



UCGE Reports  
Number 20166

Department of Geomatics Engineering

**Adjustment of Satellite-Based Ranging Observations  
for Precise Positioning and Deformation Monitoring**

(URL: <http://www.geomatics.ucalgary.ca/links/GradTheses.html>)

by

**Robert S. Radovanovic**

**October 2002**



THE UNIVERSITY OF CALGARY

Adjustment of Satellite-Based Ranging Observations for Precise Positioning and  
Deformation Monitoring

by

Robert Slobodan Radovanovic

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

OCTOBER, 2002

© Robert Slobodan Radovanovic 2002

## ABSTRACT

This dissertation presents a novel way of using measurements made by satellite-based ranging systems, such as GPS, for the purposes of precise positioning. Traditional methodologies and assumptions are discarded, resulting in a view of code and carrier phase measurements as simple spatial distances in a three-dimensional trilateration network of extremely large extent. Although the example navigation system used is GPS, the findings have general applicability to all Global Navigation Satellite Systems, such as GLONASS and GALILEO.

Traditional dependence on the single, fixed base station for GPS processing is eliminated by investigation of the underlying datum problem and an improved method based on the constraint of the network centre of mass is presented. This strategy has advantages in terms of minimally-biased ambiguity solution and positioning accuracy homogeneity across the network.

A novel approach to variance-covariance modelling for GPS error sources is developed, which separates the total error in terms of noise, multipath and tropospheric and ionospheric errors and models each individually. The models are developed using theoretical considerations coupled with empirical determination of key model parameters. Temporal and inter-frequency correlations are discussed as well.

The results of the research were used to create a GPS processing system named PADRES-GPS. It was shown that this system is capable of effectively detecting deformations of 10 mm at the 95% level on small-extent networks when data spans of only 10 minutes were used. The availability of the position solutions was shown to be much higher than that achievable by a commercial processing package due to the application of a novel “partial-fix” approach to ambiguity resolution documented herein. The PADRES-GPS system was also shown to improve positioning on larger networks extents via modelling of spatial correlations. Finally, a method of compressing dual-frequency data into a single optimal linear combination is described and its performance in kinematic network positioning is demonstrated.

## **ACKNOWLEDGEMENTS**

First and foremost, I must thank my family for providing me the unwavering support and encouragement necessary to pursue a goal to which there is no clear end. If I have built anything in this work, it was possible thanks to the solid foundation provided by my father and mother – their determination and sacrifice have left a lasting impression. The years of camaraderie and understanding provided by my brother and sister are greatly appreciated as an all important resource often used.

The roles of my supervisors, Drs. W.F. Teskey and N. El-Sheimy are gratefully noted. In particular, I would like to thank Dr. Teskey for providing me with the unrestrained freedom to explore many, many areas of research and Dr. El-Sheimy for gently curtailing it to ensure something was documented. I was truly lucky to have two supervisors whose doors were always open and who had faith, time and time again, when I promised that things would in fact get done.

In addition, my colleagues in the Department of Geomatics Engineering at The University of Calgary provided me with a solid resource in terms of both references, ideas and distractions. Many people helped me along the way, but I would like to specifically thank Georgia Fotopolous, Sandra Kennedy, and Ingrid Fjeld for help in the tedious tasks of proofreading and formatting this dissertation.

Finally, the financial support provided by several sources is gratefully acknowledged for allowing me freedom in my methods, timelines and topics of research. The continual support of Challenger Geomatics was particularly significant. In addition, the Killiam Memorial Foundation, the National Sciences and Engineering Research Council of Canada, the Alberta Heritage Scholarships Program, the L.R. Newby Foundation and the University of Calgary Teaching and Research Assistantships Program all ensured I was able to focus my attention on research and publication.

# TABLE OF CONTENTS

<b>APPROVAL PAGE</b> .....	ii	
<b>ABSTRACT</b> .....	iii	
<b>ACKNOWLEDGEMENTS</b> .....	iv	
<b>TABLE OF CONTENTS</b> .....	v	
<b>LIST OF TABLES</b> .....	x	
<b>LIST OF FIGURES</b> .....	xi	
<b>LIST OF ACRONYMS</b> .....	xiv	
<b>LIST OF COMMONLY USED SYMBOLS</b> .....	xv	
 <b>CHAPTER</b>		
<b>1.0 INTRODUCTION</b> .....	1	
<b>1.1 Traditional Application of GPS to Geodesy</b> .....	2	
<b>1.2 Research Objectives</b> .....	4	
<b>1.3 Outline of this Work</b> .....	5	
 <b>2.0 BASICS OF SPACE-BASED RADIO RANGING</b> .....		8
<b>2.1 Distance By Radio Ranging</b> .....	8	
<b>2.2 Satellite Time Frames</b> .....	10	
<b>2.3 Receiver Time Frames</b> .....	13	
<b>2.4 Satellite Orbits</b> .....	15	
<b>2.5 Geometrical Model of Earth-Space Ranging</b> .....	18	
2.5.1 - Modification for Transmission Time and Satellite Motion .....	18	
2.5.2 - Modification for a Rotating Earth .....	19	
2.5.3 - Modification for Carrier Phase Measurement .....	21	

<b>3.0 GENERAL SOLUTION TO THE SPACE-BASED TRILATERATION</b>	
<b>PROBLEM</b> .....	23
<b>3.1 Unknowns and Measurements in a Space-Based             Trilateration Network</b> .....	23
<b>3.2 Formulation of the Least-Squares Solution</b> .....	25
3.2.1 - Modification for Incorrect Apriori Stochastic Assumptions .....	28
<b>3.3 Form of the Linearized Quantities</b> .....	29
3.3.1 - Derivative with Respect to the Unknown Positions ....	31
3.3.2 - Derivatives with Respect to the Unknown Clock Offsets .....	31
3.3.3 - Derivatives with Respect to the Unknown Ambiguities .	33
3.3.4 - Derivative with Respect to Other Model Parameters .....	33
<b>3.4 Differential GPS</b> .....	35
<b>3.5 Convergence and Linearization Issues</b> .....	37
<b>4.0 DATUM ISSUES IN SATELLITE-BASED NAVIGATION</b>	
<b>SYSTEM</b> .....	41
<b>4.1. Overview of Traditional Datum Definition Solutions</b> ....	41
4.1.1 - Problems with Moving Control Points .....	43
<b>4.2 Effects of Datum Overconstraint on Positioning Accuracies</b> ....	44
<b>4.3 Datum Problem In Differential GPS</b> .....	48
4.3.1 - Ill-condition of the DGPS Problem .....	49
4.3.2 - Datum Visibility in DGPS .....	52
4.3.3 - Regularization of the DGPS Problem .....	53
4.3.4 - Deformation Monitoring with DGPS .....	59

<b>5.0 AMBIGUITY RESOLUTION FOR PRECISE POSITIONING</b> ....	63
<b>5.1 Float Ambiguity Resolution</b> .....	63
5.1.1 - Regularization and Ambiguity Resolution .....	64
5.1.2 - Base Satellite Changeovers .....	71
<b>5.2 Integer Ambiguity Resolution</b> .....	73
5.2.1 - Principle of Integer Ambiguity Resolution .....	74
5.2.2 - Ambiguity Success Rate .....	76
5.2.3 - Ambiguity Decorrelation .....	78
<b>5.3 Summary of the Ambiguity Resolution Process and         Precise Position Determination</b> .....	79
<b>6.0 ANALYSIS OF ERROR SOURCES IN GLOBAL NAVIGATION SATELLITE SYSTEMS AND STOCHASTIC MODELLING</b> ....	84
<b>6.1 Covariance of GNSS Error Sources</b> .....	85
6.1.1 - Effect of Double-Differencing on the Variance-Covariance Matrix .....	88
6.1.2 - Developing Variance-Covariance Models for GNSS ....	91
<b>6.2 Noise Errors</b> .....	94
6.2.1 - Signal Power and Theoretical Noise Variance Modelling .....	95
6.2.2 - Empirical Assessment of the Noise Variance Behaviour	99
<b>6.3 Multipath Error</b> .....	106
6.3.1 - Theoretical Multipath Variance Modelling .....	107
6.3.2 - Empirical Assessment of the Multipath Variance Behaviour .....	109
<b>6.4 Tropospheric Error</b> .....	111
6.4.1 - Theoretical Tropospheric Variance Modelling ....	113
6.4.2 - Empirical Assessment of the Tropospheric Variance Behaviour .....	117

<b>6.5 Ionospheric Error</b> .....	121
6.5.1 - Klobuchar Ionospheric Model .....	123
6.5.2 - Theoretical Modelling of the Ionospheric Error ...	125
6.5.3 - Empirical Assessment of Ionospheric Variance Behaviour .....	127
<b>6.6 Treatment of Temporal Correlations</b> .....	130
<b>6.7 A Note on Inter-Frequency Correlation</b> .....	134
<b>6.8 Practical Determination of Variance-Covariance Models</b> ...	136
<b>7.0 APPLICATION OF GLOBAL NAVIGATION SATELLITE SYSTEMS TO NETWORK POSITIONING</b> .....	139
<b>7.1 Application to Short Range Deformation Monitoring</b> ...	140
7.1.1 - Principles of a GPS Based Deformation Monitoring System .....	140
7.1.2 - Test of the PADRES-GPS System .....	146
7.1.3 - Deformation Detection Performance .....	148
<b>7.2 Application to Medium Range Static Positioning and Results</b>	156
<b>7.3 Application to Kinematic Network Positioning and     Optimal Linear Combination Determination</b> .....	162
7.3.1 - Linear Phase Combinations .....	163
7.3.2 - “Optimal” Linear Phase Combinations .....	165
7.3.3 - Determination of Optimal Linear Phase Combinations From Reference Network Data .....	168
7.3.4 - Positioning with an Optimal Linear Phase Combination	169



<b>8.0 CONCLUSIONS AND FUTURE WORK</b> .....	176
<b>8.1 Key Findings by Chapter</b> .....	177
<b>8.2 Thesis Contributions</b> .....	181
<b>8.3 Recommendations for Future Work</b> .....	182
<b>9.0 REFERENCES</b> .....	185

## LIST OF TABLES

<b>Table 6.1.</b>	Estimated Tropospheric Model Parameters. ....	119
<b>Table 7.1.</b>	Zenith Multipath/Noise Standard Deviations for Stations Involved in the Short Range Test. ....	147
<b>Table 7.2.</b>	Observed and Predicted Positioning 1- $\sigma$ Accuracies of Stations During Short-Range Monitoring Test. ....	151
<b>Table 7.3.</b>	Deformation Report Produced During Short-Range Test and Actual Movements. ....	154
<b>Table 7.4.</b>	Summary of Deformation Monitoring Performance for Stable and Deformed Stations. ....	155
<b>Table 7.5.</b>	Position Standard Deviations of Various Processing Modes. ....	171
<b>Table 7.6.</b>	Position Standard Deviations of Various Processing Modes with an Improper Stochastic Model Assumed. ....	173

## LIST OF FIGURES

<b>Figure 2.1.</b>	Satellite Clock Offset. ....	12
<b>Figure 2.2.</b>	Broadcast Satellite Clock Offset Error. ....	12
<b>Figure 2.3.</b>	Receiver Clock Offsets for Two GPS Receivers. ....	14
<b>Figure 2.4.</b>	Receiver Clock Drift for Two GPS Receivers. ....	14
<b>Figure 2.5.</b>	Comparison of Precise and Broadcast Ephemerides. ...	17
<b>Figure 2.6.</b>	Satellite Velocities Relative to WGS-84. ....	17
<b>Figure 2.7.</b>	Rotating Reference Frame Problem. ....	20
<b>Figure 3.1.</b>	Form of the Jacobian of Satellite and Receiver Clock Offsets.	32
<b>Figure 3.2.</b>	Effect of Convergence Threshold on Adjustment Accuracy and Iteration Pathway. ....	39
<b>Figure 4.1.</b>	Effect of Satellite Coordinate Error on Receiver Position Accuracy. ....	45
<b>Figure 4.2.</b>	Correlation Coefficient of Network Position Solutions. ....	47
<b>Figure 4.3.</b>	Effect of Satellite Coordinate Error on Relative Positioning Accuracy. ....	48
<b>Figure 4.4.</b>	Effect of Network Size on Condition Number. ....	52
<b>Figure 4.5.</b>	Effect of Satellite Coordinate Error on Receiver Position Accuracy. ....	53
<b>Figure 4.6.</b>	Contributions to the Regularization Error due to Noise and Bias for Various $1/\alpha$ . ....	57
<b>Figure 4.7.</b>	Minimum Variances Achievable Under Various Assumptions for Initial Estimate Accuracies. ....	57
<b>Figure 4.8.</b>	Contributions to the Regularization Error due to Noise and Bias at Various Network Extents. ....	58
<b>Figure 4.9.</b>	Contributions to the Detection Regularization Error due to Noise and Bias for Various $1/\alpha$ . ....	60
<b>Figure 4.10.</b>	Deformation Detection Accuracy. ....	61

<b>Figure 5.1.</b>	Average Position Accuracy for Various Values of $1/\alpha$ and Initial Estimate Accuracies. ....	67
<b>Figure 5.2.</b>	Average Ambiguity Accuracy for Various Values of $1/\alpha$ and Initial Estimate Accuracies. ....	67
<b>Figure 5.3.</b>	Average Position Accuracy for Various Values of $1/\alpha$ and Initial Estimate Accuracies using Centre of Mass Regularization. ....	70
<b>Figure 5.4.</b>	Average Ambiguity Accuracy for Various Values of $1/\alpha$ and Initial Estimate Accuracies using Centre of Mass Regularization. ....	70
<b>Figure 6.1.</b>	Atmospheric Attenuation Versus Elevation Angle. ....	97
<b>Figure 6.2.</b>	Minimum Received L1 Signal Power. ....	98
<b>Figure 6.3.</b>	L1 Antenna Gain Patterns for Three Varieties of Antennas. ....	99
<b>Figure 6.4.</b>	Comparison of double difference standard deviations for L1 signals. ....	100
<b>Figure 6.5.</b>	Comparison of double difference standard deviations for L2 signals. ....	100
<b>Figure 6.6.</b>	Observed versus Theoretic L1 Carrier Noise Standard Deviations for NovAtel Antennas 501, 502 and 503 and OEM 3 receivers. ...	103
<b>Figure 6.7.</b>	Comparison of Various Noise Mapping Functions for Two Antennas. ....	105
<b>Figure 6.8.</b>	Multipath Effect and Resulting Multipath Error. ....	106
<b>Figure 6.9.</b>	Comparison of Observed and Theoretical L1 Multipath Variance for Two Receivers. ....	111
<b>Figure 6.10.</b>	Network Configuration for Tropospheric Study. ....	117
<b>Figure 6.11.</b>	Double-Differenced Residual Error Standard Deviations for Various Baseline Lengths. ....	119
<b>Figure 6.12.</b>	Comparison of Observed and Theoretical Zenith Tropospheric Variances for Various Separations. ....	121
<b>Figure 6.13.</b>	Relationship between TEC and VTEC. ....	123
<b>Figure 6.14.</b>	Vertical Zenith Delay Estimates from Klobuchar Ionospheric Model ....	124
<b>Figure 6.15.</b>	Total and Residual Ionospheric Error as a Function of Local Time. ....	129
<b>Figure 6.16.</b>	Comparison of Observed and Theoretical Zenith Ionospheric Variances for Various Separations in Winter and Summer. ....	129

<b>Figure 7.1.</b>	Relationship between Update Interval, Observation Span and Sampling Rate. ....	141
<b>Figure 7.2.</b>	Typical Overhead Skyplot Showing Effect of Polar Caps. ....	146
<b>Figure 7.3.</b>	Position Estimates of Station A in Short Range Network Test. ...	149
<b>Figure 7.4.</b>	Position Estimates of Station B in Short Range Network Test. ...	149
<b>Figure 7.5.</b>	Position Estimates of Station C in Short Range Network Test. ...	150
<b>Figure 7.6.</b>	Position Estimates of Station E in Short Range Network Test. ...	150
<b>Figure 7.7.</b>	Observed and Predicted Position Accuracies for Medium Extent Network using Centre-of-Mass Regularization. ....	158
<b>Figure 7.8.</b>	Observed and Predicted Position Accuracies for Medium Extent Network using Fixing of Station BRAN. ....	159
<b>Figure 7.9.</b>	Possible Baseline Configurations. ....	160
<b>Figure 7.10.</b>	Observed and Predicted Position Accuracies for Medium Extent Network Neglecting Spatial Correlations. ....	161
<b>Figure 7.11.</b>	Selection of an Optimal Linear Combination. ....	167
<b>Figure 7.12.</b>	Candidate combination variances in cycles <sup>2</sup> and m <sup>2</sup> . ....	169
<b>Figure 7.13.</b>	Height Variation for Various Processing Modes. ....	171
<b>Figure 7.14.</b>	Height Variation for Various Processing Modes using an Incorrect Stochastic Model. ....	173

## LIST OF ACRONYMS

ASR	Ambiguity Success Rate
ECEF	Earth Centred, Earth Fixed
DGPS	Differential GPS
GALILEO	European Space Agency GNSS
GLONASS	Russian Space Program GNSS
GPS	Global Positioning System
GNSS	Global Navigation Satellite System
ICD	Interface Control Document
IGS	International GPS Service
LAMBDA	Least Squares Ambiguity Decorrelation Adjustment
L1, L2	Referring to the 1575.42 MHz and 1227.60 MHz carrier
OLPC	Optimal Linear Phase Combination
PADRES-GPS	Position and Deformation Reporting System using GPS
SA	Selective Availability
SCIGN	Southern California Integrated GPS Network
SV	Space Vehicle
TEC	Total Electron Content
TRANSIT	United States Navy Navigation Satellite System
WGS-84	World Geodetic System
VTEC	Vertical Electron Content

## LIST OF COMMONLY USED SYMBOLS

The number of symbols used in this dissertation are of a great number. For convenience, the root forms of the more commonly used symbols are provided here for reference. At all times, ambiguity is eliminated by in-text references to the particular symbols used.

$\mathbf{a}$	... regularization weight
$\mathbf{a}, \mathbf{A}$	... azimuth
$a, b$	... OLPC parameters
$\mathbf{A}, \mathbf{B}$	... Jacobian with respect to the unknowns and observations, respectively
$B$	... attenuation factor
$c$	... speed of light in vacuo, 299 792 458 m/s
$\mathbf{C}_l$	... variance-covariance matrix of the observations
$\mathbf{C}_x$	... variance-covariance matrix of the initial estimates
$\mathbf{C}_{x^*}, \mathbf{C}_{x_c}$	... variance-covariance matrix of the estimated parameters before and after application of constraints
$C/No$	... carrier to noise ratio in decibels
$\tilde{\mathbf{N}}\mathbf{D}$	... double-differencing operator
$\mathbf{d}$	... least-squares corrections to the initial estimates of the parameters
$\mathbf{d}^*, \mathbf{d}_c$	... solution for parameter corrections before and after application of constraints
$D$	... tropospheric/ionospheric correlation distance
$\mathbf{D}$	... diagonal matrix
$\mathbf{d}$	... displacement vector
$\mathbf{D}\mathbf{d}$	... error in displacement vector
$d$	... distance
$d_e$	... electromagnetic distance

$d_g$	... geometric distance
$e$	... noise
$e$	... elevation angle
$e$	... error vector
$f$	... temporal correlation factor
$f_{L1}, f_{L2}$	... frequency of the carrier phase at the GPS frequencies, 1575.42 MHz and 1227.60 MHz, respectively
$G$	... antenna gain
$\mathbf{G}$	... constraints matrix
$I$	... total ionospheric error
$dI$	... residual ionospheric delay
$\mathbf{I}$	... identity matrix
$l$	... wavelength of carrier
$\mathbf{l}$	... vector of observations
$m(\cdot)$	... mapping function
$n, e, h$	... coordinates in a local-level system (northing, easting, up)
$N$	... integer ambiguity in cycles
$\mathbf{N}$	... normal matrix
$\mathbf{N}^*, \mathbf{N}_{\text{reg}}$	... regularized normal matrix
$w$	... rotation rate of the Earth, $7\,292\,115.1467 \cdot 10^{-11}$ rad / s
$W$	... tropospheric correlation angle
$f$	... measured carrier phase
$p$	... measured pseudorange
$P$	... power in decibels
$\mathbf{P}$	... regularization matrix
$\mathbf{r}$	... residual vector
$\mathbf{r}_s, \mathbf{r}_r$	... position vectors of the satellite and receiver
$R_E$	... average radius of the Earth, approximately 6371 km
$\mathbf{s}$	... standard deviation



<b>S</b>	... S-transform matrix
<b>q</b>	... line-of-sight separation angle
<b>t</b>	... transit time
<i>t</i>	... time
<i>t'</i>	... corrected time
<i>t<sub>r</sub></i>	... reception time
<i>t<sub>s</sub></i>	... transmission time
<i>t<sub>oc</sub></i>	... reference time for satellite clock model
<b>Dt<sub>r</sub></b>	... receiver clock offset
<b>Dt<sub>s</sub></b>	... satellite clock offset
<b>dt<sub>s</sub></b>	... satellite clock offset error
<b>T</b>	... total tropospheric error
<b>dT</b>	... residual tropospheric delay
<b>v</b>	... normal vector
<b>w</b>	... misclosure vector
<b>w<sub>c</sub></b>	... constraints misclosure
<b>x</b>	... vector of unknowns
<b>x<sub>0</sub></b>	... initial estimates of the unknowns
<b>x<sub>LS</sub></b>	... least-squares estimate of the unknowns
<b>x<sub>T</sub></b>	... true values of the unknowns
<b>Dx</b>	... error in estimate of the unknowns
<i>x,y,z</i>	... coordinates in a Cartesian system
<i>z</i>	... zenith angle
<b>Z</b>	... decorrelating matrix

## **1.0 - INTRODUCTION**

The Global Positioning System (GPS) was designed by the United States Department of Defence for the dual purposes of “creating a system to navigate cheaply” and “dropping five bombs into the same hole” (Parkinson, 1996).

Since then, GPS has been thoroughly exploited by the civilian sector for uses of far greater variety and finesse than intended by the military planners at the Joint Program Office. As a navigation system, GPS is in use world-wide, from aiding the navigation of oil tankers and aeroplanes, to informing hikers of where they are. Indeed, much current research revolves around the refinement of GPS for navigation purposes, in particular in terms of its integration with other navigation sensors. Furthermore, the success of Global Navigation Satellite Systems (GNSS) has led to the development of parallel systems, such as the operating Russian GLONASS system, and the planned European GALILEO project.

Early on, the potential of applying these satellite-based navigation systems to surveying applications was recognized. For example, in the 1960's, the TRANSIT satellite system was successfully applied to static surveying by using several days worth of data and post-mission precise orbital information to achieve accuracies of less than one metre (Hofmann-Wellenhoff et al, 1994). Despite the design of the GPS system as a military navigation tool, sub-decimetre surveying accuracies quickly became achievable primarily through the development of differential techniques (Counselman et al, 1972), the use of carrier phase measurements (Collins, 1982), and the development of robust techniques to obtain dual-frequency measurements for ionospheric modelling (Ashjaee and Lorenz, 1992). Dropping receiver costs in the last two decades have made high-quality GPS receivers reasonable available to the geodetic community and hence encouraged their use in many areas.

The appeal of applying satellite-based navigation systems to static applications largely stems from the autonomous operation of the system and the lack of a requirement for surveyed sites to be intervisible. Both of these issues have been serious limitations to traditional surveying techniques applied to precise positioning and deformation monitoring. For example, the interval between position updates in networks of even modest size is often limited to months, if not years, due to the costs associated with the requirements for skilled personnel to make angular measurements in a theodolite-based network. Recent advances in terms of using robotic tacheometers have improved the rate at which observations can be made, but the requirements for the stations to be intervisible still constrains network design and extent. On the other hand, a network of GPS receivers can collect unlimited data at a very high data rate, in all weather. The network design is considerably loosened by eliminating the need for lines-of-sight, and as a result the design can focus on other considerations, such as ensuring that geo-technically significant points are observed. This also implies that inter-station distances can be considerable extended. Finally, once the initial costs of the receivers are absorbed, further operating costs are minimal, compared to the costs of trained survey personnel.

### ***1.1 - Traditional Application of GPS to Geodesy***

Given the benefits of satellite-based systems described above, it is perhaps not surprising that GPS has been applied to a wide variety of geodetic and geo-technical problems, from national scale crustal deformation monitoring in Japan (Tsuji et al, 1996), to the study of building sway under wind loading (Guo and Ge, 1997). Indeed, every major geodetic conference is certain to exhibit a large number of papers dedicated to the application of GPS.

However, the author believes that the current state of understanding regarding the geodetic applications of satellite-based navigation systems retains artefacts of the early research done on these systems, and their navigation heritage. This is particularly true in areas of variance-covariance modelling and network adjustment, where many simplifications necessary in early research are still retained in typical processing methodologies. Of

course, these areas are critical for the rigorous application of GPS (and indeed any observation system) to deformation monitoring and precise positioning.

During the early period of GPS, the high cost of receivers limited most institutions to a pair of receivers, which forced research to be performed using a *baseline* methodology, where one receiver was considered the base station and the other a remote. A network of points could then be surveyed by moving the remote receiver to various stations in the network and thus generating a set of radiating vectors. This paralleled developments in navigation-based research, which relied on the use of a base station with known coordinates to provide differential corrections to a remote station.

Interestingly, although it was obviously realized that the GPS errors are spatially correlated (and hence the returns provided by differencing observations), these spatial correlations were neglected in the diagonal variance-covariance matrix typically used in the processing of the observations. This can be explained by the fact that early knowledge of the actual spatial correlations were poor. Also, limitations of early computing power implied that use of a diagonal variance-covariance matrix had significant advantages in terms of formation and solution of the normal equations. Furthermore, since the GPS observation results were significantly better than results achievable using terrestrial methods, errors stemming from non-rigorous variance modelling were undetectable. Recently, studies have begun to apply more realistic covariance models and these are discussed in Chapter 6. However, most commercial processors have yet to implement their results.

The pioneering baseline processing methods also left their mark on the prevailing network adjustment procedures of today. In most network adjustments involving GPS, the baselines of the network are calculated separately, along with the variance-covariance matrices of the individual baselines (which themselves do not take into account a rigorous statistical model of the component error sources). These vectors are then treated as pseudo-observations in a subsequent adjustment for the network, usually constraining the central base station at its apriori coordinates. Obviously this is somewhat of a boot strapping procedure for determining the final coordinates of the network receivers and the statistics

of the adjustment are expected to become seriously distorted (and usually very over-optimistic). In practice, this effect is masked by long observation times (several hours for most geodetic applications) and the higher precision of GPS observations when compared to their terrestrial counterparts, as well as a lack of suitable “comparison” systems.

### ***1.2 - Research Objectives***

The goal of this thesis was to determine how to treat GPS observations in the same manner as any conventional geodetic measurement, without reliance on the assumptions inherent in today’s conventional processing. This required the re-evaluation of the following areas :

- Adjustment of Observations under Constraints and the GPS Mathematical Model
- Datum Definition
- Ambiguity Resolution
- Variance-covariance Modelling

The result of this research is a GPS processor that is capable of using data observed by a network of receivers, and performing a truly rigorous least-squares adjustment of the data *at the observation level*. In this way, the GPS observations become nothing more than spatial distances, much like those measured in conventional trilateration networks, albeit over much larger distances. The same principles can then be simply applied to other GNSS, such as GLONASS and GALILEO.

The appeal of such research is that, by studying GPS observations as regular geodetic observations, existing geodetic theories on network adjustment and deformation monitoring can be simply applied. For example, proper variance-covariance modelling, studied in Chapter 6, allows a greater confidence in the parameter statistics reported, a significant departure from the traditional “multiply the standard deviations by 10” rule of thumb commonly used in industry. This in turn is crucial for the identification of significant deformations, a topic studied in Chapter 7. Similarly, datum definition is an area long studied as applied to terrestrial networks, but has been largely ignored in satellite

based networks due to the reliance on the known base station. Although some studies have used the concept of inner constraints for processing a network of individually processed GPS baselines, this thesis treats proper datum definition as the starting point of the subsequent observation adjustment, as discussed in Chapter 4. Even the basic process of adjusting GPS observations is investigated, and a rigorous method of applying integer ambiguity constraints and correctly propagating the statistics of the estimated parameters is demonstrated in Chapters 3 and 5.

