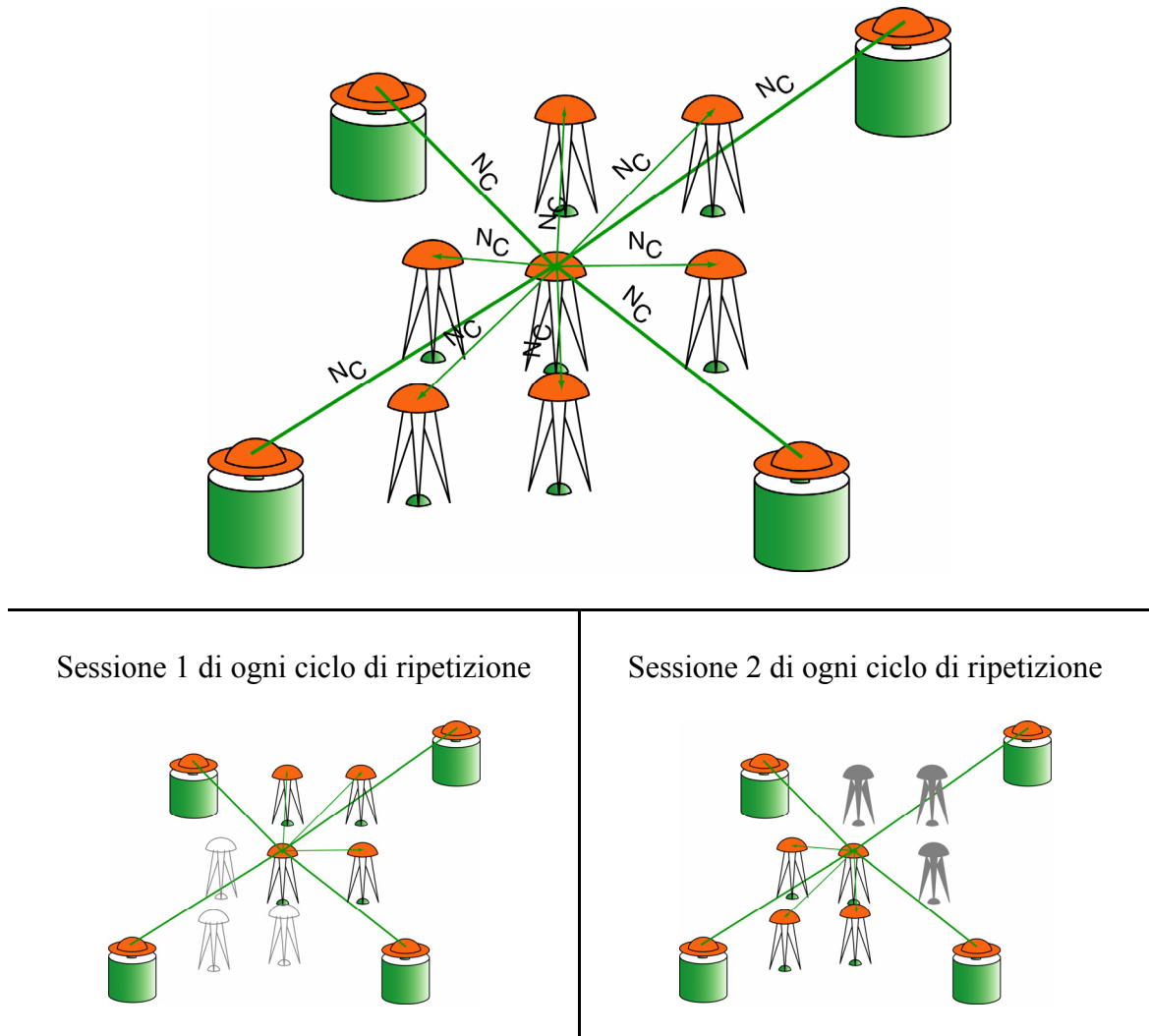


Figura 8.3. L'inquadratura di una rete locale di 7 punti rispetto a 4 stazioni permanenti secondo lo schema gerarchico a stella. Sopra, lo schema finale delle basi, sotto l'esempio di rilievo nel caso di 4 ricevitori. L'esempio garantisce ridondanza $r = N_c(7 - 1 + 4) - 7 = 10N_c - 7$, ove N_c è il numero di cicli.

	Ripetizione			Poligono chiuso			Triangolazione			
N_{PI}	N_B	R	r	N_B	R	r	N_T	N_B	R	r
2	4	2	2	3	1.5	1	1	3	1.5	1
3	6	2	3	4	1.3	1	2	5	1.7	2
5	10	2	5	6	1.2	1	5	10	2.0	5
8	16	2	8	9	1.1	1	8	16	2.0	8

Tabella 8.1. Esempi di schemi di rilievo per reti con diverso numero di punti.
La legenda è nella pagina seguente.

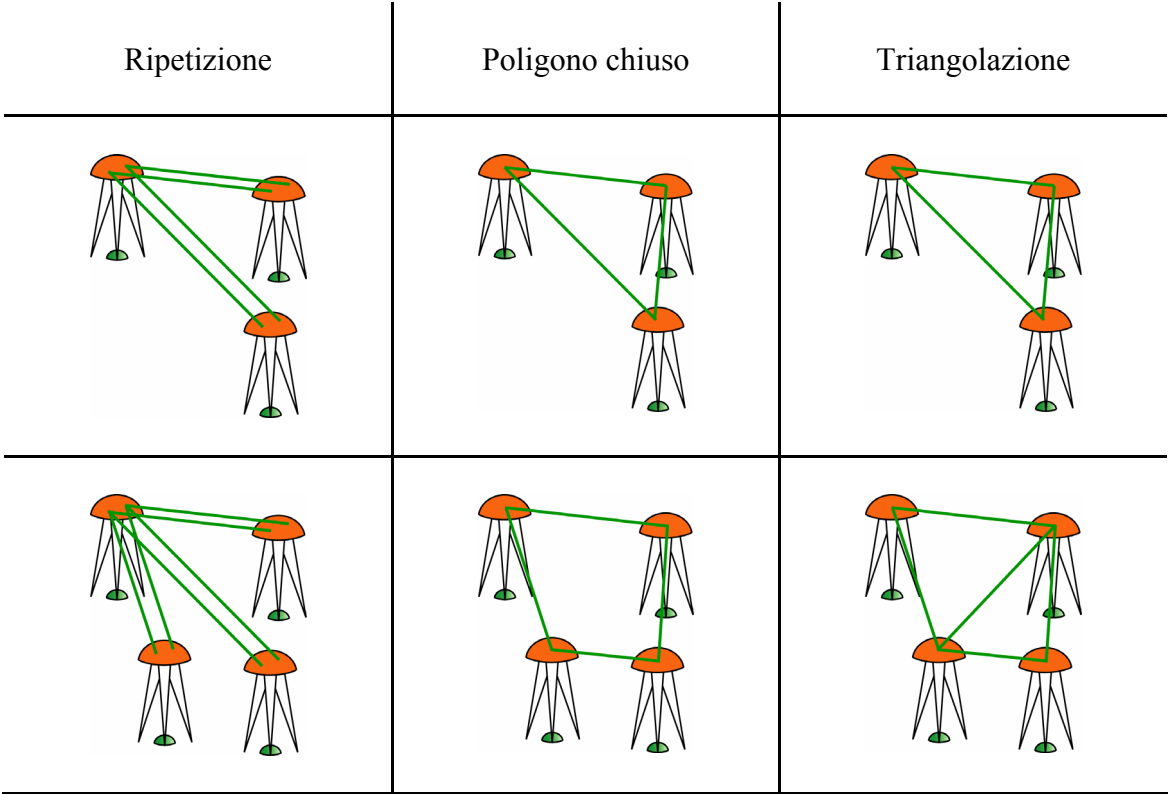


Figura 8.4.a. Esempi di schemi di rilievo per reti con diverso numero di punti:
la legenda è nella pagina seguente.

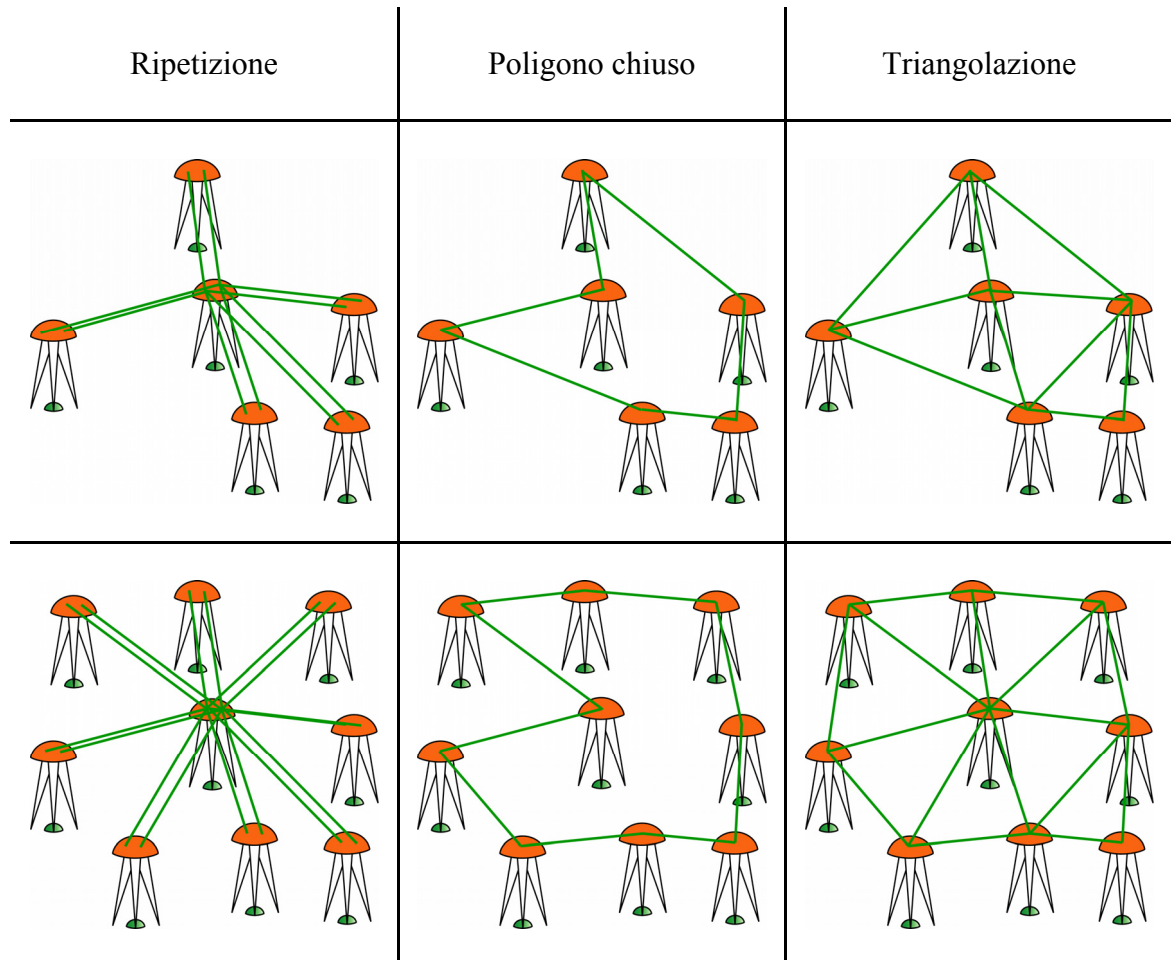
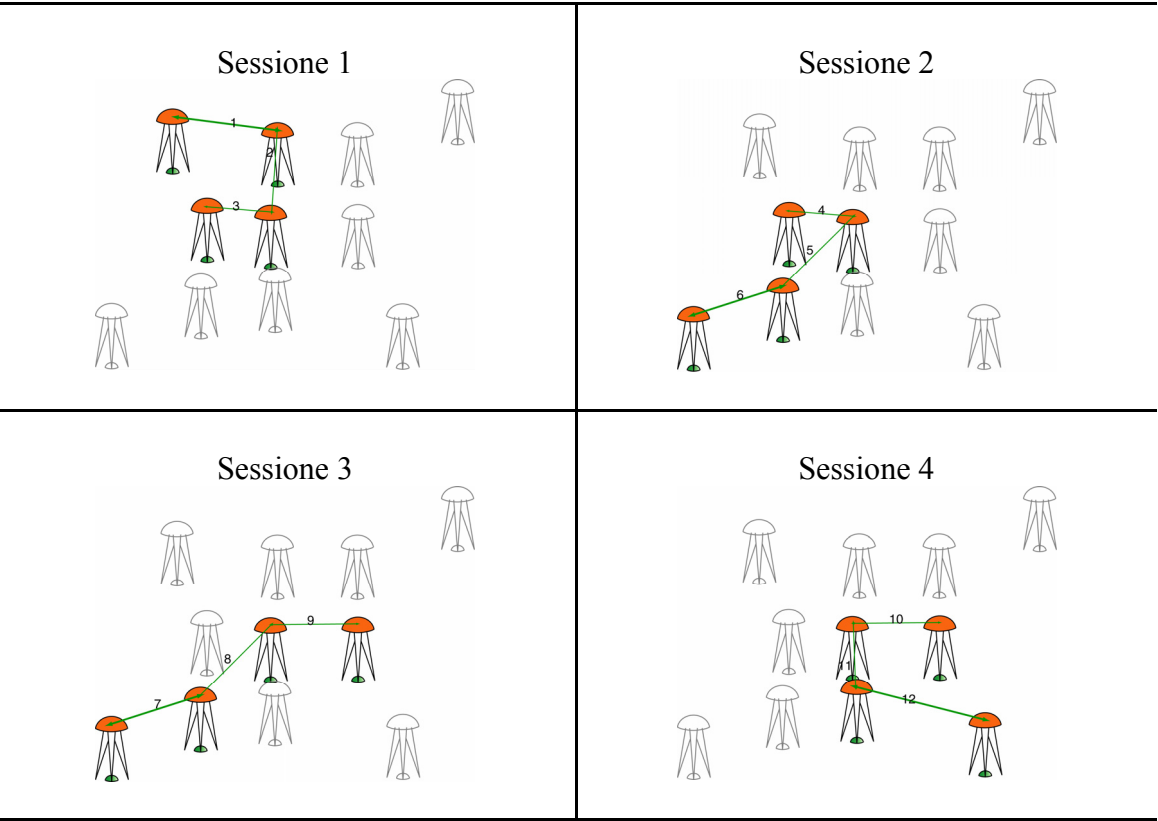


Figura 8.4.b Esempi di schemi di rilievo per reti con diverso numero di punti. Ripetizione: schema di rilievo basato sulla ripetizione due volte di tutte le basi indipendenti della rete. Poligono chiuso: schema basato sulla chiusura del poligono della rete. Triangolazione: schema basato sul rilievo delle basi ottenute da triangolazione di Delaunay della rete; in questo caso il numero N_T di triangoli dipende anche dalla forma della rete e non solo dal numero dei suoi punti (eccettuato il caso di $N_p = 3 \Rightarrow N_T = 1$!); N_B : numero di basi rilevate; $R = N_B / (N_p - 1)$, rapporto fra basi e punti incogniti; $r = N_B - N_{PI}$: ridondanza della compensazione.



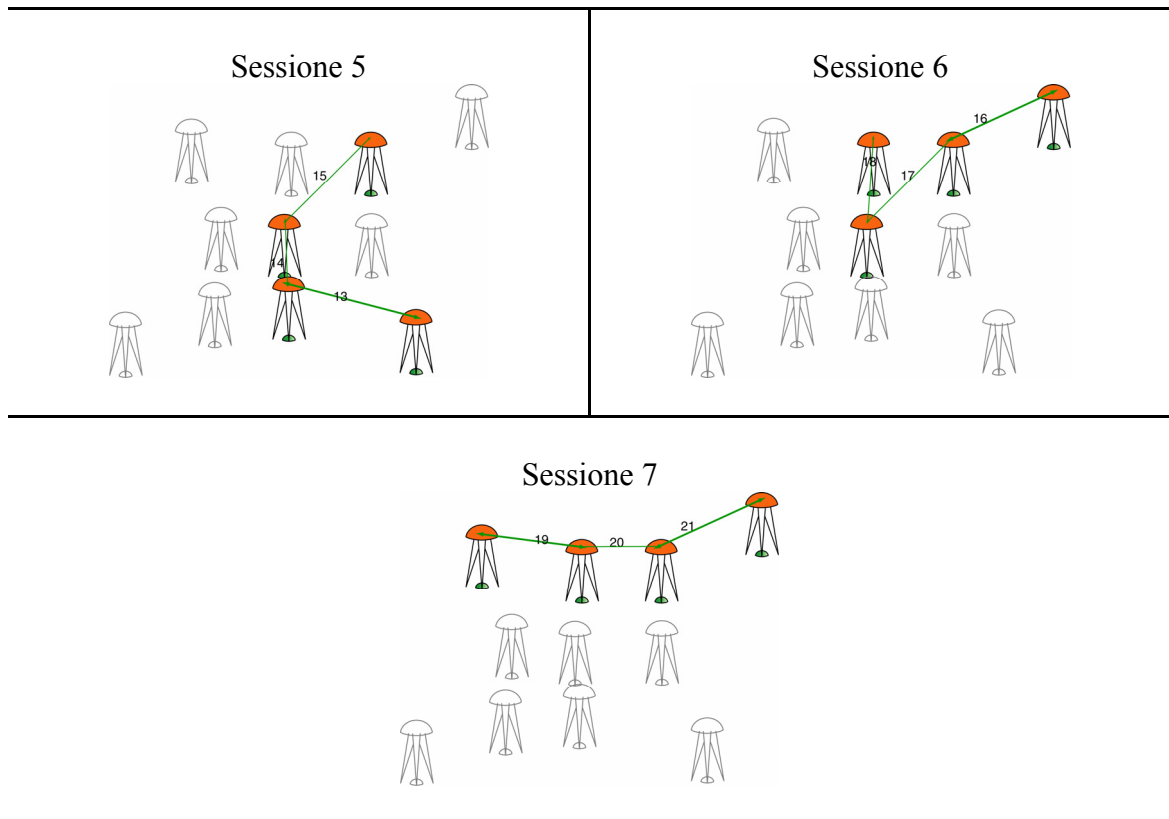


Figura 8.5. Uno schema di rilievo con ricevitori mobili e due ripetizioni di tutte le basi, ove le stazioni di inquadratura appartengano a una rete statica. Nella parte inferiore della figura, un esempio di concatenazione delle sessioni disponendo di quattro ricevitori; fra una sessione e la successiva si muovono da uno a tre ricevitori: è fondamentale il sincronismo di tutti gli spostamenti, in modo da garantire la contemporaneità di tutti i rilievi; per garantire le due rioccupazioni indipendenti di ogni punto, i ricevitori che non si muovono fra due sessioni devono comunque inizializzare nuovamente lo stazionamento. La base 20 non è necessaria per completare le ripetizioni ma viene inclusa perché coinvolge due ricevitori che hanno comunque rilevato. L'esempio garantisce ridondanza $r = 21 - 7$, mentre $R = 21/10$.