In total, the result of this work is a system of adjusting GPS observations for network positioning and deformation analysis where the traditional assumptions made during the transformation of the original observations to the final reported positions have been minimized. Not only does this create a system that is more transparent, where the effects of individual adjustment steps can be directly analyzed, but the resulting estimates are improved, both in terms of their actual and reported accuracies. As shall be seen, the lessons learned from this dissertation can be applied to a wide variety of applications, ranging from short-range deformation monitoring, to kinematic network positioning.

### ***1.3 - Outline of this Work***

This dissertation begins with a basic investigation into the operation of the satellite-based navigation systems and the nature of the observations they provide. From this starting point, aspects of the rigorous adjustment of said observations are individually studied, resulting in a complete processing methodology. Finally, the success of applying this processing methodology to various real-world applications is demonstrated.

Specifically, this dissertation is divided up into the following chapters :

Chapter 2 provides an introduction into satellite-based navigation systems and their operation, using GPS as an operational example. The peculiarities of satellite-based time-of-flight measurement are discussed, in particular in terms of the concept of *moving reference points*.

Chapter 3 reviews the principles of least-squares adjustment, but reviews the theories pertinent to adjustment with additional constraints, which plays a pivotal role in ambiguity resolution. The rigorous propagation of parameter statistics is emphasised at all stages. Finally, the details of the linearization of the satellite-based positioning problem are shown, and proper iteration termination is discussed

Chapter 4 delves into the problems of datum definition for satellite-based ranging networks. It is shown that the space-based trilateration problem is identical to that of a terrestrial one, with the moving satellites acting as control points. However, it is also shown that the datum defined by these satellites is *poorly visible*, requiring regularization of the positioning problem.

Chapter 5 continues the discussion of datum regularization, but focuses on its effects on ambiguity resolution. Also, details of the ambiguity resolution process are discussed, from the proper treatment of base satellite changeovers to integer ambiguity resolution using the LAMBDA method. The integer ambiguity problem is shown to be a problem of least-squares with constraints and a novel *partial-fix* method is presented that allows more stable positioning performance over time.

Chapter 6 introduces new concepts in variance-covariance modelling for observations made by GNSS. Individual error sources are studied, including noise, multipath, tropospheric and ionospheric effects. In particular, this section features theoretical models developed to describe the variance properties of these errors and their propagation into the double-differences formed during processing. Methods of establishing these models using collected data are also discussed, as are the inter-frequency and temporal correlations of GNSS errors.

Chapter 7 applies the theories developed in the preceding chapters to three distinct application areas. Firstly, a robust deformation monitoring system applicable to short ranges is developed, and its performance using real data is assessed. Secondly, the limits of

precise positioning on networks with extents of 10 to 150 kilometres is studied. Finally, the variance-covariance studies described are used to develop a unique method of optimal linear combination determination for efficient kinematic network based positioning.

Chapter 8 contains a summary of the discoveries made on a chapter by chapter basis, discusses their possible applications in various fields and presents recommendations for future work.



## 2.0 - BASICS OF SPACE-BASED RADIO RANGING

All space-based navigation systems used today rely on the transmission of ranging signals from space-based platforms (satellites) to receivers on or above the Earth. This results in a trilateration problem where the satellites play the role of control points and the receivers are unknown stations – a problem analogous to a terrestrial network consisting solely of electromagnetic distance measurement observations.

However, the majority of space-based systems in use are one-way systems, whereby the receiver does not transmit back to the satellite. While this allows an unlimited number of users to access the navigation system, the resulting problem of relating satellite and receiver time systems must be dealt with. As well, the motion of the satellites and the Earth itself results in issues not encountered in terrestrial trilateration networks.

This section investigates the mathematical formulation of the three-dimensional space-based trilateration problem. The results are generally applicable to all space-based ranging systems, such as GLONASS and GALILEO, but the discussion focuses on the GPS as an example.

### 2.1 -Distance By Radio Ranging

Assume a situation whereby satellites in orbit send signals to a ground-based receiver. If a given satellite emits a signal at time  $t_s$  and it is received by the receiver at time  $t_r$ , the distance,  $d$ , the signal has traveled is given by

$$d = c \cdot (t_r - t_s) \tag{2.1}$$

where  $c$  is the speed of light in vacuo.

Eq. (2.1) neglects the effects of the atmosphere on the signal and assumes that the transmission and reception time refer to an absolute time frame, see Section 2.2. In practice, the signal transmit time is implicitly transmitted in the coding of the signal and refers to the satellite's unique time frame. Similarly, the receive time is based on the receiver's local oscillator. Let the offset of the receiver and satellite time frames from some absolute frame be denoted as  $Dt_r$  and  $Dt_s$ . Note that the clock offsets are not themselves constant over time in general. The electromagnetic distance,  $d_e$ , between the satellite at transmission time and the receiver at receive time can then be written as :

$$d_e = c \cdot (t_r - t_s) + c \cdot (\Delta t_s - \Delta t_r) \quad (2.2)$$

The bracketed term farthest to the right acts as a correction to the actual measurement, which is simply the measured travel time multiplied by the speed of light. It is important to realize that  $t_r$  and  $t_s$  represent the *measured* transmit and receive times.

The *geometric* distance can be derived by taking into account the effects of the signal's passage through the electrically-charged ionosphere and the neutral atmosphere, which not only affects the velocity of the wave, but also causes it to follow a curved path due to ray-bending (Saastamoinen, 1973). These effects will be more closely studied in Chapter 6. At the moment these quantities will be simply denoted as  $I$  and  $T$  for ionospheric and tropospheric effects, respectively. Incorporating these effects results in the following expression :

$$d_g = c \cdot (t_r - t_s) + c \cdot (\Delta t_s - \Delta t_r) - T - I \quad (2.3)$$

The sign of  $T$  and  $I$  imply that they are delays, making the electromagnetic distance longer than the true geometric distance, as expected.

Finally, since measurements cannot be made with perfectly, the time of arrival of the signal has a certain error associated with it. This can be incorporated into Eq. (2.3) to yield:

$$c \cdot (t_r - t_s) = p = d_g - c \cdot (\Delta t_s - \Delta t_r) + T + I + \mathbf{e} \quad (2.4)$$

where  $p$  is termed the *measured pseudorange*. No assumptions regarding the error statistics of the  $\mathbf{e}$  term have been made, although these will be thoroughly addressed in Chapter 6.

## 2.2 -Satellite Time Frames

As discussed above, the transmission time of a ranging signal is usually embedded in the transmitted signal itself. However, this time refers only to the satellite's onboard clock since it is the clock which drives the frequency synthesizer and code generator required to generate the signal itself (ICD-GPS-200C, 1993). This means that signals simultaneously received from several satellites not only have different transmission times, but that the transmission times reported actually belong to different time frames. In addition, since clocks slowly drift over time, the difference between two satellites' individual time frames is not constant over time. This effect must be taken into account when deriving geometric distances via Eq. (2.4).

In the case of GPS, an "absolute" time frame is derived by averaging the master atomic clocks at five monitoring ground stations (Francisco,1996). The deviations of individual satellite clocks from this "ensemble clock" are continually monitored by the monitor stations. Finally, these deviations are used to create a prediction model for each satellite's clock error using the following two degree polynomial (ICD-GPS-200C, 1993):

$$\Delta t_s = a_o + a_1(t - t_{oc}) + a_2(t - t_{oc})^2 \quad (2.5)$$

where  $\Delta t_s$  is the satellite clock offset,  $a_o, a_1$ , and  $a_2$  are broadcast polynomial coefficients,  $t$  is the time in the ensemble frame and  $t_{oc}$  is the epoch to which the coefficients refer to.

Rigorously, the true time is required to properly determine the correction. Fortunately, since the  $a_1$  and  $a_2$  terms are small, the apparent satellite time can be substituted into Eq. (2.5) with negligible effects.

The satellite transmit time must also be corrected for relativistic effects due to the high altitude of the satellites and their velocity, this results in a corrected transmission time of (ICD-GPS-200C, 1993) :

$$t'_s = t_s - \Delta t_s + \frac{2\mathbf{m}^{1/2}}{c^2} e \cdot A^{1/2} \cdot \sin E_k \quad (2.6)$$

where  $t'_s$  is the corrected satellite time,  $\mathbf{m}$  is the gravitational constant,  $c$  is speed of light in vacuo, and  $A$ ,  $e$ ,  $E_k$  are the semi-major axis, eccentricity and mean anomaly of satellite orbit.

In the sequel, the last two terms of Eq. (2.6) will be lumped together into a single clock correction term. The relativistic correction alone can cause delays of up to 70 ns.

Using Eq. (2.6) one can determine the transmission time of a signal in the ensemble frame. Figure 2.1 shows the magnitude and variation of the correction term for three satellites over a six hour period. For clarity, the plots have been shifted by the following amounts : SV 4 – 650  $\mu$ s, SV 16 – 62  $\mu$ s, SV 30 – (-)3  $\mu$ s. An almost linear trend is apparent, although the direction and magnitude of the drift is satellite dependent. By differencing subsequent epochs, one determines that the size of the drift is typically below 10 ps / s. The data in Figure 2.1 was derived from observations collected by monitoring stations in the International GPS Service (IGS), and is not available in real-time. For real-time positioning, the user must use the transmitted correction model, which is of lower accuracy since it *predicts* the clock correction. Figure 2.2 shows the difference between the predicted and actual corrections for the same time period. The 10 ns bias in the plot is likely due to a bias in the reference clock used to generate the precise ephemerides. Furthermore, the jump in the error plot for SV 30 is due to a difference in the clock correction model uploaded to the

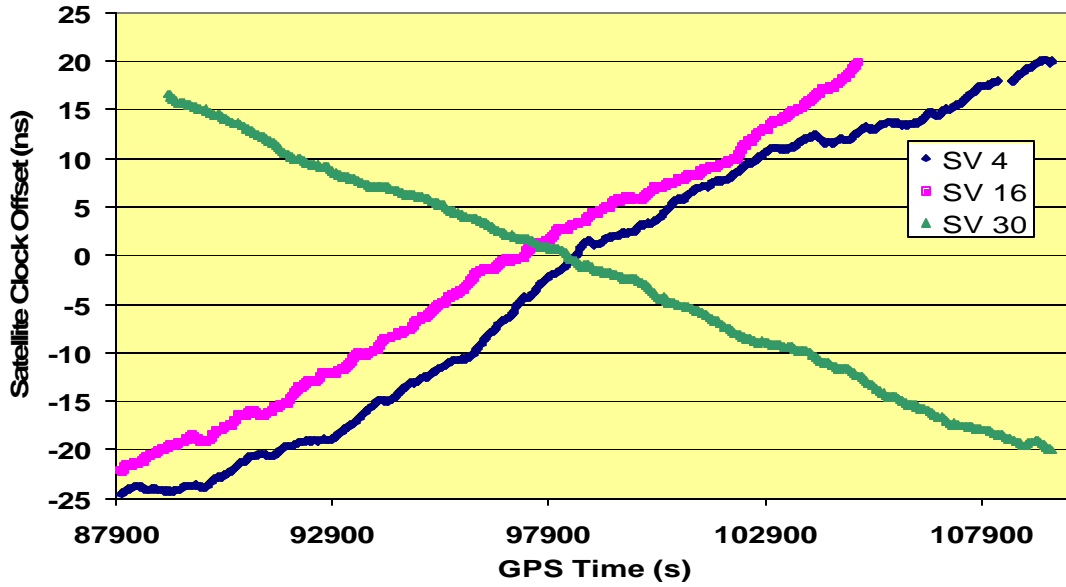


Figure 2.1. Satellite Clock Offset. Offsets have been shifted for clarity by :  
SV 4 – 650  $\mu$ s, SV 16 – 62  $\mu$ s, SV 30 – (-)3  $\mu$ s .

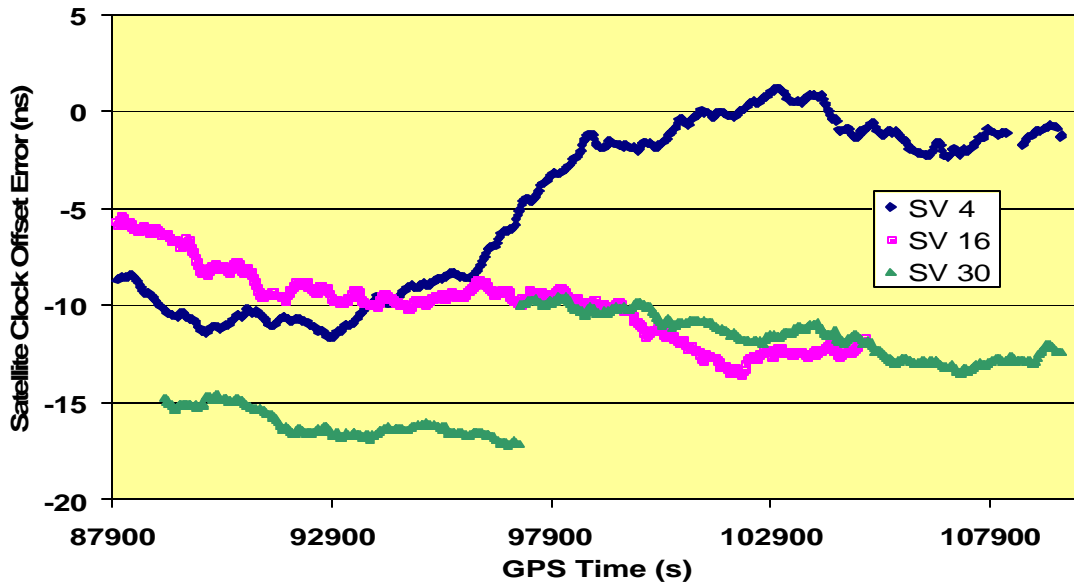


Figure 2.2. Broadcast Satellite Clock Offset Error (compared to precise ephemerides).

satellite. These changes typically occur every two hours as the ephemerides are refreshed. Note that the data examined is free of the effects of Selective Availability (SA), a satellite clock error source historically injected into the timing signal by the U.S. Department of Defence. SA was turned “off” on May 1, 2000 (U.S.A. Office of Science and Technology Policy, May 2000).

Using the clock corrections, one can modify Eq. (2.4) to read :

$$c \cdot (t_r - t_s + \Delta t_s) = p + \Delta t_s = d_g - c \cdot (dt_s - \Delta t_r) + T + I + e \quad (2.7)$$

where  $dt_s$  indicates the error in calculating the satellite clock offset. If one neglects the biases evident in Figure 2.2, the standard deviations of the predicted clock offset errors are at the 5 ns level, or 150 centimetres.

### ***2.3 -Receiver Time Frames***

Just as the transmitted signal transmission time refers to the satellite’s unique time frame, the measured receive time refers to the unique time frame of the receiver. However, whereas the clocks of the satellites are high-quality atomic timepieces, receivers typically employ low cost quartz oscillators. As a result, receiver clock offsets can be large and have higher drift rates than those of the satellites.

Figure 2.3 shows the receiver clock offsets (in distance units) for two models of receivers over a two hour time frame. The difference in the offsets is quite obvious, with the NovAtel OEM3 offsets remaining close to zero and the Trimble 4000 SSI offsets drifting to 1 ms before suddenly shifting back to zero. The reason for these differing offset behaviours lies in the receiver implementation. In the case of the NovAtel, the receiver calculates its clock offset internally and attempts to “steer” the clock to minimize its clock offset. Thus,

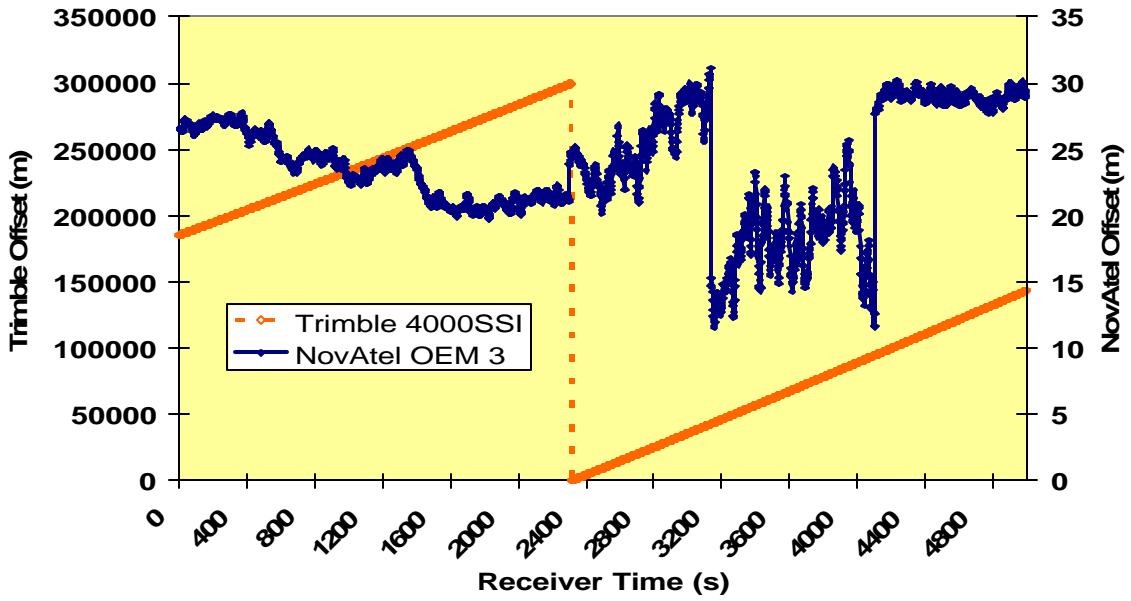


Figure 2.3. Receiver Clock Offsets for Two GPS Receivers.

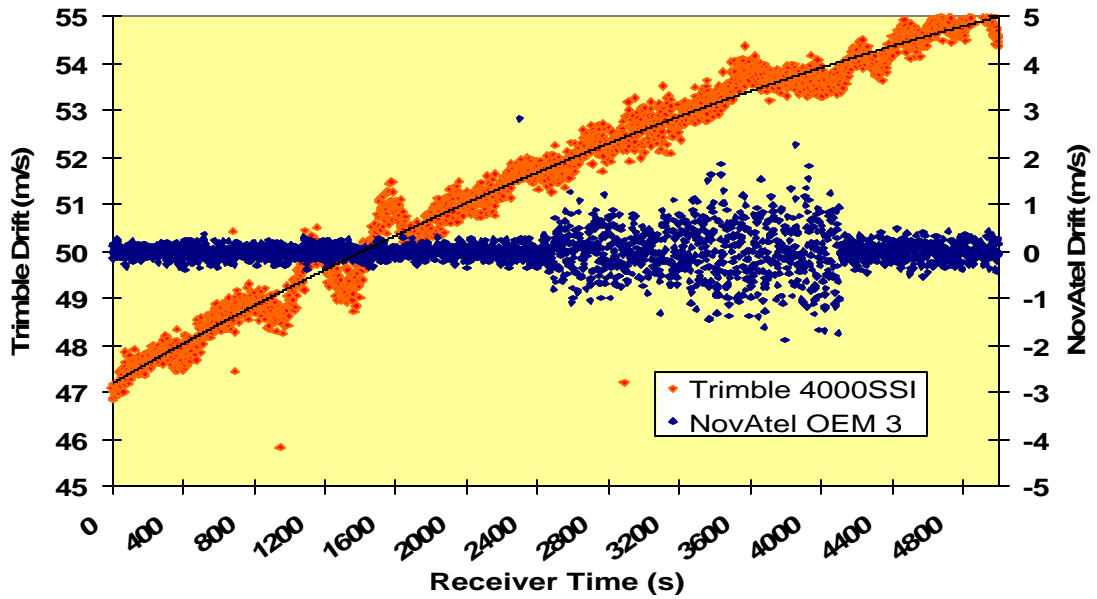


Figure 2.4. Receiver Clock Drift for Two GPS Receivers.

the offsets remain close to zero. However, the Trimble receiver allows the drift to grow to 1 ms before resetting the clock. The drift properties of the receiver clocks are also well illustrated by Figure 2.4, which show the clock drifts for the two receivers. The drift of the NovAtel receiver remains very close to zero, with an average value of 3mm/s. The Trimble receiver, on the other hand, has a significant drift, and the drift rate itself varies over time.

In modern GPS receivers, the receiver makes observations to all satellites simultaneously. As a result, the clock offset affecting each pseudorange is identical. This allows the receiver clock offset to be solved for as an unknown parameter along with the receiver's position. As the clock offset and its drift changes from epoch to epoch a new offset must be solved for each epoch. Alternatively, the observations collected at a receiver can be *differenced*, a concept discussed in detail in Chapter 3. Note that since the clock offsets are either eliminated or estimated, the size of the clock offsets are not significant in static, post-processed applications. However, in real-time applications, large clock offsets *are* undesirable due to the need for measurements to be made synchronously at different sites.

#### ***2.4 -Satellite Orbits***

The coordinate system of GPS (WGS-84) is defined by the combined coordinates of 5 monitor stations located at Hawaii, Colorado, Ascension Island, Diego Garcia and Kwajalein (Hofmann-Wellenhof et al, 1994). However, as the user can only make observations to orbiting satellites, the coordinates of these satellites must be provided to allow access to the WGS-84 frame. The direct analogy in terrestrial networks is that of using second-order control to define the datum for an engineering network, where the coordinates of the second-order control have been previously derived from a survey campaign linking them to a sparse first-order control set. Datum definition is more thoroughly discussed in Chapter 4.

The difference between GPS and terrestrial networks lies in the fact that the “observable” control points are continuously moving. This results in several peculiarities. For example, to allow for real-time operation, the satellite positions must be *predicted*. This prediction is



carried out in two steps (Russell and Schaibly, 1980). Initially, data collected at the five monitor stations is processed via a Kalman filter to calculate actual satellite positions, velocities, solar pressure constants and other satellite states every 15 minutes for a one week period. This week's worth of data is then used to predict a satellite orbit (or *ephemeris*) for a one day span. However, due to bandwidth limitations, neither the full orbital model nor the coordinates of the satellites themselves are transmitted to the user, but rather the parameters of a simplified 15 element perturbed-Keplerian orbital model, derived using 4 hours of the 24 hr predicted base ephemeris. The user must then calculate the satellite position for their particular epoch using the best available broadcast ephemeris, which are typically updated every two hours (Beutler et al, 1998).

Just as in the case of the satellite clock model, actual observed orbits are provided by several agencies, including the IGS. These are known as *precise ephemerides*. Figure 2.5 shows the difference between the precise ephemerides and the broadcast orbit for several satellites over a 10 hour period. The along track agreement is the worst, as might be expected since this is along the direction of the satellite's motion. The accuracy of the IGS orbits used are at the 5 centimetre level, although a variety of IGS orbit products are available with accuracies dependant on their latencies (Roulston et al, 2000). As a result, the data shown in Figure 2.5 implies a broadcast accuracy of 1.2 m radial, 2.4 m across track and 4.5 m along track. These are in close agreement to those obtained by Zumberge and Bertiger (1994).

In general, the GPS satellites follow an almost circular orbit with a radius of roughly 26560 km. As well, the orbital period is one half of a sidereal day, which implies a repeating satellite geometry every 23hr 56min (due to the rotating Earth). Figure 2.6 shows the velocities of observed satellites over a 10 hour period. The velocities shown are with respect to the WGS-84 frame and, since the frame rotates with the Earth, the shown velocities are lower than the "true" satellite velocities. In an inertial frame, the average satellite velocity is 3860 m/s. Due to the great velocities of the GPS satellites, the true transmission time of the ranging signals must be accurately known in order to properly calculate the position of the satellite at transmission, as discussed in Section 2.5.1.

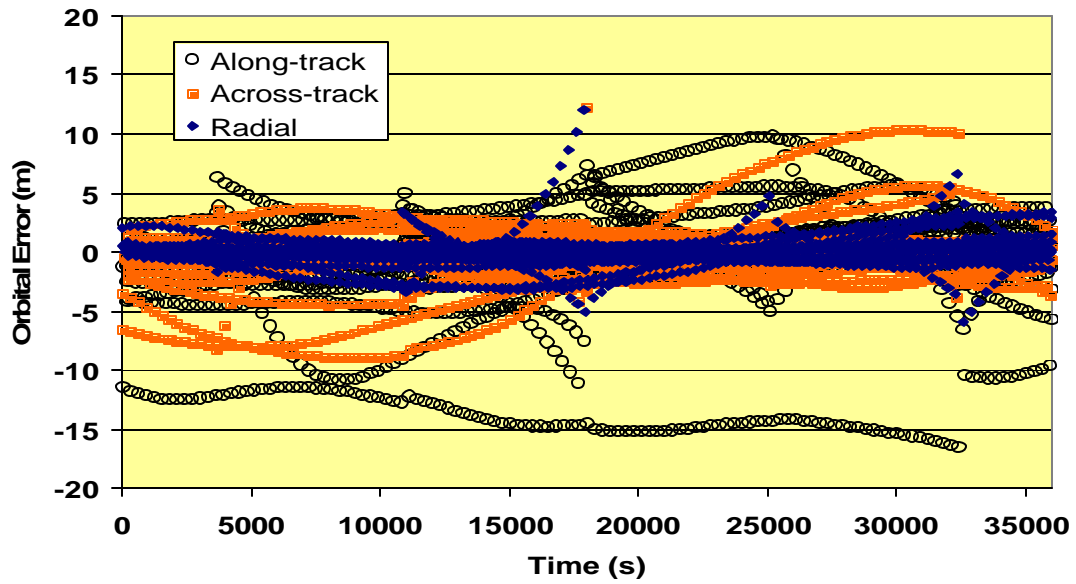


Figure 2.5. Comparison of Precise and Broadcast Ephemerides.

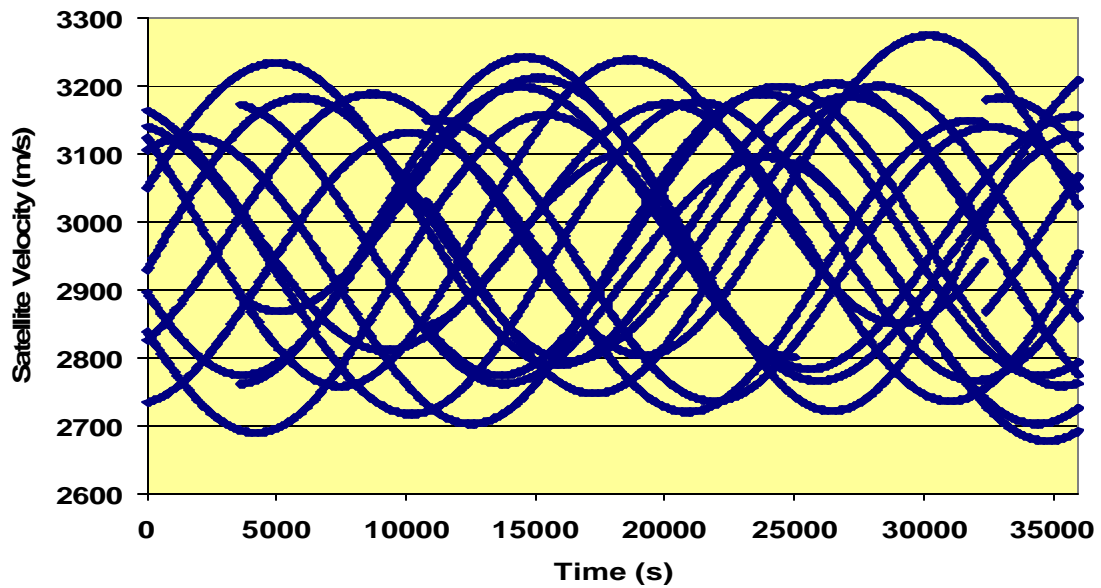


Figure 2.6. Satellite Velocities Relative to WGS-84.

## 2.5 -Geometrical Model of Earth-Space Ranging

The purpose of space-borne trilateration systems is to determine the three-dimensional coordinates of one or more receivers in a given reference plane. For our purposes, the reference frame used will be an Earth-centred, Earth-fixed (ECEF) frame, such as WGS-84. Once coordinates of receivers have been determined in this frame, they can be converted into any other frame through a suitable transformation.

A three-dimensional distance between a satellite and receiver is simply described as,

$$d_g = \|r_s(t_s') - r_r(t_r')\| = \sqrt{(x_s(t_s') - x_r(t_r'))^2 + (y_s(t_s') - y_r(t_r'))^2 + (z_s(t_s') - z_r(t_r'))^2} \quad (2.8)$$

where  $\mathbf{r}_s, \mathbf{r}_r$  are the position vectors of the satellite and receiver at transmission and reception time, respectively in a fixed (non-rotating) reference frame and  $t_s', t_r'$  are the *actual* transmission and reception times.

### 2.5.1 - Modification for Transmission Time and Satellite Motion

Eq. (2.8) requires the coordinates of the receiver and satellites to be known at the actual transmission and reception times. Since only the reception time is explicitly known, one must derive the transmission time. If a measurement to a satellite is logged by a receiver at time  $t_r$  (apparent receiver time), then the true time of measurement is  $t_r - \Delta t_r$ , with  $\Delta t_r$  being the receiver clock offset. Furthermore, if the electromagnetic distance to the satellite is  $d_e$ , then the true transmission time must be

$$t_s' = t_r - \Delta t_r - d_e \quad (2.9)$$

By inserting Eq. (2.2) into the above and using the definition of a pseudorange, one can derive the true transmission time in terms of the measured pseudorange,  $p$  :

$$t_s' = t_r - \Delta t_r - \left( \frac{P}{c} + (\Delta t_s - \Delta t_r) \right) = t_r - \frac{P}{c} - \Delta t_s \quad (2.10)$$

where  $\Delta t_s$  is the estimated satellite clock offset and known (from the broadcast ephemeris) to approximately 5 ns. The pseudorange measurement noise is at the several metre level, corresponding to a timing error of about 33 ns. Note that the tropospheric and ionospheric delays are implicitly incorporated in the measured pseudorange, as is the receiver clock offset. Combined, the resulting error in transmit time calculation contributes less than one centimetre to the satellite position error. However, given that GPS satellites have an orbital radius of roughly 26 560 km, it can be shown that transit times vary from approximately 67 ms to 90 ms (Remondi, 1984). As a result, if the correction for transmission time was not applied, a satellite position error of up to 350 m could result, based on the satellite velocities presented in Section 2.4.

### 2.5.2 -Modification for a Rotating Earth

Eq. (2.8) holds true only if the coordinates of the receiver and satellite (which refer to different epochs) relate to the same reference frame. Thus, if one wishes to work with positions in an ECEF frame, a correction must be made since the coordinate system to which the satellite coordinates refer to is rotated with respect to the coordinate system of the receiver. This phenomenon is illustrated in Figure 2.7. The amount of rotation is solely dependant on the transmission time and as a result different for each satellite observed. This “frame rotation” can induce up to 175 m of error into the satellite coordinates if not taken into account.

One can modify Eq. (2.8) to explicitly take into account the transit time by adding a rotation about the  $z$ -axis, resulting in the following expression for the geometric distance between the satellite and receiver at transmission and reception times, respectively, in a frame referring to the true reception time :

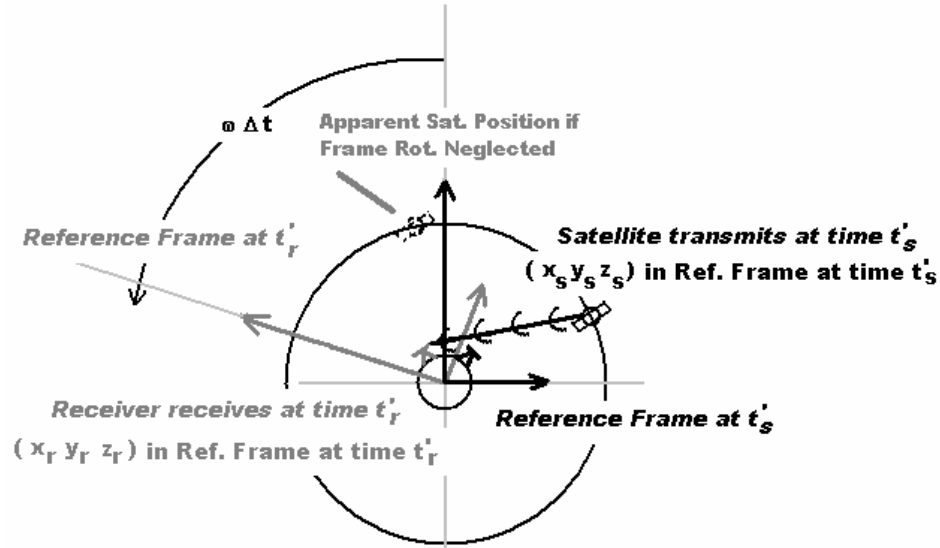


Figure 2.7. Rotating Reference Frame Problem.

$$\begin{aligned}
 d_g &= \|R_z(\mathbf{w} \cdot \mathbf{t}) \cdot r_s(t_s') - r_r(t_r')\| \\
 &= \sqrt{(x_s(t_s') \cos \mathbf{w}t + y_s(t_s') \sin \mathbf{w}t - x_r(t_r'))^2 + (y_s(t_s') \cos \mathbf{w}t - x_s(t_s') \sin \mathbf{w}t - y_r(t_r'))^2} \\
 &\quad \sqrt{+ (z_s(t_s') - z_r(t_r'))^2}
 \end{aligned}
 \tag{2.11a}$$

where

$$R_z(\mathbf{q}) = \begin{bmatrix} \cos \mathbf{q} & \sin \mathbf{q} & 0 \\ -\sin \mathbf{q} & \cos \mathbf{q} & 0 \\ 0 & 0 & 1 \end{bmatrix}
 \tag{2.11b}$$

and  $t_s'$  and  $t_r'$  refer to the true transmission and reception time of the signal,  $\mathbf{t}$  is the transit time of the signal and  $\mathbf{w}$  is the rotation rate of the Earth. The transit time is equal to the electromagnetic distance and thus can be derived from Eq. (2.2) using the measured pseudorange and the estimated receiver and satellite clock offsets. Assuming no knowledge of the receiver clock offset (and thus a maximum 0.5 ms effect), the worst-case resulting error in satellite orbit due to frame rotation is 1 m, and so within the accuracy of the broadcast orbits themselves.

### 2.5.3 -Modification for Carrier Phase Measurement

GPS provides two measurement types – code and carrier phase. The code measurement is simply a time-of-flight measurement of the signal transit time made by correlating the satellite generated ranging code to an identical copy generated inside of the receiver (but shifted due to the transit time delay and the satellite and receiver clock offsets). Details of the implementation of this measurement mode can be found in Ward (1996) and in Spilker (1980). Carrier phase measurement, on the other hand, is a more precise measurement mode created by comparing the carrier frequencies of the transmitted signal and the internally generated replica. The Doppler shift between the two frequencies is accumulated over time, resulting in a range - see Remondi (1984) and Ward (1996) for implementation details. The principal advantage of using the carrier phase method is that observations of this accumulated Doppler are precise to within roughly 1% of the wavelength of the carrier (Hofmann-Wellenhof et al, 1994). Thus in the case of GPS, where the L1 carrier wavelength is 19 cm, this translates into a measurement noise of several mm. Conversely, the pseudorange measurement noise is currently at the several decimetre level (Langley, 1997).

However, using the carrier phase measurement introduces several modifications to the ranging model. Since the carrier phase is actually an accumulated Doppler from some initial epoch, the starting range from the receiver to the satellite is unknown. More specifically, while the receiver can measure the fractional phase offset of the transmitted and generated carrier phases at the initial epoch, the number of full cycles between the receiver and satellite is not known. This is termed the *integer ambiguity*. As a result, Eq. (2.4) describing the pseudorange measurement can be modified to yield :

$$\mathbf{f} = \frac{d_g - c \cdot (\Delta t_s - \Delta t_r) + T + I}{\mathbf{l}} + \mathbf{e}_f + N \quad (2.12)$$

where  $\mathbf{f}$  is the carrier phase measurement (in cycles),  $\mathbf{l}$  is the carrier wavelength,  $N$  is the integer ambiguity and  $\mathbf{e}_f$  is the carrier phase noise.

As shall be seen, a crucial step in making the carrier phase measurement useful for positioning is the solution of the integer ambiguity, which is different for each satellite observed. In addition, while the ambiguity is constant while the satellite is continuously observed, if for some reason signal lock to the satellite is lost (i.e. due to poor signal-to-noise ratio, high dynamics, or signal blockage), a new ambiguity for the satellite must be estimated. Detection of such cycle-slips is not a trivial problem, as the difference in the old and new ambiguities can range from millions of cycles to only a few. This is particularly true if the receiver is moving, or if measurement errors are at high levels. Often, Kalman filtering of the raw carrier observations is performed as a pre-processing step to flag and correct these slips, but many other techniques exist; see Han (1995); Binath (2000); Kim and Langley (2002) for the development of the current state-of-the-art.

### **3.0 - GENERAL SOLUTION TO THE SPACED-BASED TRILATERATION PROBLEM**

In general, the final goal of a space-based navigation system is to provide the user with the positions of one or several receivers, at a particular epoch and in a specified reference frame. As discussed in the previous chapter, modern space-based navigation systems, such as GPS, are based on measuring ranges from orbiting satellites to receivers on the Earth. As a result, they are defined as trilateration networks, and, when comprised of sufficient and appropriate observations, these networks can be solved via the methods of least-squares.

This section presents a general solution to the spaced-based trilateration problem. A review of least-squares theory is presented, with special focus on the case of added parameter constraints. The measurements and unknowns encountered in the GPS positioning problem are discussed, as are the mathematical model and the linearized quantities involved in the adjustment problem. The concept of *differential* techniques is introduced and the mathematical properties of the differencing procedure are illustrated. Finally, the problem of iteration and convergence is investigated, as well as some implications of the *curvature* of the GPS mathematical model.

#### ***3.1 - Unknowns and Measurements in a Space-Based Trilateration Network***

Assume the situation of a number of receivers in an engineering network simultaneously observing a constellation of GPS satellites. In most applications, the sole purpose of the measurement campaign is to establish the three-dimensional coordinates of the receivers at a particular epoch or over a particular time-span.

For this purpose, each receiver makes a number of measurements to each of the visible satellites. In the GPS system, these measurements can take two basic forms – the pseudorange and the carrier phase. Currently, GPS satellites broadcast information on two frequencies – 1575 MHz (L1) and 1227 MHz (L2). As will be discussed in Chapter 6, this



allows for the extraction of the first-order ionospheric effect, which is frequency dependant. By 2005, an extra frequency is expected to be available at 1176.45 MHz (Fontana et al, 2001). Finally, for each frequency observed, the Doppler shift due to the satellite motion relative to the receiver is measured. While this measurement was an important aspect of measurement with early satellite-based systems such as TRANSIT, it is rarely used in static applications due to its noisiness and therefore will not be considered.

Ideally, one would use the available measurements and by linearizing Eq. (2.8), which describes the geometric distance between the receivers and the satellites, arrive at a least-squares solution for the position unknowns. However, several factors complicate the solution. Firstly, as discussed in Section 2.2 and 2.3, the receiver and satellite time frames are not synchronized. Furthermore, the individual receiver time frames are not synchronized to one another and are also not stable over time. In addition, as presented in Section 2.5.3, each carrier phase measurement carries with it an unknown integer offset which is stable over time, so long as lock to the satellite is maintained. Finally, since the measured ranges refer to electromagnetic distances, Eq. (2.4) indicates that they are corrupted by tropospheric and ionospheric delays, as well as noise.

Several methods exist to deal with these complications. Firstly, one can attempt to *solve* for these added factors as *nuisance parameters*. A second possibility is to treat their effects as *noise*, coupled with an appropriate stochastic model. Finally, if their effects are identical (or very similar) between pairs of observations, then the factor can be effectively *removed* from the problem by *differencing* the observations. In general, the decision as to which of these approaches is most appropriate depends on the nature of the effect. For example, the ambiguity terms, since they are large, of unpredictable magnitude, and are constant over time, are treated as additional unknowns. Satellite clock offsets, on the other hand, are usually treated by differencing if multiple receivers are available since they change from epoch to epoch, but are identical for a given satellite. The same applies for receiver clock offsets, with the modification that the offset is identical for all observations at a given receiver. Finally, tropospheric and ionospheric errors are often treated with a hybrid approach of differencing, estimation, and stochastic modeling.

### 3.2 - Formulation of the Least-Squares Solution

In any static GPS campaign, the vector of unknowns,  $\mathbf{x}$ , can be described as follows

$$\mathbf{x}^T = \left[ \mathbf{x}_{pos}^T \quad \mathbf{x}_{rx}^T \quad \mathbf{x}_{sat}^T \quad \mathbf{x}_{mod}^T \quad \mathbf{x}_{amb}^T \right] \quad (3.1)$$

where  $\mathbf{x}_{pos}$  is the vector of unknown receiver positions,  $\mathbf{x}_{rx}$  is the vector of receiver clock offsets at each epoch,  $\mathbf{x}_{sat}$  is the vector of satellite clock offsets at each epoch,  $\mathbf{x}_{mod}$  is the vector of additional model parameters (i.e. tropospheric parameters) and  $\mathbf{x}_{amb}$  is the vector of unknown ambiguities (if carrier phases have been measured). The observations made over all epochs at all receivers can be collected into a vector,  $\mathbf{l}$ , including pseudorange and carrier phase observations made at all frequencies.

The observations are linked to the unknown parameters through the mathematical model used. For example, a carrier phase measurement can be described in terms of the unknown receiver coordinates and ambiguity term and the known satellite coordinates, as in Eq. (2.13). If the mathematical models for all the individual observations are collected, and linearized with respect to the observations and the unknowns, the linear math model can be expressed as

$$\mathbf{A} \cdot \mathbf{d} + \mathbf{B} \cdot \mathbf{r} + \mathbf{w} = \mathbf{0} \quad (3.2)$$

where  $\mathbf{d}$  is the vector of corrections to the initial estimates of the unknown parameters,  $\mathbf{r}$  is a vector of corrections to the observations and  $\mathbf{w}$  is known as the misclosure vector.  $\mathbf{A}$  and  $\mathbf{B}$  are the Jacobian matrices of the mathematical model with respect to the unknowns and observations, respectively. The Jacobian matrices are evaluated about the initial estimates of the unknowns,  $\mathbf{x}_0$  and the observations,  $\mathbf{l}$ . The notation presented follows that of Vanicek and Krakowsky (1986).

The mathematical model described above can be augmented if additional information regarding the relationships between the unknowns is available. These are termed *parameter constraints*, and can in general be expressed as

$$\mathbf{G}^T \cdot \mathbf{d} + \mathbf{w}_c = 0 \quad (3.3)$$

where  $\mathbf{G}$  is the linearized constraint matrix and  $\mathbf{w}_c$  is the constraint misclosure. For a solution to exist,  $\mathbf{G}$  must have linearly independent columns – i.e. any constraint can not be expressed as a function of two other constraints. Consequently, this implies that the rank of  $\mathbf{G}$  must equal the number of columns in  $\mathbf{G}$ .

The objective of the least-squares problem is then to find a solution to  $\mathbf{x}$  that minimizes (Mikhail, 1972) :

$$\mathbf{q} = \mathbf{r}^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r} + \mathbf{d}^T \cdot \mathbf{C}_x^{-1} \cdot \mathbf{d} + 2 \cdot \mathbf{k}_1^T \cdot (\mathbf{A} \cdot \mathbf{d} + \mathbf{B} \cdot \mathbf{r} + \mathbf{w}) + 2 \cdot \mathbf{k}_2^T \cdot (\mathbf{G}^T \cdot \mathbf{d} + \mathbf{w}_c) \quad (3.4)$$

where  $\mathbf{C}_1$  is the variance-covariance matrix relating the stochastic properties of the observations,  $\mathbf{C}_x$  is the variance-covariance matrix containing the stochastic properties of the initial estimates of the unknowns and  $\mathbf{k}_1$ ,  $\mathbf{k}_2$  are Lagrange multipliers required to ensure that the solution not only minimizes the quadratic terms, but also satisfies Eq. (3.2) and Eq. (3.3).

Leick (1990) presents a very useful, stepwise solution to the minimization problem. First, a least-squares solution for the problem without constraints is calculated as :

$$\begin{aligned} \mathbf{d}_* &= -\left(\mathbf{A}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot \mathbf{A} + \mathbf{C}_x^{-1}\right)^{-1} \cdot \mathbf{A}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot \mathbf{w} \\ &= -\left(\mathbf{N} + \mathbf{C}_x^{-1}\right)^{-1} \cdot \mathbf{u}_* = -\mathbf{N}_*^{-1} \cdot \mathbf{u}_* \end{aligned} \quad (3.5a)$$

$$\mathbf{r}_* = -\mathbf{C}_1 \cdot \mathbf{B}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot (\mathbf{A} \cdot \mathbf{d}_* + \mathbf{w}) \quad (3.5b)$$

where  $\mathbf{d}_*$  is the solution of the least-squares problem without constraints, and  $\mathbf{r}_*$  is the resulting residual vector.

Given a suitable constraint matrix,  $\mathbf{G}$ , the solution for the problem incorporating the constraints is then given as :

$$\mathbf{d}_c = \mathbf{d}_* - \mathbf{N}_*^{-1} \cdot \mathbf{G} \cdot (\mathbf{G}^T \cdot \mathbf{N}_*^{-1} \cdot \mathbf{G})^{-1} \cdot (\mathbf{G}^T \cdot \mathbf{d}_* + \mathbf{w}_c) \quad (3.6a)$$

$$\mathbf{r}_c = \mathbf{r}_* - \mathbf{C}_1 \cdot \mathbf{B}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot \mathbf{A} \cdot (\mathbf{d}_c - \mathbf{d}_*) \quad (3.6b)$$

where  $\mathbf{d}_c$  and  $\mathbf{r}_c$  refer to the solutions after the constraints have been applied.

Two additional quantities of note are the weighted sum-of-squares of residuals for the two cases. In the case of no constraints, the quadratic form can be solved as

$$\mathbf{r}_*^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_* = -\mathbf{u}_*^T \cdot \mathbf{N}_*^{-1} \cdot \mathbf{u}_* + \mathbf{w}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot \mathbf{w} \quad (3.7a)$$

whereas in the constrained case this changes to

$$\mathbf{r}_c^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_c = \mathbf{r}_*^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_* + (\mathbf{G}^T \cdot \mathbf{d}_* + \mathbf{w}_c)^T \cdot (\mathbf{G}^T \cdot \mathbf{N}_*^{-1} \cdot \mathbf{G})^{-1} \cdot (\mathbf{G}^T \cdot \mathbf{d}_* + \mathbf{w}_c) \quad (3.7b)$$

Note that the sum-of-squares in the constrained case is always greater, since the constraints effectively move the solution away from the local minimum.

Lastly, the variance-covariance matrices of the estimated parameters in the two situations are given by

$$\mathbf{C}_{\mathbf{x}^*} = \mathbf{N}_*^{-1} \quad (3.8a)$$

$$\mathbf{C}_{\mathbf{x}} = \mathbf{N}_* - \mathbf{N}_*^{-1} \cdot \mathbf{G} \cdot (\mathbf{G}^T \cdot \mathbf{N}_*^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \cdot \mathbf{N}_*^{-1} \quad (3.8b)$$

Note that in this case the estimated accuracy of the unknowns always increases, since we are adding information to the solution in the form of constraints.

An advantage of using the above approach in solving the constrained least-squares adjustment problem lies in the fact that it is easy to update the solution if additional constraints are to be applied at a later point. So long as the new constraints,  $\mathbf{G}_2$ , are independent of the initial constraints  $\mathbf{G}_1$  (implying that the rank of the total constraint matrix is still equal to the number of constraints), the updated quantities of interest are calculated as

$$\mathbf{d}_{c2} = \mathbf{d}_c - \mathbf{C}_{xc} \cdot \mathbf{G}_2 \cdot (\mathbf{G}_2^T \cdot \mathbf{C}_{xc} \cdot \mathbf{G}_2)^{-1} \cdot (\mathbf{G}_2 \cdot \mathbf{d}_c + \mathbf{w}_{c2}) \quad (3.9a)$$

$$\mathbf{r}_{c2} = \mathbf{r}_* - \mathbf{C}_1 \cdot \mathbf{B}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot \mathbf{A} \cdot (\mathbf{d}_{c2} - \mathbf{d}_*) \quad (3.9b)$$

$$\mathbf{r}_{c2}^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_{c2} = \mathbf{r}_c^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_c + (\mathbf{G}_2^T \cdot \mathbf{d}_c + \mathbf{w}_{c2})^T \cdot (\mathbf{G}_2^T \cdot \mathbf{C}_{xc} \cdot \mathbf{G}_2)^{-1} \cdot (\mathbf{G}_2^T \cdot \mathbf{d}_c + \mathbf{w}_{c2}) \quad (3.9c)$$

$$\mathbf{C}_{xc2} = \mathbf{C}_{xc} - \mathbf{C}_{xc} \cdot \mathbf{G}_2 \cdot (\mathbf{G}_2^T \cdot \mathbf{C}_{xc} \cdot \mathbf{G}_2)^{-1} \cdot \mathbf{G}_2^T \cdot \mathbf{C}_{xc} \quad (3.9d)$$

This will become a key development in Chapter 5, when the concept of ambiguity resolution is discussed.

### 3.2.1 -Modification for Incorrect Apriori Stochastic Assumptions

The preceding developments only hold if the  $\mathbf{C}_x$  matrix used actually corresponds to the actual variance-covariance matrix of the initial estimates, or is neglected entirely. In the case that an arbitrary matrix  $\mathbf{P}$  is used in the place of  $\mathbf{C}_x$ , the least-squares estimate of the parameters is still given by Eq. (3.5), but the associated variance-covariance matrix of the estimated parameters becomes

$$\mathbf{C}_{x*} = (\mathbf{I} - \mathbf{N}_*^{-1} \cdot \mathbf{N}) \cdot \mathbf{C}_x \cdot (\mathbf{I} - \mathbf{N}_*^{-1} \cdot \mathbf{N})^T + \mathbf{N}_*^{-1} \cdot \mathbf{N} \cdot \mathbf{N}_*^{-1} \quad (3.10a)$$

where

$$\mathbf{N}_*^{-1} = (\mathbf{N} + \mathbf{P}^{-1})^{-1} \quad (3.10b)$$

Krakiwsky (1975) shows that Eq. (3.10a) reduces to (3.8a) if  $\mathbf{P}^{-1} = \mathbf{C}_x$ . Otherwise, the error variance of the parameters is composed of a bias component dependant on the accuracy of the initial estimates and a noise component incorporating the accuracy of the observations and the geometry of the network. This relationship is revisited in Chapter 4.

If the solution is updated by the introduction of a constraints matrix,  $\mathbf{G}$ , the resulting solution is still given by Eq. (3.6a). However, the variance-covariance matrix of this solution is not simply calculated via Eq. (3.8b). Rather, while the  $\mathbf{C}_{xc}$  is calculated using Eq. (3.8b), the *actual* error variance is calculated by adding the propagated corresponding noise and bias components, which themselves are determined via :

$$\mathbf{C}_{\text{noise}} = \mathbf{C}_{xc} \cdot \mathbf{N} \cdot \mathbf{C}_{xc} \quad (3.11a)$$

$$\mathbf{C}_{\text{bias}} = \left( \mathbf{I} - \mathbf{C}_{xc} \cdot \mathbf{N} - \mathbf{N}_*^{-1} \cdot \mathbf{G} \cdot (\mathbf{G}^T \cdot \mathbf{N}_*^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \right) \cdot \mathbf{C}_x \cdot \left( \mathbf{I} - \mathbf{C}_{xc} \cdot \mathbf{N} - \mathbf{N}_*^{-1} \cdot \mathbf{G} \cdot (\mathbf{G}^T \cdot \mathbf{N}_*^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \right)^T \quad (3.11b)$$

If additional constraints are added, the solution proceeds as presented in the previous section, but the variance-covariance matrices are propagated using the above equations, substituting  $\mathbf{C}_{xc2}$  for  $\mathbf{C}_{xc}$  and  $\mathbf{C}_{xc}$  for  $\mathbf{N}_*^{-1}$  where appropriate.

### 3.3 -Form of the Linearized Quantities

The classical method of solving an overdetermined problem via least-squares as discussed above requires the mathematical model to be linearized. Essentially, the actual mathematical model surface  $\mathbf{f}(\mathbf{x}, \mathbf{l}) = \mathbf{0}$  is approximated by a plane normal to the surface at initial estimates  $\mathbf{x}_0, \mathbf{l}$ . The derivatives of the mathematical model with respect to all of the unknowns and observations are then calculated, thereby forming the Jacobian matrices  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. This section briefly presents the form of the derivatives, as they will be useful to understand some particular aspects of the GPS positioning problem.

Eqs. (2.8), (2.4) and (2.12) form the basis of our mathematical model. For completeness they will be repeated here :

$$d_g = \|r_s(t_s') - r_r(t_r')\| = \sqrt{(x_s(t_s') - x_r(t_r'))^2 + (y_s(t_s') - y_r(t_r'))^2 + (z_s(t_s') - z_r(t_r'))^2} \quad (2.8)$$

$$p = d_g - c \cdot (\Delta t_s - \Delta t_r) + T + I + \mathbf{e} \quad (2.4)$$

$$\mathbf{f} = \frac{d_g - c \cdot (\Delta t_s - \Delta t_r) + T + I}{\mathbf{1}} + \mathbf{e}_f + N \quad (2.12)$$

where Eq. (2.8) represents the geometrical model, Eq. (2.4) represents the pseudorange measurement model and Eq. (2.12) is the carrier phase measurement model. One such model can be established for every observation between a receiver and satellite. These models can then be collected into a vector  $\mathbf{f}$ .

The partial derivative of  $\mathbf{f}$  with respect to the unknowns and observations results in the necessary matrices  $\mathbf{A}$  and  $\mathbf{B}$ . According to the partitioning of the vector of unknowns used in Eq. (3.1), the  $\mathbf{A}$  and  $\mathbf{B}$  matrices can be similarly written as

$$\mathbf{A} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{\text{pos}}} & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{\text{rx}}} & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{\text{sat}}} & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{\text{mod}}} & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{\text{amb}}} \end{bmatrix} \quad (3.12a)$$

$$\mathbf{B} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{l}} \end{bmatrix} \quad (3.12b)$$

Note that since there is one observation for each element in  $\mathbf{f}$ , the  $\mathbf{B}$  matrix will be identity.

### 3.3.1 -Derivative with Respect to the Unknown Positions

The derivatives of Eq. (2.4) with respect to the unknown receiver positions are

$$\frac{\partial p_{ij}}{\partial x_{ri}} = \frac{-(x_{sj} - x_{ri})}{\sqrt{(x_{sj} - x_{ri})^2 + (y_{sj} - y_{ri})^2 + (z_{sj} - z_{ri})^2}} \quad (3.13a)$$

$$\frac{\partial p_{ij}}{\partial y_{ri}} = \frac{-(y_{sj} - y_{ri})}{\sqrt{(x_{sj} - x_{ri})^2 + (y_{sj} - y_{ri})^2 + (z_{sj} - z_{ri})^2}} \quad (3.13b)$$

$$\frac{\partial p_{ij}}{\partial z_{ri}} = \frac{-(z_{sj} - z_{ri})}{\sqrt{(x_{sj} - x_{ri})^2 + (y_{sj} - y_{ri})^2 + (z_{sj} - z_{ri})^2}} \quad (3.13c)$$

where it is assumed that frame synchronization and transmission time effects have been taken into account, and  $\langle x_{sj} \ y_{sj} \ z_{sj} \rangle$  is the vector of coordinates for the  $j^{\text{th}}$  satellite and  $\langle x_{ri} \ y_{ri} \ z_{ri} \rangle$  is the vector of coordinates for the  $i^{\text{th}}$  receiver. By collecting the derivatives in Eq. (3.13) into a single vector, it can be seen that this vector represents the normal vector from the satellite to the receiver. If carrier phases are used, then the above derivatives must be divided by the wavelength of the carrier, in accordance with Eq. (2.12). Obviously, if a particular receiver does not appear in an observation, then the derivatives of the observation with respect to that receiver's coordinates are zero. Finally, as the satellite coordinates are continuously changing, the above derivatives must be recalculated for every epoch.