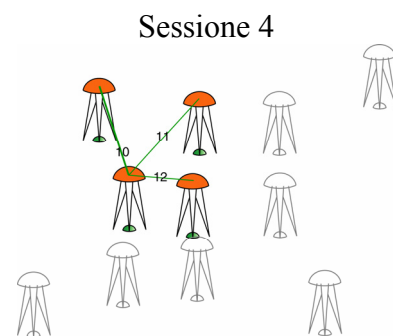
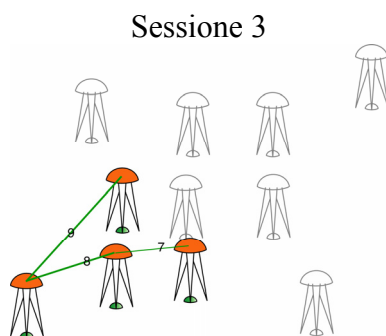
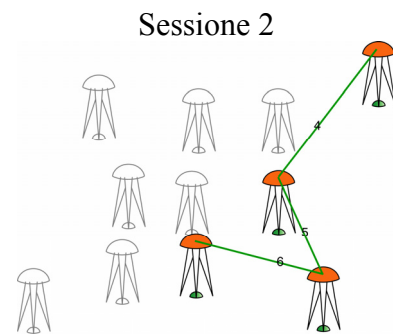
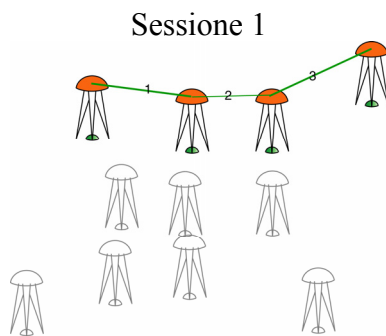
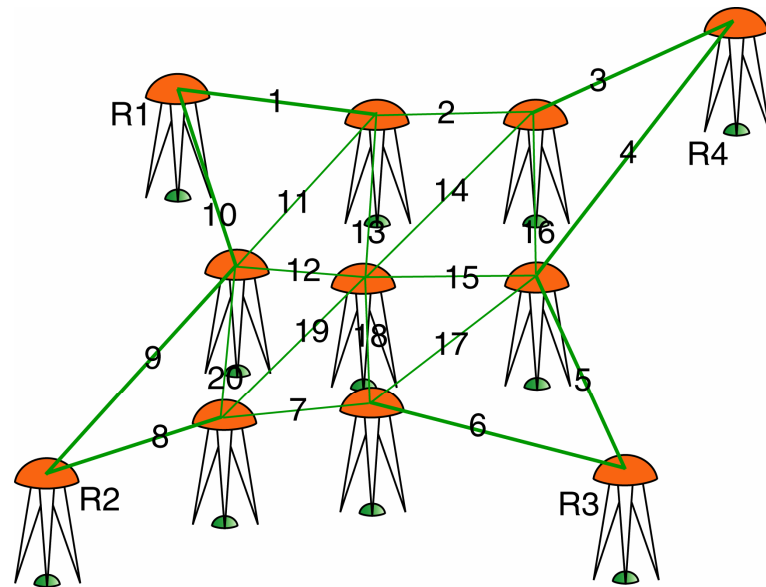


Figura 8.6: continua in pagina seguente

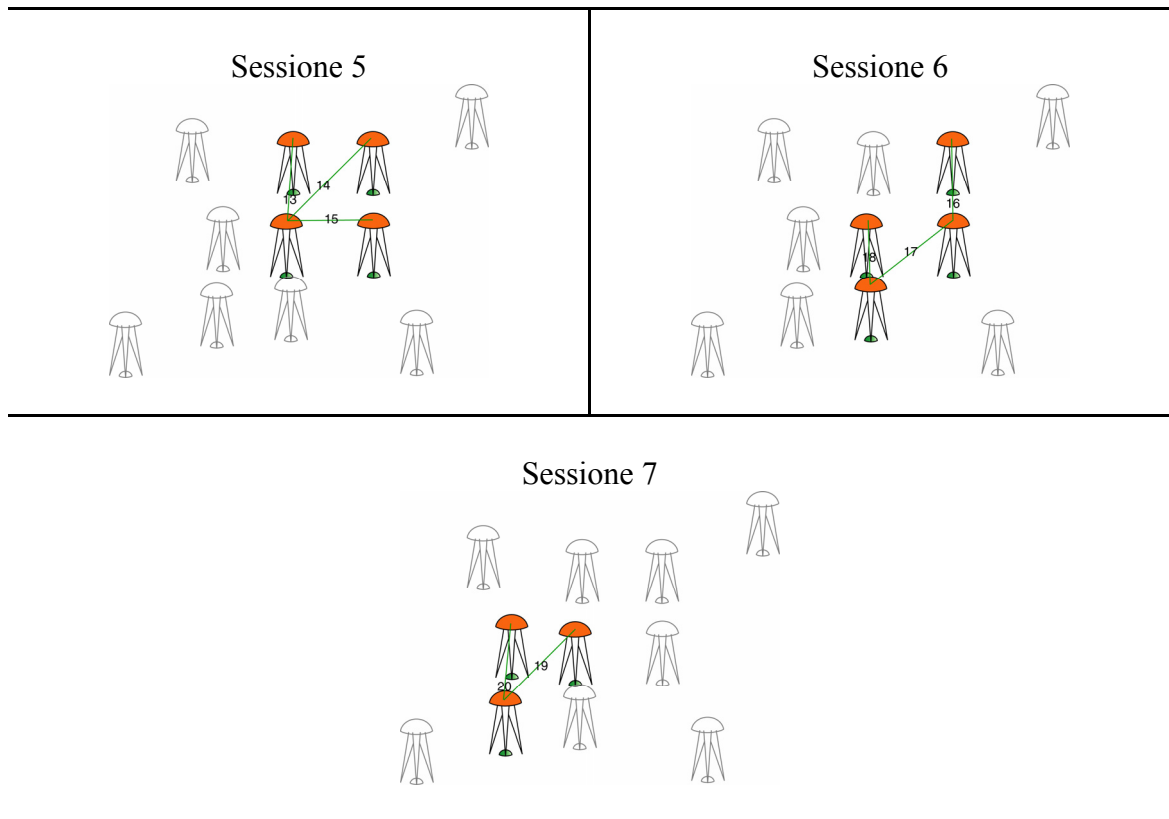


Figura 8.6. Uno schema di rilievo a poligoni chiusi, in questo caso triangoli, ove le stazioni di inquadramento appartengano a una rete statica, con la numerazione progressiva delle basi. Nella parte inferiore della figura, un esempio di concatenazione delle sessioni disponendo di quattro ricevitori; fra una sessione e la successiva si muovono in genere tre ricevitori: è fondamentale il sincronismo di tutti gli spostamenti, in modo da garantire la contemporaneità di tutti i rilievi; per garantire le due rioccupazioni indipendenti di ogni punto i ricevitori che non si muovono fra due sessioni devono comunque inizializzare il rilievo. L'esempio garantisce ridondanza $r = 20 - 7$, $R = 20/10$.

Nome Progetto:		Nome vertice:		Località:	
Data □□/□□/□□	Giorno GPS □□□	N. sessione □□□	Inizio □□.□□ Fine □□.□□	Cut Off: _____ Rate [s]: _____	
Schema punto:			Operatore/i: _____ _____		
			Tipo Ricevitore: _____		
			Numero: _____		
			Tipo Antenna: _____ Numero: _____		
Foto Nord:		Foto Sud:		h. antenna [m]: _____(I) h. antenna [m]: _____(F) modalità di misura: _____	
Foto Est:		Foto Ovest:		Nome File: _____ N.sat: □□	

Figura 8.7. Esempio di foglio di campagna

Dalle	Alle	Unità 1	Unità 2	Unità 3	Unità 4
8.30	9.30	R001	P001	P002	R002
10.00	11.00	P004	R003	P003	R002
11.30	12.30	P004	R004	P005	P006
13.00	14.00	P007	R001	P001	P006
14.30	15.30	P007	P002	P001	P003
16.00	17.00	P007	P002	P004	P003
17.30	18.30	P007	P005		P006

Tabella 8.2. Esempio di tabella di marcia per il rilievo schematizzato in Fig. 8.6: Rxxx indica un punto di riferimento, Pxxx un punto incognito. Le ipotesi sui tempi di spostamento danno ad ogni unità 5 minuti per smontare la stazione, 20 minuti di spostamento e 10 minuti per stazionare il nuovo punto.

8.3 La postelaborazione delle basi

I dettagli tecnici relativi alla postelaborazione di una base sono già stati discussi nel Cap. 5: si vuole qui dare qualche indicazione su come gestire operativamente l'elaborazione di una rete. Innanzitutto si deve costruire l'ordine delle basi da elaborare: poiché nell'elaborazione di una base è necessario conoscere con la maggior accuratezza possibile le coordinate del punto indietro, si può procedere nel seguente modo. Vengono dapprima elaborate le basi che spiccano da tutti i punti di inquadramento, di cui vengono vincolate le coordinate ai valori monografati; la stima delle basi fornisce quindi la stima dei relativi punti avanti, che si utilizzano come punti indietro per ulteriori elaborazioni di basi, in modo da propagare le stime; il processo procede in questo modo sino alla stima di tutte le basi, da utilizzare poi nella compensazione. Una via alternativa consiste nel propagare le stime di base da un unico punto di inquadramento, sino a completare la stima di tutte le basi: tale approccio, che è quello tipicamente implementato nei programmi commerciali, dovrebbe fornire risultati sostanzialmente identici al precedente ma può risentire di eventuali problemi sul punto di emanazione scelto.

Si consideri ora una sessione in cui abbiano operato N_R ricevitori: da tale sessione si devono costruire e elaborare le basi indipendenti previste dallo schema che era stato definito nel progetto del rilievo. Può però accadere che una o alcune fra le basi previste evidenzino problemi nell'elaborazione: in tal caso è opportuno provare a elaborare schemi alternativi di basi indipendenti, nella speranza di superare i problemi, poiché è evidentemente meglio compensare basi relative a un disegno di rete non ottimale piuttosto che basi distribuite in modo ottimale ma affette da problemi significativi. Tipicamente i programmi commerciali non effettuano in automatico questo tipo di analisi, che richiede quindi l'intervento manuale dell'operatore.

8.4 Introduzione all'analisi di rete

La postelaborazione in relativo di un rilievo GPS viene condotta mediante programmi dedicati e produce la stima delle 3 componenti della base fra punto indietro e punto avanti e la relativa matrice di covarianza:

$$\Delta \hat{\mathbf{x}}_{i,j}, \mathbf{C}_{\Delta \hat{\mathbf{x}}_{ij}} = \hat{\sigma}_{\Delta \hat{\mathbf{x}}_{ij}}^2 \mathbf{Q}_{\Delta \hat{\mathbf{x}}_{ij}} \quad (8.8)$$

Nel seguito tali stime vengono considerate come osservazioni - e relativo modello stocastico - prescindendo dal metodo di elaborazione dei dati grezzi con cui sono state ottenute. La prassi usuale è quella di considerare incorrelate le stime delle basi in ingresso all'analisi di rete; tale ipotesi di lavoro costituisce in generale una semplificazione, che è significativamente errata nel caso di basi rilevate in una stessa sessione e che condividono un ricevitore poiché esse hanno condiviso anche parte delle osservazioni utilizzate nell'elaborazione. Peraltro solo programmi molto specialistici – utilizzati prettamente in ambito universitario – calcolano correttamente le correlazioni fra le basi di una sessione e l'analisi di correlazione fra sessioni diverse è ancora ambito di ricerca; tutti i programmi commerciali elaborano le basi in modo indipendente e non forniscono in output una stima delle loro correlazioni: perciò si discuterà l'ipotesi semplificata

$$\mathbf{C}_{\Delta \hat{\mathbf{x}}_i, \Delta \hat{\mathbf{x}}_j} = \mathbf{0} \quad \forall i, j; i \neq j$$

8.5 L'analisi delle ripetibilità e delle chiusure

Si consideri un punto P_j sul quale si siano rilevate in un'unica sessione m basi a partire da altrettanti punti di inquadramento; si possono calcolare le stime di posizione $\hat{\mathbf{x}}_{j_i}, i = 1, \dots, m$ del punto

$$\hat{\mathbf{x}}_{j_i} = \mathbf{x}_{Ri} + \Delta \hat{\mathbf{x}}_{i,j} \quad (8.9)$$

il controllo delle ripetibilità fra le stime di posizione permette di evidenziare eventuali differenze significative che possono essere dovute alla presenza di significativi errori nella stima delle basi oppure nelle coordinate di vincolo delle stazioni di inquadramento: il controllo applicato a un singolo punto non permette di distinguere fra le due possibili cause; però il confronto incrociato fra i risultati relativi a punti diversi permette la discriminazione; siano infatti dati due punti, sia $\hat{\mathbf{x}}_{1i} = \mathbf{x}_{Ri} + \Delta \hat{\mathbf{x}}_{i,1}$ un risultato problematico, ovvero con scarti elevati rispetto agli altri risultati per il primo punto; se $\hat{\mathbf{x}}_{2i} = \mathbf{x}_{Ri} + \Delta \hat{\mathbf{x}}_{i,2}$ presenta analoghi problemi per il secondo punto, è probabile che vi sia

un errore nelle coordinate del punto di inquadramento R_i ; se $\hat{\mathbf{x}}_{2_i}$ fornisce risultati accettabili, è probabile che i problemi siano nella base $\Delta\hat{\mathbf{x}}_{i,1}$.

Ulteriori analisi riguardano le ripetibilità delle basi e le chiusure dei poligoni: innanzitutto due ripetizioni della stessa base Δ_{ij} dovrebbero dare il medesimo risultato:

$$\begin{bmatrix} \Delta\hat{X}_{i,j} \\ \Delta\hat{Y}_{i,j} \\ \Delta\hat{Z}_{i,j} \end{bmatrix}_1 = \begin{bmatrix} \Delta\hat{X}_{i,j} \\ \Delta\hat{Y}_{i,j} \\ \Delta\hat{Z}_{i,j} \end{bmatrix}_2$$

inoltre si consideri un qualsiasi poligono per il quale sono state rilevate e stimate in modo indipendente le basi perimetrali: trattandosi di un percorso chiuso, per ciascuna componente, (X,Y,Z) , la somma sulle basi deve essere nulla. Ad esempio per un qualunque triangolo $P1, P2, P3$ dell'esempio di Fig. 8.6, si dovrebbe avere:

$$\begin{bmatrix} \Delta\hat{X}_{1,2} \\ \Delta\hat{Y}_{1,2} \\ \Delta\hat{Z}_{1,2} \end{bmatrix} + \begin{bmatrix} \Delta\hat{X}_{2,3} \\ \Delta\hat{Y}_{2,3} \\ \Delta\hat{Z}_{2,3} \end{bmatrix} + \begin{bmatrix} \Delta\hat{X}_{3,1} \\ \Delta\hat{Y}_{3,1} \\ \Delta\hat{Z}_{3,1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

In una più conveniente notazione vettoriale si ha quindi per due ripetizioni di una base

$$\Delta\mathbf{x}_{i,j_1} = \Delta\mathbf{x}_{i,j_2} \quad (8.10)$$

per un poligono Π chiuso e composto dai punti $P_{1(\Pi)}, P_{2(\Pi)}, \dots, P_{n(\Pi)}$

$$\sum_{i=1(\Pi)}^{n(\Pi)-1} \Delta\mathbf{x}_{i(\Pi), i+1(\Pi)} = \mathbf{0} \quad (8.11)$$

La presenza di errori di stima comporta uno scarto nella ripetibilità:

$$\Delta\hat{\mathbf{x}}_{i,j_1} - \Delta\hat{\mathbf{x}}_{i,j_2} = \mathbf{v}_{i,j} \quad (8.12)$$

o nelle chiusure:

$$\sum_{i=1(\Pi)}^{n(\Pi)-1} \Delta\hat{\mathbf{x}}_{i(\Pi), i+1(\Pi)} = \mathbf{v}_{\Pi} \quad (8.13)$$

Un primo controllo visivo degli scarti permette di identificare l'esistenza di eventuali problemi grossolani nelle stime delle basi; a tale riguardo si noti che l'accettabilità o meno degli scarti è anche funzione del tipo di accuratezze che si attendono dal rilievo eseguito: se, ad esempio, si stanno considerando rilievi statico rapidi su basi di circa 10 km, scarti dell'ordine di 2-4 cm sono del tutto probabili e accettabili; i medesimi scarti sono del tutto sospetti se si stanno analizzando basi di 1-5 km con occupazioni di diverse ore.

L'analisi degli scarti non sempre permette di identificare la stima che genera il problema. Ad esempio si consideri una base rilevata due volte secondo uno schema a stella gerarchica: la presenza di uno scarto inaccettabile segnala l'esistenza di un problema ma non indica di per se quale sia la stima problematica; analogamente nel caso di un triangolo chiuso i cui lati non siano condivisi da altri poligoni: un errore di chiusura non permette comunque di capire quale delle tre basi abbia avuto problemi. La situazione migliora significativamente con configurazioni più ridondanti: il confronto a coppie di tre ripetizioni della stessa base permette di isolare l'eventuale stima problematica; se i tre lati di un triangolo sono condivisi uno a uno con altrettanti poligoni è possibile trovare la base problematica mediante controllo incrociato degli errori di chiusura; un esempio in tal senso è schematizzato nella Fig. 8.7.