### 3.3.2 -Derivatives with Respect to the Unknown Clock Offsets

The derivative of an observation made at a particular epoch involving the  $i^{\text{th}}$  receiver and the  $j^{\text{th}}$  satellite with respect to the clock offsets of the receiver and satellite are expressed as

$$\frac{\partial p_{ij}}{\partial \Delta t_{ri}} = c \quad (3.14a)$$

$$\frac{\partial p_{ij}}{\partial \Delta t_{sj}} = -c \quad (3.14b)$$



Since the numerical value of the speed of light is very great in comparison to the other derivatives, the clock offsets are typically parameterized in distance units and the  $c$  replaced by 1 in the above equations. This prevents numerical round off errors in the subsequent processing. Also, note that since the clock offsets vary with time, new unknowns are required at every epoch. Given the case of a three-receiver network observing two satellites over two epochs, the resulting  $\mathbf{A}$  matrix for the clock offset terms is shown in Figure 3.1.

$$\mathbf{A}_{\text{clk}} = \begin{array}{c} \begin{array}{cccccc|cccc} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} \end{array} \\ \begin{array}{l} \text{Epoch 1} \\ \text{Epoch 2} \end{array} \\ \begin{array}{cc} \text{Rx Clock} & \text{Sat Clock} \end{array} \end{array}$$

**Figure 3.1.** Form of the Jacobian of Satellite and Receiver Clock Offsets.

As the data span used for positioning increases, it becomes evident that the number of clock terms to be estimated increases linearly, resulting in unwieldy matrices to manipulate. Fortunately the process of differencing allows us to remove these unknowns from the solution.

### 3.3.3 -Derivatives with Respect to the Unknown Ambiguities

In the case that carrier phase measurements are used, the additional unknown ambiguity terms must be estimated. The resulting derivative of a carrier phase observation with respect to the ambiguity term is simply

$$\frac{\partial \mathbf{f}_{ij}}{\partial N_{ij}} = 1 \quad (3.15)$$

As the ambiguity remains constant for a particular satellite-receiver observation so long as lock to the satellite is maintained by the receiver, the number of unknown ambiguities and the resulting size of the  $\mathbf{A}$  matrix of the ambiguities remains at a manageable size, for normally encountered time spans. However, a peculiarity of solving for the ambiguities lies in that the mathematical model as specified does not contain information regarding the integer nature of the ambiguities. As a result, the least-squares solution will yield real-valued ambiguities. This is known as the *float* solution. Incorporating the integer nature of the ambiguities is known as *ambiguity-resolution* and is treated in Chapter 5.

### 3.3.4 -Derivative with Respect to Other Model Parameters

In Eq. (3.1), the vector of unknowns is partitioned into a vector of position unknowns, unknown clock offsets, ambiguity unknowns, and a vector of additional model parameters. These model parameters are any models that augment the basic mathematical model contained in Eq. (2.4).

For example, the residual tropospheric effect is often modelled as (Mendes, 1999)

$$T_{ij} = \Delta z \cdot m(\mathbf{e}_{ij}) \quad (3.16)$$

where  $\Delta z$  is the unknown zenith tropospheric delay, and  $m(\cdot)$  is the known tropospheric mapping function which itself is a function of the elevation angle of the satellite as seen by the receiver,  $\mathbf{e}_{ij}$ . The desire is to improve the positioning accuracy by modelling the

tropospheric error through estimation of the zenith tropospheric delay. Depending on the parameterization,  $\Delta z$  may be considered to be constant, or vary according to a random walk process. Schüler (2001) provides an excellent review of techniques to model the tropospheric error.

Incorporating Eq. (3.16) into the mathematical model for the pseudoranges, the derivative of Eq. (2.4) with respect to the unknown zenith tropospheric delay is

$$\frac{\partial p_{ij}}{\partial \Delta z} = \frac{\partial T_{ij}}{\partial \Delta z} = m(\mathbf{e}_{ij}) \quad (3.17)$$

However, a problem arises since the tropospheric error, which requires the elevation angle of the satellite, will as a result depend on the coordinates of the receiver, vis.

$$\frac{\partial T_{ij}}{\partial x_{ri}} = \Delta z \cdot \frac{\partial m(\mathbf{e}_{ij})}{\partial x_{ri}} = \Delta z \cdot \frac{\partial m(\cdot)}{\partial \mathbf{e}} \cdot \frac{\partial \mathbf{e}_{ij}}{\partial x_{ri}} \quad (3.18)$$

Depending on the tropospheric mapping function used, the above derivatives may not be trivial to compute. Fortunately, due to the great distance between the receivers and satellites, the elevation angle is fairly insensitive to slight changes in the receiver coordinates – 100m difference will result in an angle change of 1". This further implies that the derivative of the tropospheric error with respect to the receiver coordinates is very close to zero and so can be neglected.

Many other models can be added to the basic positioning model. For example, the ionospheric error can be parameterized in numerous ways, as can receiver and satellite clock offsets. However, it is important when implementing these models to account for interdependencies in all the unknowns considered.

### 3.4 -Differential GPS

When a set of observations is made at a given receiver, the receiver clock offset affecting the set is identical. Thus, selecting one range in the set as a “base” and subtracting all the others from this base observation will result in a set of pseudo-observations which are not affected by the receiver clock offset. A similar cancelling of a satellite clock offset is seen in the differencing of a set of observations to a single satellite. This is the basis of the concept of differencing in GPS. In this work, this procedure will also be known as processing in “differential GPS” or DGPS mode. Aside from removing unknowns from the solution, differencing also removes errors correlated between observations, as will be shown in Chapter 6.

A practical description of differencing can be given as such: given a set of receivers observing to a constellation of commonly-visible satellites, one satellite and one receiver are chosen as the “bases.” Then the observations made at each receiver are subtracted from observation at that receiver to the base satellite. These *single-differenced* pseudo-observations are then subtracted from their corresponding single-differences calculated at the base receiver, to create *double-differences*. For a set of observations collected at three receivers observing four satellites, the corresponding differencing matrix,  $\tilde{\mathbf{N}}\mathbf{D}$ , is given by

$$\nabla\Delta = \begin{bmatrix} \mathbf{1} & -\mathbf{1} & 0 & 0 & -\mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & -\mathbf{1} & 0 & -\mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & 0 & -\mathbf{1} & -\mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ \mathbf{1} & -\mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & -\mathbf{1} & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & -\mathbf{1} & 0 & 0 & 0 & 0 & 0 & -\mathbf{1} & 0 & \mathbf{1} & 0 \\ \mathbf{1} & 0 & 0 & -\mathbf{1} & 0 & 0 & 0 & 0 & -\mathbf{1} & 0 & 0 & \mathbf{1} \end{bmatrix} \quad (3.19a)$$

where the original observations have been ordered as

$$\mathbf{I}^T = [p_{1A} \quad p_{2A} \quad p_{3A} \quad p_{4A} \quad p_{1B} \quad p_{2B} \quad p_{3B} \quad p_{4B} \quad p_{1C} \quad p_{2C} \quad p_{3C} \quad p_{4C}] \quad (3.19b)$$

and the subscripted letters refer to the three receivers and the numerals refer to the satellites. The resulting vector of pseudo-observations (or double-differences),  $\mathbf{l}_{DD}$ , is:

$$\mathbf{l}_{DD} = \nabla\Delta \cdot \mathbf{l} \quad (3.20)$$

In general, the rank of the differencing matrix must be equal to the number of pseudo-observations produced. This implies all of the pseudo-observations must be linearly independent. A simple way to ensure this and generate the maximum number of possible pseudo-observations is to ensure that the differencing matrix always has a form similar to that of Eq. (3.19a); namely, selecting one base receiver and one base satellite for the entire set of observations. In this way, for a set of observations made at  $n$  receivers to  $m$  satellites, the maximum number of resulting double-differences is  $(n-1)(m-1)$ .

Due to the use of the differencing operator to map the observations into a subspace, the corresponding Jacobian matrices,  $\mathbf{A}$  and  $\mathbf{B}$ , must be modified accordingly, and the resulting expressions are simply

$$\mathbf{A}_{DD} = \nabla\Delta \cdot \mathbf{A} \quad (3.21a)$$

$$\mathbf{B}_{DD} = \nabla\Delta \cdot \mathbf{B} \quad (3.21b)$$

The least-squares solution then proceeds according to Eq. (3.5), with the appropriate substitutions.

Double-differencing has significant implications on the efficiency of solving the GPS positioning problem. Section 3.3.2 revealed that the number of clock offset terms to estimate grows linearly with the number of epochs observed. While the number of parameters to estimate grows linearly, the number of computations required to solve them grows cubically (Press et al, 1992), due to the inversion of the normal matrix required. Thus for a three receiver epoch observing an average of 4 satellites for 20 minutes at a rate of 1 observation per 5 seconds, the number of clock terms to estimate is 1680, requiring

significant storage space and processing time to solve. On the other hand, applying the differencing operator to the Jacobian of the clock offsets shown in Figure 3.1, coupled with an appropriate reordering of the unknowns to satisfy the form of Eq. (3.19a), shows that the Jacobian is reduced to a null matrix. Thus the clock offsets become inestimable due to the differencing operator, and as a result are removed from the problem entirely. This significantly improves the situation, as the problem now becomes to invert a 9 x 9 matrix of unknown positions, in this three receiver problem.

A second effect of the differencing operator becomes evident when carrier phases are used, and their associated ambiguities must be solved for. In the case of the three receiver network observing 4 satellites without loss of lock, there are 12 ambiguities to be resolved (assuming single frequency observations), and so the Jacobian matrix of the unknown ambiguities is a 12 x 12 identity matrix. However, once the differencing operator is applied, the Jacobian of the ambiguities becomes identical to Eq. (3.19a), which is rank deficient. This means that the original ambiguities can not be solved for, but that instead  $(n-1).(m-1)$ , or 6, double-differenced ambiguities *can* be solved for. So long as the main quantities of interest are the positions of the receivers, rather than the ambiguities of the observations, this is of little concern. Also, note that the differencing operator, since it is composed of integer elements, does not affect the integer nature of the ambiguities, a fact that will be taken advantage of in Chapter 5. Finally, whereas the original ambiguities were functionally independent, the double-differenced ambiguities are mathematically *correlated*. This has important implications in the ambiguity resolution process and will also be further discussed in chapter 5.

### 3.5 - Convergence and Linearization Issues

Once a solution for the corrections to the unknowns and observations,  $\mathbf{d}$  and  $\mathbf{r}$ , have been obtained, the new estimates of these quantities are calculated as

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{d}_i \quad (3.22a)$$

$$\mathbf{l}_i = \mathbf{l}_0 + \mathbf{r}_i \quad (3.22b)$$

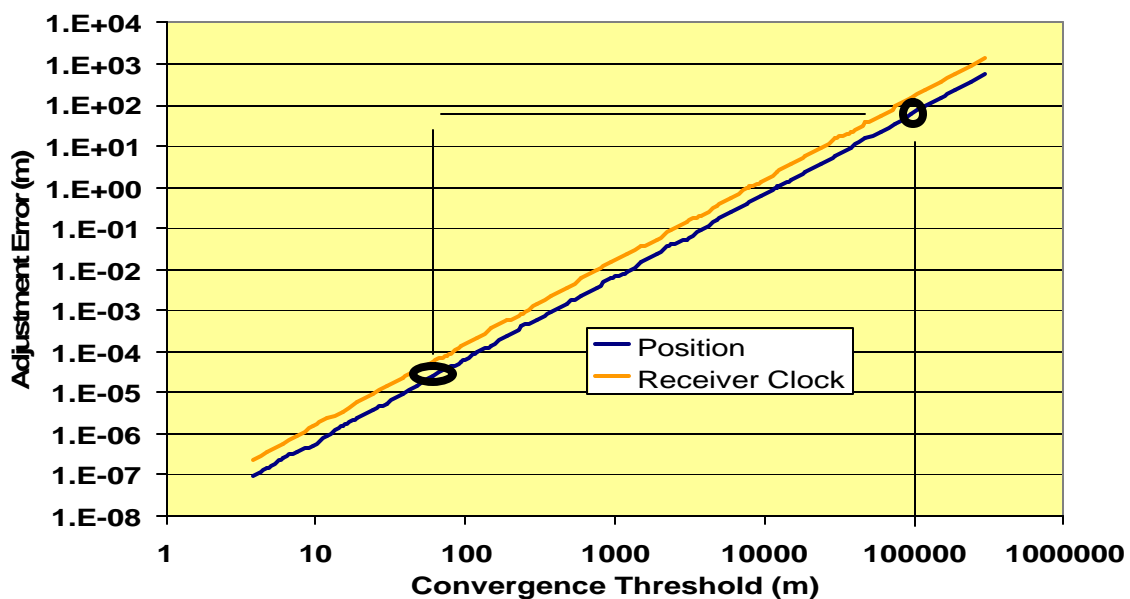
where the subscripts denote the current iteration. Note that the observations are updated differently from the unknowns; this is due to the definition of the residual vector and is discussed thoroughly (along with many other convergence issues) in Pope (1972).

A problem is considered *non-linear* if any of the terms in the Jacobian matrices **A** and **B** depend on the estimates of the unknowns or observations. In this case, the normal vector to the solution surface  $\mathbf{f}(\mathbf{x},\mathbf{l}) = 0$  will change between the estimate  $i$  and the updated estimate  $i+1$ . As a result, the problem is solved by *iteration*, with iteration continuing until the  $\mathbf{l}$  vector reaches some arbitrarily small size, and the Jacobian matrices remain practically constant between estimate and update.

GPS positioning is a particularly interesting problem in that the solution is typically non-linear only due to the position unknowns – the model derivatives with respect to the clock offsets, ambiguities and observations are all constants. In addition, as shown in Section 3.3.1, the derivatives of the mathematical model with respect to the unknown positions produce the direction vector from the satellite to the receiver. Due to the large separation of the satellites from the receivers, it can be inferred that the normal vector is insensitive to updates in the initial estimates, given modestly accurate initial positions. As a result, very few iterations are required to converge to a solution, and no initial estimates at all are required regarding the clock offset and ambiguity unknowns.

Of particular interest is exactly when convergence can be terminated. To determine the effect of convergence thresholds on positioning accuracy, the following test was constructed. Given a typical five satellite geometry and a known receiver position, a set of simulated range measurements was generated. The initial estimate of the receiver position was then chosen with a particular error level corresponding to the convergence threshold tested. For example, to test the effect of terminating the iteration when the corrections are at the metre level, the initial estimates were chosen to be one metre in error from the true position.

The  $d$  vector was calculated using this initial estimate and the updated solution compared to the known receiver position. The resulting error in positioning is the error that could be expected by using that convergence threshold. Figure 3.2 shows the adjustment error that can be expected by terminating the iteration when the  $d$  vector (for positions) is of a given magnitude. During the test, two types of unknowns were solved for – the position of the receiver and the receiver clock offset. Importantly, both the linear and the non-linear unknowns are affected by early termination. Note that the slightly worse accuracy of the receiver clock offset is simply a result of the positioning geometry. It is also crucial to note that the errors shown in Figure 3.2 are *only* errors due to early termination – the ultimate positioning accuracy depends on the quality of the observations themselves.



**Figure 3.2.** Effect of Convergence Threshold on Adjustment Accuracy and Iteration Pathway.

Figure 3.1 also reveals an important aspect of the practical adjustment of GPS observations, namely, that the mathematical model is relatively flat, even with respect to the position unknowns. For example, assume that the initial estimate of a receiver position is known only to 100 km, a very reasonable assumption in most static applications. From Figure 3.1, one can expect that after one iteration, the positioning accuracy will be at the 100m level, neglecting any measurement errors. A second iteration brings the positioning accuracy to



the micron level. Thus, if millimetre-level accuracy is reasonably expected from the measurements and only one iteration is to be used, initial position estimates at the 300 m level are required.

An often overlooked issue is that all the statistical measures provided by least-squares are at their root based on the law of propagation of errors and thus a flatness of the mathematical surface is implicitly assumed by the use of the matrix forms. Of course, this flatness is only true in fact in the immediate neighbourhood of the solution. Furthermore, the solution under constraints shown in Eq. (3.6) is also based on the assumption that the mathematical surface is flat. Fortunately, the results contained in Figure 3.1 indicate that as long as the perturbations in the position estimates are within several hundred metres, these assumptions are valid. For practically all applications this is indeed the case.

A final point to consider is the effect of iteration termination on the linear parameters. It was shown that premature iteration termination of the non-linear terms will cause errors in the linear terms. However, repeating the test described above with the modification that the initial estimates of the *linear* terms are in error shows that only one iteration is required to achieve perfect accuracy of both the linear and non-linear terms, regardless of the size of the initial estimate error. Thus the adjustment of a problem involving linear and non-linear unknowns is insensitive to the accuracy of the initial estimates of the linear terms, which implies that the update of the linear terms is not necessary until the final iteration. This is particularly significant when carrier phase ambiguities are involved, as their magnitude can range from several to millions of cycles and no initial estimate of their value is typically available.

## **4.0 - DATUM ISSUES IN SATELLITE-BASED NAVIGATION SYSTEMS**

A common problem in all positioning problems, regardless of application or accuracy, is that of *datum definition*. In this work, the problem is more precisely referred to as reference-frame definition. Simply put, the reference frame is the spatial framework to which the reported positions refer. As such, a reference frame must have an origin, an orientation, and a scale. In the case of GPS, the default reference frame used is WGS-84. As discussed in Chapter 2, this reference frame is defined by the coordinates of five reference stations located around the world. Note that the actual *datum* is completely defined by further setting the parameters of the WGS-84 ellipsoid and specified value for the gravity potential (Hoffmann-Wellenhof et al, 1994).

User access to this reference frame is provided via the transmission of the satellite coordinates in the ephemeris. Again, this is the same as in the case of terrestrial networks, where the datum is provided to the user typically via the publication of the coordinates of second-order stations in the vicinity of a project. Also, in direct analogy to the terrestrial case, the observations made by receivers in a network are typically of higher accuracy than the coordinates of the observed control points. This section investigates the effects and methods of datum constraint as applied to satellite-based navigation systems and their effects on positioning accuracies and deformation detection.

### ***4.1 - Overview of Traditional Datum Definition Solutions***

In terrestrial engineering networks, the problem of the datum definition has been solved largely through the use of *minimal* and *inner* constraints (Blaha, 1971). Consider the situation of a terrestrial trilateration network consisting of a set of stations which observe distances to a separate set of targets with published coordinates. The set of observations alone creates a rigid geometric figure that defines the spatial relationships between the targets and the stations. The rigidity of the figure is defined solely by the geometry and the

accuracy of the distance measurements (Falkenberg and Schnadelbach, 1987). Furthermore, if the published coordinates of the targets are neglected, this figure is free to translate and rotate in space, the scale being defined by the scale of the distance measurements themselves. If the internal deformation of the figure as a whole is the only quantity to be studied, a comparison of the internal geometries can still be undertaken from epoch to epoch from the available data.

However, in general it is desirable to refer the coordinates to some tangible reference frame. The problem arises when the figure created by the high-accuracy observations is mated to the constraints imposed by the lower-accuracy published coordinates. If the coordinates of the target points are simply considered known and held fixed, inaccuracy of the published coordinates will be forced into the residuals for the observations. This will cause masking of the high observational accuracy and large residuals may be flagged where none actually exist. In addition, any assessments of the network accuracy will become unreliable.

Several solutions exist to this problem. Firstly, the most accurate point can be chosen to establish the location of the network in space, with two pieces of rotation information used to define its orientation. This is known as the implementation of a *minimal constraint*. Obviously, one short coming of this method is that it does not use the published values of the other points as anything more than a check and it may be difficult to gauge exactly which target point is the most accurate. In addition, the accuracy of the network points can be expected to degrade with distance from the fixed point due to the accumulation of errors (Vanicek and Krakowsky, 1986).

A second possibility is to use all of the coordinate information and constrain the network such that its centre of mass does not move and that the overall rotation of the network is zero. This is known as applying *inner constraints* and is discussed thoroughly by Biacs (1989), Koch, (1988), and Blaha (1971), among others. Advantages of this method include that the accuracy of the network becomes more homogenous and that the solution guarantees that the sum-of-squares of the position displacements (from the published

coordinates) is a minimum. For this reason it is known as a minimum-trace solution. Note that in both these cases the target point coordinates are now treated as unknowns to be solved.

Finally, if apriori information regarding the coordinate accuracies is available, then the solution can be augmented by using the published coordinates as observations. If the accuracies of the published coordinates are available, this method is the most desirable since it uses the maximum information available. However, if this information is incorrect, biases in the estimates may result due to improper weighting of the “observed” initial positions (Biacs, 1989).

#### *4.1.1 - Problems with Moving Control Points*

In the terrestrial trilateration network just considered, the total number of unknown stations to be solved for equalled the number of observing stations added to the number of target stations. The problem with satellite-based systems lies in the fact that the available control points are continuously moving. As a result, to rigorously implement any of the datum definition systems previously described, the coordinates of the satellites *at each epoch* would have to be added as unknowns with apriori estimates. Obviously this would cause serious computational burden even if short time spans were considered. For example, if ten satellites were observed for 20 minutes at a sample rate of one sample per 30 seconds, the resulting number of position unknowns would equal 6000, not including the position of the receivers, or any additional parameters to be solved.

A possible solution would be to model each satellite’s motion over time as a Keplerian arc and solve for a reduced set of observations describing key parameters of the arc (Parrot, 1989). This is a concept known as *orbit fixing* and essentially equivalent to deriving a new ephemeris applicable to the particular data set collected. However, introduction of a dynamic model significantly complicates the resulting mathematical models and, due to the short time spans typically encountered, solution instabilities may arise. As a result, this method is not typically used in engineering networks but does find applicability in global scale networks.

Unfortunately, the most convenient method of dealing with the datum problem in satellite-based ranging systems is to assume the coordinates of the satellites as absolute and thus *overconstrain* the network. As a result, it becomes imperative to study the effects of fixing the satellite coordinates on the resulting solution for the receiver positions.

#### 4.2 -Effects of Datum Overconstraint on Positioning Accuracies

In cases where the datum cannot be properly defined, but rather over constrained, it is possible to assess the effects of the network overconstraint via simulation. For the case of a GPS network, given the approximate locations of the receivers and a suitable ephemeris, the network geometry can be simulated by calculating the theoretic ranges from the receivers to the satellites. Furthermore, the Jacobian with respect to the unknown receiver positions can be generated for all the simulated observations using Eq. (3.11). In Section 3.3 it was noted that the Jacobian with respect to the observed ranges was the identity matrix, as the only observations are the ranges themselves. However, in the current case, not only will the ranges be considered observed, but the coordinates of the observed satellites will also be considered as observations, with an associated accuracy. Thus, the resulting **B** matrix for a particular range from a satellite to a receiver takes the form :

$$\mathbf{B} = \begin{bmatrix} 1 & \frac{(x_r - x_s)}{d_{rs}} & \frac{(y_r - y_s)}{d_{rs}} & \frac{(z_r - z_s)}{d_{rs}} \end{bmatrix} \quad (4.1)$$

where  $x_s$ ,  $y_s$  and  $z_s$  refer to the coordinates of the satellite,  $x_r$ ,  $y_r$ ,  $z_r$  refer to the coordinates to the receiver,  $d_{rs}$  is the range from the satellite to the receiver. The standard deviation of the range itself is constrained to one millimetre, whereas the standard deviation of the satellite coordinates is assumed to be 6m, based on the values presented in Section 2.4 and an isotropic assumption.

Once the **A** and **B** matrices for the entire set of simulated observations is generated, the resulting estimated accuracy of the receiver positions is given by Eq. (3.8), which is rewritten below including Eq. (3.5) :

$$\mathbf{C}_{x^*} = \left( \mathbf{A}^T \cdot (\mathbf{B} \cdot \mathbf{C}_1 \cdot \mathbf{B}^T)^{-1} \cdot \mathbf{A} \right)^{-1} \quad (4.2)$$

Since the standard deviations of the observed ranges are significantly lower than those of the observed satellite positions, the resulting receiver position accuracies are due to the overconstraint of the network.

Such a simulation was conducted on a four receiver network observing a single epoch. The baseline lengths were varied from 2 to 2000 kilometres and the number of satellites observed from 4 to 7. Figure 4.1 shows the resulting average three dimensional position accuracies for the receiver positions. Perhaps counter to intuition, the accuracy of the position estimates *improves* as the baseline lengths increase. However, it must be realized that the accuracies shown are due solely to the effects of the inaccuracies of the satellite coordinates and that no differencing of the observations has been performed.

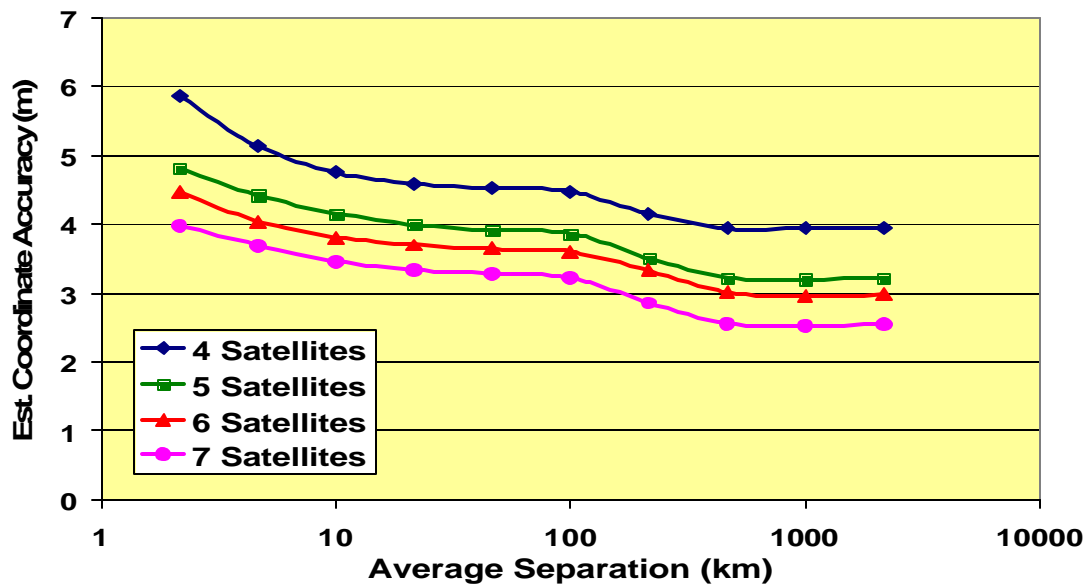


Figure 4.1. Effect of Satellite Coordinate Error on Receiver Position Accuracy.

When the extent of the network is small, all lines of sights from the receivers to the satellites are nearly identical in orientation – i.e. within 2' for baseline separations of 10 km. This implies a very poor intersection geometry at the satellite and so any error in the satellite's position becomes undetectable as it results in a nearly common bias among the affected ranges. Conversely, if the baseline lengths are extended, the angle of intersection between observations to a satellite grows, and so the external precision improves as well. This improvement in the positioning geometry is reflected in the higher positioning accuracy. When additional satellites are added to the solution, the receiver positioning accuracy improves as well due to the averaging out of the datum constraint errors.

The rationale provided for the results of Figure 4.1 implies that the errors resulting from the satellite coordinate errors should be highly correlated between stations when their separation is small. Essentially, *the satellite coordinate errors cause a common shift in all the receiver positions*. Figure 4.2 shows the average correlation coefficient between the position estimates for the receivers in the network described above. As expected, the correlation between the position solutions decreases as the network extent grows, primarily because a given satellite coordinate error causes different ranging errors depending on the orientation of the receiver-satellite line-of-sight vector. Regardless, for all baseline lengths tested, the correlation coefficient does not deviate from 1.00 by more than 1%.

This result is very significant, as it implies that, while the external accuracy of the network is poor due to the relatively low accuracy of the satellite orbits, the precision of the position estimates with respect to one another may be much better. To test this hypothesis, the simulated receiver networks were reprocessed, with the added constraint that the centre of mass of the receiver network was assumed known and fixed. The new position standard deviations of the network points were calculated using Eq. (3.9d), rewritten here for clarity:

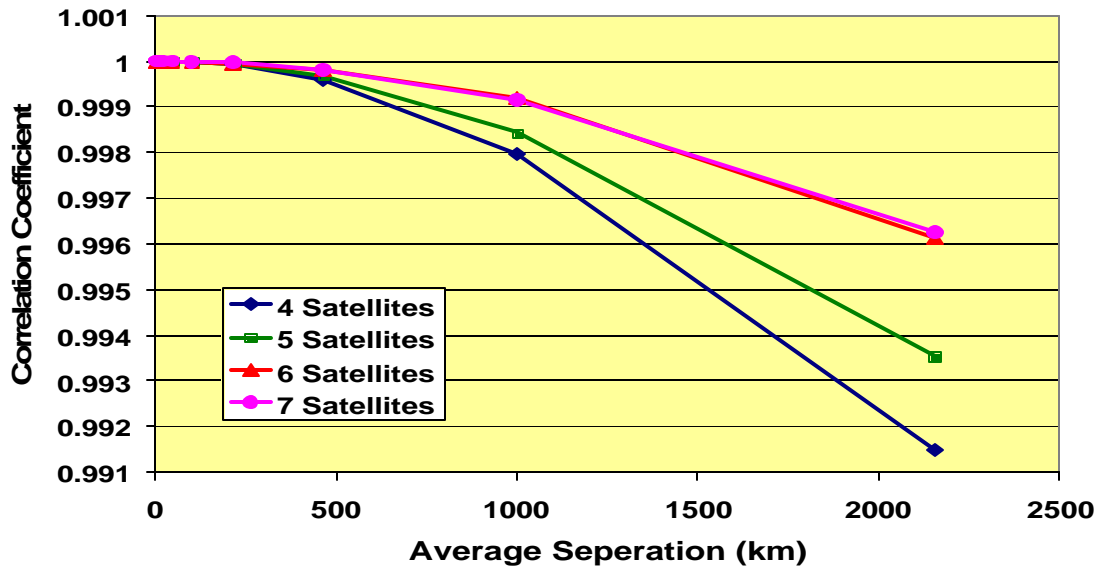


Figure 4.2. Correlation Coefficient of Network Position Solutions.

$$\mathbf{C}_{\mathbf{x}} = \mathbf{C}_{\mathbf{x}^*} - \mathbf{C}_{\mathbf{x}^*} \cdot \mathbf{G} \cdot (\mathbf{G}^T \cdot \mathbf{C}_{\mathbf{x}^*} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \cdot \mathbf{C}_{\mathbf{x}^*} \quad (3.9d)$$

with

$$\mathbf{G}^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 1 & \dots \end{bmatrix} \quad (4.3)$$

The resulting position accuracy estimates are shown in Figure 4.3. The results clearly now show the expected degradation in relative positioning accuracy with baseline separation, as the orbital errors effects decorrelate with distance. Examination of the resulting correlation coefficients shows that the correlation between position estimates is at the 10% level with no appreciable dependence on baseline length. As more satellites are added to the solution, the relative position accuracy improves since the errors in the satellite orbits tend to average out. The same effect occurs when longer time spans are used, as the orbital errors



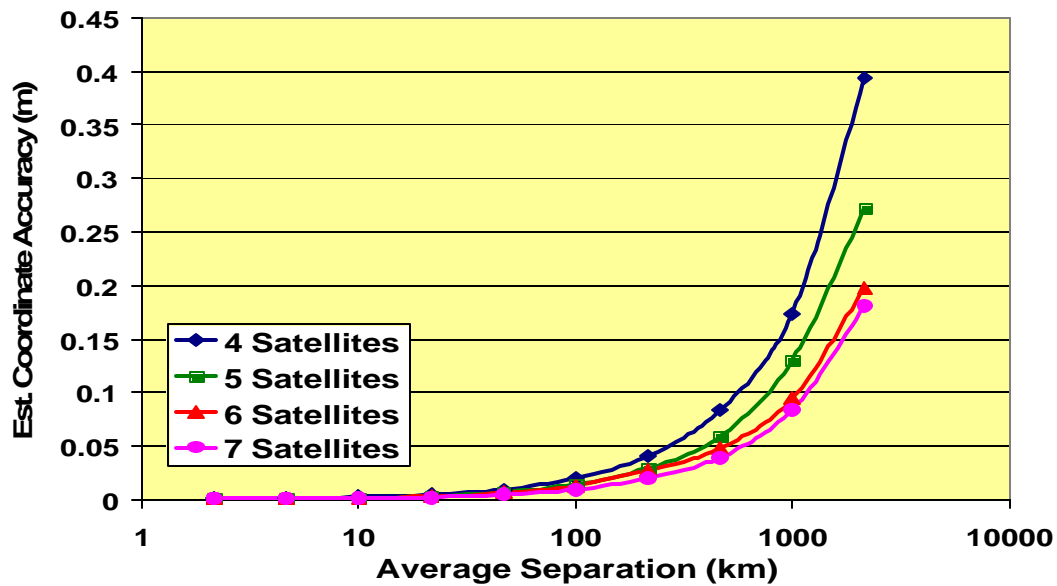


Figure 4.3. Effect of Satellite Coordinate Error on Relative Positioning Accuracy

change over time. Finally, improving the satellite coordinate accuracy directly improves the positioning accuracy, since this is equivalent to scaling the variance-covariance matrix of the observed satellite coordinates.

#### 4.3 -Datum Problem In Differential GPS

In Section 3.4, the concept of differential GPS (DGPS) was introduced and its utility in reducing the number of nuisance parameters to be solved was demonstrated. Furthermore, the previous section revealed that overconstraining the datum via fixing of the satellite coordinates, while not optimal, preserves the internal accuracy of the network to levels much lower than the accuracy of the orbits themselves, and even the absolute accuracy of the network remains better than the orbital accuracy. As a result, a reasonable conclusion is that the process of differencing should not affect the datum problem, as the datum is sufficiently defined prior to the differencing operation. However, this section shows that in practice this is not the case.

### 4.3.1 -Ill-condition of the DGPS Problem

The partial derivatives of a measured range with respect to the unknown parameters were presented in Section 3.3.1 and it was shown that the triad of derivatives formed constituted the normal vector from the satellite to the receiver involved. By reparameterizing the position quantities to refer to the local-level frame and making appropriate substitutions, the Jacobian of the measured range can be rewritten as

$$\begin{bmatrix} \frac{\partial p}{\partial n} & \frac{\partial p}{\partial e} & \frac{\partial p}{\partial h} \end{bmatrix} = -[\cos \mathbf{a} \cdot \cos \mathbf{e} \quad \sin \mathbf{a} \cdot \cos \mathbf{e} \quad \sin \mathbf{e}] = \mathbf{v} \quad (4.4)$$

where  $n, e, h$  refer to the northing, easting and height of the receiver coordinates in a local level frame centred at the receiver.  $\mathbf{a}$  and  $\mathbf{e}$  refer to the azimuth and elevation angle of the receiver-satellite line of sight, where the azimuth is measured clockwise from geodetic north and the elevation angle is measured from the horizon.  $\mathbf{v}$  is the normal vector from the satellite to the receiver.

In the case of three receivers observing two satellites, the resulting Jacobian is given by :

$$\mathbf{A} = \begin{bmatrix} \mathbf{v}_A^1 & 0 & 0 \\ \mathbf{v}_A^2 & 0 & 0 \\ 0 & \mathbf{v}_B^1 & 0 \\ 0 & \mathbf{v}_B^2 & 0 \\ 0 & 0 & \mathbf{v}_C^1 \\ 0 & 0 & \mathbf{v}_C^2 \end{bmatrix} \quad (4.5)$$

where  $\mathbf{v}_j^i$  indicates the normal vector between satellite  $i$  and receiver  $j$ . Upon applying the double differencing operator,  $\tilde{\mathbf{N}}\mathbf{D}$ , using satellite 1 and receiver A as the bases, the double-differenced Jacobian,  $\mathbf{A}_{DD}$ , becomes :

$$\mathbf{A}_{DD} = \begin{bmatrix} \mathbf{v}_A^2 - \mathbf{v}_A^1 & \mathbf{v}_B^1 - \mathbf{v}_B^2 & \mathbf{0} \\ \mathbf{v}_A^2 - \mathbf{v}_A^1 & \mathbf{0} & \mathbf{v}_C^1 - \mathbf{v}_C^2 \end{bmatrix} \quad (4.6)$$

Since the value of the normal vector is relatively insensitive to the position of the receiver position due to the great separation between the receivers and satellites, the values of the non-zero elements in columns will be very similar. For example, if the three receivers are separated by 100 kilometres, a typical example of the  $\mathbf{A}_{DD}$  matrix is

$$A_{DD} = \begin{bmatrix} -1.6960 & 0.8795 & 0.3765 & 1.6959 & -0.8813 & -0.3767 & 0 & 0 & 0 \\ -1.6960 & 0.8795 & 0.3765 & 0 & 0 & 0 & 1.6972 & -0.8789 & -0.3768 \end{bmatrix}$$

This result indicates that the columns of the  $\mathbf{A}_{DD}$  matrix are nearly *linearly dependent*. This has serious consequences for the numerical stability of the least-squares solution, since the resulting  $\mathbf{N}$  matrix will be nearly rank-deficient. As the baseline separation decreases, the columns of  $\mathbf{A}$  become more linearly dependant until the  $\mathbf{N}$  matrix becomes singular to the working precision of the computer. This is known as an *ill-conditioned* system.

Calculation of the *condition number* is a useful tool in studying how ill-conditioned a system is. The condition number is the ratio of the largest to the smallest singular values of a matrix. The singular values, in turn, are the result of the decomposition of a matrix  $\mathbf{Q}$  into the orthogonal matrices  $\mathbf{U}$  and  $\mathbf{V}$ , and the diagonal matrix  $\mathbf{W}$  such that

$$\mathbf{Q} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T \quad (4.7a)$$

where

$$\mathbf{U}^T \cdot \mathbf{U} = \mathbf{V} \cdot \mathbf{V}^T = \mathbf{I} \quad (4.7b)$$

with  $\mathbf{I}$  as the identity matrix. The singular values are the elements of the diagonal matrix  $\mathbf{W}$ . The determination of the matrices  $\mathbf{U}$ ,  $\mathbf{V}$  and  $\mathbf{W}$  is beyond the scope of this thesis, but

many routines are documented in the literature (Press et al, 1992). The inverse of  $\mathbf{Q}$  (if it exists) can be calculated via

$$\mathbf{Q}^{-1} = \mathbf{V} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}^T \quad (4.8)$$

If any of the values in the diagonal matrix  $\mathbf{W}$  are very close to zero, then the inverse of the matrix  $\mathbf{Q}$  will be dominated by error in the calculating the reciprocal of the very small number. In the limit, if any of the singular values is zero, then the matrix  $\mathbf{Q}$  is singular in fact. Thus the condition number gives a good estimate of how noisy the inverse of the matrix can be expected to be.

The end result of ill-conditioning in a system is that the solution to the system become *unstable*. For example, in the overdetermined system  $\mathbf{A}\mathbf{x} = \mathbf{b}$  it is expected that small changes in  $\mathbf{b}$  result in small changes in  $\mathbf{x}$ . However, if the system is ill-conditioned, then the matrix inverse  $(\mathbf{A}^T \cdot \mathbf{A})$ , will contain very large, imprecise values. Consequently, a small perturbation in the  $\mathbf{b}$  vector will be amplified and cause large, incorrect variation in the  $\mathbf{x}$  vector. Obviously, this is an undesirable feature, as the errors in  $\mathbf{x}$  are not due to errors in  $\mathbf{b}$ , but rather due to the limited precision in calculating the large values of the required inverse.

Figure 4.4 illustrates results for a network of receivers of varying numbers and separation. The condition number of the normal matrix for each tested network configuration was calculated and plotted against its average extent. As the receiver separation increases, the condition number improves. This is expected, as increased receiver separations cause larger differences in the line-of-sight vectors between a particular satellite and the various receivers. Thus the columns are less linearly dependant and the condition numbers improve. A similar situation occurs as the number of receivers used improves since adding receivers implies that it becomes more difficult to *exactly* express one column as the sums of the others. Note that the ill-conditioning is a *numerical* problem, and not a datum problem, since the problem lies with the limited precision of the computer used.

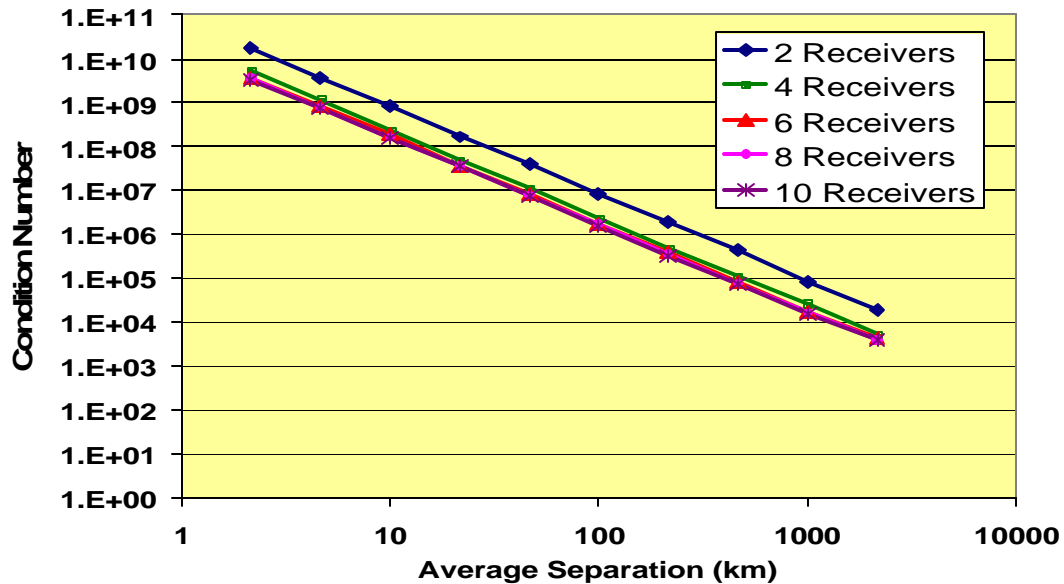


Figure 4.4. Effect of Network Size on Condition Number

#### 4.3.2 - Datum Visibility in DGPS

A different effect that also causes problems when double-differencing is applied concerns the *visibility* of the datum constraints. In the undifferenced case, it was shown that the effects of satellite coordinate inaccuracy increased as the baseline separation decreased due to the inability of a smaller network to detect errors in the satellite coordinates. This behaviour is amplified when double-differencing is applied, since as the receiver separations decrease, the lines-of-sight between the receivers and satellites become more coincident and any satellite coordinate errors increasingly cancel. Thus the translation constraints implied by the fixing of the satellite coordinates become progressively weaker and absolute positioning accuracy degrades.

To gauge the effects of the satellite orbit error on positioning accuracy in a double-differenced mode, the simulation performed in Section 4.2 was repeated, but the **A** and **B** Jacobian matrices were modified to account for the double differencing. Figure 4.5 shows the resulting absolute positioning accuracies assuming a 6 metre orbit standard deviation for a range of receiver network extents all containing 5 receivers and 6 satellites. The accuracies resulting from processing in an undifferenced mode are shown for comparison.

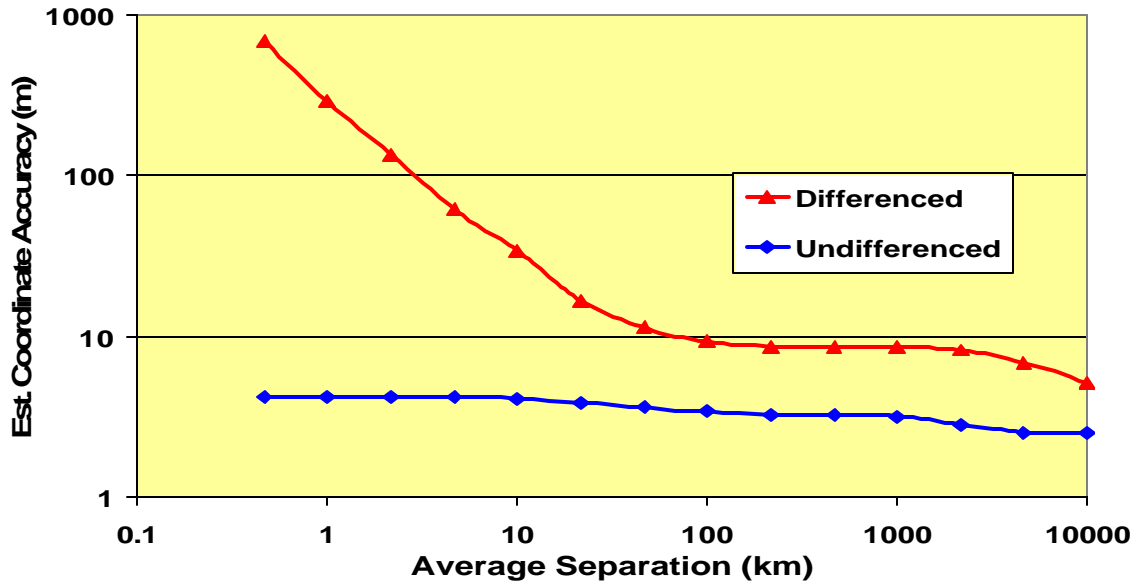


Figure 4.5. Effect of Satellite Coordinate Error on Receiver Position Accuracy.

Whereas the absolute positioning accuracy of the undifferenced mode converges to a limit as the network extent is reduced, the position variances in the differenced mode continue to grow unbounded as the effective datum defined by the fixed satellite coordinates becomes increasingly deficient. Note that the orientation of the network and the internal precision are still maintained at the levels shown in Figure 4.3.

#### 4.3.3 -Regularization of the DGPS Problem

Regularization refers to methods traditionally used to alleviate problems with ill-conditioning. It also implicitly resolves the problem of poorly visible datum constraints. The root of all regularization methods is to essentially add constraints on the possible parameter solutions that limit the amplification of noise, while still maintaining sensitivity to the data contained within the observations (i.e. reducing biases due to a priori assumptions). The desired effect can be expressed as a minimization of the function :

$$\mathbf{g} = \mathbf{d}^T \cdot \mathbf{P} \cdot \mathbf{d} + \mathbf{r}^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r} \quad (4.9)$$

subject to the conformity of  $\mathbf{d}$  and  $\mathbf{r}$  to the mathematical models used, where  $\mathbf{d}$  is the vector of corrections to the initial estimates,  $\mathbf{r}$  is the vector of corrections to the observations (residuals),  $\mathbf{C}_1$  is the variance-covariance matrix of the observations and  $\mathbf{P}$  is known as the regularization matrix.

Note that if the matrix  $\mathbf{P}^{-1}$  is chosen to be  $\mathbf{C}_x$ , the problem reduces to that of the least-squares problem with apriori information on the parameters, the solution to which was presented in Section 3.1. However, the problem remains as to what approach is appropriate when no apriori information is available, as in the case of an initial epoch of a deformation monitoring network.

It is possible to describe the true positions of a set of receivers by a vector  $\mathbf{x}_T$ . As a result, the true ranges between these receivers and observed satellites can be denoted by the vector  $\mathbf{l}_T$  given by

$$\mathbf{l}_T = \mathbf{A} \cdot \mathbf{x}_T \quad (4.10)$$

The actual observations,  $\mathbf{l}_o$ , are contaminated by a noise vector  $\mathbf{e}$ , which will be assumed to have zero mean and known stochastic properties as contained within the variance-covariance matrix  $\mathbf{C}_1$ . Furthermore, denote a vector  $\mathbf{x}_o$ , as the vector of initial estimates of the positions of the receivers, sufficiently close to  $\mathbf{x}_T$  that linearity considerations can be ignored. As a result, the misclosure vector,  $\mathbf{w}$ , can be calculated as

$$\mathbf{w} = \mathbf{A} \cdot \mathbf{x}_o - \mathbf{l}_o = \mathbf{A} \cdot (\mathbf{x}_o - \mathbf{x}_T) - \mathbf{e} \quad (4.11)$$

The least-squares solution for the receiver positions is then calculated via Eq. (3.5), resulting in a correction vector which is added to the vector of initial approximates. Assuming the  $\mathbf{B}$  matrix as identity and substituting Eq. (4.11), this results in the expression

$$\begin{aligned}\mathbf{x}_{LS} &= \mathbf{x}_0 + \mathbf{d} = \mathbf{x}_0 - (\mathbf{A}^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot \mathbf{w} \\ &= \mathbf{x}_T + \mathbf{N}^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot \mathbf{e}\end{aligned}\quad (4.12)$$

where  $\mathbf{x}_{LS}$  is the least-squares solution, with an expected value of  $\mathbf{x}_T$  and a variance of  $\mathbf{N}^{-1}$ . Note that the solution is independent of  $\mathbf{x}_0$ , as expected due to the assumption of linearity.

Due to the ill-conditioning of the  $\mathbf{N}$  matrix in the double-differenced positioning mode, direct application of Eq. (4.12) is impossible since the inaccuracies in calculating  $\mathbf{N}^{-1}$  result in amplification of errors in  $\mathbf{e}$ . One solution to this problem is to add a small diagonal matrix to  $\mathbf{N}$  to produce a matrix  $\mathbf{N}_{reg}$  which has stable inversion properties. This is known as the process of *Tikhonov regularization* (Hansen, 1998). The regularized solution is then calculated as

$$\begin{aligned}\mathbf{x}_{reg} &= \mathbf{x}_0 - \mathbf{N}_{reg}^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot [\mathbf{A} \cdot (\mathbf{x}_0 - \mathbf{x}_T) - \mathbf{e}] \\ &= (\mathbf{I} - \mathbf{N}_{reg}^{-1} \cdot \mathbf{N}) \cdot \mathbf{x}_0 + \mathbf{N}_{reg}^{-1} \cdot \mathbf{N} \cdot \mathbf{x}_T + \mathbf{N}_{reg}^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot \mathbf{e}\end{aligned}\quad (4.13a)$$

with

$$\mathbf{N}_{reg} = \mathbf{N} + a^2 \cdot \mathbf{I}\quad (4.13b)$$

where  $a$  is a selected regularization weight and  $\mathbf{I}$  is the identity matrix.

The error in the regularized solution,  $\Delta \mathbf{x}_{reg}$ , is thus given by

$$\Delta \mathbf{x}_{reg} = \mathbf{x}_{reg} - \mathbf{x}_T = (\mathbf{I} - \mathbf{N}_{reg}^{-1} \cdot \mathbf{N}) \cdot (\mathbf{x}_0 - \mathbf{x}_T) + \mathbf{N}_{reg}^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot \mathbf{e}\quad (4.14)$$

which can be separated into a bias component dependant on the initial estimates and a noise component. The variance of the regularization error due to the bias and noise is a



function of the regularization parameter, the observational noise level, and the accuracy of the initial estimates related together through the following equation :

$$\mathbf{C}_{\Delta\mathbf{x}} = (\mathbf{I} - \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N}) \cdot \mathbf{C}_{\mathbf{x}} \cdot (\mathbf{I} - \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N})^T + \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N} \cdot \mathbf{N}_{\text{reg}}^{-1} \quad (4.15)$$

where  $\mathbf{C}_{D\mathbf{x}}$  is the variance-covariance matrix of the regularization error and  $\mathbf{C}_{\mathbf{x}}$  expresses the accuracy of the initial estimates. Note the similarity in form to Eq. (3.10a). As the regularization parameter grows,  $\mathbf{N}_{\text{reg}}$  becomes increasingly different from  $\mathbf{N}$ , and to the bias term begins to dominate. However, if the regularization error is small, then the noise term grows due to the large elements in  $\mathbf{N}^{-1}$ . This behaviour is illustrated in Figure 4.6, which shows the resulting regularization error for various values of  $1/\alpha$  in the case of a 20 kilometre, five receiver network. Only orbital errors are considered and the accuracy of the initial estimates is assumed to be 50 metres. In general, the optimal value for  $1/\alpha$  is near the accuracy of the initial estimates. Disappointingly, in this case the accuracy of the regularized solution is limited to roughly the accuracy of the initial estimates averaged over the network, and not the accuracy of the undifferenced solution, which is at the 4 metre level.

Figure 4.7 shows the minimum error variance for a range of network extents under initial estimate accuracies of 1, 50 and 100 metres. In addition, the accuracies of the undifferenced and unregularized solutions are shown for comparison. The horizontal dashed lines indicate the standard deviation of the initial estimates averaged over the entire network. An interesting behaviour is evident in the regularized solutions. As the network separation decreases, the unregularized solution accuracy degrades, for reasons discussed above. However, the regularized solution accuracy converges to the accuracy of the initial estimates. When the network expands, the differenced solution becomes more stable, and the solution accuracies converge to the accuracy of the undifferenced solution.

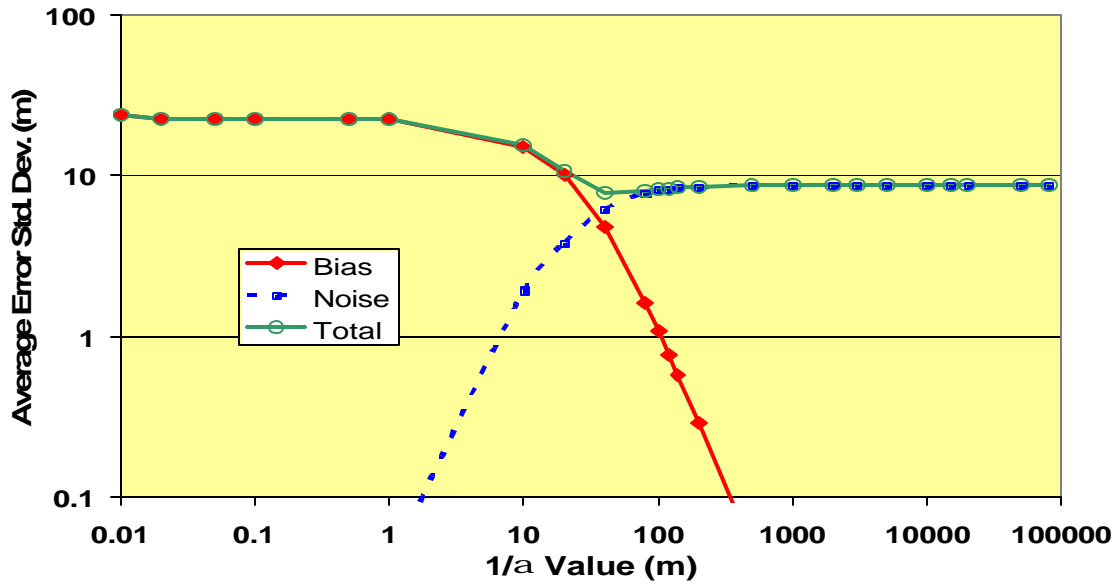


Figure 4.6. Contributions to the Regularization Error due to Noise and Bias for Various  $1/\alpha$ .

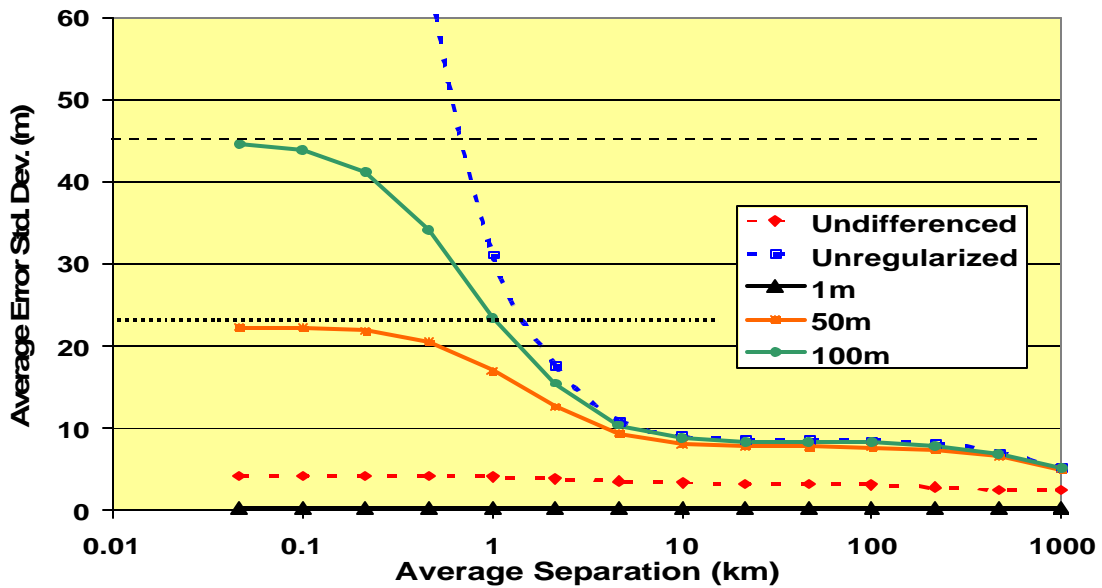


Figure 4.7. Minimum Variances Achievable Under Various Assumptions for Initial Estimate Accuracies.