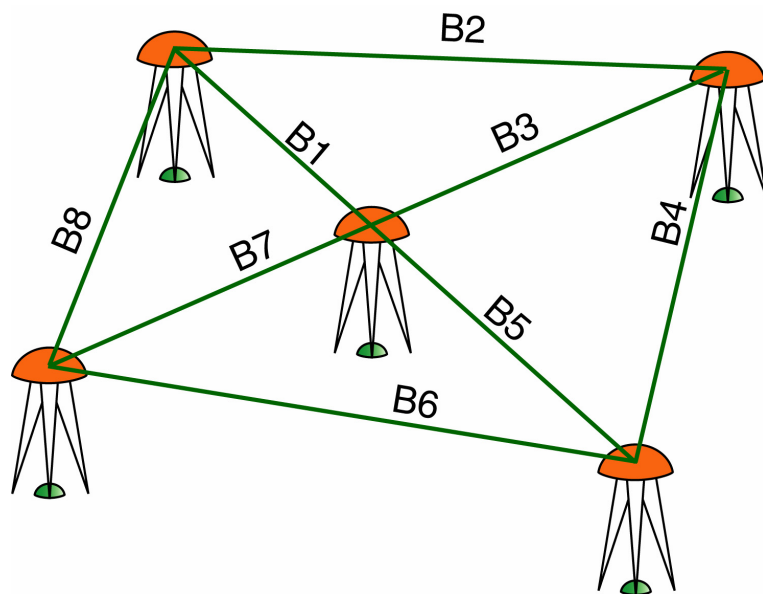


Figura 8.7. Un esempio di analisi delle chiusure. Triangolo $B1-B2-B3$: piccoli errori di chiusura, ovvero relative basi stimate correttamente; triangolo $B3-B4-B5$: piccoli errori di chiusura, ovvero relative basi stimate correttamente; triangolo $B5-B6-B7$: errori di chiusura elevati: poiché la base $B5$ ha già superato un test, le basi problematiche possono essere la $B6$ o la $B7$; triangolo $B1-B7-B8$: errori di chiusura elevati, ovvero le basi $B7$ oppure la $B8$ hanno errori grossolani; per dirimere si analizza il quadrilatero $B1-B5-B6-B8$: in caso di piccoli errori di chiusura la base

sospetta di errori grossolani è la $B7$; in caso di errore elevato, il dubbio fra $B6, B7, B8$ rimane irrisolto e richiede ulteriori analisi.

8.6 La compensazione di rete

Sia stato adottato uno schema di rilievo ridondante su una rete GPS; siano state effettuate le analisi di chiusura e la relativa reiezione di errori grossolani. E' ora il momento di effettuare la compensazione e l'inquadramento della rete mediante applicazione dei minimi quadrati: i dati in ingresso sono le posizioni dei punti di inquadramento della rete, le stime delle basi rilevate e le relative matrici di covarianza: queste ultime vengono considerate osservazioni, ovvero

$$\begin{aligned}\Delta \hat{\mathbf{x}}_{i,j} &\Rightarrow \Delta \mathbf{x}_{i0} \\ \mathbf{C}_{\Delta \mathbf{x}_{i,j}} &\Rightarrow \sigma_{0l}^2 \mathbf{Q}_l\end{aligned}\tag{8.14}$$

Le coordinate dei punti di inquadramento possono essere utilizzate in due modi alternativi: possono essere vincolate, ovvero entrano nel termine noto e in tal caso solo le coordinate della rete locale entrano nel vettore dei parametri incogniti; oppure possono essere introdotte come osservazioni, con proprie varianze, relative ad altrettante posizioni da stimare in aggiunta a quelle dei punti della rete locale. Da un punto di vista formale il secondo approccio è l'unico corretto: le coordinate a priori dei punti di inquadramento sono frutto di stime precedenti e quindi caratterizzate da un proprio errore, di cui si deve tener conto. Da un punto di vista operativo, se le coordinate a priori di un punto di inquadramento presentano un errore grossolano, il loro vincolo rigido distorce significativamente le stime della rete locale; la loro introduzione come osservazioni invece permette un controllo a posteriori sui valori introdotti. Ciò è particolarmente importante quando i punti di inquadramento appartengono a una rete statica, magari rilevata diversi anni prima: in tal caso è infatti molto probabile che i valori monografati siano meno accurati e coerenti dei risultati forniti dal rilievo. Naturalmente è importante associare accuratezze realistiche alle posizioni dei punti di inquadramento: in tal senso diamo le seguenti indicazioni.

1. E' ragionevole associare ai valori monografati dei caposaldi IGM95 accuratezze di 2-3 cm in planimetria, 3-5 cm in quota: fanno eccezione naturalmente quei punti che negli ultimi venti anni abbiano subito cedimenti anomali, cosa che va verificata mediante analisi dei residui al termine della compensazione.
2. Le stime delle coordinate delle stazioni di RDN hanno oggi accuratezze migliori di 1 cm in tutte le componenti: con il passare del tempo, se non si introdurrà un effettivo monitoraggio continuo della rete, le accuratezze gradualmente degraderanno.
3. Le stime delle coordinate di una stazione permanente continuamente e correttamente monitorata hanno accuratezze dell'ordine di 2-5 mm e ciò vale sicuramente per le stazioni IGS e EPN, le cui accuratezze sono peraltro pubblicate insieme alle coordinate stesse.

4. I servizi di posizionamento e le stazioni permanenti singolarmente gestite in Italia hanno prassi di monitoraggio delle coordinate molto eterogenee e altrettanto lo sono le relative accuratèzze: l'utente deve richiedere la relativa documentazione all'ente gestore.

Se si adotta il metodo di vincolo stocastico delle coordinate si ha il seguente modello:

$$\mathbf{y}_0 = \mathbf{A}\mathbf{x} + \mathbf{v}, \quad (8.15)$$

ove il vettore delle osservazioni è dato dalle basi $\Delta\mathbf{x}_{l_0}$ e dalle coordinate dei punti di riferimento $\mathbf{x}_{Ri}$

$$\mathbf{y}_0 = \begin{bmatrix} \Delta\mathbf{x}_0 \\ \mathbf{x}_0 \end{bmatrix}, \quad \Delta\mathbf{x}_0 = \begin{bmatrix} \Delta\mathbf{x}_{10} \\ \Delta\mathbf{x}_{20} \\ \dots \\ \Delta\mathbf{x}_{NB0} \end{bmatrix}, \quad \Delta\mathbf{x}_{l_0} = \Delta\mathbf{x}_{l_0}(P_i, P_j), \quad \mathbf{x}_0 = \begin{bmatrix} \mathbf{x}_{R_{10}} \\ \mathbf{x}_{R_{20}} \\ \dots \\ \mathbf{x}_{N_{PR0}} \end{bmatrix}, \quad \mathbf{x}_{Ri0} = \mathbf{x}(P_{Ri}) + \mathbf{v}_{Ri}$$

I parametri incogniti sono le coordinate dei punti della rete locale e dei punti di riferimento:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_I \\ \mathbf{x}_R \end{bmatrix}, \quad \mathbf{x}_I = \begin{bmatrix} \mathbf{x}_{P_1} \\ \mathbf{x}_{P_2} \\ \dots \\ \mathbf{x}_{P_{N_{PI}}} \end{bmatrix}, \quad \mathbf{x}_R = \begin{bmatrix} \mathbf{x}_{R_1} \\ \mathbf{x}_{R_2} \\ \dots \\ \mathbf{x}_{R_{N_{PR}}} \end{bmatrix}$$

Per il modello deterministico si ha

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_\Delta \\ \mathbf{A}_x \end{bmatrix}, \quad \mathbf{A}_\Delta = \begin{bmatrix} \mathbf{A}_{\Delta_{11}} & \mathbf{A}_{\Delta_{12}} & \dots & \mathbf{A}_{\Delta_{1N_P}} \\ \mathbf{A}_{\Delta_{21}} & \mathbf{A}_{\Delta_{22}} & \dots & \mathbf{A}_{\Delta_{2N_P}} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}_{\Delta_{NB1}} & \mathbf{A}_{\Delta_{NB2}} & \dots & \mathbf{A}_{\Delta_{NB_{N_P}}} \end{bmatrix}, \quad \mathbf{A}_x = \begin{bmatrix} \mathbf{A}_{x_{11}} & \mathbf{A}_{x_{12}} & \dots & \mathbf{A}_{x_{1N_P}} \\ \mathbf{A}_{x_{21}} & \mathbf{A}_{x_{22}} & \dots & \mathbf{A}_{x_{2N_P}} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}_{x_{N_{PR}1}} & \mathbf{A}_{x_{N_{PR}2}} & \dots & \mathbf{A}_{x_{N_{PR}N_P}} \end{bmatrix}$$

Le sottomatrici $\mathbf{A}_{\Delta_{ij}}$ hanno dimensione 3×3 : se

$$\Delta\mathbf{x}_{l_0} = \Delta\mathbf{x}_{i,j0} (= \mathbf{x}_j - \mathbf{x}_i + \mathbf{v}_{\Delta_i})$$

allora

$$\mathbf{A}_{\Delta_{l,j}} = -\mathbf{I}, \quad \mathbf{A}_{\Delta_{l,j}} = \mathbf{I}, \quad \mathbf{A}_{\Delta_{l,m}} = \mathbf{0} \quad \forall m \neq i, j$$

Le sottomatrici $\mathbf{A}_{x_{ij}}$ hanno nuovamente dimensione 3×3 ; se

$$\mathbf{x}_{i_0} = \mathbf{x}_{Rj} + \mathbf{v}_i$$

allora

$$\mathbf{A}_{x_{i,N_{PI}+j}} = \mathbf{I}, \quad \mathbf{A}_{x_{i,m}} = \mathbf{0}, \quad m \neq N_{PI} + j$$

Per il modello stocastico si ha infine

$$\begin{aligned} \mathbf{C}_{yy}^{3(N_B+N_{PR}) \times 3(N_B+N_{PR})} &= \sigma_0^2 \begin{bmatrix} \mathbf{Q}_{\Delta\Delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{RR} \end{bmatrix} \\ \mathbf{Q}_{\Delta\Delta}^{3N_B \times 3N_B} &= \begin{bmatrix} \frac{\sigma_0^2}{\sigma_0^2} \mathbf{Q}_{\Delta_1\Delta_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{\sigma_0^2}{\sigma_0^2} \mathbf{Q}_{\Delta_2\Delta_2} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{\sigma_0^2}{\sigma_0^2} \mathbf{Q}_{\Delta_{N_B}\Delta_{N_B}} \end{bmatrix} \\ \mathbf{Q}_{RR}^{3N_{PR} \times 3N_{PR}} &= \begin{bmatrix} \frac{\sigma_0^2}{\sigma_0^2} \mathbf{Q}_{R_1R_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{\sigma_0^2}{\sigma_0^2} \mathbf{Q}_{R_2R_2} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{\sigma_0^2}{\sigma_0^2} \mathbf{Q}_{N_{PR},N_{PR}} \end{bmatrix} \end{aligned} \quad (8.16)$$

ove tutte le $\mathbf{Q}_{\Delta_i\Delta_i}$ e le $\mathbf{Q}_{R_iR_i}$ sono matrici 3×3 .

Come usuale, la ridondanza del sistema è definita come la differenza fra numero di osservazioni e numero di incognite; in particolare si ha:

$$\begin{aligned}
 r &= \dim(\mathbf{y}) - \dim(\mathbf{x}) = 3(N_B + N_{PR} - N_P) = 3(N_B + N_{PR} - N_{PI} - N_{PR}) \\
 &= 3(N_B - N_{PI})
 \end{aligned}
 \tag{8.17}$$

Talvolta, per semplicità espositiva, si trascura il fatto che ogni punto e ogni base hanno tre componenti, e si pone semplicemente

$$r = N_B - N_{PI} \tag{8.18}$$

Il processo di compensazione avviene secondo lo schema esposto in App. 1 e non appare qui rilevante darne una descrizione; è però utile che al termine del calcolo, vengano riportati e controllati i seguenti indici statistici necessari per l'analisi dei risultati:

1. il risultato del test sul modello globale, ai tipici valori di affidabilità;
2. i residui di osservazione delle basi in ingresso, possibilmente espressi in componenti locali Est, Nord e quota;
3. i corrispondenti residui normalizzati;
4. le ridondanze locali e le affidabilità interne delle osservazioni;
5. le affidabilità esterne delle stime delle coordinate dei punti.

Si propongono alcuni elementi operativi, sostanzialmente basati sull'esperienza: la compensazione di una rete GPS è in genere un processo iterativo ove, ad ogni iterazione, si tenta la soluzione di un particolare problema; inoltre, nel caso di compensazioni relative a campagne di rilievo molto complesse, ovvero con molti punti, molte basi e molte sessioni può essere opportuno dividere la rete in sottoblocchi: si compensano in modo indipendente i singoli sottoblocchi al fine di isolare eventuali problemi con più facilità e da ultimo si effettua la compensazione congiunta dei sottoblocchi. In ogni caso, al termine di ogni passo di compensazione, si deve controllare il superamento o meno del test sul modello globale, al riguardo del quale si possono avere tre esiti:

- I. il test viene ben superato: in tal caso non vi sono problemi da risolvere;
- II. il test non viene superato di poco: in tal caso probabilmente vi sono errori sparsi nel modello stocastico (causa IIA) mentre è improbabile la presenza di un errore grossolano isolato nelle osservazioni (che indichiamo come causa IIB);
- III. il test non viene superato di molto: in tal caso sicuramente vi sono significativi errori di modello stocastico (causa IIIA), oppure uno o più errori grossolani in una o più osservazioni (causa IIIB).

Innanzitutto si controllano i residui di osservazione sulle basi: se una o più basi mostrano residui inaccettabili rispetto alle accuratezze previste si è nelle condizioni IIB o IIIB, ma non sempre è banale identificare la base con errori grossolani: infatti i minimi quadrati non sono robusti, ovvero l'effetto di errori isolati tende a propagarsi su tutti i residui. Il Leave One Out può costituire un possibile approccio per meglio

identificare univocamente il problema: si ripete la compensazione iterativamente, escludendo ad ogni iterazione, ciclicamente, una e una sola osservazione fra quelle sospette: l'eventuale compensazione che porta a un significativo miglioramento del test del modello globale dovrebbe identificare la base problematica; si ricorda ancora una volta che il concetto di inaccettabilità dei residui è connesso al tipo di rilievo oggetto della compensazione e che quindi residui accettabili nella compensazione di un rilievo statico rapido non lo sono nel caso di un rilievo statico prolungato.

Si sia invece nel caso in cui il test sul modello globale non viene superato ma non si trovano evidenze di errori realmente inaccettabili; probabilmente si è nei casi IIA o IIIA, che normalmente si manifestano con residui normalizzati superiori ai valori limite in modo sparso (IIA) o generalizzato (IIIA): questo è un caso molto tipico ed è dovuto alla sovrastima delle accuratezze di stima della basi fornite dalla postelaborazione delle basi: non è necessaria la reiezione di alcune osservazioni poiché non esistono veri errori grossolani, ma può essere opportuno effettuare la ripesatura generale delle stesse, in modo da permettere una valutazione più realistica delle accuratezze delle stime finali delle coordinate dei punti. Il processo di ripesatura deve essere condotto iterativamente, sino alla convergenza del test sul modello globale: le regole generali per condurlo sono esposte in App. 1, qui si puntualizza solo che il processo richiede una certa dimestichezza e esperienza.

Sia la reiezione degli errori grossolani, sia la ripesatura delle osservazioni possono compromettere le ridondanze locali e le affidabilità delle stime rispetto a quanto previsto nella progettazione del rilievo: infatti la prima esclude osservazioni, la seconda modifica il loro peso relativo. Per tale motivo al termine della compensazione è necessario controllare che ridondanze locali e affidabilità siano ancora omogenee per tutte le osservazioni e tutte le incognite; quando ciò non si verifichi, può essere opportuno ripetere il rilievo di alcune basi, in modo da ottenere nuovamente gli equilibri previsti inizialmente.

8.7 Le regioni di confidenza per le stime dei punti

Al termine della compensazione, si vuole descrivere rigorosamente in senso probabilistico la regione di confidenza per la stima della posizione di ogni punto della rete, ovvero la regione dello spazio tridimensionale alla quale il punto appartiene con un certo livello di probabilità: a tal fine si estrae dalla matrice di covarianza globale di tutti i punti la sottomatrice di covarianza del punto stesso, $\mathbf{C}_{\hat{\mathbf{x}}_i \hat{\mathbf{x}}_i}$; la regione di confidenza per il punto, ad un certo livello di probabilità $1 - \alpha$ è data dalla quadratica o ellissoide in tre dimensioni:

$$(\mathbf{x}_i - \hat{\mathbf{x}}_i)^T \mathbf{C}_{\hat{\mathbf{x}}_i \hat{\mathbf{x}}_i}^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}_i) \leq F_{3, m-n}(\alpha)$$

ove $m-n$ è la ridondanza della compensazione che ha portato alle stime, $F_{3,m-n}(\alpha)$ è il valore della funzione F di Fisher a $3, m-n$ gradi di libertà corrispondente alla probabilità $1-\alpha$: in genere si scelgono i valori $\alpha = 0.01, 0.05$, ovvero $P(\alpha) = 99\%, 95\%$; talvolta viene anche utilizzato l'ellissoide di errore standard, che è definito dalla scelta $F_{3,(M-N)}(\alpha) = 1$, e che per $m-n > 10$ corrisponde a $\alpha \cong 0.81$.

8.7.1 L'ellissi di errore in planimetria

In genere, nell'analisi dei risultati di compensazione di una rete GPS, per ogni punto si scinde il problema in planimetria e altimetria: l'ellissi di errore in planimetria è data dalla

$$(\mathbf{x}_i - \hat{\mathbf{x}}_i)^T \mathbf{C}_{\hat{\mathbf{x}}_i \hat{\mathbf{x}}_i}^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}_i) \leq F_{2,m-n}(\alpha) \quad (8.19)$$

ove, in questo caso si ha

$$\mathbf{x}_i = \begin{bmatrix} E \\ N \end{bmatrix}_i, \mathbf{C}_{\hat{\mathbf{x}}_i \hat{\mathbf{x}}_i} = \begin{bmatrix} \sigma_E^2 & \sigma_{EN}^2 \\ \sigma_{EN} & \sigma_N^2 \end{bmatrix}$$

In particolare per l'ellissi di errore standard si sceglie $F_{2,m-n}(\alpha) = 1$, che in questo caso per $m-n > 10$ corrisponde a $\alpha \cong 0.61$.

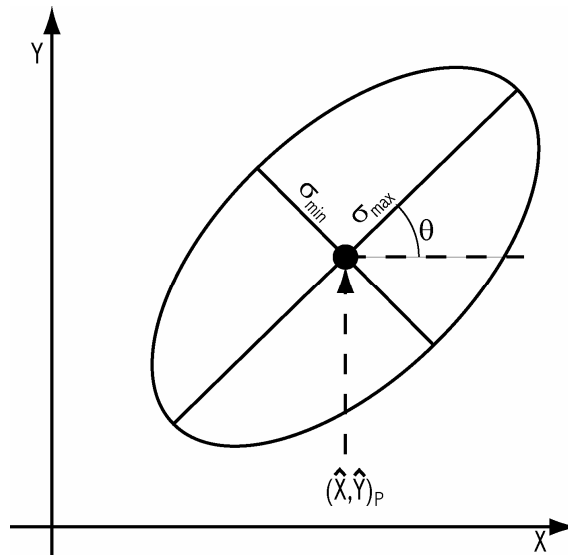


Figura 8.8. I parametri geometrici dell'ellissi di errore planimetrica.

L'ellisse è caratterizzata da 3 parametri: la lunghezza dei semiassi maggiore e minore e l'angolo di direzione del semiasse maggiore, convenzionalmente quello antiorario con la direzione Est lungo il parallelo. Per calcolare i parametri geometrici dell'ellisse di errore si procede nel seguente modo: i semiassi $\sigma_{\max, \min}$ sono dati dalle radici degli autovalori $\lambda_{\max}, \lambda_{\min}$ della matrice $C_{\hat{x}\hat{x}}$, ovvero le soluzioni del sistema

$$\det(C_{\hat{x}\hat{x}} - \lambda \mathbf{I}) = \det \begin{bmatrix} \hat{\sigma}_E^2 - \lambda & \hat{\sigma}_{EN} \\ \hat{\sigma}_{EN} & \hat{\sigma}_N^2 - \lambda \end{bmatrix} = 0$$

si ha dunque

$$\lambda^2 - (\hat{\sigma}_E^2 + \hat{\sigma}_N^2)\lambda + \hat{\sigma}_E^2 \hat{\sigma}_N^2 - \hat{\sigma}_{EN}^2 = 0$$

ovvero

$$\lambda_{\max, \min} = \frac{\hat{\sigma}_E^2 + \hat{\sigma}_N^2}{2} \pm \frac{1}{2} \sqrt{(\hat{\sigma}_E^2 - \hat{\sigma}_N^2)^2 + 4\hat{\sigma}_{EN}^2}$$

corrispondenti a

$$\sigma_{STD \max}, \sigma_{STD \min} = \sqrt{\lambda_{\max}, \lambda_{\min}} \quad (8.20)$$

Indichiamo con θ l'angolo antiorario fra Est e il semiasse maggiore dell'ellisse d'errore; il versore

$$\mathbf{v} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$$

coincide con l'autovettore $\mathbf{e}_{\max}$ corrispondente a $\lambda_{\max}$, ovvero con il vettore soluzione del sistema

$$C_{\hat{x}\hat{x}} \mathbf{e}_{\max} = \lambda_{\max} \mathbf{e}_{\max}; \quad (C_{\hat{x}\hat{x}} - \lambda_{\max} \mathbf{I}) \mathbf{e}_{\max} = 0$$

ovvero

$$(\hat{\sigma}_E^2 - \lambda_{\max}) e_{1\max} + \hat{\sigma}_{EN} e_{2\max} = 0$$

$$\hat{\sigma}_{EN} e_{1\max} + (\hat{\sigma}_N^2 - \lambda_{\max}) e_{2\max} = 0$$

Si ha dunque

$$\tan \theta = \frac{e_2}{e_1} = \frac{\lambda_{\max} - \hat{\sigma}_E^2}{\hat{\sigma}_{EN}}$$

e, dalla relazione trigonometrica

$$\tan 2\theta = 2 \tan \theta / (1 - \tan^2 \theta),$$

si ha

$$\tan 2\theta = \frac{\hat{\sigma}_E^2 - \hat{\sigma}_N^2}{2\hat{\sigma}_{EN}}$$

ovvero

$$\theta = \frac{1}{2} \arctg \frac{\hat{\sigma}_E^2 - \hat{\sigma}_N^2}{2\hat{\sigma}_{EN}} \quad (8.21)$$

I parametri di qualunque altra ellissi, con $F_{2,(m-n)}(\alpha) \neq 1$, sono dati dalla

$$\begin{aligned} \sigma_{\max} &= \sqrt{F_{2,(m-n)}(\alpha)} \cdot \sigma_{STD\max} \\ \sigma_{\min} &= \sqrt{F_{2,(m-n)}(\alpha)} \cdot \sigma_{STD\min} \end{aligned}$$

Appendice 1. I minimi quadrati

Siano date m osservazioni

$$\mathbf{y}_0 = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix}_0$$

per cui valga

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{y} + \mathbf{v} \\ E[\mathbf{y}_0] &= \mathbf{y} \\ \mathbf{C}_{yy} &= \mathbf{C}_{vv} = \sigma_0^2 \mathbf{Q} \end{aligned}$$

ove $\mathbf{y}$ è detto vettore delle osservabili; $\mathbf{v}$ è il vettore degli errori di osservazione; σ_0^2 è la varianza a priori, $\mathbf{Q}$ è la matrice dei cofattori. Sia $\mathbf{x}$ un vettore contenente n parametri incogniti, con $n \leq m$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

Sia noto il modello deterministico del problema, ovvero la relazione funzionale fra $\mathbf{x}$ e $\mathbf{y}$

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{A1.1}$$

$\mathbf{A}$ è detta matrice disegno. Si cercano $\hat{\mathbf{x}}$ e $\hat{\mathbf{y}}$ consistenti, con $\hat{\mathbf{y}}$ a minima distanza da $\mathbf{y}_0$; ovvero $\hat{\mathbf{x}}$ e $\hat{\mathbf{y}}$ tali che

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{A}\hat{\mathbf{x}} \\ (\mathbf{y}_0 - \mathbf{y})^T \mathbf{Q}^{-1} (\mathbf{y}_0 - \mathbf{y}) &= \min \end{aligned} \tag{A1.2}$$

Si noti che in genere il modello viene proposto nella forma

$$\boldsymbol{\eta} = \mathbf{A}\mathbf{x} + \mathbf{b}$$

ove $\mathbf{b}$ è un vettore di costanti, detto termine noto; la semplice sostituzione $\mathbf{y}_0 = \boldsymbol{\eta}_0 - \mathbf{b}$, $\mathbf{C}_{yy} = \mathbf{C}_{\eta\eta}$ riconduce al modello (A1.1), senza perdita di generalità e più semplice nella notazione. Dalle (A1.1), (A1.2) si ricava il sistema normale

$$\mathbf{N}\hat{\mathbf{x}} = \mathbf{A}^T \mathbf{Q}^{-1}(\mathbf{y}_0 - \mathbf{b}),$$

ove $\mathbf{N}$ è definita matrice normale, $\mathbf{N} = \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{A}$. Prima di discuterne la soluzione, introduciamo il problema nel caso non lineare.

A1.1 I minimi quadrati per il caso non lineare

Non esiste una formulazione dei minimi quadrati per il caso non lineare

$$\mathbf{y}_O = \mathbf{y} + \mathbf{v} = \mathbf{f}(\mathbf{x}) + \mathbf{v}$$

ove

$$\mathbf{f}(\mathbf{x}) = \begin{cases} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \dots \\ f_m(x_1, x_2, \dots, x_n) \end{cases}$$

Per risolvere il problema generale è prima necessaria una linearizzazione. Si devono conoscere dei valori approssimati per i parametri incogniti:

$$\tilde{\mathbf{x}}^T = [\tilde{x}_1, \dots, \tilde{x}_n] \cong \mathbf{x}^T : \quad \tilde{x}_1 \cong x_1, \dots, \tilde{x}_n \cong x_n;$$

è allora possibile linearizzare la relazione $\mathbf{y} = \mathbf{f}(\mathbf{x})$ mediante uno sviluppo di Taylor arrestato al primo ordine nell'intorno di $\tilde{\mathbf{x}}$:

$$y_1 \cong f_1(\tilde{\mathbf{x}}) + \frac{\partial f_1}{\partial x_1}(\tilde{\mathbf{x}}) \cdot (x_1 - \tilde{x}_1) + \dots + \frac{\partial f_1}{\partial x_n}(\tilde{\mathbf{x}}) \cdot (x_n - \tilde{x}_n)$$

...

$$y_m \cong f_m(\tilde{\mathbf{x}}) + \frac{\partial f_m}{\partial x_1}(\tilde{\mathbf{x}}) \cdot (x_1 - \tilde{x}_1) + \dots + \frac{\partial f_m}{\partial x_n}(\tilde{\mathbf{x}}) \cdot (x_n - \tilde{x}_n)$$

il problema originale diviene quindi

$$\boldsymbol{\eta}_o = \boldsymbol{\eta} + \mathbf{v} = \mathbf{A}\boldsymbol{\xi} + \mathbf{v}$$

ove

$$\boldsymbol{\eta} = \mathbf{y} - \mathbf{f}(\tilde{\mathbf{x}}) = \begin{bmatrix} y_1 - f_1(\tilde{\mathbf{x}}) \\ \dots \\ y_m - f_m(\tilde{\mathbf{x}}) \end{bmatrix}, \quad \boldsymbol{\xi} = \mathbf{x} - \tilde{\mathbf{x}} = \begin{bmatrix} x_1 - \tilde{x}_1 \\ \dots \\ x_n - \tilde{x}_n \end{bmatrix}, \quad \dim[\mathbf{A}] = m \times n, \quad A_{ij} = \frac{\partial f_i}{\partial x_j}(\tilde{\mathbf{x}})$$

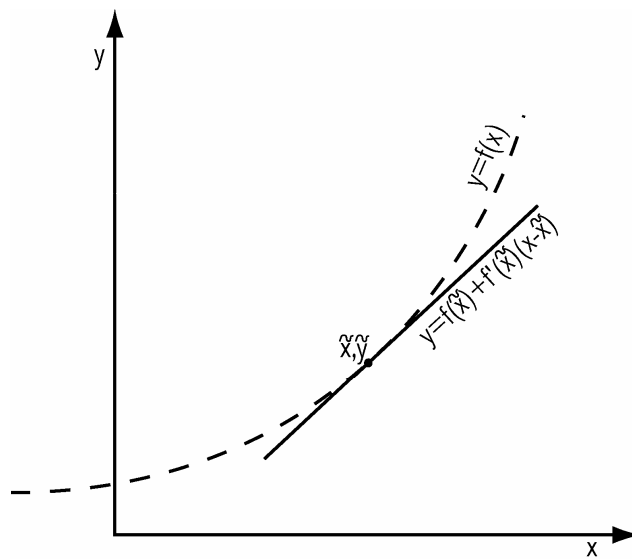


Figura A1.1. La linearizzazione della funzione

A1.1.1 La linearizzazione delle osservazioni di distanza

Sia P un punto di posizione incognita:

$$P = \begin{bmatrix} X_P \\ Y_P \\ Z_P \end{bmatrix}$$

siano invece P^1, P^2, P^3, P^4 quattro punti di posizione nota:

$$P^i = \begin{bmatrix} \bar{X}^i \\ \bar{Y}^i \\ \bar{Z}^i \end{bmatrix}, i = 1, \dots, 4.$$

Da P sono state misurate le distanze ai quattro punti, ottenendo i valori $\rho_{P_0}^1, \rho_{P_0}^2, \rho_{P_0}^3, \rho_{P_0}^4$; si indichi con $\mathbf{p}_0$ il vettore delle osservazioni di distanza. Sono note le coordinate approssimate di P

$$\tilde{P} = \begin{bmatrix} \tilde{X}_P \\ \tilde{Y}_P \\ \tilde{Z}_P \end{bmatrix}$$

si vuole stimare la posizione di P : per far ciò si consideri la generica equazione di osservazione

$$\rho_P^i = \sqrt{(X_P - \bar{X}^i)^2 + (Y_P - \bar{Y}^i)^2 + (Z_P - \bar{Z}^i)^2}$$

Si hanno quindi 4 osservazioni per 3 incognite; per la linearizzazione della generica osservazione di distanza si ha

$$\begin{aligned} \rho_P^i &\cong \sqrt{(\tilde{X}_P - \bar{X}^i)^2 + (\tilde{Y}_P - \bar{Y}^i)^2 + (\tilde{Z}_P - \bar{Z}^i)^2} + \\ &+ \frac{(\tilde{X}_P - \bar{X}^i)}{\sqrt{(\tilde{X}_P - \bar{X}^i)^2 + (\tilde{Y}_P - \bar{Y}^i)^2 + (\tilde{Z}_P - \bar{Z}^i)^2}} (X_P - \tilde{X}_P) + \\ &+ \frac{(\tilde{Y}_P - \bar{Y}^i)}{\sqrt{(\tilde{X}_P - \bar{X}^i)^2 + (\tilde{Y}_P - \bar{Y}^i)^2 + (\tilde{Z}_P - \bar{Z}^i)^2}} (Y_P - \tilde{Y}_P) + \\ &+ \frac{(\tilde{Z}_P - \bar{Z}^i)}{\sqrt{(\tilde{X}_P - \bar{X}^i)^2 + (\tilde{Y}_P - \bar{Y}^i)^2 + (\tilde{Z}_P - \bar{Z}^i)^2}} (Z_P - \tilde{Z}_P) \end{aligned}$$

ovvero

$$\rho_P^i \cong \tilde{\rho}_P^i + \tilde{\mathbf{e}}_P^i \cdot \boldsymbol{\xi} \quad (\text{A1.3})$$

ove

$$\tilde{\rho}_P^i = \sqrt{(\tilde{X}_P - \bar{X}^i)^2 + (\tilde{Y}_P - \bar{Y}^i)^2 + (\tilde{Z}_P - \bar{Z}^i)^2}$$

è la distanza calcolata nelle coordinate approssimate di P ,

$$\tilde{\mathbf{e}}_P^i = \frac{1}{\tilde{\rho}_P^i} \begin{bmatrix} \tilde{X}_P - \bar{X}^i \\ \tilde{Y}_P - \bar{Y}^i \\ \tilde{Z}_P - \bar{Z}^i \end{bmatrix}$$

è il versore approssimato da P^i a P ,

$$\xi = \begin{bmatrix} X_P - \tilde{X}_P \\ Y_P - \tilde{Y}_P \\ Z_P - \tilde{Z}_P \end{bmatrix}$$

è la correzione da apportare alle coordinate approssimate. Il problema assume dunque la forma

$$\begin{bmatrix} \rho_{P_0}^1 \\ \rho_{P_0}^2 \\ \rho_{P_0}^3 \\ \rho_{P_0}^4 \end{bmatrix} = \begin{bmatrix} \tilde{\rho}_P^1 \\ \tilde{\rho}_P^2 \\ \tilde{\rho}_P^3 \\ \tilde{\rho}_P^4 \end{bmatrix} + \begin{bmatrix} \tilde{e}_X^1 & \tilde{e}_Y^1 & \tilde{e}_Z^1 \\ \tilde{e}_X^2 & \tilde{e}_Y^2 & \tilde{e}_Z^2 \\ \tilde{e}_X^3 & \tilde{e}_Y^3 & \tilde{e}_Z^3 \\ \tilde{e}_X^4 & \tilde{e}_Y^4 & \tilde{e}_Z^4 \end{bmatrix} \begin{bmatrix} \xi_X \\ \xi_Y \\ \xi_Z \end{bmatrix}$$

ovvero

$$\boldsymbol{\eta}_O = \boldsymbol{\eta} + \mathbf{v}$$

$$\boldsymbol{\eta} = \boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}$$

$$\boldsymbol{\eta} = \mathbf{A}\xi$$

risolvibile mediante minimi quadrati.

A1.2 La deficienza di rango: il problema e le soluzioni

Si hanno due casi: se $\mathbf{A}$ è di rango pieno, ovvero le sue colonne sono linearmente indipendenti, si ha $\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$; in questo caso il problema non presenta deficienza di rango e $\mathbf{N}$ è invertibile: la discussione della soluzione è nel paragrafo successivo; se invece $\mathbf{A}$ non è di rango pieno, alcune sue colonne sono linearmente dipendenti dalle altre:

$$\mathbf{Ax} = \mathbf{0} \text{ per qualche } \mathbf{x} \neq \mathbf{0}$$

In questo caso il problema presenta deficienza di rango e $\mathbf{N}$ non è invertibile. Definiamo il nucleo di $\mathbf{A}$ come

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{x}_0 \mid \mathbf{A}\mathbf{x}_0 = \mathbf{0}\} \quad (\text{A1.4})$$

evidentemente se $\hat{\mathbf{x}}$ è soluzione di $\mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{y}}$, anche $\hat{\mathbf{x}} + \mathbf{x}_0$ lo è; infatti
 $\mathbf{A}(\hat{\mathbf{x}} + \mathbf{x}_0) = \mathbf{A}\hat{\mathbf{x}} + \mathbf{A}\mathbf{x}_0 = \hat{\mathbf{y}} + \mathbf{0} = \hat{\mathbf{y}}$

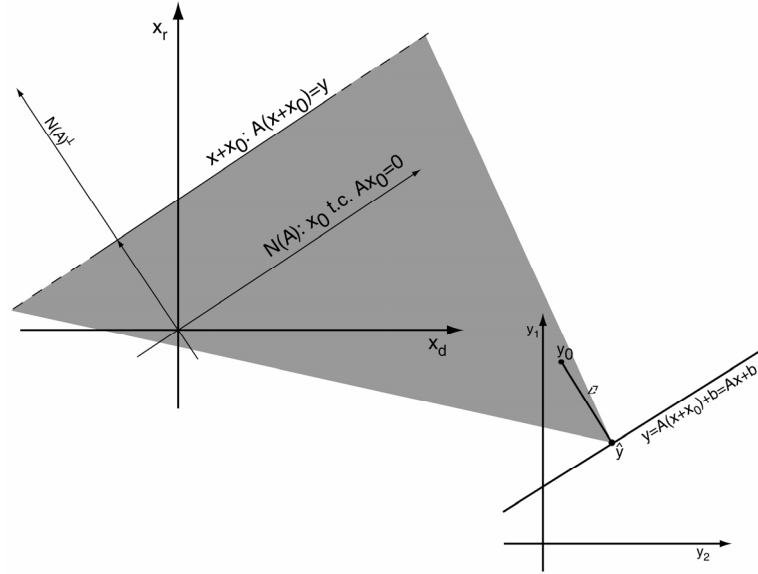


Figura A1.2. Relazioni fra nucleo di A e spazio delle osservazioni

In sostanza le osservazioni non contengono abbastanza informazione per stimare tutti i parametri e tale caratteristica non dipende dall'assenza di ridondanza ma dal disegno del problema. In altri termini, è sempre possibile ottenere una stima ottimale delle osservabili $\mathbf{y}$, ma a questa corrispondono infinite soluzioni per i parametri incogniti.