The reason for this behaviour lies in the fact that for small network extents, the noise of the unregularized solution is very great, and so the solution must rely on the initial estimates as an indication of the absolute location of the network. Thus the solutions are expected to be highly biased. Conversely, when the network extent is greater, the datum as defined by the satellites becomes more visible, and so the reliance on the initial estimates is lessened and the bias shrinks. This behaviour is confirmed by Figure 4.8, which shows the relative contributions to the total error from bias and noise sources for various baseline lengths, under the assumption of initial accuracies accurate to 50 metres. At the smallest receiver separations, the orbital errors largely cancel out, and as a result the noise contribution is very small. As the network expands, the decorrelation of the orbital error causes instabilities in the solution which are regulated by the bias in the initial estimate. Finally, as the geometry improves further, the oscillations due to the orbital noise are damped as well, resulting in error variance behaviour similar to the undifferenced case. Of course, when the accuracy of the initial estimates better than the accuracy of the satellites, the solution essentially fixes the datum deficiency by using the initial estimates, ignoring the satellite overconstraint altogether. Again, it must be stressed that the accuracies discussed are absolute, and that the relative accuracies are limited by the values shown in Figure 4.3.

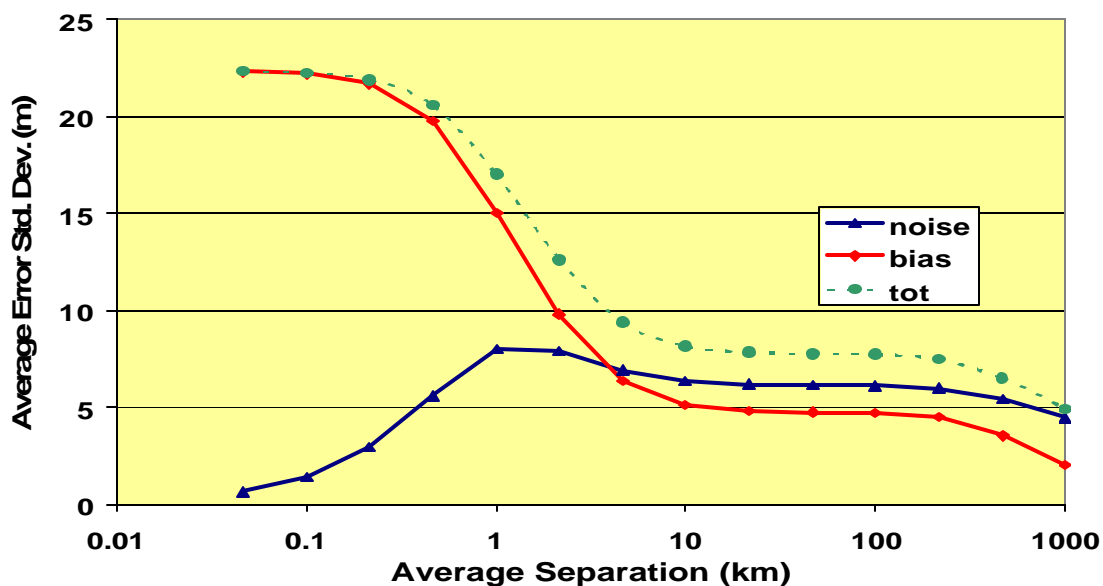


Figure 4.8. Contributions to the Regularization Error due to Noise and Bias at Various Network Extents.

#### 4.3.4 -Deformation Monitoring with DGPS

The results of the previous section indicate that the *absolute* (global) accuracy of DGPS positioning remains at the several metre level, due to the inaccuracy of the satellite orbits acting as control points and the poor visibility of the translation component definition to receivers on the ground. However, it was also shown that the *relative* precision of DGPS is quite good. In deformation monitoring, the quantities of interest are the changes in the position of points in the network. Furthermore, in most applications, the points of interest do not move large distances, usually on the order of several centimetres or decimetres.

As an example, the true positions of a station at two epochs will be defined as  $\mathbf{x}_{T1}$  and  $\mathbf{x}_{T2}$ . In processing the two epochs, it is assumed that the same initial estimate,  $\mathbf{x}_0$ , for the receiver positions have been used. In addition, to simplify the discussion, the observation geometry is assumed to be identical for the two sessions, which is true for observation spans of identical length separated by one sidereal day. In the case that the observation spans are not separated by 23hr 56min, the above assumptions hold if the positioning accuracies and observation error levels are similar, which will be assumed true in the sequel. Using the expression for the regularized position estimates, the resulting regularized estimates for the deformation vector,  $\mathbf{d}_{reg21}$ , is then given by ;

$$\begin{aligned} \mathbf{d}_{reg21} &= \mathbf{x}_{reg2} - \mathbf{x}_{reg1} \\ &= (\mathbf{I} - \mathbf{N}_{reg}^{-1} \cdot \mathbf{N}) \cdot (\mathbf{x}_0 - \mathbf{x}_0) + \mathbf{N}_{reg}^{-1} \cdot \mathbf{N} \cdot (\mathbf{x}_{T2} - \mathbf{x}_{T1}) + \mathbf{N}_{reg}^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot (\mathbf{e}_2 - \mathbf{e}_1) \end{aligned} \quad (4.16)$$

Since the true deformation vector is  $\mathbf{x}_{T2} - \mathbf{x}_{T1}$ , the error in the deformation estimate is

$$\Delta \mathbf{d}_{reg21} = (\mathbf{I} - \mathbf{N}_{reg}^{-1} \cdot \mathbf{N}) \cdot (\mathbf{x}_{T2} - \mathbf{x}_{T1}) - \mathbf{N}_{reg}^{-1} \cdot \mathbf{A}^T \mathbf{C}_1^{-1} \cdot (\mathbf{e}_2 - \mathbf{e}_1) \quad (4.17)$$

which does not depend on the biases of the regularized solutions, but is biased now by the magnitude of the deformation to be measured. The error variance-covariance of the regularized deformation vector is

$$\mathbf{C}_{\Delta d} = (\mathbf{I} - \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N}) \cdot \mathbf{C}_d \cdot (\mathbf{I} - \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N})^T + 2 \cdot (1 - c) \cdot \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N} \cdot \mathbf{N}_{\text{reg}}^{-1} \quad (4.18)$$

where  $\mathbf{C}_d$  is the variance-covariance matrix of the expected deformations and  $c$  is a constant between 0 and 1 expressing the percentage of correlation between the observational errors between the two epochs.

For deformation monitoring to be effective, the error variance of the regularized deformation vector must be smaller than the variance of the deformation itself. For example, if the solution is not regularized at all, the deformation estimate becomes unbiased, but excessively noisy, making it a useless measure. However, if the solution is over-regularized, then the error of the deformation estimate is at the same level as the magnitude of the deformation itself. Figure 4.9 shows the error standard deviation of the regularized deformation estimate averaged over a five receiver network with a 20 km extent under various values of  $1/\alpha$ . A deformation level of 1 cm was assumed as well as a 6 m satellite orbit error with no correlation between observation spans. The behaviour is very similar to that of the absolute case shown in Figure 4.6.

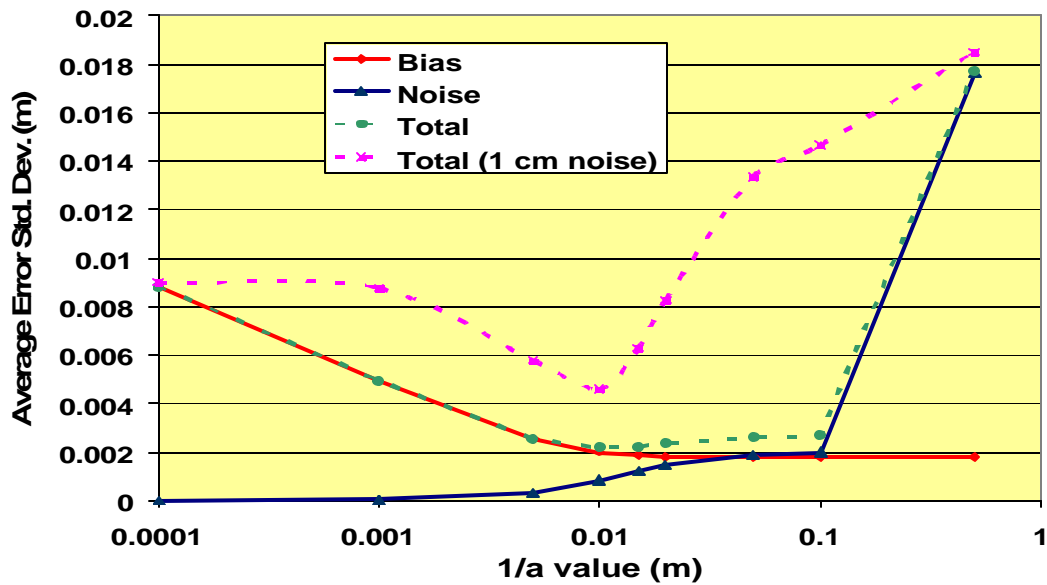


Figure 4.9. Contributions to the Detection Regularization Error due to Noise and Bias for Various  $1/\alpha$ .

However, unlike the case of absolute positioning, where the value of  $1/\alpha$  is typically around the value of the initial estimate accuracy due to its very large variance, in the deformation monitoring mode, the value of  $1/\alpha$  depends closely on the ratio of the observational variance and the expected deformation magnitude. For example, Figure 4.9 also shows the regularization error when a white noise with a standard deviation of 1 cm is added to the system. Fortunately, given proper stochastic modelling of the observations, it is possible to determine the optimal value of  $1/\alpha$  via application of Eq. (4.18).

Figure 4.10 shows the maximum deformation monitoring accuracy achievable under various baseline lengths. The results assume a deformation levels of 5 mm, 1 cm and 2 cm and considers the effects of the satellite coordinate error and a 1 cm noise level. When compared to the absolute accuracies achievable as shown in Figure 4.7, the accuracy of the deformation detection is striking. As expected, the detection sensitivity decreases as the network separation increases, due to the increased decorrelation of the orbital errors. Also, the sensitivity of the deformation detection decreases as the deformation magnitudes increase, due to biasing of the solution. However, at all distances analysed, the sensitivity is still below the magnitude of the deformations expected.

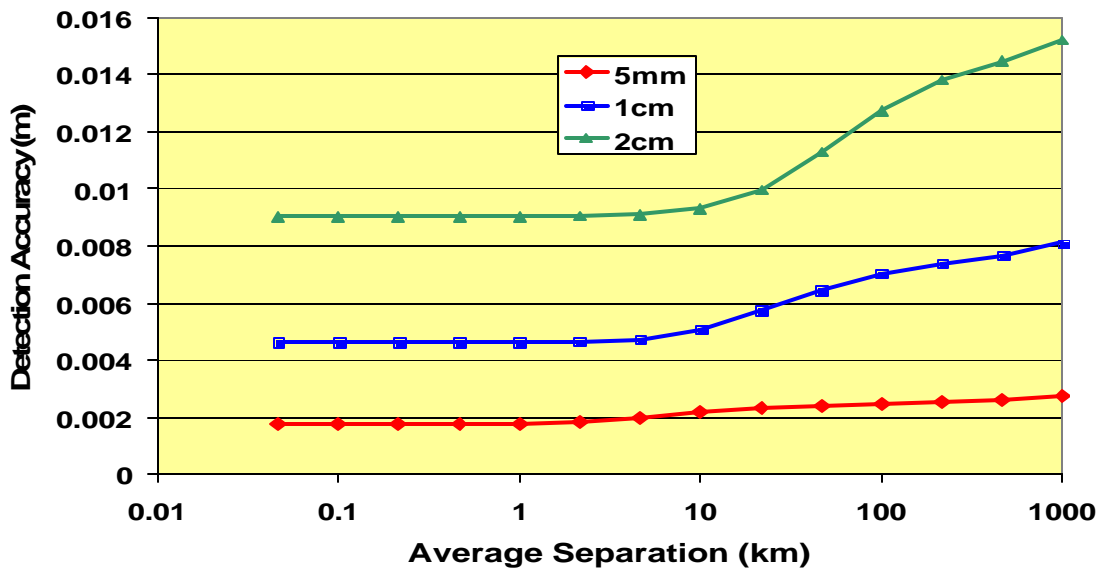


Figure 4.10. Deformation Detection Accuracy.

Under actual conditions, the error levels would depend on the actual errors present and network extents, as discussed in Chapter 6, and the length of time the network has been observed. Thus the preceding discussion is only intended to study the principles of datum definition in satellite-based networks. An investigation of the actual accuracies, both absolute and relative, is presented in Chapter 7. Also, the study so far has limited itself to the case where only position unknowns are considered. In reality, to achieve the highest levels of positioning accuracy, it is necessary to use the carrier phase measurement. The complications involved due to the added ambiguity unknowns are discussed further in Chapter 5, including their effects on the datum resolution problem.

## **5.0 - AMBIGUITY RESOLUTION FOR PRECISE POSITIONING**

Two types of measurements are typically used in satellite-based navigation systems – code (or pseudorange) and carrier phase. As discussed in Chapter 2, the code measurement is simply a time-of-flight measurement made using a ranging code superimposed on a carrier signal. The carrier phase is essentially a measurement of the integrated Doppler shift of the carrier signal itself. As a result, it is corrupted by an unknown integer ambiguity due to the inability to sense the number of cycles between the receiver and satellite at the moment of signal acquisition. This ambiguity remains constant as long as lock is maintained to the satellite.

Despite the added unknowns incorporated by utilising the carrier phase, this measurement is crucial for precise positioning applications. This is because the process of measuring the carrier phase is inherently more accurate than the code measurement process. In fact, the carrier phase is often two orders of magnitude more accurate, as will be discussed in Chapter 6. As a result, for high precision applications, use of the carrier phases is required and a methodology for dealing with the added ambiguity unknowns must be established. This chapter investigates these issues, including the effects of datum constraint on the ambiguity resolution process and changing satellite geometry, and reviews the LAMBDA method of resolving the integer nature of these ambiguities.

### ***5.1 - Float Ambiguity Resolution***

In Chapter 3, the observations were shown to be linear with respect to the unknown ambiguity terms. Given a set of  $m$  satellites observed at  $n$  receivers, the total number of observations is  $mn$ , which is also equal to the number of unknown ambiguities. If positions are to be solved for as well, it becomes evident that the problem is underdetermined unless additional information is included (i.e. a priori information regarding the receiver positions) or more than one epoch is observed without loss of lock.



A further complication is introduced due to the presence of the unknown satellite and receiver clock offsets. For example, the receiver clock offset affects all ranges measured at the receiver equally, lengthening or shortening them by a quantity  $c \cdot DT$ . However, the unknown ambiguities also directly shorten or lengthen the ranges by a quantity  $N$ . Thus, given measurements made over several epochs, it is impossible to separate the effects of the clock offsets and the ambiguities, since the average clock offset can be absorbed by adding this value to all the ambiguities measured at the receiver. A similar argument can be made when considering the satellite clock offsets. Hence, the system of equations becomes effectively singular and no solution is possible unless modelling of the clock offsets is preformed, which is not possible in typical applications.

Double-differencing provides the solution to this problem. Discussed in Chapter 3, double differencing removes all the clock offsets from the solution. However, an added effect is that the process of double differencing also causes the *entire* set of individual ambiguities to become inestimable. For example, assume that  $nm$  ambiguities are to be solved. The original Jacobian with respect to these ambiguities is an  $nm$  by  $nm$  identity matrix. Denoting the double differencing matrix as  $\tilde{N}D$ , the resulting normal matrix for the ambiguities becomes :

$$\mathbf{N}_{\text{amb}} = \mathbf{I} \cdot \nabla \Delta^T \cdot (\nabla \Delta \cdot \mathbf{C}_1 \cdot \nabla \Delta^T)^{-1} \cdot \nabla \Delta \cdot \mathbf{I} \quad (5.1)$$

where  $\mathbf{I}$  is an identity matrix and  $\mathbf{C}_1$  is the variance-covariance matrix of the observations.  $\mathbf{N}_{\text{amb}}$  is singular since the dimensions of  $\tilde{N}D$  are  $(n-1)(m-1)$  by  $nm$  as per Section 3.4. Fortunately, the ambiguities are only nuisance parameters, and thus *only the double-differenced ambiguities are solved for*.

### 5.1.1 -Regularization and Ambiguity Resolution

The results of Chapter 4 indicated that the datum of the a receiver network was actually over constrained by the fixing of the satellite coordinates, but that additional regularization had to be performed due to the poor visibility of the datum. It was also shown that if

absolute accuracy was desired, that the value of the regularization parameter,  $1/\alpha$ , should be chosen near to the accuracy of the initial estimates. In small-extent networks, the absolute accuracy is limited to the accuracy of the initial estimates, whereas on larger extents, accuracy limit converges to the accuracy of the satellite orbits. However, if deformation monitoring is the end result, the  $1/\alpha$  parameter should be chosen near the magnitude of the expected deformations. This will result in very biased results, but sensitive deformation detection.

In the case when carrier phases have been measured, it is important to consider the effect of the additional unknowns on the regularization scheme. A similar development to that used in Chapter 4 will be followed. Firstly, the vector of parameters will be extended such that

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{\text{pos}} \\ \mathbf{x}_{\text{amb}} \end{bmatrix} \quad (5.2)$$

where  $\mathbf{x}_{\text{pos}}$  refers to the unknown position quantities and  $\mathbf{x}_{\text{amb}}$  refers to the unknown ambiguities. Assume that the accuracy of the initial estimates of the position quantities can be expressed by the variance-covariance matrix  $\mathbf{C}_x$  and that the accuracy of the initial estimates of the ambiguities is undefined. In addition, we assume that the regularization matrix  $\mathbf{P}$  is only applied to the position estimates. The normal matrix for the unknowns can be partitioned into components referring to the positions and ambiguities respectively, namely:

$$\mathbf{N} = \begin{bmatrix} \mathbf{N}_{\text{pos}} & \mathbf{N}_{\text{amb/pos}} \\ \mathbf{N}_{\text{pos/amb}} & \mathbf{N}_{\text{amb}} \end{bmatrix} \quad (5.3)$$

and the regularized normal matrix is therefore given by

$$\mathbf{N}_{\text{reg}} = \begin{bmatrix} \mathbf{N}_{\text{pos}} + a^2 \cdot \mathbf{P} & \mathbf{N}_{\text{amb/pos}} \\ \mathbf{N}_{\text{pos/amb}} & \mathbf{N}_{\text{amb}} \end{bmatrix} \quad (5.4)$$

where the matrix  $\mathbf{P}$  is an arbitrary regularization matrix, chosen to be the identity matrix in the case of Tikhonov regularization and  $\mathbf{N}_{\text{amb}}$  is always invertible.

The total regularization error is finally given by Eq. (4.14)

$$\mathbf{C}_{\Delta x} = (\mathbf{I} - \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N}) \cdot \begin{bmatrix} \mathbf{C}_{x^*} & 0 \\ 0 & 0 \end{bmatrix} \cdot (\mathbf{I} - \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N})^T + \mathbf{N}_{\text{reg}}^{-1} \cdot \mathbf{N} \cdot \mathbf{N}_{\text{reg}}^{-1} \quad (5.5)$$

modified to take into account the fact that the ambiguities are of unknown accuracy.

An experiment was conducted using a 10 receiver network with an average separation of 40 kilometres. A twenty minute segment of data was used to generate the normal matrix corresponding to the 10 unknown receiver positions and the 72 unknown ambiguity terms. The accuracies of the initial estimates of the unknowns were chosen to be 1m, 60m and 200m respectively, and the observational stochastic models followed those presented in Chapter 6, using standard values. The resulting average position accuracies for various values of  $1/\alpha$  are shown in Figure 5.1, along with the noise contribution to the total error. For very small values of  $1/\alpha$ , the noise contribution is slight, but the bias component grows with the degradation in the initial estimate accuracy. If the  $1/\alpha$  value is high, the bias error is small, but noise is amplified, a result already encountered in Chapter 4. Figure 5.2 shows the average ambiguity accuracy for the tested  $1/\alpha$  values as well as the noise contribution to the overall accuracy. Large biases become apparent as the  $1/\alpha$  values decrease and the initial estimate variances increase. The reason for this is due to the fact that the double-differenced ambiguities are an *internal* quantity of the network.

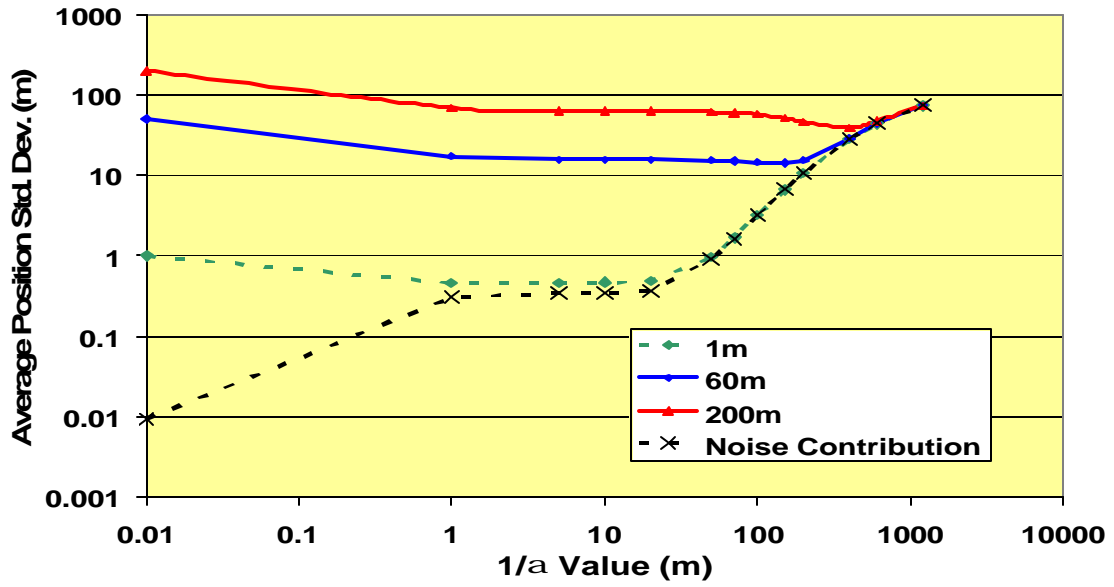


Figure 5.1. Average Position Accuracy for Various Values of  $1/\alpha$  and Initial Estimate Accuracies.

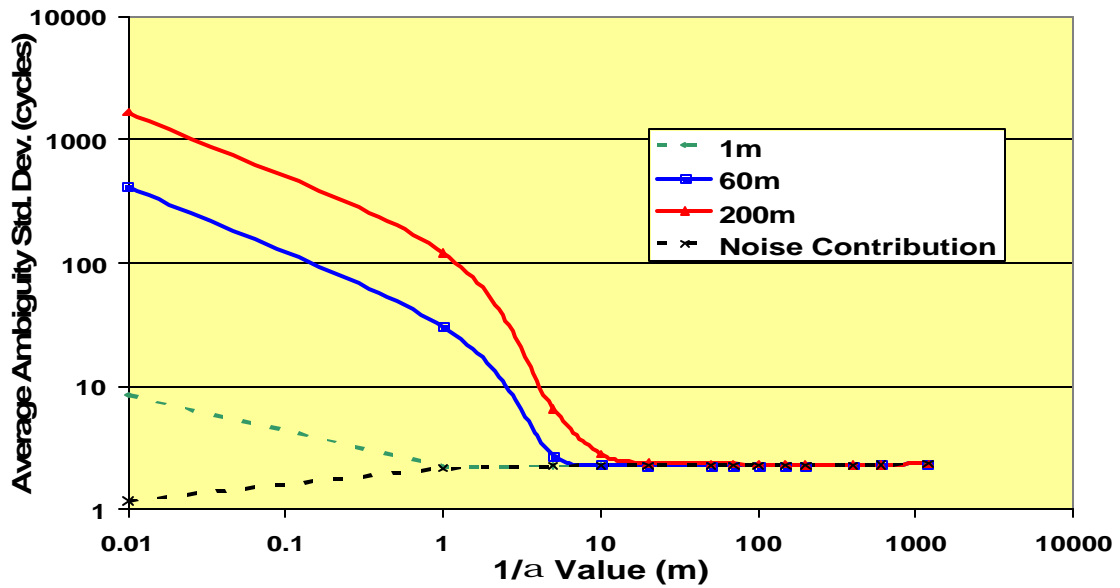


Figure 5.2. Average Ambiguity Accuracy for Various Values of  $1/\alpha$  and Initial Estimate Accuracies.

Due to the geometry of the double differences, the entire network can be translated in space without changing the values of the double differences significantly. However, the double-differences *are* very sensitive to the relative position of the receivers within the network. Thus, if good initial estimates are available, the accuracy of the ambiguity resolution improves, as evidenced by the decreased noise contribution for small values of  $1/\alpha$  when the initial estimates are weighted heavily. In addition, if the initial estimates are very poor and the regularization light, then the ambiguities are not affected by the large biases in the position estimates, since they are resolved using the actual data available – resulting in poor accuracy, but low bias. The worst situation occurs when the initial estimates are poor, and the regularization is heavy. Basically, this corresponds to distorting the internal geometry of the network, and the ambiguity solution suffers as a consequence.

In Chapter 4, it was shown that a heavy regularization is desirable for deformation monitoring since, although the absolute position estimates are heavily biased, the resulting deformation detection vector is very sensitive. Unfortunately, a review of Eq. (4.17) shows that the deformation vector is sensitive to the bias in the difference between the unknowns between the two epochs. When the position unknowns only are considered, this is not a problem, since the order of the deformations is at the centimetre-level. However, when carrier phases are used, the differences between the ambiguities for the two sessions can be millions of cycles and this propagates into the deformation sensitivity for the positions. As a result, it is important to ensure that the ambiguities are as unbiased as possible. This is particularly important if the ambiguities are to be resolved as integers, discussed in Section 5.2.

A consideration of the reason for the ambiguity bias suggests the solution to this problem. The ambiguities become biased when the internal network geometry is deformed due the overweighting of inaccurate a priori position estimates. Furthermore, the need for regularization stems from the poor visibility of the translation component of the datum definition, previously discussed in Chapter 4. Thus the goal is to determine a regularization method that solves the translation problem (i.e. the centre of mass of the network) while

allowing the internal geometry of the network to be solely determined from the observed data. One method of accomplishing this is through an appropriate change in the regularization matrix used.

An identity regularization matrix is equivalent to a weighted observation of each point as its initial estimate. If instead, the centre of mass of the network is considered the “observed” quantity, it can be derived by averaging all the initial estimates for the various points in the network. This can be expressed via the following matrix

$$\mathbf{G}^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 1 & \dots \end{bmatrix} \quad (5.6)$$

The variance of each centre of mass “observation” is equal to the variance of the initial estimates divided by the number of points in the network. Using the above matrix as a regularization matrix results in the following expression

$$\mathbf{N}_{\text{reg}} = \begin{bmatrix} \mathbf{N}_{\text{pos}} + \frac{a^2}{n} \cdot \mathbf{G} \cdot \mathbf{G}^T & \mathbf{N}_{\text{amb/pos}} \\ \mathbf{N}_{\text{pos/amb}} & \mathbf{N}_{\text{amb}} \end{bmatrix} \quad (5.7)$$

where  $n$  is the number of points in the network. Note that this is very similar to the method of inner constraints, with the exception that the constraints are weighted, rather than absolute and the  $\mathbf{N}_{\text{pos}}$  matrix is merely ill-conditioned, rather than mathematically singular.