L'esempio del triangolo di livellazione

Sia dato un triangolo di caposaldi (Fig. A1.3) connessi mediante un anello di livellazione, composto dalle osservazioni $dH_{120}, dH_{230}, dH_{310}$; si suppongano le osservazioni incorrelate e di uguale precisione; si vogliano determinare le quote dei tre punti. La posizione del problema è la seguente

$$\mathbf{y}_0 = \begin{bmatrix} dH_{120} \\ dH_{230} \\ dH_{310} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} + \mathbf{v} = \mathbf{A}\mathbf{x} + \mathbf{v}; \quad \mathbf{C}_{yy} = \sigma_0^2 \mathbf{I}$$

$$\text{ove } \mathbf{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix}$$

il problema è intuitivamente deficiente di rango; in questo caso particolarmente semplice è immediato identificare il nucleo della matrice disegno:

$$\mathcal{N}(\mathbf{A}) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} H, \forall H \right\}, \mathbf{N} = \mathbf{A}^T \mathbf{A}, \det(\mathbf{N}) = 0$$

Casi più complessi richiedono viceversa la caratterizzazione algebrica del nucleo.

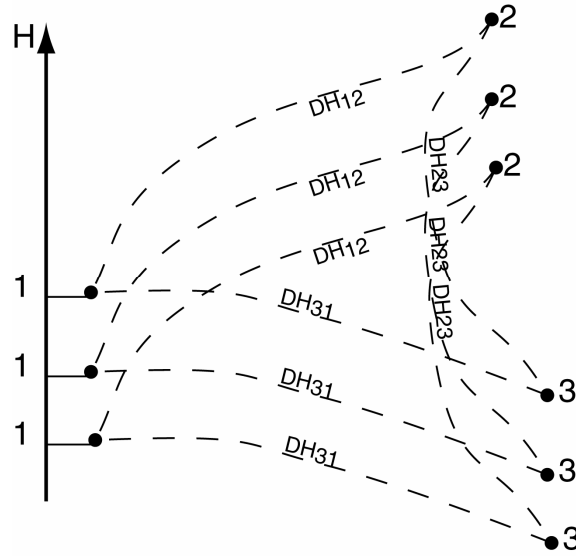


Figura A1.3. Il triangolo di livellazione e la sua deficienza di rango

A1.2.1 L'identificazione del nucleo della matrice disegno

Sia $\mathbf{A}$ di rango r , con $r < n$; ne segue che r colonne sono linearmente indipendenti, le restanti colonne $n - r = d$ sono combinazione delle prime; eventualmente riordinando i parametri incogniti si può scrivere

$$\mathbf{A}_{m \times n} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}_r & \mathbf{A}_d \end{bmatrix} = \begin{bmatrix} \mathbf{A}_r & \mathbf{A}_r \cdot \mathbf{D} \end{bmatrix} = \mathbf{A}_r \begin{bmatrix} \mathbf{I}_r & \mathbf{D} \end{bmatrix}$$

ove $\mathbf{D}$ è una opportuna matrice di coefficienti. E' facile dimostrare che

$$\mathcal{N}(\mathbf{A}) = \begin{bmatrix} -\mathbf{D} \\ \mathbf{I}_d \end{bmatrix} \mathbf{x}_d \quad \forall \quad \mathbf{x}_d \quad (A1.5)$$

Si ha infatti

$$\mathbf{A} \begin{bmatrix} -\mathbf{D} \\ \mathbf{I}_d \end{bmatrix} \mathbf{x}_d = \mathbf{A}_r [\mathbf{I}_r \quad \mathbf{D}] \begin{bmatrix} -\mathbf{D} \\ \mathbf{I}_d \end{bmatrix} \mathbf{x}_d = \mathbf{A}(-\mathbf{D} + \mathbf{D})\mathbf{x}_d = \mathbf{0} \quad \forall \quad \mathbf{x}_d$$

Per dimostrare il viceversa sia $\mathbf{Ax} = \mathbf{0}$; si utilizzi la decomposizione

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_d \end{bmatrix}$$

Si ha dunque $\mathbf{A}_r [\mathbf{I}_r \quad \mathbf{D}] \mathbf{x} = \mathbf{0} \Rightarrow [\mathbf{I}_r \quad \mathbf{D}] \begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_d \end{bmatrix} = \mathbf{0}$ perché $\mathbf{A}_r$ è di rango pieno; quindi

$\mathbf{x}_r + \mathbf{D}\mathbf{x}_d = \mathbf{0} \Rightarrow \mathbf{x}_r = -\mathbf{D}\mathbf{x}_d$, ovvero se $\mathbf{x}$ è tale che $\mathbf{Ax} = \mathbf{0}$, allora

$$\mathbf{x} = \begin{bmatrix} -\mathbf{D}\mathbf{x}_d \\ \mathbf{x}_d \end{bmatrix} = \begin{bmatrix} -\mathbf{D} \\ \mathbf{I}_d \end{bmatrix} \mathbf{x}_d \quad \forall \quad \mathbf{x}_d \quad (A1.6)$$

A1.2.2 La soluzione di un problema con deficienza di rango

Il sistema normale viene riscritto nella forma

$$\begin{bmatrix} \mathbf{I}_r \\ \mathbf{D}^T \end{bmatrix} \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{A}_r [\mathbf{I}_r \quad \mathbf{D}] \begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_d \end{bmatrix} = \begin{bmatrix} \mathbf{I}_r \\ \mathbf{D}^T \end{bmatrix} \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{y}_o$$

ovvero

$$\begin{cases} \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{A}_r \mathbf{x}_r + \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{A}_r \mathbf{D} \mathbf{x}_d = \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{y}_o \\ \mathbf{D}^T \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{A}_r \mathbf{x}_r + \mathbf{D}^T \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{A}_r \mathbf{D} \mathbf{x}_d = \mathbf{D}^T \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{y}_o \end{cases} \quad (A1.7)$$

La seconda equazione è combinazione della prima e quindi non aggiunge informazione al sistema; la prima, ricordando che $\mathbf{A}_r$ è di rango pieno, può essere scritta come

$$\mathbf{x}_r = \mathbf{N}_r^{-1} \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{y}_o - \mathbf{D} \mathbf{x}_d \quad (\text{A1.8})$$

ove

$$\mathbf{N}_r = \mathbf{A}_r^T \mathbf{Q}^{-1} \mathbf{A}_r$$

$r \times r$

La componente $\mathbf{x}_r$ del vettore dei parametri incogniti è funzione di $\mathbf{x}_d$ e non vi è possibilità di stimare contemporaneamente le due componenti. Vi sono differenti possibili alternative per risolvere un problema che presenti deficienza di rango: per la loro rilevanza nel caso della elaborazione dei dati GPS e della compensazione di reti geodetiche, si propongono qui la soluzione a minimi vincoli, la soluzione di Tykonov e la soluzione sovravincolata.

A1.2.2.1 La soluzione a minimi vincoli deterministici

Per eliminare la deficienza di rango si impongono vincoli della forma:

$$\mathbf{H}^T \mathbf{x} = \mathbf{0}, \text{ ove } \dim(\mathbf{H}) = n \times v$$

Si noti che la precedente implica: $\mathbf{x} \in \mathcal{N}(\mathbf{H}^T)$, $\mathbf{x} \perp \mathcal{R}(\mathbf{H})$; i vincoli sono detti minimi quando

$$\begin{cases} \mathcal{N}(\mathbf{A})^\perp + \mathcal{R}(\mathbf{H}) = \mathfrak{R}^n \\ \mathcal{N}(\mathbf{A})^\perp \cap \mathcal{R}(\mathbf{H}) = \mathbf{0} \end{cases}$$

il che implica $v = d$. Per ottenere la stima si cerca la soluzione

$$\hat{\mathbf{x}}_H = \underset{H^T \mathbf{x}_H = \mathbf{0}}{\text{Min}} \|\mathbf{y}_o - \mathbf{A} \mathbf{x}_H\|$$

Con diversi calcoli si dimostra il pieno rango del nuovo sistema, e in particolare si ottiene la soluzione

$$\begin{aligned} \hat{\mathbf{x}}_H &= \mathbf{R}^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_o \\ \mathbf{C}_{x_H x_H} &= \sigma_0^2 \mathbf{R}^{-1} \mathbf{N} \mathbf{R}^{-1} \end{aligned} \quad (\text{A1.9})$$

ove $\mathbf{R} = \mathbf{N} + \mathbf{H} \mathbf{H}^T$.

A1.2.2.2 La soluzione a minimi vincoli stocastici

Si considerano gli stessi vincoli del caso precedente, aggiungendoli però come osservazioni con una propria varianza: si risolve cioè il sistema

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{A}\mathbf{x} + \mathbf{v}_y, & \mathbf{C}_{yy} &= \sigma_0^2 \mathbf{Q} \\ 0 &= \mathbf{H}^T \mathbf{x} + \mathbf{v}_{hh}, & \mathbf{C}_{hh} &= \sigma_{hh}^2 \mathbf{Q}_{hh} = \sigma_0^2 \frac{\sigma_{hh}^2}{\sigma_0^2} \mathbf{Q}_{hh} = \frac{\sigma_0^2}{\lambda} \mathbf{Q}_{hh} \end{aligned} \quad (\text{A1.10})$$

si ottiene la soluzione

$$\begin{aligned} \hat{\mathbf{x}}_S &= \mathbf{R}^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_O, & \mathbf{R} &= (\mathbf{N} + \lambda \mathbf{H} \mathbf{Q}_{hh}^{-1} \mathbf{H}^T) \\ \mathbf{C}_{x_S x_S} &= \sigma_0^2 \mathbf{R}^{-1} \end{aligned} \quad (\text{A1.11})$$

Si può dimostrare che la stima a minimi vincoli stocastici non dipende dalla varianza attribuita alle osservazioni aggiunte e in particolare minimi vincoli stocastici e deterministici forniscono esattamente le stesse stime dei parametri incogniti; cambia invece la covarianza, che nei minimi vincoli stocastici dipende dalla covarianza attribuita ai vincoli. Spesso i vincoli imposti sono nella forma $\boldsymbol{\eta}_0 = \mathbf{H}^T \mathbf{x}$, con $\boldsymbol{\eta}_0 \neq 0$; in tal caso la soluzione è data dalla

$$\hat{\mathbf{x}}_H = \mathbf{R}^{-1} (\mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_O + \lambda \mathbf{H}^T \mathbf{Q}_{hh}^{-1} \boldsymbol{\eta}_0) \quad (\text{A1.12})$$

A1.2.2.3 La soluzione di Tykonov

In questo caso si aggiungono al sistema originario pseudoosservazioni sull'intero vettore dei parametri incogniti; si risolve cioè il sistema

$$\begin{aligned} \mathbf{y}_O &= \mathbf{A}\mathbf{x} + \mathbf{v}_y, & \mathbf{C}_{yy} &= \sigma_0^2 \mathbf{Q} \\ 0 &= \mathbf{x} + \mathbf{v}_{xx}, & \mathbf{C}_{xx} &= \frac{\sigma_0^2}{\lambda} \mathbf{Q}_{xx} \end{aligned} \quad (\text{A1.13})$$

si ha il nuovo sistema normale

$$(\mathbf{N} + \lambda \mathbf{Q}_{xx}^{-1}) \hat{\mathbf{x}}_T = \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0$$

l'aggiunta di valori positivi in diagonale garantisce il pieno rango alla matrice normale; si ha quindi

$$\begin{aligned}\hat{\mathbf{x}}_T &= (\mathbf{N} + \lambda \mathbf{Q}_{xx}^{-1})^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0 \\ \mathbf{C}_{x_T x_T} &= \sigma_0^2 (\mathbf{N} + \lambda \mathbf{Q}_{xx}^{-1})^{-1}\end{aligned}\quad (\text{A1.14})$$

A differenza della soluzione a minimi vincoli, la stima di Tykonov dipende dalla varianza attribuita alle osservazioni aggiunte; in particolare se $\sigma_{xx}^2 \rightarrow 0$, si ha evidentemente

$$\lambda \rightarrow \infty, \hat{\mathbf{x}}_T \rightarrow 0$$

Per contro, se σ_{xx}^2 è troppo grande, $\lambda \rightarrow 0$ e il sistema normale torna deficiente di rango. Da ultimo si noti che l'utilizzo di pseudoosservazioni $\boldsymbol{\eta}_0 = \mathbf{x}$, con $\boldsymbol{\eta}_0 \neq 0$, porta alla soluzione

$$\hat{\mathbf{x}}_T = (\mathbf{N} + \lambda \mathbf{Q}_{xx}^{-1})^{-1} (\mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0 + \lambda \mathbf{Q}_{xx}^{-1} \boldsymbol{\eta}_0) \quad (\text{A1.15})$$

A1.2.2.4 Applicazione delle soluzioni al triangolo di livellazione

Si consideri ora l'esempio del triangolo. Per analizzare meglio le proprietà geometriche delle diverse soluzioni si consideri il caso di assenza di errori nelle osservazioni, ovvero

$$\mathbf{y}_0 = \begin{bmatrix} dH_{120} \\ dH_{230} \\ dH_{310} \end{bmatrix} = \begin{bmatrix} H_2 - H_1 \\ H_3 - H_2 \\ H_1 - H_3 \end{bmatrix} \Rightarrow \mathbf{v} = \mathbf{0}$$

In questo caso la più semplice soluzione di minimi vincoli è quella che considera un punto della rete come punto di zero: ad esempio $H_1 = 0$; se consideriamo i vincoli stocastici, il sistema da risolvere è dunque

$$\begin{cases} \mathbf{y}_0 = \begin{bmatrix} H_2 - H_1 \\ H_3 - H_2 \\ H_1 - H_3 \end{bmatrix}, & \mathbf{C}_{yy} = \sigma_0^2 \mathbf{I} \\ H_1 = 0, & \sigma_H^2 = \sigma_0^2 / \lambda \end{cases}$$

ovvero

$$\begin{cases} \mathbf{y}_0 = \mathbf{A}\mathbf{x}, \mathbf{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \\ 0 = \mathbf{H}^T \mathbf{x}, \mathbf{H}^T = [1 \quad 0 \quad 0] \end{cases}$$

Si ha quindi

$$\mathbf{R} = \begin{bmatrix} 2+\lambda & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \mathbf{R}^{-1} = \frac{1}{3\lambda} \begin{bmatrix} 3 & 3 & 3 \\ 3 & 3+2\lambda & 3+\lambda \\ 3 & 3+\lambda & 3+2\lambda \end{bmatrix}$$

$$\hat{\mathbf{x}}_H = \frac{1}{3\lambda} \begin{bmatrix} 0 \\ \lambda(2dH_{120} - dH_{230} - dH_{310}) \\ \lambda(-2dH_{310} + dH_{230} + dH_{120}) \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 \\ H_2 - H_1 \\ H_3 - H_1 \end{bmatrix}$$

le stime non dipendono quindi dalla covarianza del vincolo e sono date dalla differenza fra le quote “vere” e la quota di vincolo.

Per la covarianza della soluzione si ottiene

$$\mathbf{C}_{x_H x_H} = \begin{bmatrix} \sigma_H^2 & \sigma_H^2 & \sigma_H^2 \\ \sigma_H^2 & \sigma_H^2 & \sigma_H^2 \\ \sigma_H^2 & \sigma_H^2 & \sigma_H^2 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma_0^2 & \sigma_0^2 \\ 0 & \sigma_0^2 & \sigma_0^2 \end{bmatrix}$$

che invece dipende dalla varianza attribuita al vincolo. In particolare se si è attribuito scarso peso ai vincoli, ovvero si è adottato σ_H^2 “grande”, la covarianza delle stime cresce; se si è adottato un σ_H^2 piccolo o nullo (equivalente ai minimi vincoli deterministici) la covarianza è quella minima. A titolo di esercizio, si provi a ripetere il calcolo con il vincolo

$$\bar{H} = \frac{1}{3}(H_1 + H_2 + H_3) = 0 \Rightarrow H^T = \frac{1}{3}[1 \quad 1 \quad 1]$$

ovvero si vincoli a zero il baricentro della rete; questa è detta anche soluzione interna o a minima norma e si dovrebbe ottenere

$$\hat{\mathbf{x}} = \begin{bmatrix} H_1 - \bar{H} \\ H_2 - \bar{H} \\ H_3 - \bar{H} \end{bmatrix}$$

Si ipotizzi ora di costruire il sistema per la soluzione di Tykonov: in questo caso si ha

$$\left\{ \begin{array}{l} \mathbf{y}_0 = \begin{bmatrix} H_2 - H_1 \\ H_3 - H_2 \\ H_1 - H_3 \end{bmatrix}, \quad \mathbf{C}_{yy} = \sigma_0^2 \mathbf{I} \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} H_{10} \\ H_{20} \\ H_{30} \end{bmatrix}, \quad \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_0^2 / \lambda \end{array} \right.$$

ovvero il sistema

$$\left\{ \begin{array}{l} \mathbf{y}_0 = \mathbf{A}\mathbf{x}, \quad \mathbf{C}_{yy} = \sigma_0^2 \mathbf{I} \\ \mathbf{0} = \mathbf{x}, \quad \mathbf{C}_{xx} = \sigma_0^2 \mathbf{I} / \lambda \end{array} \right.$$

In questo caso

$$\mathbf{N} + \lambda \mathbf{I} = \begin{bmatrix} 2 + \lambda & -1 & -1 \\ -1 & 2 + \lambda & -1 \\ -1 & -1 & 2 + \lambda \end{bmatrix},$$

$$(\mathbf{N} + \lambda \mathbf{I})^{-1} = \frac{1}{\lambda(3 + \lambda)} \begin{bmatrix} \lambda + 1 & 1 & 1 \\ 1 & \lambda + 1 & 1 \\ 1 & 1 & \lambda + 1 \end{bmatrix} = \frac{1}{3 + \lambda} \mathbf{I} + \frac{1}{\lambda(3 + \lambda)} \mathbf{ii}^T$$

Osservando che $\mathbf{ii}^T \mathbf{A}^T = \mathbf{0}$ si ha in definitiva

$$\hat{\mathbf{x}}_T = \frac{1}{3 + \lambda} \begin{bmatrix} dH_{310} - dH_{120} \\ dH_{120} - dH_{230} \\ dH_{230} - dH_{310} \end{bmatrix} = \frac{3}{3 + \lambda} \begin{bmatrix} H_1 - \bar{H} \\ H_2 - \bar{H} \\ H_3 - \bar{H} \end{bmatrix}$$

La covarianza è data dalla

$$\mathbf{C}_{x_T x_T} = \sigma_0^2 (\mathbf{N} + \lambda \mathbf{I})^{-1} = \sigma_0^2 \frac{1}{3 + \lambda} (\mathbf{I} + \frac{1}{\lambda} \mathbf{ii}^T)$$

Sia le stime delle incognite sia la stima della covarianza dipendono da λ ; in particolare si hanno le seguenti

$$\lambda \rightarrow \infty: \hat{\mathbf{x}}_T \rightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{C}_{x_T x_T} \rightarrow \mathbf{0}$$

ovvero attribuendo peso molto alto alle pseudoosservazioni, le stime tendono a zero e così la loro covarianza;

$$\lambda \rightarrow 0: \hat{\mathbf{x}}_T \rightarrow \frac{3}{3} \begin{bmatrix} H_1 - \bar{H} \\ H_2 - \bar{H} \\ H_3 - \bar{H} \end{bmatrix} = \begin{bmatrix} H_1 - \bar{H} \\ H_2 - \bar{H} \\ H_3 - \bar{H} \end{bmatrix}, \mathbf{C}_{x_T x_T} \rightarrow \infty$$

ovvero attribuendo peso molto basso alle pseudoosservazioni, le stime tendono alla stima di minima norma; però la covarianza delle stime esplode, perché facendo tendere a zero il peso delle pseudoosservazioni si è implicitamente tornati al sistema (A1.1), che era deficiente di rango.

A1.2.3 La soluzione sovravincolata, ovvero l'inquadrimento di sottoreti

Si consideri un sistema sovravincolato, ovvero del tipo:

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{A}\mathbf{x} + \mathbf{v}_y, \quad \mathbf{C}_{yy} = \sigma_0^2 \mathbf{Q} \\ \boldsymbol{\eta}_0 &= \boldsymbol{\Xi}\mathbf{x} + \mathbf{v}_{hh}, \quad \boldsymbol{\Xi} = \begin{bmatrix} 1 & 0 & \dots & \dots \\ \dots & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & 1 & \dots \end{bmatrix}, \quad \mathbf{C}_{hh} = \frac{\sigma_0^2}{\lambda} \mathbf{Q}_{hh} \end{aligned} \quad (\text{A1.16})$$

con $\dim(\boldsymbol{\eta}_0) = p, \dim(\boldsymbol{\Xi}) = p \times n, p < n$.

Si tratta di un sistema simile a quello di Tykonov, ove però le pseudoosservazioni vengono poste solo su un sottoinsieme dei parametri incogniti; la scelta delle pseudoosservazioni deve essere tale che da ottenere il pieno rango per la nuova matrice disegno, ovvero

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{\Xi} \end{bmatrix} \mathbf{x}_0 = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

per parlare di soluzione sovravincolata deve però essere superata la semplice condizione di minimi vincoli, il che implica perlomeno $p > d$: ad esempio, tale scelta corrisponde al vincolo stocastico di due punti nel triangolo di livellazione. Si ottiene la soluzione

$$\begin{aligned} \hat{\mathbf{x}}_C &= \mathbf{R}_C^{-1} (\mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0 + \lambda \mathbf{\Xi}^T \mathbf{Q}_{hh}^{-1} \boldsymbol{\eta}_0) \\ \mathbf{C}_{xx} &= \sigma_0^2 \mathbf{R}_C^{-1} \end{aligned} \quad (\text{A1.17})$$

ove $\mathbf{R}_C = (\mathbf{N} + \lambda \mathbf{\Xi}^T \mathbf{Q}_{hh}^{-1} \mathbf{\Xi})$.

Tale soluzione presenta le medesime caratteristiche della soluzione di Tykonov: le stime delle osservabili e delle incognite tendono ad adattarsi ai vincoli stocastici al decrescere di σ_{hh}^2 . L'approccio sovravincolato è quello usualmente adottato per la compensazione e l'inquadramento di reti geodetiche locali in reti di ordine superiore: si inseriscono nella sottorete punti noti in numero superiore a quelli strettamente necessari per superare i problemi di deficienza di rango, e le coordinate a priori dei punti noti vengono inserite con una propria varianza, ovvero come osservazioni.

A1.2.4 Due definizioni e una nota

Si definisce deficienza di rango intrinseca di una rete per un certo tipo di osservabili quella non diminuibile modificando il disegno della rete, ovvero aggiungendo nuove osservabili del medesimo tipo al disegno stesso. Si definisce Sistema di Riferimento per una rete di osservabili un sistema di coordinate alle quali sia imposta una scelta arbitraria di vincoli tale da eliminare la deficienza di rango intrinseca della rete. Nella compensazione di reti geodetiche si devono adottare soluzioni che eliminino la deficienza di rango intrinseca: nel caso di nostro interesse, ovvero quello delle reti locali, tipicamente uno o più punti vengono vincolati alle coordinate a priori fornite da precedenti stime.

A1.3 La stima dei minimi quadrati

Dopo aver risolto l'eventuale deficienza di rango, i minimi quadrati forniscono le seguenti stime:

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0 && \text{parametri incogniti,} \\ \hat{\mathbf{y}} &= \mathbf{A} \hat{\mathbf{x}} = \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0 = \mathbf{P}_A \mathbf{y}_0 && \text{osservazioni,} \\ \mathbf{v} &= \mathbf{y}_0 - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{P}_A) \mathbf{y}_0 = \mathbf{P}_{A^\perp} \mathbf{y}_0 && \text{residui di osservazione,} \end{aligned} \quad (\text{A1.18})$$

$$\hat{\sigma}^2 = \frac{\hat{\mathbf{v}}^T \mathbf{Q}^{-1} \hat{\mathbf{v}}}{R} \quad \text{stima a posteriori della varianza;}$$

ove $R = m - n$ è definito ridondanza, o numero di gradi di libertà. Si hanno quindi le stime delle matrici di covarianza:

$$\begin{aligned} \mathbf{C}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} &= \hat{\sigma}^2 \mathbf{N}^{-1} && \text{incognite,} \\ \mathbf{C}_{\hat{\mathbf{y}}\hat{\mathbf{y}}} &= \hat{\sigma}^2 \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T && \text{osservazioni,} \\ \mathbf{C}_{\hat{\mathbf{v}}\hat{\mathbf{v}}} &= \hat{\sigma}^2 (\mathbf{Q} - \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T) && \text{scarti.} \end{aligned} \quad (\text{A1.19})$$

A1.3.1 La stima sequenziale

Al riguardo della stima dei parametri incogniti, si noti il seguente aspetto tecnico: si supponga che il vettore di osservazioni possa esser scomposto in due insiemi di osservazioni scorrelate, ovvero

$$\mathbf{y}_0 = \begin{bmatrix} \mathbf{y}_{10} \\ \mathbf{y}_{20} \end{bmatrix}, \mathbf{Q}_{yy} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{22} \end{bmatrix}$$

si abbia il seguente modello deterministico fra le osservabili e il vettore dei parametri incogniti:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{x} \quad (\text{A1.20})$$

si supponga infine di aver stimato il vettore dei parametri incogniti a partire dal solo sottoinsieme $\mathbf{y}_1$ in base allo stimatore:

$$\hat{\mathbf{x}}_1 = \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{Q}_1^{-1} \mathbf{y}_{10} \quad (\text{A1.21})$$

ove $\hat{\mathbf{x}}_1$ indica che la stima di tutto $\mathbf{x}$ è stata fornita dal sottoinsieme $\mathbf{y}_1$, mentre

$$\mathbf{N}_1 = \mathbf{A}_1^T \mathbf{Q}_1^{-1} \mathbf{A}_1$$

La stima di $\mathbf{x}$ a partire da $\mathbf{y}_0$ completo è data dalla

$$\hat{\mathbf{x}} = \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{y}_0 = (\mathbf{N}_1 + \mathbf{N}_2)^{-1} (\mathbf{A}_1^T \mathbf{Q}_1^{-1} \mathbf{y}_{10} + \mathbf{A}_2^T \mathbf{Q}_2^{-1} \mathbf{y}_{20})$$

la precedente può però essere scritta come

$$\begin{aligned}
 \hat{\mathbf{x}} &= (\mathbf{N}_1 + \mathbf{N}_2)^{-1} (\mathbf{N}_1 \hat{\mathbf{x}}_1 + \mathbf{A}_2^T \mathbf{Q}_2^{-1} \mathbf{y}_{20}) \\
 &= (\mathbf{N}_1 + \mathbf{N}_2)^{-1} ((\mathbf{N}_1 + \mathbf{N}_2) \hat{\mathbf{x}}_1 - \mathbf{N}_2 \hat{\mathbf{x}}_1 + \mathbf{A}_2^T \mathbf{Q}_2^{-1} \mathbf{y}_{20}) \\
 &= \hat{\mathbf{x}}_1 + (\mathbf{N}_1 + \mathbf{N}_2)^{-1} (\mathbf{A}_2^T \mathbf{Q}_2^{-1} \mathbf{y}_{20} - \mathbf{A}_2^T \mathbf{Q}_2^{-1} \mathbf{A}_2 \hat{\mathbf{x}}_1) \\
 &= \hat{\mathbf{x}}_1 + (\mathbf{N}_1 + \mathbf{N}_2)^{-1} \mathbf{A}_2^T \mathbf{Q}_2^{-1} (\mathbf{y}_{20} - \mathbf{A}_2 \hat{\mathbf{x}}_1)
 \end{aligned} \tag{A1.22}$$

per $\mathbf{A}, \mathbf{D}$ quadrate di rango pieno si ha la

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D}^{-1} + \mathbf{CA}^{-1} \mathbf{B})^{-1} \mathbf{CA}^{-1}$$

che applicata al nostro caso, ricordando che $\mathbf{N}_2 = \mathbf{A}_2^T \mathbf{Q}_2^{-1} \mathbf{A}_2$, fornisce

$$\begin{aligned}
 (\mathbf{N}_1 + \mathbf{N}_2)^{-1} &= \mathbf{N}^{-1} = \mathbf{N}_1^{-1} - \Delta \mathbf{N} \\
 \Delta \mathbf{N} &= \mathbf{N}_1^{-1} \mathbf{A}_2^T (\mathbf{Q}_2^{-1} + \mathbf{A}_2 \mathbf{N}_1^{-1} \mathbf{A}_2^T)^{-1} \mathbf{A}_2 \mathbf{N}_1^{-1}
 \end{aligned} \tag{A1.23}$$

ovvero, la stima ottenibile dall'insieme completo delle osservazioni $\mathbf{y}_0$ è ottenuta sommando alla stima ottenuta dal primo sottoinsieme $\mathbf{y}_{10}$ un termine correttivo calcolato sul secondo sottoinsieme; la nuova stima richiede l'inversione solo della matrice $\mathbf{Q}_2$ e della matrice $\mathbf{Q}_2^{-1} + \mathbf{A}_2 \mathbf{N}_1^{-1} \mathbf{A}_2^T$, che ha le stesse dimensioni di $\mathbf{y}_2$: si noti inoltre che

$$(\mathbf{N}_1 + \mathbf{N}_2)^{-1} = \mathbf{N}^{-1} = \mathbf{Q}_{xx}$$

ovvero la (A1.23) fornisce anche la stima della matrice dei cofattori delle incognite dopo l'aggiornamento.

Tale approccio di calcolo prende nome di minimi quadrati sequenziali ed è particolarmente utile per tutti i processi di stima in cui le osservazioni vengono approntate rese disponibili sequenzialmente e non in contemporanea: la tipica applicazione è evidentemente quella delle elaborazioni in tempo reale.

A1.4 La verifica del modello

Si ricorda che il metodo dei minimi quadrati fornisce stime corrette e di minima varianza per i parametri incogniti; inoltre le stime sono indipendenti dal valore di σ_0^2 , che quindi non è necessario conoscere a priori.

Gli errori nel modello deterministico, ovvero nella matrice disegno $\mathbf{A}$, o nel modello stocastico, ovvero nella matrice dei cofattori $\mathbf{Q}$, causano errori di stima: si vogliono ricordare le principali proprietà dello stimatore e i principali metodi per verificare la presenza o l'assenza di errori di modello. Si ricorda che la verifica statistica di ipotesi è

un'operazione che consente di stabilire se, statisticamente, ovvero con una certa probabilità di errore, vale una certa ipotesi H_0 . Tipicamente si costruisce una statistica campionaria che, sotto l'ipotesi H_0 debba seguire una distribuzione nota; che viceversa, qualora H_0 sia sbagliata, vada ad assumere valori “grandi”, ovvero non accettabili statisticamente; si confronta quindi la statistica campionaria con i valori limite ammessi dalla sua distribuzione teorica. Nell'elaborazione dei dati grezzi GPS e nella compensazione delle reti GPS si pone tipicamente l'obiettivo di verificare la correttezza del modello globale, poi si individuano eventuali errori grossolani isolati, infine si corregge il modello stocastico.

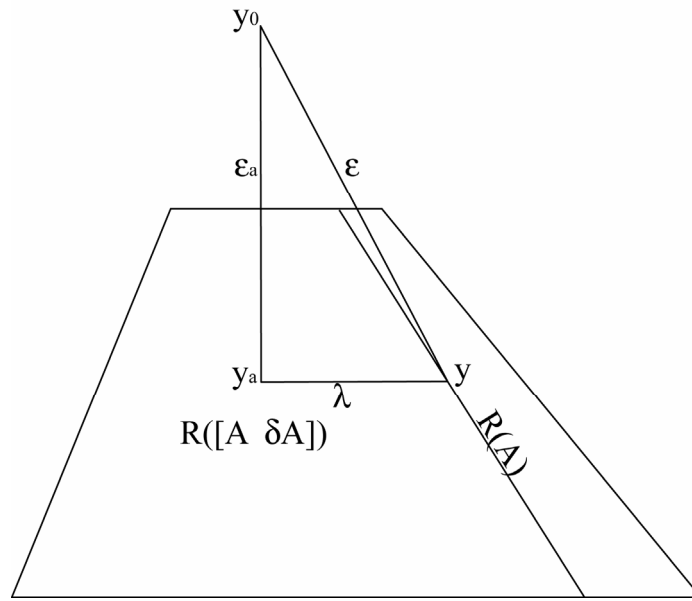


Figura A1.4. La geometria dei minimi quadrati nello spazio delle osservabili.

Si pongono le ipotesi alternative

$$H_0 : \underset{m \times 1}{\mathbf{y}} = \underset{m \times n}{\mathbf{A}} \underset{n \times 1}{\mathbf{x}}, \quad H_a : \underset{m \times 1}{\mathbf{y}} = \begin{bmatrix} \underset{m \times n}{\mathbf{A}} & \underset{m \times b}{\delta \mathbf{A}} \end{bmatrix} \begin{bmatrix} \underset{n \times 1}{\mathbf{x}} \\ \underset{b \times 1}{\delta \mathbf{x}} \end{bmatrix} \quad (\text{A1.24})$$

ove

$$\mathcal{R}\{\begin{bmatrix} \mathbf{A} & \delta \mathbf{A} \end{bmatrix}\} \supset \mathcal{R}\{\mathbf{A}\}$$

Si definisce

$$\boldsymbol{\lambda} = \delta \mathbf{A}^T \mathbf{Q}^{-1} (\mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{Q}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}^{-1}) \mathbf{y}_0 = \delta \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{P}_{\mathbf{A}^\perp} \mathbf{y}_0, \quad \|\boldsymbol{\lambda}\|^2 = \boldsymbol{\lambda}^T \mathbf{Q}_{\lambda\lambda}^{-1} \boldsymbol{\lambda} \quad (\text{A1.25})$$

ove

$$\mathbf{Q}_{\lambda\lambda} = \delta \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{P}_{A^\perp} \delta \mathbf{A}$$

Inoltre si pone

$$\lambda_b = \frac{\delta \mathbf{x}^T \mathbf{Q}_{\lambda\lambda} \delta \mathbf{x}}{\sigma_0^2}$$

si dimostra che sotto H_0 , vale la

$$E \left\{ \frac{\|\lambda\|^2}{\sigma_0^2 b} \right\} = 1, \quad \frac{\|\lambda\|^2}{\sigma_0^2 b} \sim F_{b,\infty} = \frac{\chi_b^2}{b}, \quad (\text{A1.26})$$

sotto H_a invece vale la

$$E \left\{ \frac{\|\lambda\|^2}{\sigma_0^2 b} \right\} = 1 + \frac{\lambda_b}{b}, \quad \frac{\|\lambda\|^2}{\sigma_0^2 b} \sim F_{b,\infty,\lambda_b} \quad (\text{A1.27})$$

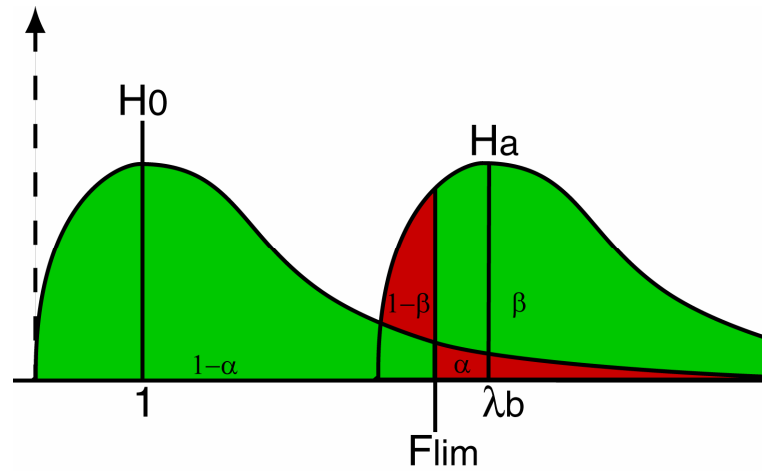


Figura A1.5. La verifica di modello: l'area verde $1-\alpha$ indica la probabilità di corretta accettazione di H_0 quando è vera H_0 ; l'area rossa α indica la probabilità di errore nel rifiutare H_0 quando è vera H_0 ; l'area rossa $1-\beta$ è la probabilità di errore nell'accettare H_0 quando è vera H_a ; l'area verde β indica la probabilità di corretto rifiuto di H_0 quando è vera H_a . Nota: le aree $1-\beta$ e β si estendono verso il basso sino all'asse delle ascisse.