Figure 5.3 shows the resulting position accuracies when the regularization scheme shown above is implemented. Figure 5.4 illustrates the resulting accuracies of the ambiguities. Interestingly, the position regularization error becomes less dependent on the degree of regularization when the centre of mass regularization technique is used. In addition, further inspection of the resulting errors shows that the predominant error is a translation of the entire network, with very little distortion even at extremely high regularization levels. This

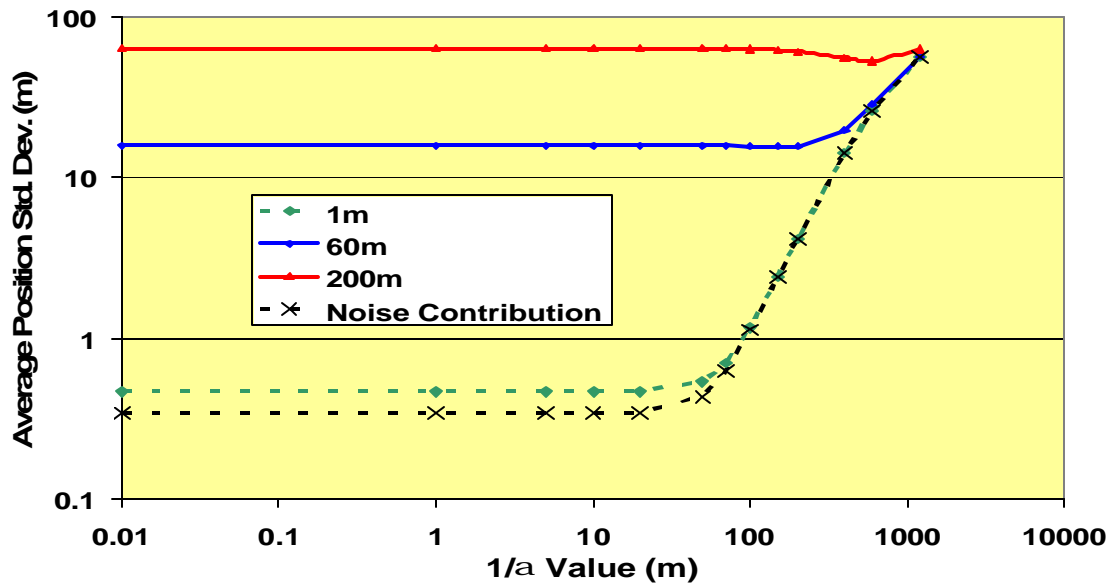


Figure 5.3. Average Position Accuracy for Various Values of  $1/\alpha$  and Initial Estimate Accuracies using Centre of Mass Regularization.

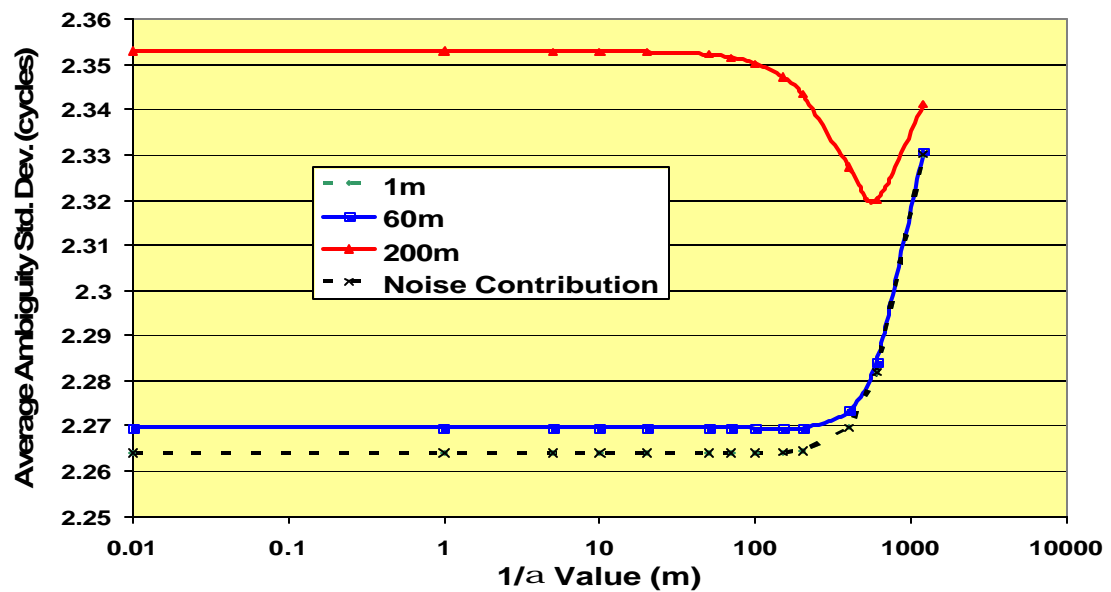


Figure 5.4. Average Ambiguity Accuracy for Various Values of  $1/\alpha$  and Initial Estimate Accuracies using Centre of Mass Regularization.

is because a high regularization level only implies that the centre of mass of the network is rigidly fixed, and that the points are still allowed to move in relation to one another. The higher noise contribution component is thus explained, since the regularization does not attenuate the noise of the observations themselves.

The greatest effect of the centre of mass regularization is in the ambiguity domain. As shown in Figure 5.4, the ambiguity regularization error becomes very insensitive to changes in  $1/\alpha$  (note the change in axis scale) and that the bias is relatively small, even for very poor initial estimates. This is extremely important for deformation monitoring applications, as it implies that the  $1/\alpha$  parameter can be very small, resulting in high deformation detection accuracy, without significantly biasing the ambiguities. However, even if the initial estimates are very good, only constraining the centre of mass does not improve the ambiguity resolution since no information of the internal geometry is added, which is what the double-differenced ambiguities depend on.

Another unfortunate consequence of only constraining the centre of mass of the network is that translations of the entire network from epoch to epoch become impossible to detect. Regardless, the improvement in the relative deformation detection outweighs this shortcoming. In general, if the initial estimates of the receiver positions are very good (i.e. at the centimetre-level), then the first regularization technique should be used as it improves the internal geometry of the network, allowing for improved ambiguity solution. However, if the initial estimates are poor (i.e. at the first epoch of a monitoring campaign) then the second technique should be used to minimize the biases in the ambiguities. Subsequent epochs can then be processed using the a posteriori variance-covariance matrix of the position estimates. This will result in a series of translated solutions, with increasingly precise internal geometry, which can be easily analysed for relative deformations.

### *5.1.2 -Base Satellite Changeovers*

One issue in satellite-based positioning is that the constellation of observed satellites is continually changing. In addition, the basis of double-differencing relies on constant visibility to a single base satellite, since that the ambiguity of that observation is implicitly



contained in all the double-differenced ambiguities formed. Unfortunately, even for short time spans, it may be possible that the base satellite descends out of view, or below a certain threshold elevation. As a result, a new base satellite is chosen and consequently all the ambiguities must be re-evaluated. Since the accuracies of the ambiguities are dependant on the length of time that they have been observed for (Teunissen, 1996), this implies that two observation spans of equal duration may have significant differences in accuracy if a base satellite changeover occurs during one of them.

Fortunately, under certain conditions, it may be possible to prevent the loss in accuracy associated with base satellite changeovers. Given a set of three satellites, the double differenced ambiguities formed can be written as

$$\begin{aligned}\nabla\Delta N_{AB} &= \Delta N_A - \Delta N_B \\ \nabla\Delta N_{CB} &= \Delta N_C - \Delta N_B\end{aligned}\tag{5.8}$$

where  $D$  refers to the single-differences between receivers formed by satellites  $A, B$  and  $C$  and  $B$  is the base satellite. Further assume that at some epoch the base satellite  $B$  falls below some threshold elevation and that satellite  $C$  becomes the new base satellite and that no cycle slips occur during the base satellite changeover. The new double-differenced ambiguities are then given by

$$\begin{aligned}\nabla\Delta N_{AC}^2 &= \Delta N_A - \Delta N_C = \nabla\Delta N_{AB}^1 - \nabla\Delta N_{CB}^1 \\ \nabla\Delta N_{BC}^2 &= \Delta N_B - \Delta N_C = -\nabla\Delta N_{CB}^1\end{aligned}\tag{5.9}$$

where the superscripts refer to a particular observation span before and after the base satellite changeover. Note that if a cycle slip occurs on satellite  $C$ , neither relation holds, and that if a cycle slip occurs on satellite  $A$  or  $B$  only one of the above relations is invalidated – i.e. if satellite  $B$  is not observed after the changeover, it does not affect the relationship written for the  $N_{AC}$  ambiguity.

The implication of the above relations is that for each base satellite changeover, equations of the form of Eq. (5.9) can be written for the entire set of satellite pairs observed after the changeover that did not suffer from cycle slips. These can then be gathered into a set of constraint equations and used to improve the float solution using Eq. (3.6), repeated here for convenience:

$$\mathbf{d}_c = \mathbf{d}_* - \mathbf{C}_{x^*}^{-1} \cdot \mathbf{G} \cdot (\mathbf{G}^T \cdot \mathbf{C}_{x^*}^{-1} \cdot \mathbf{G})^{-1} \cdot (\mathbf{G}^T \cdot \mathbf{d}_* + \mathbf{w}_c) \quad (3.6)$$

where  $\mathbf{d}_*$  is the initial float solution and  $\mathbf{d}_c$  is the solution after applying constraints.  $\mathbf{G}$  is the matrix of constraint equations of the linearized form of Eq. (5.9) consisting of ones and zeros and  $\mathbf{w}_c$  is a null vector.  $\mathbf{C}_{x^*}$  is equal to the variance-covariance matrix of the initial estimated parameters. Note that both the position and ambiguity unknowns will be modified due to the application of the constraint and that the accuracy of both positions and ambiguities will improve, in accordance with Eq. (3.8b). A final effect of the application of the constraints is that the number of independent ambiguities will be reduced by the number of constraints applied, since each constraint defines one ambiguity as the difference of two others.

For example, for a data segment that contains a single base satellite changeover, the average ambiguity variance can be expected to be twice as large as that for a similar segment that contains no changeovers but similar geometries. This is simply due to the fact that twice as many unknowns to be solved. The application of the constraints significantly improves the result, since essentially the number of independent ambiguities to be solved are reduced in half. The effects of applying the base satellite changeover diminish as length of the data segment increases but even in this case the reduction in the number of independent ambiguities is significant for integer ambiguity resolution.

## ***5.2 -Integer Ambiguity Resolution***

The ambiguities affecting the observed carrier phases are integer, and the process of double-differencing maintains this integer nature. Unfortunately, the classic least-squares

solution is incapable of incorporating this information, and as a result the solved ambiguities are real-valued numbers. Not only does this imply that the ambiguities absorb some of the observation residuals into the real-valued component of their solution, but positioning accuracy suffers since valuable information is not added to the solution.

The process of ambiguity resolution seeks to determine the most likely integer values for the ambiguities and then use this information to improve the solution for the position, by constraining ambiguities to their integer values. The methodology implemented herein follows the LAMBDA method pioneered by Teunissen (1993) and only a cursory overview of this non-trivial topic is provided here, with focus limited to peculiarities of ambiguity resolution for high precision applications. For further details of the LAMBDA method the excellent and practical review by De Jonge and Tiberius (1996) is suggested.

### *5.2.1 -Principle of Integer Ambiguity Resolution*

The process of integer ambiguity resolution begins with an initial solution for the position of receivers in a network and the float ambiguity solution, along with their associated variance-covariance matrix. Assume that the true integer values of the ambiguities was known, and denoted by the vector  $\mathbf{w}_{\text{amb}}$ . These known ambiguities can be used to form a constraint equation of the form

$$N_A = w_A \tag{5.10}$$

where  $N_A$  an unknown (but estimated) ambiguity, and  $w_A$  is the known integer value for the ambiguity. In general, the ambiguities do not have to be double differenced, but are assumed to be so in the sequel. An updated solution for the positions and ambiguities can then be calculated as a step-wise addition of constraints, set out in Eq. (3.6). In the case that not all the ambiguities have known integer values, the unknown ambiguity estimates will be improved by the knowledge of the other ambiguities. All the known ambiguities will be fixed to their integer values. The resulting sum-of-squares of errors for the updated solution will be given by Eq. (3.7b), repeated here:

$$\mathbf{r}_c^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_c = \mathbf{r}_*^T \cdot \mathbf{C}_1^{-1} \cdot \mathbf{r}_* + (\mathbf{G}^T \cdot \mathbf{d}_* + \mathbf{w}_c)^T \cdot (\mathbf{G}^T \cdot \mathbf{C}_{x*}^{-1} \cdot \mathbf{G})^{-1} \cdot (\mathbf{G}^T \cdot \mathbf{d}_* + \mathbf{w}_c) \quad (3.7b)$$

where  $\mathbf{r}_c$  and  $\mathbf{r}_*$  are the residuals after and before the application of the constraints,  $\mathbf{G}$  is the constraint matrix formed by the left side of Eq. (5.10),  $\mathbf{d}_*$  is the solution for the positions and the ambiguities prior to the application of the ambiguity constraints,  $\mathbf{w}_c$  is the constraint vector formed by the right side of Eq. (5.10), and  $\mathbf{C}_{x*}$  is the variance-covariance matrix of the position and ambiguity unknowns prior to ambiguity fixing.

The constrained sum-of-squares of residuals is always larger than the non-constrained case. In addition, selection of a different  $\mathbf{w}_c$  vector (i.e. changing what the integer ambiguities are assumed to be) will change the constrained sum-of-squares. Since the goal of least-squares is to minimize the overall sum-of-squares of residuals, the goal of the integer ambiguity process is to determine the integer vector  $\mathbf{w}_c$  that minimizes the value of Eq. (3.7b), or minimizes the change in the original sum-of-squares of residuals. Inspection of the right side of Eq. (3.7c) shows that the rightmost term is minimized when the distance from the vector  $\mathbf{w}_c$  to the float ambiguity solution is smallest, calculated in the norm defined by  $\mathbf{G}^T \mathbf{C}_{x*} \mathbf{G}$ . Once this integer vector is determined, it can be used as a constraint on the ambiguity values to update the estimates of both the positions and all the ambiguities using Eq. (3.6).

In general, the optimal integer ambiguity vector is determined by establishing a *search-region* around the float ambiguity vector, selecting integer ambiguity vectors inside of this search-region and calculating the distances between the float and integer solutions. The size of the search-region must be carefully selected, since an inappropriate selection may exclude the actual minimizing integer vector, whereas too little restriction will result in too many candidates being examined for the process to be efficient. Typically, the error ellipsoid surrounding the float ambiguity estimate is used to define the shape of the search region, and this region is scaled sufficiently to include at least one integer vector. A further difficulty lies in actually selecting integer elements inside of the error ellipsoid. In practice, the semi-major axis of the ellipsoid is used to establish the range of possible integers for

the ambiguities to be determined, which in the case of a highly elongated ellipsoid results in an overly large volume of space to be investigated.

A key assumption in the application of Eq. (3.6) lies in the fact that the integer ambiguity vector  $\mathbf{w}_c$  is *known*. Obviously, since this vector is the quantity to be determined, it is not known. However, as shown by Teunissen (2002), given that the float ambiguities are of sufficient accuracy, the integer vector, once properly selected, can be assumed known with little adverse effect on the resulting statistics. A more serious problem lies in the fact that the minimizing integer vector will be the vector that lies “closest” to the initial float ambiguity vector. As a result, the presence of any biases in the float vector may cause the integer vector to become biased as well. However, since the integer vector is assumed correct, the resulting statistical estimates after applying the integer constraint will be grossly optimistic despite the fact that the resulting position estimates will be also biased. For this reason, it is crucial that the float ambiguity vector remain as unbiased as possible, which required appropriate modelling of observational errors and careful selection of the regularization technique, as discussed previously.

Lastly, if any base satellite changeover constraints have been applied in the determination of the float solution, the integer ambiguity search must be restricted to the independent ambiguities if the constraints are to apply after the resolution process. This creates a situation where some ambiguities are actually partially fixed, since one of the parent ambiguities is fixed and the other is not.

### 5.2.2 -Ambiguity Success Rate

According to Joosten and Tiberius (2000), the lower-bound probability that a set of  $n$  unbiased ambiguities can be resolved as their true integer values is given by

$$ASR = \prod_{i=1}^n \left[ 2 \cdot \Phi \left( \frac{1}{2 \cdot \mathbf{s}_{i|l}} \right) - 1 \right] \quad (5.11a)$$

with

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\mathbf{p}}} e^{-\frac{1}{2}z^2} dz \quad (5.11b)$$

where  $ASR$  is the lower-bound probability that the ambiguity set can be resolved as integer,  $\Phi$  is the standard normal cumulative probability distribution and  $\mathbf{s}_{i|l}$  is the standard deviation of the  $i^{\text{th}}$  ambiguity conditioned on the solution of all the previous ambiguities. The values of  $\mathbf{s}_{i|l}$  are obtained from the diagonal elements of the  $\mathbf{D}$  matrix resulting from the triangular decomposition of the variance-covariance matrix of the ambiguities,  $\mathbf{C}_{\text{amb}}$ , written as

$$\mathbf{C}_{\text{amb}} = \mathbf{L}^T \cdot \mathbf{D} \cdot \mathbf{L} \quad (5.12)$$

For the assumptions inherent in Eq. (3.7b) to hold, this probability must be sufficiently high such that the integer ambiguities can be considered exact. A different view is to consider that *any* ambiguity can be resolved as *some* integer, regardless of its accuracy. However, a high success rate ensures that the ambiguities are being resolved to their correct values, neglecting the effects of biases in the float ambiguity solution. Ambiguity success rates of 99% are usually used for this purpose since an incorrectly resolved ambiguity will cause undetectable biases in the position estimates.

In a typical observation campaign, some ambiguities will be determined to a high accuracy, whereas others, perhaps due to low elevation of the satellite, or great separation of the receivers, will be estimated with a lower accuracy. As a result, it is usually very difficult for the entire set of ambiguities to be resolved as integer to a sufficiently high probability. Traditionally, this would result in no ambiguities being resolved as integers, resulting in a *float solution*. However, a more effective procedure is to order the ambiguities in order of decreasing conditional accuracy (i.e. by inspection of the  $\mathbf{D}$  matrix of the  $\mathbf{L}^T\mathbf{D}\mathbf{L}$  decomposition), and to only select the set of ambiguities that result in an

ambiguity success rate of 99%. These ambiguities are then determined as integer, and their constraint improves both the position solution and the remaining float ambiguities. The main advantage of using this strategy for deformation monitoring is that the accuracy of subsequent epochs is made more consistent, since a single bad satellite or receiver does not cause the entire network to revert to a float mode solution.

### *5.2.3 -Ambiguity Decorrelation*

The main problem with a naïve search for an integer vector  $\mathbf{w}_c$  based from the float ambiguity solution is that the float ambiguities are typically highly correlated. For this reason, the error ellipsoid surrounding the float ambiguity estimate is highly elongated and as a result, the search for the integer ambiguities is very inefficient (Teunissen et al, 1994). In general, the most efficient searching occurs when the ambiguities are uncorrelated, and if the ambiguities are completely decorrelated, the solution for the most likely integer ambiguities simplifies to a rounding of the float estimates to their nearest integer.

Teunissen (1993) shows that it is possible to develop a transformation that transforms the correlated ambiguities into an almost uncorrelated equivalent set, while maintaining the integer nature of these ambiguities. Determination of the integer matrix  $\mathbf{Z}$  that accomplishes this partial decorrelation is the heart of the LAMBDA method, and details on the procedure for determining the  $\mathbf{Z}$  matrix is given by De Jonge and Tiberius (1996).

Application of the  $\mathbf{Z}$  matrix to the original ambiguities results in a transformed set, with an associated variance covariance matrix that has the same volume as the original set. Thus the accuracy of the ambiguities is not changed, but merely their correlation. The integer vector that minimizes the weighted distance to the transformed float ambiguity is then chosen as the most likely integer. Due to the reduced correlation of the transformed ambiguities, the efficiency of this process is greatly improved, often reducing the number of candidates to be tested by several orders of magnitude. The resulting most-likely integer vector in the original system is then solved using the inverse of  $\mathbf{Z}$ , which is always square and invertible. Finally, this integer vector is used as a constraint and the solution for the positions and ambiguities is updated.

Note that in the case that constraints on the ambiguities have been applied, such as those due to base satellite changeovers, only the independent ambiguities are considered. As well, from the subset of independent ambiguities, only the ambiguity set that allows for a sufficiently high ambiguity success rate is estimated.

Advantages of decorrelating the ambiguities goes beyond increasing the efficiency of the integer vector determination. Eq. (5.11a) provides a lower bound on the ambiguity success rate. Teunissen (2000) shows that this lower bound becomes increasingly sharp as the ambiguities become increasingly decorrelated. Although the decorrelating transformation does *not* improve the overall accuracy of the ambiguities to be solved, the decorrelation of the ambiguities raises ambiguity success rate estimate for a given set of ambiguities, which in turn allows more ambiguities to be added to maintain the success rate estimate at a particular level.

### ***5.3 - Summary of the Ambiguity Resolution Process and Precise Position Determination***

This section provides a step-by-step summary of the solution for the positions of receivers in a static network and the ambiguities associated with a particular observation session.

*Step 1. Establish the regularization method to be used.*

If no prior information on the receiver positions is available, use the single point solutions of the receivers to establish the centre-of-mass of the network and utilize a high  $1/\alpha$  value (i.e. 0.001 m). Otherwise, use the variance-covariance matrix of the estimated parameters from the prior epoch and their estimates. This results in the regularized normal equations

$$\mathbf{N}_{\text{reg}} = \begin{bmatrix} \mathbf{N}_{\text{pos}} + a^2 \cdot \mathbf{P} & \mathbf{N}_{\text{amb/pos}} \\ \mathbf{N}_{\text{pos/amb}} & \mathbf{N}_{\text{amb}} \end{bmatrix} \quad (5.4)$$



*Step 2 Determine the regularized float solution of the ambiguities and the receiver positions.*

This results in the quantities  $\mathbf{d}_{\text{reg}}$  and  $\mathbf{C}_{\text{xreg}}$ , calculated using the methods of Chapter 3, with the forms

$$\mathbf{d}_{\text{reg}} = \begin{bmatrix} \mathbf{d}_{\text{pos}} \\ \mathbf{d}_{\text{amb}} \end{bmatrix}, \mathbf{C}_{\text{xreg}} = \begin{bmatrix} \mathbf{C}_{\text{pos}} & \mathbf{C}_{\text{pos/amb}} \\ \mathbf{C}_{\text{amb/pos}} & \mathbf{C}_{\text{amb}} \end{bmatrix} \quad (5.13)$$

where  $\mathbf{d}_{\text{pos}}$  and  $\mathbf{d}_{\text{amb}}$  are the corrections to the initial estimates of the position and ambiguity unknowns and  $\mathbf{C}_{\text{pos}}$ ,  $\mathbf{C}_{\text{amb}}$  are the variance covariance matrices of the estimated positions and ambiguities and  $\mathbf{C}_{\text{pos/amb}}$  is the cross-covariance matrix of the position and ambiguity unknowns.

*Step 3. Apply any base satellite changeover constraints to improve the receiver position and ambiguity estimates.*

The updated quantities are calculated via :

$$\mathbf{d}_{\text{base}} = \mathbf{d}_{\text{reg}} - \mathbf{C}_{\text{xreg}} \cdot \mathbf{G}_{\text{base}} \cdot (\mathbf{G}_{\text{base}}^T \cdot \mathbf{C}_{\text{xreg}} \cdot \mathbf{G}_{\text{base}})^{-1} \cdot (\mathbf{G}_{\text{base}} \cdot \mathbf{d}_{\text{reg}}) \quad (3.9a)$$

$$\mathbf{C}_{\text{base}} = \mathbf{C}_{\text{xreg}} - \mathbf{C}_{\text{xreg}} \cdot \mathbf{G}_{\text{base}} \cdot (\mathbf{G}_{\text{base}}^T \cdot \mathbf{C}_{\text{xreg}} \cdot \mathbf{G}_{\text{base}})^{-1} \cdot \mathbf{G}_{\text{base}}^T \cdot \mathbf{C}_{\text{xreg}} \quad (3.9d)$$

where  $\mathbf{G}_{\text{base}}$  is the matrix consisting of the base satellite changeover constraints. Note that both the position and ambiguity estimates are affected due to their cross-correlation at the float estimation stage.

The constraint matrix,  $\mathbf{G}_{\text{base}}$  can be used to establish a matrix  $\mathbf{Z}_{\text{ind}}$  which collects the remaining independent ambiguities. These will be the ambiguities that have not yet been expressed as the sum of any other ambiguities. Thus the independent ambiguities are collected via :

$$\mathbf{a}_{\text{ind}} = \mathbf{Z}_{\text{ind}} \cdot \mathbf{a}_{\text{tot}} \quad (5.14a)$$

where  $\mathbf{a}_{\text{ind}}$  are the independent ambiguities and  $\mathbf{a}_{\text{tot}}$  are the total ambiguities.

Application of Eq. (5.14) also results in the modified variance-covariance matrix of the ambiguities given by

$$\mathbf{C}_{\text{ind}} = \mathbf{Z}_{\text{ind}} \cdot \mathbf{C}_{\text{amb}} \cdot \mathbf{Z}_{\text{ind}}^T \quad (5.14b)$$

*Step 4. Form the  $\mathbf{L}^T \mathbf{D} \mathbf{L}$  decomposition of the variance-covariance matrix of the independent float ambiguities.*

This decomposition is necessary for the subsequent decorrelation step and inversion of the variance-covariance matrix of the ambiguities. Algorithms for the decomposition are widely available, for example in Press et al (1992).

*Step 5. Decorrelate the independent ambiguities.*

The decorrelation process relies on the determination of the integer decorrelating matrix  $\mathbf{Z}$ , as per de Jonge and Tiberius (1996). The development of the  $\mathbf{Z}$  matrix therein operates on the  $\mathbf{L}^T \mathbf{D} \mathbf{L}$  decomposition and also reorders the elements such that the transformed ambiguities are arranged from lowest to highest accuracy. The decorrelation matrix  $\mathbf{Z}$  creates a transformed ambiguity vector,  $\mathbf{a}_{\text{decor}}$ , and its associated variance covariance matrix, given by

$$\mathbf{a}_{\text{decor}} = \mathbf{Z}_{\text{decor}} \cdot \mathbf{a}_{\text{ind}} \quad (5.15a)$$

$$\mathbf{C}_{\text{decor}} = \mathbf{Z}_{\text{decor}} \cdot \mathbf{C}_{\text{ind}} \cdot \mathbf{Z}_{\text{decor}}^T \rightarrow \mathbf{C}_{\text{decor}} = \mathbf{L}_{\text{decor}}^T \cdot \mathbf{D}_{\text{decor}} \cdot \mathbf{L}_{\text{decor}} \quad (5.15b)$$

*Step 6. Select the ambiguities that allow for an ambiguity success rate of at least 99%.*

This is accomplished by going through the  $\mathbf{D}_{\text{decor}}$  matrix from the most precise to the least precise ambiguity and calculating successive ambiguity success rates until

the rate drops below 99%. A matrix  $\mathbf{Z}_{\text{sel}}$  then selects the ambiguities to be solved as integer.

$$\mathbf{a}_{\text{sel}} = \mathbf{Z}_{\text{sel}} \cdot \mathbf{a}_{\text{decor}} \quad (5.16a)$$

$$\mathbf{C}_{\text{sel}} = \mathbf{Z}_{\text{sel}} \cdot \mathbf{C}_{\text{decor}} \cdot \mathbf{Z}_{\text{sel}}^T \rightarrow \mathbf{L}_{\text{sel}} = \mathbf{Z}_{\text{sel}} \cdot \mathbf{L}_{\text{decor}} \cdot \mathbf{Z}_{\text{sel}}^T, \mathbf{D}_{\text{sel}} = \mathbf{Z}_{\text{sel}} \cdot \mathbf{D}_{\text{decor}} \cdot \mathbf{Z}_{\text{sel}}^T \quad (5.16b)$$

*Step 7. Calculate the inverse of the variance-covariance matrix of the selected ambiguities.*

This is necessary to calculate the distance between the various integer vectors to be tested and the transformed float solution. Since  $\mathbf{L}_{\text{sel}}$  is lower triangular and  $\mathbf{D}_{\text{sel}}$  is diagonal, the inverse is efficiently calculated via

$$\mathbf{C}_{\text{sel}}^{-1} = \mathbf{L}_{\text{sel}}^{-1} \cdot \mathbf{D}_{\text{sel}}^{-1} \cdot \mathbf{L}_{\text{sel}}^{-1 T} \quad (5.17)$$

*Step 8. Select the most likely integer for the selected ambiguities.*

This process relies on calculating the distances from a set of integer vectors to the transformed selected ambiguities and selecting the one that minimises the distance under the transformed variance-covariance matrix for the subset. This results in the integer vector  $\mathbf{w}_c$ .