Per eseguire il test si fissa la significatività α ; questa determina $F_{\lim}$ tale che

$$P\left\{\frac{\|\lambda\|^2}{\sigma_0^2 b} > F_{\lim} \mid H_0\right\} < \alpha$$

Se $\frac{\|\lambda\|^2}{\sigma_0^2 b} > F_{\lim}$, H_0 viene rigettata, altrimenti è accettata. Si noti che esiste un valore β tale che

$$P\left\{\frac{\|\lambda\|^2}{\sigma_0^2 b} \leq F_{\lim} \mid H_a\right\} = 1 - \beta$$

ovvero esiste un rischio uguale a β di accettare H_0 anche se, in effetti, è vera H_a ; α è detta significatività del test e rappresenta il rischio di sbagliare rigettando H_0 ; β è detta potenza del test e rappresenta il rischio di sbagliare accettando H_0 : da un punto di vista qualitativo la potenza di un test è ovviamente tanto più alta quanto più l'ipotesi alternativa è dissimile da H_0 .

La formulazione appena introdotta è del tutto generale, quindi generica, e l'esecuzione di un test completo presupporrebbe l'analisi di tutti i possibili errori di modello indicati dalla (A1.24); tipicamente nell'analisi dei risultati si effettuano il test sul modello globale e il test sulle singole osservazioni, sotto descritti.

A1.4.1 Il test del χ^2 o test globale sul modello

L'ipotesi alternativa è che gli errori di modello siano distribuiti su tutta la matrice disegno e quindi si ha

$$\begin{aligned} H_0 : \mathbf{y} &= \mathbf{A}\mathbf{x} \\ H_a : \mathbf{y} &= [\mathbf{A} \quad \delta\mathbf{A}] \begin{bmatrix} \mathbf{x} \\ \delta\mathbf{x} \end{bmatrix}, \quad \mathcal{R}\{[\mathbf{A} \quad \delta\mathbf{A}]\} = R^m \end{aligned} \tag{A1.28}$$

Sotto H_0

$$\frac{\|\lambda\|^2}{\sigma_0^2(m-n)} = \frac{\boldsymbol{\varepsilon}^T \mathbf{Q}^{-1} \boldsymbol{\varepsilon}}{\sigma_0^2(m-n)} \sim \frac{\chi_{(m-n)}^2}{m-n}$$

La statistica di test è quindi

$$\frac{\hat{\sigma}_0^2}{\sigma_0^2}(m-n) = \chi_{sp}^2 \quad (A1.29)$$

Sia α la significatività del test e $\chi_{lim}^2 = \chi_{m-n}^2(\alpha)$ il valore limite tale che

$$P(0 \leq \chi_{m-n}^2 \leq \chi_{lim}^2) = 1 - \alpha$$

se $\chi_{sp}^2 \leq \chi_{lim}^2$, H_0 viene accettata; se $\chi_{sp}^2 > \chi_{lim}^2$ H_0 viene rigettata.

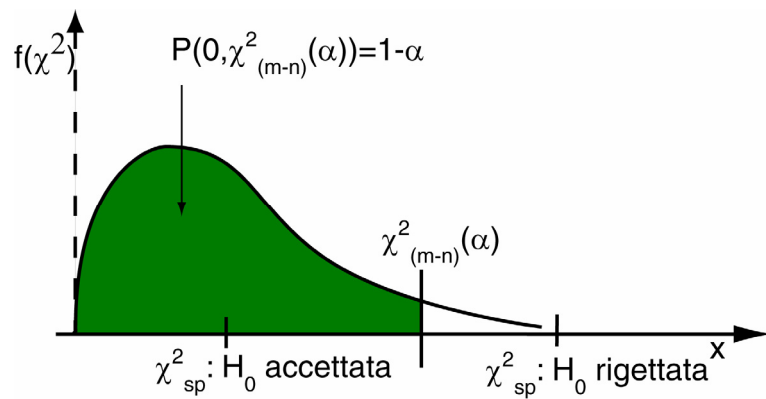


Figura A1.6. Il Test del χ^2 .

A1.4.2 Il test locale sulla singola osservazione

Ci si pone l'obiettivo di identificare errori di modello su una singola osservazione y_{i_0} , e quindi l'ipotesi fondamentale è

$$H_0 : \mathbf{y} = \mathbf{A}\mathbf{x}$$

l'ipotesi alternativa è

$$H_a : \mathbf{y} = [\mathbf{A} \quad \mathbf{e}_i] \begin{bmatrix} \mathbf{x} \\ \delta_i \end{bmatrix}, \quad \mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i \quad (A1.30)$$

Sotto H_0

$$\frac{\|\lambda\|^2}{\sigma_0^2(m-n)} \sim \chi_{(1)}^2$$

ovvero

$$\sqrt{\frac{\|\lambda\|^2}{\sigma_0^2(m-n)}} = z_{sp} \sim \mathcal{N}[0,1] = Z$$

Usualmente nell'esecuzione del test si adotta l'approssimazione $\mathbf{Q}_{\lambda\lambda} \cong \delta \mathbf{A}^T \mathbf{P}_{A^\perp} \delta \mathbf{A}$, poiché allora la precedente diviene semplicemente

$$z_{sp} = \frac{\hat{\varepsilon}_i}{\sigma_{\varepsilon_i}} \sim Z \quad (\text{A1.31})$$

Il test è a due code, ovvero si devono rigettare sia scarti negativi sia scarti positivi in modulo troppo grandi: quindi, definito z_{lim} il valore teorico tale che

$$P(0 \leq |z| \leq z_{\text{lim}}) = 1 - \alpha$$

$$P(|z| > z_{\text{lim}}) = \alpha$$

se $|z| \leq z_{\text{lim}}$ H_0 viene accettata;

se $|z| > z_{\text{lim}}$ H_0 viene rigettata.

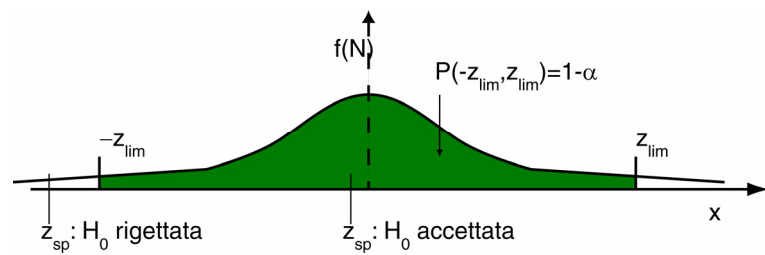


Figura A1.7. Il test a due code sulla normale standardizzata.

A1.4.3 L'affidabilità interna del test sulla singola osservazione

Sotto H_a si ha, per l'osservazione y_i

$$z_{sp} \sim \mathcal{N}\left[\frac{\sqrt{Q_{\lambda\lambda ii}} \delta_i}{\sigma_0}, 1\right] \quad (\text{A1.32})$$

Per effettuare il test si è fissato α e, da questo si è derivato $z_{\lim}$: per ogni errore δ_i è anche possibile calcolare

$$\int_{-\infty}^{z_{\lim}} \mathcal{N}\left[\frac{\sqrt{Q_{\lambda\lambda ii}} \delta_i}{\sigma_0}, 1\right] = 1 - \beta,$$

ovvero la potenza del test rispetto a δ_i . In alternativa si può procedere in modo opposto ovvero si fissa β e si ricava numericamente il valore δ_i tale che

$$\int_{-\infty}^{z_{\lim}} \mathcal{N}\left[\frac{\sqrt{Q_{\lambda\lambda ii}} \delta_i}{\sigma_0}, 1\right] = 1 - \beta$$

δ_i è il massimo errore nascosto, ovvero non rilevabile con potenza β , nell'osservazione i-esima; anche senza svolgere il calcolo esplicito, è possibile ricavare qualche indicazione qualitativa su di esso. Dalla (A1.32) risulta che δ_i è funzione di α e β , ma è anche inversamente proporzionale a $\sqrt{Q_{\lambda\lambda ii}}$, ovvero

$$\delta_i = f(\alpha, \beta) \frac{\sigma_0}{\sqrt{Q_{\lambda\lambda ii}}} = f(\alpha, \beta) \frac{\sigma_0}{\sqrt{\mathbf{e}_i^T \mathbf{Q}^{-1} (\mathbf{I} - \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1}) \mathbf{e}_i}}$$

Adottando la semplificazione

$$\mathbf{e}_i^T \mathbf{Q}^{-1} (\mathbf{I} - \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1}) \mathbf{e}_i \cong \mathbf{e}_i^T (\mathbf{I} - \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1}) \mathbf{e}_i = P_{A^\perp ii}$$

si definisce la ridondanza locale dell'osservazione i-esima $r_i = P_{A^\perp ii}$ e si ottiene

$$\delta_i = f(\alpha, \beta) \frac{\sigma_{y_i}}{\sqrt{r_i}}$$

si hanno due casi limite

$$r_i \rightarrow 0 : \delta_i \rightarrow \infty$$

la ridondanza locale è uguale a zero quando l'osservazione è completamente proiettata sul $\mathcal{R}\{\mathbf{A}\}$: in tal caso il massimo errore nascosto tende all'infinito.

$$r_i \rightarrow 1 : \delta_i \rightarrow \min$$

la ridondanza locale è uguale a uno quando l'osservazione è completamente proiettata su $\mathcal{R}\{\mathbf{A}\}^\perp$: in tal caso il massimo errore nascosto è minimo.

Si definisce affidabilità interna dell'osservazione i -esima con significatività α e potenza β il massimo errore nascosto in essa, δ_i . La peggiore affidabilità interna fra tutte le osservazioni è data dalla:

$$\delta = \max_i \delta_i \quad (\text{A1.33})$$

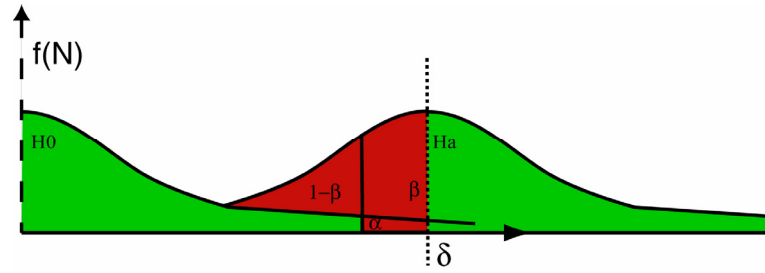


Figura A1.8. Il massimo errore nascosto sulla singola osservazione. Viene rappresentato solo il semispazio positivo.

A1.4.4 L'affidabilità esterna del test sulla singola osservazione

Dato un errore nascosto δ_i nell'osservazione i -esima, ci si chiede come esso influisca sulla stima dei parametri incogniti. Si consideri

$$\delta \hat{\mathbf{x}}(\delta_i) = \mathbf{N}^{-1} \mathbf{A}^T \mathbf{Q}^{-1} \mathbf{e}_i \delta_i$$

Si hanno nuovamente due casi limite: nel caso

$$r_i = 0 : \mathbf{P}_{Aii} = 1 : \mathbf{e}_i \delta_i \in \mathcal{R}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^T)$$

ovvero l'errore si proietta nelle stime delle incognite. Viceversa, nel caso

$$r_i = 1 : \mathbf{P}_{Aii} = 0 : \mathbf{e}_i \delta_i \in \mathcal{N}(\mathbf{A}^T) \perp \mathcal{R}(\mathbf{A})$$

l'errore non si proietta nelle stime delle incognite.

L'affidabilità esterna del parametro x_j è definita mediante la

$$\delta \hat{x}_j = \max_i (\delta \hat{x}_j(\delta_i)) \quad (\text{A1.34})$$

La peggior affidabilità esterna su tutte le incognite è

$$\delta \hat{x} = \max_j (\delta \hat{x}_j) \quad (\text{A1.35})$$

A1.4.5 La prassi di esecuzione del test locale sulla singola osservazione

Se si suppone di non conoscere σ_0^2 si può usare $\hat{\sigma}_0^2$: in questo caso la statistica di test è data dalla

$$\frac{\hat{\varepsilon}_i}{\hat{\sigma}_0 \sqrt{Q_{eeii}}} \sim \tau_{(m-n)}$$

se H_0 è vera

$$\tau_{sp} \sim \tau_{(m-n)}$$

ove $\tau_{(m-n)}$ è la distribuzione di Thomson a $m-n$ gradi di libertà. In effetti la presenza anche di un solo errore grossolano tipicamente modifica gli scarti delle altre osservazioni e quindi può influire sulla stima del $\hat{\sigma}_0^2$, che in definitiva risulta non attendibile: in tale situazione si costruiscono i cosiddetti residui pseudo-normalizzati

$$\frac{\hat{\varepsilon}_i}{\sqrt{Q_{eeii}}} \text{ oppure } \frac{\hat{\varepsilon}_i}{\sigma_0 \sqrt{Q_{eeii}}}$$

Anziché confrontare i residui pseudo-normalizzati con un valore limite dato da una statistica, li si confronta in senso relativo: se ve ne è uno significativamente più grande degli altri, si elimina l'osservazione corrispondente, ricompensando i dati. Nel caso di persistenza di problemi, si procede quindi iterativamente, provando a escludere ulteriori osservazioni e verificando ad ogni esclusione la convergenza del test sul modello globale: al termine si devono infine controllare le osservazioni eliminate, calcolando i loro scarti rispetto alla soluzione che le ha escluse, per decidere se eliminarle definitivamente o reintrodurle.

A1.5 La ripesatura delle osservazioni

Qualora il test sul modello globale non sia stato superato ma vi sia una situazione con scarti normalizzati omogenei, può esservi semplicemente un problema di sottostima delle varianze a priori delle osservazioni in ingresso, ovvero una sovrastima delle loro accuratezze. In genere l'errore di modello stocastico è su gruppi di osservazioni piuttosto che su osservazioni singole: si pone il problema, a seguire la stima, di ristimare le accuratezze delle osservazioni. Siano

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ \dots \\ y_{ip} \end{bmatrix}, \mathbf{C}_{ii} = \sigma_i^2 \mathbf{Q}_{ii}, \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \dots \\ \mathbf{y}_q \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \sigma_1^2 \mathbf{Q}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_q^2 \mathbf{Q}_{qq} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} \mathbf{A}_1 \\ \dots \\ \mathbf{A}_q \end{bmatrix} \mathbf{x} \quad (\text{A1.36})$$

ove si ritengano incogniti, oltre al vettore $\mathbf{x}$, anche i diversi σ_i^2 . Si pone un certo σ_0^2 a priori e si risolve il problema ai minimi quadrati utilizzando

$$\mathbf{Q} = \begin{bmatrix} \frac{\sigma_1^2}{\sigma_0^2} \mathbf{Q}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\sigma_q^2}{\sigma_0^2} \mathbf{Q}_{qq} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{qq} \end{bmatrix} \quad (\text{A1.37})$$

Esiste un modello rigoroso per la stima a posteriori dei σ_i^2 , che però risulta molto complesso: un compromesso fra correttezza e semplicità di implementazione tipicamente adottato nei processi di ripesatura è quello dato dalla seguente formula:

$$\sigma_i^2 = \frac{\hat{\mathbf{\epsilon}}_i^T \mathbf{K}_{ii}^{-1} \hat{\mathbf{\epsilon}}_i}{m_i - \text{tr}\{\mathbf{A}_i \mathbf{N}^{-1} \mathbf{A}_i^T \mathbf{K}_{ii}^{-1}\}}$$

che può essere scritta come

$$\sigma_i^2 = \frac{\hat{\mathbf{\epsilon}}_i^T \mathbf{K}_{ii}^{-1} \hat{\mathbf{\epsilon}}_i}{\sum_{j \in i} r_{jj}} \quad (\text{A1.38})$$

Effettuata la stima delle covarianze è necessario stimare nuovamente le incognite e quindi ristimare le covarianze: il processo è tipicamente iterativo e si arresta alla convergenza dei risultati.

A1.6 L'accuratezza dei parametri stimati

Siano stati eseguiti il test globale sul modello e il filtraggio delle osservazioni, con esito finale positivo. Si considera ora la stima dei parametri, $\hat{\mathbf{x}}$ e della relativa matrice di covarianza $\mathbf{C}_{\hat{\mathbf{x}}} = \hat{\sigma}_0^2 \mathbf{N}^{-1}$: la regione di confidenza per il vettore dei parametri incogniti ad un certo livello di probabilità $1 - \alpha$ è data dalla

$$(\mathbf{x} - \hat{\mathbf{x}})^T \mathbf{C}_{\hat{\mathbf{x}}}^{-1} (\mathbf{x} - \hat{\mathbf{x}}) \leq F_{n, (m-n)}(\alpha) \quad (\text{A1.39})$$

ove $F_{n, (m-n)}(\alpha)$ è il valore della funzione F di Fisher a $n, m-n$ gradi di libertà, corrispondente alla probabilità $P = 1 - \alpha$; in genere si scelgono i valori $\alpha = 0.01/0.05/0.10$, ovvero $P = 99\%/95\%/90\%$ rispettivamente. Tipicamente si è interessati alla regione di confidenza per un sottoinsieme di parametri incogniti, ξ , con $\dim(\xi) = r$: per analizzare la regione di confidenza di ξ si estrae dal vettore $\hat{\mathbf{x}}$ il relativo sottovettore $\hat{\xi}$; quindi si estrae dalla matrice di covarianza totale $\mathbf{C}_{\hat{\mathbf{x}}}$ la matrice $\mathbf{C}_{\hat{\xi}\hat{\xi}}$. La regione di confidenza con probabilità $P = 1 - \alpha$ per il vettore ξ è data dalla

$$(\xi - \hat{\xi})^T (\mathbf{C}_{\hat{\xi}\hat{\xi}})^{-1} (\xi - \hat{\xi}) \leq F_{r, (m-n)}(\alpha) \quad (\text{A1.40})$$

Tale quadratica rappresenta in generale un iperellissoide in $\mathbb{R}^r$; si ha quindi un ellissoide per le stime di coordinate di un punto in tre dimensioni, un'ellissi per posizioni sul piano e un semplice intervallo per stime di quote.

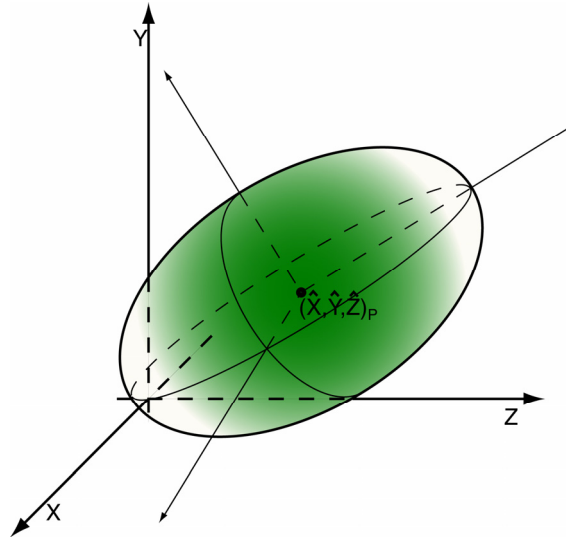


Figura A1.9. L'ellissoide di confidenza in tre dimensioni

Bibliografia, acronimi e siti

- AAVV, *Networked Transport of RTCM via Internet Protocol (NTRIP), version 1.0*, http://igs.ifag.de/index_ntrip.htm, 2004
- Abidin H. Z., *On-the-fly ambiguity resolution: formulation and results*, Manuscripta Geodaetica, vol. 18(6), 1993
- Adam J., et al, *The European Reference System coming of age*, Geodesy Beyond 2000, IAG General Assembly, 1999, IAG Symposia, Vol. 121, Springer-Verlag, 1999
- Altamimi Z., Collilieux X., Legrand J., Garayt B., Boucher C., *ITRF2005: A new release of the International Terrestrial Reference Frame based on time series of station positions and Earth Orientation Parameters*, Journal of Geophysical Research, Vol. 112, 2007
- Aoudia A., Barzaghi R., Borghi A., Sabadini R., Marotta A. M., Panza G., Troisi C., Manzino A. M., Roggero M., Lucchetta A., Carraro C., Zampedri G., Laffi R., Crotta S., De Donatis S., Gerbino P. G., Sguerso D., Voelksen C., Drewes H., Valpersdorf A., Zivcic M., *Il progetto ALPS-GPSQuakenet*, Atti della 8a Conferenza Nazionale ASITA, Roma, 2004
- Bancroft S., *An Algebraic Solution of the GPS Equations*, IEEE Trans. Aerospace and Electronic Systems, vol. AES-21(1), 1985
- Barbarella M., *Reti di stazioni permanenti per il rilievo in tempo reale*, Bollettino SIFET No 2, 2008
- Baroni, L., Cauli F., Donatelli D., Farolfi G., Maseroli R., *La rete dinamica nazionale ed il nuovo sistema di riferimento ETRF2000*, Servizio Geodetico - Istituto geografico Militare – Firenze, online sul sito IGM, 2009
- Barzaghi R., Borghi A., Crespi M., Pietrantonio G., Riguzzi F., *GPS Permanent Network Solution: the Impact of Temporal Correlations*, V Hotine-Marussi Symposium on Mathematical Geodesy, IAG Symposia, Vol. 127, Springer-Verlag, 2004
- DeMets et al., *Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions*, Geophys. Res. Lett. Vol. 21 No. 20, 1994
- Beutler G., Rothacher M., Schaer S., Springer T.A., Kouba J., Neilan R.E., *The International GPS Service (IGS): An Interdisciplinary Service in Support of Earth Sciences*, Adv. Space Res. Vol. 23, No. 4, 1999
- Biagi L., Caldera S., Mazzoni A., *Compensazione della rete dinamica nazionale dell'IGM: procedure applicate e risultati ottenuti dal gruppo G3*, Relazione consegnata a IGM e in stampa sul Bollettino SIFET, 2009.

- Biagi L., Caldera S., Visconti M. G., *A first estimate of the transformation from the global IGS and the Italian ETRF89-IGM95 RF for the Italian Peninsula*, GRF2006, IAG Symposia Vol. 134, 2009
- Biagi L., Crespi M., Benciolini B., Biagi L., Crespi M., Manzino A., Roggero M., *Reference frames for GNSS positioning services: some problems and proposals*, Journal of Applied Geodesy, No 2, 2008
- Biagi L., Sansò F., *Sistemi di riferimento in Geodesia: algebra e geometria dei minimi quadrati per un modello con deficienza di rango*, - Parti I, II, III, Bollettino di Geodesia e Scienze Affini, anno LXII, 2003, N°. 4, anno LXIII, Nn 1 e 2, 2004
- Biagi L., Sansò F. Edizioni., *Il Servizio di Posizionamento in Regione Lombardia e la prima sperimentazione sui servizi di rete in tempo reale*, Bollettino SIFET No 3, 2006
- Biagi L., Sansò F. Edizioni., *I Servizi di posizionamento satellitare per l'e-government*, Geomatics Workbooks, Vol. 7, 2007
- Blewitt, G., Lavallée D., *Effects of annual signals on geodetic velocity*, Journal of geophysical research, Vol. 107, NO B7, 2002
- Boucher, C., Altamimi Z., Memo, *Specifications for reference frame fixing in the analysis of a EUREF GPS campaign*, V6, <http://lareg.ensg.ign.fr/EUREF/>, 2007
- Brown N., Keenan R., Richter B., Troyer L., *Advances in Ambiguity Resolution for RTK Applications Using the New RTCM V3.0 Master-Auxiliary Messages*, Proceedings of ION GNSS 2005, 2005
- Brunner F. K., Gu M., *An improved model for the dual frequency ionospheric correction of GPS observations*, Manuscripta Geodaetica 16(3), 1991
- Cina A., *GPS, Principi, modalità e tecniche di Posizionamento*, Celid, 2001
- Cina A., Manzino A., Piras M., *Le reti NRTK alla luce del nuovo formato RTCM3*, Bollettino SIFET No 2, 2008
- Crespi M., *A software Package for the adjustment and the analysis of GPS control networks* In Reports on Survey and Geodesy in Memory of Proff. A. Gubellini and G. Folloni, Ed Nautilus, Bologna, 1996
- Crespi M., *Il trattamento dei dati GPS: vantaggi e limiti degli approcci tradizionali*, Ricerche di Geodesia, Topografia e Fotogrammetria No 11, CLUP 1996
- Dach R., Hugentobler U., Fridez P., Meindl M., *Bernese GPS Software Version 5.0*, Astronomical Institute, University of Berne, 2007
- Davis P., Blewitt G., *Methodology for global geodetic time series estimation: a new tool for geodynamics*, Journal of Geophysical Research, vol 105, no B5, May, 10, 2000
- Dermanis, A., *On the maintenance of a Proper reference frame for VLBI and GPS Global Networks*, Geodesy, The Challenge of the third millennium, Springer, 2001

- Dermanis A., *The ITRF beyond the linear model. Choices and challenges*, IAG Symposia, Vol. 132, Springer-Verlag, 2007
- Grejner-Brzezinska D. A., Kashani I., Wielgosz P., *On accuracy and reliability of instantaneous network RTK as a function of network geometry, station separation, and data processing strategy*, GPS Solutions 2005, No 9, 2005
- Dousa J., *On the specific aspects of precise tropospheric path delay estimation in GPS analysis*, IAG Symposia, Vol. 125, Springer-Verlag, 2001
- Eissfeller, B., Tiberius C., Pany T., Biberger R., Schueler T., Heinrichs G., *Real-time kinematic in the light of GPS modernization and Galileo*, Proceedings of ION GPS 2001, 2002
- EPN, *Guidelines for EPN analysis centres*, online sul sito EPN
- Euler H-J., Keenan C.R., Zebhauser B.E., Wübbena G., *Study of a Simplified Approach in Utilizing Information from Permanent Reference Station Arrays*, Proceedings of ION GPS 2001, 2001
- Ferland R., Gendt. G., Schone T., *IGS Reference Frame Maintenance, Proceedings of IGS: Celebrating a decade of the International GPS Service*, Berne, March 1-5, 2004, AIUB, Berne, 2004.
- Geodesist's Handbook*, Bulletin Geodesique, 66, 1992
- Habrich, H., *Geodetic Applications of the Global Navigation Satellite System (GLONASS) and of GLONASS/GPS Combinations*, Inauguraldissertation der Philosophisch-naturwissenschaftlichen Fakultät der Universität Bern vorgelegt von Deutschland, 1999, reperibile online
- Heroux P., Kouba J., *GPS precise point positioning using IGS orbit products*, Physics and Chemistry of the Earth Part A, Vol. 26, 2001
- Jimenez-Munt I., Sabadini R., Gardi A., Bianco G., *Active deformation in the Mediterranean from Gibraltar to Anatolia inferred from numerical modeling and geodetic and seismological data*. J. Geophys. Res Vol 108, 2006
- Klobuchar J. A., *Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users*, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-23(3), 1987
- Kouba J., *A guide to using International GPS Service (IGS) products*, on line sul sito IGS
- Landau H., Vollath U., Chen X., *Virtual Reference Station Systems*, Journal of Global Positioning Systems Vol. 1, No. 2, 2002
- Leick A., *GPS Satellite Surveying*, Third Edition, Wiley & Sons, 2004
- Liu J., Cannon M.E., Alves P., Petovello M.G., Lachapelle G., MacGougan G., Degroot L., *A performance comparison of single and dual frequency GPS ambiguity resolution strategies*, GPS Solutions, Vol.7, No.2, 2003
- Manzino A., *Lezioni di Topografia*, Dipartimento di Georisorse e Territorio, 2000

- McCarthy D. D., Petit G. eds., *IERS Conventions*, IERS Technical Note No. 32, Verlag des BKG, Frankfurt am Main, 2004
- Radio Technical Commission for Maritime Services, *RTCM recommended standards for differential GNSS service*, Version 3.0, RTCM paper, 2004.
- Rothacher M., Springer T. A., Schaer S., Beutler G., *Processing strategies for regional GPS networks*, IAG Symposia Vol. 118, Springer-Verlag, 1998
- Saastamoinen J., *Atmospheric correction for the troposphere and stratosphere in radio ranging satellites*, Use of artificial satellites for geodesy, Geophysics Monograph Series, vol. 15, 1972
- Sansò F., *Quaderni di trattamento statistico dei dati*, Città Studi Edizioni, 1998
- Sansò F., *Navigazione geodetica e rilevamento cinematico*, Polipress, 2006
- Seeber G., *Satellite Geodesy*, 2nd edition, Walter de Gruyter, 2003
- Serpelloni e altri, *Data Analysis of permanent GPS networks in Italy and surrounding regions: application of a distributed processing approach*, Annals of Geophysics vol 49, n. 4/5, 2006
- Surace L., *La nuova rete geodetica nazionale IGM95: risultati e prospettive di utilizzazione*, Bollettino di Geodesia e Scienze affini, Anno LVI, N° 3, 1997
- Teunissen P. J., *Quality Control in Geodetic Networks*, Optimization and Design of Geodetic Networks, Springer-Verlag, 1985
- Teunissen, P.J., *The least squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation*, Journal of Geodesy, 70, 1995
- Teunissen P. J., Joosten P., Tiberius C., *A comparison of TCAR, CIR and Lambda GNSS Ambiguity Resolution*, Proceedings of ION GPS 2002, 2002
- Torres, J., *EUREF – The Infrastructure for geo-referencing in Europe*, GeoInformatics, No. 6, Vol. 8, 2005
- Verhagen, S., *Performance Analysis of GPS, Galileo and Integrated GPSGalileo*, Proceedings of ION GPS 2002, 2002
- Vollath U., Patra R., Chen X., Landau H., Allison T., *Galileo/Modernized GPS: A New Challenge to Network RTK*, Proceedings of ION GNSS 2004, 2004
- Williams SDP, *CATS: GPS coordinate time series analysis software*, GPS Solution 2008-12, 2008
- Wübbena G., *On the modelling of GNSS observations for high-precision position determination*, Translation of Wübbena, G., Zur Modellierung von GNSS-Beobachtungen für die hochgenaue Positionsbestimmung. Wissenschaftliche Arbeiten Fachrichtung Vermessungswesen an der Universität, Hannover, Festschrift Prof. G. Seeber zum 60. Geburtstag, Nr. 239, Hannover, on line su <http://www.geopp.de/download/>, 2001

Zumberge, J. F., Heflin, M. B., Jefferson, D. C., Watkins, M. M., Webb, F. H., *Precise point positioning for efficient and robust analysis of GPS data from large networks*, Journal of Geophysical Research, 102, 1997

Acronimi e siti

Vengono qui riportati gli acronimi utilizzati nel volume; per quelli relativi a enti o istituzioni si riporta anche il sito web di riferimento. Per alcune sigle non esiste (o non si è trovata) la dizione estesa cui essi corrispondono: ne viene semplicemente data una descrizione, evidenziata dall'uso del corsivo.

Alps-GPSQuakeNet: *rete per il monitoraggio dell'arco Alpino*

www.alps-gps.units.it/

AGNES: *Servizio di Posizionamento della Svizzera*

www.swisstopo.admin.ch

AIUB: Astronomisches Institut Universität Bern

www.aiub.unibe.ch/content/index_eng.html

ARP: Antenna Reference Point

ASI: Agenzia Spaziale Italiana

www.asi.it

BIH: Bureau International de l'Heure, vedi BIPM e IERS

BIPM: Bureau International des Poids et Mesures

www.bipm.org/en/home/

BKG: Bundesamts für Kartographie und Geodäsie

www.bkg.bund.de/

C/A: Coarse Acquisition Code

CDDIS: Crustal Dynamics Data Information System

cddis.nasa.gov/

CIO: Conventional International Origin

CODE: Center for Orbit Determination in Europe, vedi AIUB

CTP: Conventional Terrestrial Pole

DGFI: Deutsches Geodätisches Forschungsinstitut

www.dgfi.badw.de/index.php?id=dgfi

DOD: US Department of Defense,

www.defenselink.mil/

DOP: Dilution of Precision

EOP: Earth Orientation parameters

EPN: European Permanent Network

www.epncb.oma.be/

ET: Ephemeris Time

ETRF:	European Terrestrial Reference Frame
ETRS89:	European Terrestrial Reference System 1989
EUREF:	European Reference Frame IAG Subcommission www.euref-iag.net/
FOC:	Full Operational Capability
FKP:	Flächenkorrekturparameter
GALILEO:	Costellazione GNSS europea www.esa.int/esaNA/galileo.html
GDOP:	Global DOP
GeoDAF:	Geodetic Data Archiving Facility geodaf.mt.asi.it/html_old/index.html
GeoTop:	<i>rete di stazioni permanenti in Italia</i> www.geotop.it/gps_GNSS_stazione_fissa.htm
GLONASS:	Global'naja Navigacionnaja Sputnikovaja Sistema www.glonass-ianc.rsa.ru/
GMST:	Greenwich Mean Sidereal Time
GNSS:	Global Navigation Satellite Systems
GPRS:	General Packet Radio Service
GPS:	Global Positioning Service www.gps.gov/
GPSLombardia:	<i>Servizio di posizionamento della Lombardia</i> www.gpslombardia.it
GRS:	Geodetic Reference System
GSM:	Global System for Mobile Communication
HDOP:	Horizontal DOP
IAG:	International Association of Geodesy www.iag-aig.org/
IAU:	International Astronomical Union www.iau.org/
IERS:	International Earth Rotation Service hpiers.obspm.fr/
IGM:	Istituto Geografico Militare Italiano www.igmi.org/
IGM95:	Rete italiana GPS di ordine zero
IGS:	International GPS Service, igscb.jpl.nasa.gov/
ILRS:	International SLR Service, ilrs.gsfc.nasa.gov/
INGV:	Istituto Nazionale di Geofisica e Vulcanologia, www.ingv.it/

IOC:	Initial Operational Capability
ITRF:	International Terrestrial Reference Frame, itrf.ensg.ign.fr/
ITRS:	International Terrestrial Reference System,
IVS:	International VLBI Service, ivscc.gsfc.nasa.gov/ivs.html
L2C:	L2 Civilian
Leica:	<i>rete di stazioni permanenti in Italia ItalPos</i> smartnet.leica-geosystems.it/SpiderWeb/frmIndex.aspx
LOD:	Length of Day
M:	Military
MAC:	Master Auxiliary Concept
NGA:	National Geospatial-Intelligence Agency, ex NIMA, www1.nga.mil/Pages/Default.aspx
NIMA:	National Imagery Mapping Agency, ora NGA
NTRIP:	Networked Transport of RTCM via Internet Protocol www.ntrip.org/
PCV:	Phase Centre Variations
PDOP:	Position DOP
PPP:	Precise Point Positioning
PPS:	Precise Positioning Service
PR:	Pseudo Random
P(Y):	Encrypted Precise Code
QZSS:	Quasi-Zenith Satellite System
RDN:	Rete Dinamica Nazionale 87.30.244.175/rdn/rdn.php
RINEX:	Receiver INdependent EXchange format
RING:	Rete Integrata Nazionale INGV ring.gm.ingv.it/
RTCM:	Radio Technical Commission for Maritime Services www.rtcn.org/
SAPOS:	Satellitenpositionierungsdienst der deutschen Landesvermessung www.sapos.de
SLR:	Satellite Laser Ranging, vedi ILRS
SPS:	Standard Point Service
SVN:	Space Vehicle Number
TAI:	International Atomic Time
TCG:	Geocentric Coordinate Time
TDOP:	Time DOP
TEC:	Total Electron Content

Trimble:	<i>rete di stazioni permanenti in Italia GPS-Assogeo</i> www.gps-assogeo.com/
UT1:	Universal Time 1
UTC:	Coordinated Universal Time
VDOP:	Vertical DOP
VLBI:	Very Long Baseline Interferometry, vedi IVS
VRS:	Virtual Reference Station
WGS84:	World Geodetic System 1984
ZTD:	Zenital Tropospheric Delay