*Step 9. Solve for the updated positions and ambiguities.*

Setting the solved transformed ambiguities as constraints, the effective constraint matrix as applied to the entire, untransformed set can be written as :

$$\mathbf{w}_c = \mathbf{Z}_{\text{sel}} \cdot \mathbf{Z}_{\text{decor}} \cdot \mathbf{Z}_{\text{ind}} \cdot \mathbf{a}_{\text{tot}} = \mathbf{Z}_{\text{tot}} \cdot \mathbf{a}_{\text{tot}} \quad (5.18)$$

The updated position and ambiguities can then be solved using this constraint and Eq. (3.9), along with the float solution after the application of the base satellite changeover constraint :

$$\mathbf{d}_{\text{int}} = \mathbf{d}_{\text{base}} - \mathbf{C}_{\text{base}} \cdot \mathbf{Z}_{\text{tot}}^T \cdot \left( \mathbf{Z}_{\text{tot}} \cdot \mathbf{C}_{\text{base}} \cdot \mathbf{Z}_{\text{tot}}^T \right)^{-1} \cdot \left( \mathbf{Z}_{\text{tot}} \cdot \mathbf{d}_{\text{base}} - \mathbf{w}_c \right) \quad (5.19a)$$

$$\mathbf{C}_{\text{int}} = \mathbf{C}_{\text{base}} - \mathbf{C}_{\text{int}} \cdot \mathbf{Z}_{\text{tot}}^T \cdot \left( \mathbf{Z}_{\text{tot}} \cdot \mathbf{C}_{\text{base}} \cdot \mathbf{Z}_{\text{tot}}^T \right)^{-1} \cdot \mathbf{Z}_{\text{tot}} \cdot \mathbf{C}_{\text{base}} \quad (5.19b)$$

where the term in parenthesis has already been calculated via Eq. (5.17). Note that if the regularization matrix used is not equal to the apriori variance-covariance matrix of the initial estimates, the final variance-covariance matrix for the unknowns and ambiguities must be calculated using the modified method of Section (3.11a). However, the actual solution procedure does not change.

Note that as a result of this process, not all the ambiguities will be fixed, since only a transformed subset are fixed as integer, thus the procedure is termed a *partial-fix mode*. However, due to the correlation of other ambiguities to the fixed subset, the overall accuracy of the solution improves significantly. Also, since not all the integers are required to be fixed as integer, it can be expected that a few bad satellites will not significantly corrupt the solution, since these will not be selected to be fixed, and as a result, their high variance will be weakly weighted in the overall position solution. This is a desirable property in any scenario where accuracies are desired to be similar from epoch to epoch, such as deformation monitoring.

The application of the base satellite changeover in particular improves the ambiguity resolution procedure by reducing the number of independent ambiguities to be solved and essentially lengthening the observation period of these ambiguities. For example, in a typical 24 hr observation span of a 10 receiver network, 1251 ambiguities were to be solved. However, after the application of the changeover constraints, only 182 remained. Of these 182, none would have been identified as fixable prior to decorrelation. After decorrelation, 132 could be solved at the 99% level. This illustrates the necessity of both tracking the base satellite changeovers and ambiguity decorrelation for successful application of carrier phase measurements to the precise positioning problem.

## **6.0 - ANALYSIS OF ERROR SOURCES IN GLOBAL NAVIGATION SATELLITE SYSTEMS AND STOCHASTIC MODELLING**

At the centre of every least-squares adjustment lies the need to relate the stochastic properties of the observations via the variance-covariance matrix. In the simplest sense, the variance covariance matrix describes the proper weighting of the observations to be used and allows the accuracies of the resulting parameters to be properly estimated. Early studies in satellite-based network positioning recognized the importance of modelling the covariances between double-differences (Beutler et al, 1986; Remondi, 1984), but focused on the mathematical correlations due to the differencing. Interestingly, these studies recognized the existence of physical correlations between the original phases, but were forced to ignore them due to a lack of suitable correlation models and difficulties in handling the large matrices required, especially when multiple baselines were involved.

Much research has been dedicated to assessing the possible accuracy achievable with satellite-based ranging systems and GPS in particular. For example, the literature abounds with studies of the noise properties of various receivers (Tiberius et al, 1999; Langley, 1997; Gerdan, 1995; Kujawa, 1998, among others) and of multipath effects in various environments (Ray, 2000; Braasch, 1998; Georgiadou and Kleusberg, 1988). As a result, a good deal of information exists regarding the magnitude of these errors. Similarly, tropospheric and ionospheric modelling has been ongoing for several decades, starting with studies by Hopfield (1963) and Klobuchar (1986), and more modern studies by Mendes (1999) and Schaer (1999) provide thorough studies of the accuracies of the models popularly used.

Unfortunately, the majority of this work has been done in an ad-hoc fashion and usually focuses on a single error source, to the neglect of others. This has led to a system of “rule-of-thumb” in estimating the achievable accuracy of GPS as a function of receiver

separation. In particular, knowledge of the spatial correlations of tropospheric and ionospheric errors is scant, as is understanding of the time variations of these errors and their statistical properties. Similarly, while work has been dedicated to minimizing noise and multipath, modelling the statistical properties of the remaining errors needs to be investigated. Overall, a systematic procedure for establishing and testing stochastic models of various error sources is needed.

This chapter seeks to analyse the main error sources affecting GPS and to model their stochastic properties. The ultimate goal of this work is to develop a method of using these models to properly create a variance-covariance matrix describing the relationship between observations, for use in a rigorous least-squares adjustment.

### 6.1 - Covariance of GNSS Error Sources

The basic model of a range measurement made between a receiver and a satellite was given in Chapter 2 as :

$$p = d_g - c \cdot (\Delta t_s - \Delta t_r) + T + I + \mathbf{e} \quad (2.4)$$

where  $p$  is the measured pseudorange,  $d_g$  is the actual geometric distance between the satellite and receiver,  $\Delta t_s$  and  $\Delta t_r$  are the satellite and receiver clock offsets,  $T$  and  $I$  are the tropospheric and ionospheric delays and  $\mathbf{e}$  is the noise. In the case of a carrier phase measurement, an ambiguity term is added. In general, a tropospheric and ionospheric model is used, and thus the  $T$  and  $I$  terms become  $dT$  and  $dI$ . In addition, the noise term can be separated into two components – namely noise and multipath. Furthermore, assuming that all the errors have zero mean expectation, the resulting error in the pseudorange measurement is denoted

$$err = p - E(p) = -\Delta t_s + \Delta t_r + dT + dI + n + m \quad (6.1)$$

where the clock offset terms have been parameterized as distances, and  $m$  denotes the multipath component of the observational noise and  $n$  denotes the remaining noise component. In the case of a carrier phase measurement, the sign of the ionospheric term is reversed and the values are divided by the wavelength of the carrier to express the error in cycles. This is further discussed in Section 6.5.

Each error term in Eq. (6.1) is due to a different physical process. Thus, for a given carrier phase, the individual error components are uncorrelated with one another. For example, there is no logical reason why the value of the satellite clock offset should at all affect the residual tropospheric delay contained within a measured carrier phase.

In addition it is important to note that we assume that the individual error components are *stationary*, or that their statistics can be meaningfully described by moments that are independent of time or space. In particular, we assume that all the errors have zero-mean. However, the *ergodicity* of the error processes is not guaranteed. For example, the receiver clock offset of a receiver at start up is random and has an expected value of zero since the clock can be ahead or behind GPS time with equal probabilities. However, for a given sample set (or *realisation*), the receiver clock offset can be constant with time at some value. Thus, simply taking the average of the receiver clock for a given realisation does not correctly imply the actual mean value of the process. On the other hand, the noise and multipath processes *are* ergodic.

Consider the situation of three receivers making simultaneous measurements to three satellites. The covariance of any two carrier phase measurements between receivers  $A$  and  $B$  and satellites  $i$  and  $j$  is expressed as

$$\begin{aligned}
\mathbf{s}_{ij}^{AB} &= E\left[\left(\mathbf{f}_i^A - E(\mathbf{f}_i^A)\right) \cdot \left(\mathbf{f}_j^B - E(\mathbf{f}_j^B)\right)\right] \\
&= E\left[\left(\frac{1}{I}\left[(\Delta t_{sati}^A - \Delta t_{recA}^i) + \mathbf{dT}_i^A - \mathbf{dl}_i^A + m_i^A + n_i^A\right]\right) \cdot \left(\frac{1}{I}\left[(\Delta t_{satj}^B - \Delta t_{recB}^j) + \mathbf{dT}_j^B - \mathbf{dl}_j^B + m_j^B + n_j^B\right]\right)\right] \quad (6.2)
\end{aligned}$$

since the expected value of  $\mathbf{f}_i^A$  is equal to the range and ambiguity of the carrier phases and the error sources have zero mean. Note that the satellite clock error is wholly dependent on the satellite, the receiver clock error is dependent on the receiver and all other errors are dependent on both the particular receiver/satellite combination.

Eq. (6.2) can be further simplified by taking into account that the error sources are physically uncorrelated among themselves. This results in the following expression :

$$\begin{aligned}
\mathbf{s}_{ij}^{AB} &= \frac{1}{I^2} E\left[(\Delta t_{sati}^A \Delta t_{satj}^B + \Delta t_{recA}^i \Delta t_{recB}^j) + \mathbf{dT}_i^A \mathbf{dT}_j^B + \mathbf{dl}_i^A \mathbf{dl}_j^B + m_i^A m_j^B + n_i^A n_j^B\right] \\
&= \mathbf{s}_{ij}^{AB}(\Delta t_{sat}) + \mathbf{s}_{ij}^{AB}(\Delta t_{rec}) + \mathbf{s}_{ij}^{AB}(\mathbf{dT}) + \mathbf{s}_{ij}^{AB}(\mathbf{dl}) + \mathbf{s}_{ij}^{AB}(m) + \mathbf{s}_{ij}^{AB}(n) \quad (6.3)
\end{aligned}$$

where the covariances of the individual error sources have been introduced.

For the nine observations implied by the three receiver, three satellite scenario, the variance-covariance matrix for the set can be written as :



$$\mathbf{C}_1 = \begin{bmatrix} \mathbf{s}_1^{A^2} & \mathbf{s}_{12}^A & \mathbf{s}_{13}^A & \mathbf{s}_1^{AB} & \mathbf{s}_{12}^{AB} & \mathbf{s}_{13}^{AB} & \mathbf{s}_1^{AC} & \mathbf{s}_{12}^{AC} & \mathbf{s}_{13}^{AC} \\ \mathbf{s}_{21}^A & \mathbf{s}_2^{A^2} & \mathbf{s}_{23}^A & \mathbf{s}_{21}^{AB} & \mathbf{s}_2^{AB} & \mathbf{s}_{23}^{AB} & \mathbf{s}_{21}^{AC} & \mathbf{s}_2^{AC} & \mathbf{s}_{23}^{AC} \\ \mathbf{s}_{31}^A & \mathbf{s}_{32}^A & \mathbf{s}_3^{A^2} & \mathbf{s}_{31}^{AB} & \mathbf{s}_{32}^{AB} & \mathbf{s}_3^{AB} & \mathbf{s}_{31}^{AC} & \mathbf{s}_{32}^{AC} & \mathbf{s}_3^{AC} \\ \mathbf{s}_1^{BA} & \mathbf{s}_{12}^{BA} & \mathbf{s}_{13}^{BA} & \mathbf{s}_1^{B^2} & \mathbf{s}_{12}^B & \mathbf{s}_{13}^B & \mathbf{s}_1^{BC} & \mathbf{s}_{12}^{BC} & \mathbf{s}_{13}^{BC} \\ \mathbf{s}_{21}^{BA} & \mathbf{s}_2^{BA} & \mathbf{s}_{23}^{BA} & \mathbf{s}_{21}^B & \mathbf{s}_2^{B^2} & \mathbf{s}_{23}^B & \mathbf{s}_{21}^{BC} & \mathbf{s}_2^{BC} & \mathbf{s}_{23}^{BC} \\ \mathbf{s}_{31}^{BA} & \mathbf{s}_{32}^{BA} & \mathbf{s}_3^{BA} & \mathbf{s}_{31}^B & \mathbf{s}_{32}^B & \mathbf{s}_3^{B^2} & \mathbf{s}_{31}^{BC} & \mathbf{s}_{32}^{BC} & \mathbf{s}_3^{BC} \\ \mathbf{s}_1^{CA} & \mathbf{s}_{12}^{CA} & \mathbf{s}_{13}^{CA} & \mathbf{s}_1^{CB} & \mathbf{s}_{12}^{CB} & \mathbf{s}_{13}^{CB} & \mathbf{s}_1^{C^2} & \mathbf{s}_{12}^C & \mathbf{s}_{13}^C \\ \mathbf{s}_{21}^{CA} & \mathbf{s}_2^{CA} & \mathbf{s}_{23}^{CA} & \mathbf{s}_{21}^{CB} & \mathbf{s}_2^{CB} & \mathbf{s}_{23}^{CB} & \mathbf{s}_{21}^C & \mathbf{s}_2^{C^2} & \mathbf{s}_{23}^C \\ \mathbf{s}_{31}^{CA} & \mathbf{s}_{32}^{CA} & \mathbf{s}_3^{CA} & \mathbf{s}_{31}^{CB} & \mathbf{s}_{32}^{CB} & \mathbf{s}_3^{CB} & \mathbf{s}_{31}^C & \mathbf{s}_{32}^C & \mathbf{s}_3^{C^2} \end{bmatrix} \quad (6.4)$$

Note the expected symmetry of the above matrix and the fact that only four basic forms of variances exist, namely  $\mathbf{s}_i^A$ ,  $\mathbf{s}_i^{AB}$ ,  $\mathbf{s}_i^B$ ,  $\mathbf{s}_i^{BC}$ . Furthermore, elements of the form  $\mathbf{s}_{ij}^A$  are equal to their corresponding elements of the form  $\mathbf{s}_{ji}^A$ . As well, elements of the form  $\mathbf{s}_i^{AB}$  are equal to elements of the form  $\mathbf{s}_i^{BA}$ . This also holds for pairs of elements with the form  $\mathbf{s}_{ij}^{AB}$  and  $\mathbf{s}_{ji}^{BA}$ , respectively. However, elements  $\mathbf{s}_{ij}^{AB}$  are *not* in general equal to elements  $\mathbf{s}_{ji}^{AB}$  or  $\mathbf{s}_{ij}^{BA}$ . Applying these equalities, it can be shown that the above matrix is symmetric, as expected.

### 6.1.1 -Effect of Double-Differencing on the Variance-Covariance Matrix

As indicated in the above section, errors between satellites are often correlated, as are errors between receivers. In Chapter 3, the satellite and receiver clock errors were treated as unknowns to be solved, which is contrary to the treatment given here, where they are considered as errors. However, the method of double-differencing, which was previously used to remove the unknown clock offsets as unknowns, can also be shown to remove the clock offsets if they are considered as errors. In addition, double differencing reduces other errors correlated between receivers, such as tropospheric and ionospheric errors. One consequence, however, is that error sources previously uncorrelated between observations, such as multipath and noise, become correlated when the double-differenced observations are formed.

For the set of observations made by three receivers to three satellites, the resulting variance-covariance matrix of the double-differenced observations can be calculated via propagation of variances, using a suitable  $\mathbf{B}$  matrix.

$$\mathbf{C}_{DD} = \mathbf{B}\mathbf{C}_1\mathbf{B}^T = \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix} \cdot \mathbf{C}_1 \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (6.5)$$

$$= \begin{bmatrix} \mathbf{s}^2(\nabla\Delta_{12}^{AB}) & \mathbf{s}(\nabla\Delta_{12}^{AB}, \nabla\Delta_{13}^{AB}) & \mathbf{s}(\nabla\Delta_{12}^{AB}, \nabla\Delta_{12}^{AC}) & \mathbf{s}(\nabla\Delta_{12}^{AB}, \nabla\Delta_{13}^{AC}) \\ \mathbf{s}(\nabla\Delta_{12}^{AB}, \nabla\Delta_{13}^{AB}) & \mathbf{s}^2(\nabla\Delta_{13}^{AB}) & \mathbf{s}(\nabla\Delta_{13}^{AB}, \nabla\Delta_{12}^{AC}) & \mathbf{s}(\nabla\Delta_{13}^{AB}, \nabla\Delta_{13}^{AC}) \\ \mathbf{s}(\nabla\Delta_{12}^{AB}, \nabla\Delta_{12}^{AC}) & \mathbf{s}(\nabla\Delta_{13}^{AB}, \nabla\Delta_{12}^{AC}) & \mathbf{s}^2(\nabla\Delta_{12}^{AC}) & \mathbf{s}(\nabla\Delta_{12}^{AC}, \nabla\Delta_{13}^{AC}) \\ \mathbf{s}(\nabla\Delta_{12}^{AB}, \nabla\Delta_{13}^{AC}) & \mathbf{s}(\nabla\Delta_{13}^{AB}, \nabla\Delta_{13}^{AC}) & \mathbf{s}(\nabla\Delta_{12}^{AC}, \nabla\Delta_{13}^{AC}) & \mathbf{s}^2(\nabla\Delta_{13}^{AC}) \end{bmatrix}$$

where  $\mathbf{C}_1$  is given by Eq. (6.4). The resultant terms in the  $\mathbf{C}_{DD}$  matrix are the variances and covariances of the double-differenced observations formed.

The variance-covariance matrix shown is fully-populated and symmetric. The explicit form of each element can be derived by completing the multiplication indicated in Eq. (6.5), given the  $\mathbf{C}_1$  matrix shown in Eq. (6.4). The results are as follows :

$$\begin{aligned}
\mathbf{s}^2(\nabla\Delta_{12}^{AB}) &= \mathbf{s}_1^{A^2} + \mathbf{s}_1^{B^2} + \mathbf{s}_2^{A^2} + \mathbf{s}_2^{B^2} - 2\mathbf{s}_{12}^A - 2\mathbf{s}_{12}^B - 2\mathbf{s}_1^{AB} - 2\mathbf{s}_2^{AB} + 2\mathbf{s}_{12}^{AB} + 2\mathbf{s}_{21}^{AB} \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}, \nabla\Delta_{13}^{AB}) &= \mathbf{s}_1^{A^2} + \mathbf{s}_1^{B^2} - \mathbf{s}_{12}^A - \mathbf{s}_{13}^A - \mathbf{s}_{12}^B - \mathbf{s}_{13}^B + \mathbf{s}_{23}^A + \mathbf{s}_{23}^B - 2\mathbf{s}_1^{AB} + \mathbf{s}_{12}^{AB} + \mathbf{s}_{21}^{AB} \\
&\quad + \mathbf{s}_{13}^{AB} + \mathbf{s}_{31}^{AB} - \mathbf{s}_{32}^{AB} - \mathbf{s}_{23}^{AB} \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}, \nabla\Delta_{12}^{AC}) &= \mathbf{s}_1^{A^2} + \mathbf{s}_2^{A^2} - \mathbf{s}_1^{AB} - \mathbf{s}_2^{AB} - \mathbf{s}_1^{AC} - \mathbf{s}_2^{AC} - 2\mathbf{s}_{12}^A + \mathbf{s}_{12}^{AB} + \mathbf{s}_{21}^{AB} + \mathbf{s}_{12}^{AC} + \mathbf{s}_{21}^{AC} \\
&\quad + \mathbf{s}_1^{bC} + \mathbf{s}_2^{bC} - \mathbf{s}_{12}^{bC} - \mathbf{s}_{21}^{bC} \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}, \nabla\Delta_{13}^{AC}) &= \mathbf{s}_1^{A^2} - \mathbf{s}_{12}^A - \mathbf{s}_{13}^A - \mathbf{s}_1^{AB} + \mathbf{s}_1^{BC} - \mathbf{s}_1^{AC} + \mathbf{s}_{23}^A + \mathbf{s}_{12}^{AB} + \mathbf{s}_{31}^{AB} - \mathbf{s}_{32}^{AB} \\
&\quad + \mathbf{s}_{21}^{AC} - \mathbf{s}_{21}^{bC} + \mathbf{s}_{13}^{AC} - \mathbf{s}_{23}^{AC} - \mathbf{s}_{13}^{bC} + \mathbf{s}_{23}^{bC}
\end{aligned} \tag{6.6}$$

where

- $\mathbf{s}(\tilde{\mathbf{N}}D_{12}^{AB}, \tilde{\mathbf{N}}D_{12}^{AB})$ ,  $\mathbf{s}(\tilde{\mathbf{N}}D_{13}^{AB}, \tilde{\mathbf{N}}D_{13}^{AB})$ ,  $\mathbf{s}(\tilde{\mathbf{N}}D_{12}^{AC}, \tilde{\mathbf{N}}D_{12}^{AC})$ , and  $\mathbf{s}(\tilde{\mathbf{N}}D_{13}^{AC}, \tilde{\mathbf{N}}D_{13}^{AC})$
- $\mathbf{s}^2(\tilde{\mathbf{N}}D_{12}^{AB}, \tilde{\mathbf{N}}D_{13}^{AB})$  and  $\mathbf{s}^2(\tilde{\mathbf{N}}D_{12}^{AC}, \tilde{\mathbf{N}}D_{13}^{AC})$
- $\mathbf{s}^2(\tilde{\mathbf{N}}D_{12}^{AB}, \tilde{\mathbf{N}}D_{13}^{AC})$  and  $\mathbf{s}^2(\tilde{\mathbf{N}}D_{13}^{AB}, \tilde{\mathbf{N}}D_{12}^{AC})$
- $\mathbf{s}^2(\tilde{\mathbf{N}}D_{13}^{AB}, \tilde{\mathbf{N}}D_{13}^{AC})$  and  $\mathbf{s}^2(\tilde{\mathbf{N}}D_{12}^{AB}, \tilde{\mathbf{N}}D_{12}^{AC})$

have similar forms, respectively.

The covariances of the double-differences do not only contain terms relating to the variances of the component observations, but also terms describing the correlation between different observations. Using the property of independence between physical processes, the variances described in Eq. (6.6) can be described as the sum of variances and covariances of each double-differenced error source. For example,

$$\begin{aligned}
\mathbf{s}^2(\nabla\Delta_{12}^{AB}) &= \mathbf{s}^2(\nabla\Delta_{12}^{AB} \Delta t_{sat}) + \mathbf{s}^2(\nabla\Delta_{12}^{AB} \Delta t_{rec}) + \mathbf{s}^2(\nabla\Delta_{12}^{AB} n) + \mathbf{s}^2(\nabla\Delta_{12}^{AB} dT) \\
&\quad + \mathbf{s}^2(\nabla\Delta_{12}^{AB} dl) + \mathbf{s}^2(\nabla\Delta_{12}^{AB} m)
\end{aligned} \tag{6.7}$$

This is a crucial fact, as it implies that the total double-differenced variance-covariance matrix can be formed using the *summation* of individual double-differenced variance-covariance matrices for each error source. This also will influence the methods used to

study each error source individually, as any one error variance can be calculated if the total variance, and the variances of the other errors sources are provided.

Finally, it is also important to note that traditionally, a diagonal  $\mathbf{C}_1$  matrix has been assumed, often with equal weights assumed for each observation. In such a case, the variance-covariance matrix of the double differences becomes

$$\mathbf{C}_{DD} = \mathbf{s}_o^2 \cdot \begin{bmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 1 & 2 \\ 2 & 1 & 4 & 2 \\ 1 & 2 & 2 & 4 \end{bmatrix} \quad (6.8)$$

where  $\mathbf{s}_o^2$  represents the variance of the observed phase multipath and noise. If Eq. (6.8) is compared to the expressions in Eq. (6.5) and (6.6), it is evident that they are equivalent only if one assumes that all covariances are zero, or if all covariances are equal.

### 6.1.2 -Developing Variance-Covariance Models for GNSS

One of the fundamental goals of this work is to establish a methodology for determining the appropriate variance-covariance matrix for use during a particular observational campaign. This is a non-trivial task, as it is well known that the error sources in GPS are not ergodic (for example, the ionospheric error magnitude varies with the time of day) and can sometimes be site specific. As a result, the variance-covariance matrix for the observations is expected to change between observation sessions. Thus a flexible model that adequately models the observational variance under a variety of conditions is required.

Traditional methods of variance-component estimation have fallen into two categories – apriori and simultaneous. Apriori modelling seeks to determine appropriate models for observations expected to be made in a given network, usually relying on laboratory

experimentation and calibration. For example, in traditional terrestrial networks, distance measurement variances are often modelled by the following model

$$\mathbf{s}_d = a + b \cdot d \quad (6.9)$$

where  $d$  is the measured distance and  $a$  and  $b$  are experimentally determined constants. The study by Nickerson (1978) is an excellent example of this form of variance modelling, and contains many model parameter values for a wide variety of equipment types and situations encountered in terrestrial networks. The advantage of apriori models lies in that they are simple to implement once all the model parameters have been established. Unfortunately, in the case of GPS, there is simply little knowledge of the stochastic properties of error sources, so no widely accepted apriori models exist. In addition, since it is known that GPS errors, such as the ionosphere, vary over time and space, it is unlikely that a useful stochastic model can be purely based on theoretical considerations

Simultaneous variance-covariance modelling seeks to simultaneously determine the value of unknown parameters of interest and the variance-covariance matrix linking the observations. The MINQUE (Minimum Norm Quadratic Unbiased Estimator) method of Rao (1971) is a classic example of this class and attempts to solve the standard least-squares problem with additional unknowns in the variance-covariance matrix itself, such that

$$\mathbf{C}_1 = \sum_{i=1}^k q_i \cdot \mathbf{T}_i \quad (6.10)$$

where  $q_i$  is an unknown variance factor and  $\mathbf{T}_i$  is a known matrix. Many alternative, but often numerically equivalent, methods exist in this class, and are extensively reviewed by Grodecki (1997). The appeal of simultaneous methods lies in the fact that the variance-models are derived from the data themselves, so they are very useful if the variances are expected to change or are unknown. For example, Wang et al (1998) used the MINQUE

method to derive the covariance matrix of a very short GPS baseline with some success and discovered that the double-differences were indeed correlated.

The major shortcoming in simultaneous methods lies in their computational complexity and numerical instability, which can sometime result in non-sensical results such as negative variances (Grodecki, 1997). These problems are amplified when one considers the sheer number of observations that must be considered in a typical GPS observation session and the fact that the variances are expected to vary from epoch to epoch. Furthermore, a more subtle limitation in the simultaneous techniques is that for stable solutions to occur, the  $\mathbf{T}_i$  matrices must be well estimated, and describe the covariance between observations. Again, this is information that is not precisely known in GPS. Of course, a large number of  $\mathbf{T}_i$  matrices, representing every conceivable pattern of correlations could be included, but then the number of unknowns becomes impractical and solution stability suffers.

Instead, the approach used in this thesis is a hybrid between the two methods. Firstly, due to the principle of physical independence of error sources, noise, multipath, tropospheric and ionospheric errors are all considered separately. Traditionally, orbital error has been considered as a ranging error, but in this work its effect has been studied instead as a datum error and its effects documented in Chapters 4 and 5.

For a given error source, the first step in developing the variance model lies in investigating the theoretical underpinnings of the error source and from these considerations, developing an appropriate error model. Empirical data is then used to gauge the validity of the theoretical model. Also, this empirical data can be used to determine key parameters of the theoretical model, which may be time-varying. This affords insight into how the stochastic properties of the error source behave over space and time. The advantage of this method is that the preliminary theoretical study essentially serves the process of determining the  $\mathbf{T}_i$  matrices required by the MINQUE method, while the empirical study allows the model to be flexible to changes in the observation environments. Work by Raquet (1998) follows a similar tact, but with less emphasis on the initial model development.

## 6.2 -Noise Errors

Noise error is an effect that is unavoidable in any measurement process. In the case of GPS, it is the error in phase and code measurements due to imperfect tracking of the GPS signal by the phase and delay lock loops. As a result, it is internal to the receiver and is usually independent between satellites since separate loops are dedicated to each signal tracked (Ward, 1996a). Furthermore, measurement noise is typically considered uncorrelated in time beyond the predetection integration period (typically 20 ms). Thus it can be treated as a white noise process, assuming the output phases and ranges are not smoothed internally by the receiver.

In static applications, the predominant error component in carrier phase noise is jitter of the phase lock loop caused by thermal noise. This can be expressed as (Ward, 1996a) :

$$\mathbf{s}_n = \mathbf{s}_{PLL} = \frac{\mathbf{l}}{2\mathbf{p}} \sqrt{\frac{B_n}{c/n_o} \left( 1 + \frac{1}{2Tc/n_o} \right)} \quad (\text{metres}) \quad (6.11)$$

where  $B_n$  is the carrier loop noise bandwidth in Hz,  $c/n_o$  is the carrier to noise power expressed as  $10^{C/No/10}$  with  $C/No$  in dB-Hz,  $T$  is the predetection integration time in seconds and  $\mathbf{l}$  is the carrier wavelength. The code measurement noise is dependent on the correlation processes internal to the receiver, and is generally at the decimetre level (Langley, 1998), compared to the millimetre level for carrier phase measurements. Since the code measurements are so noisy that the final result essentially depends on the accuracy of the carrier phases, the code measurements will not be considered here.

Since the noise is uncorrelated between satellites, the double difference covariances expressed in Eq. (6.6) can be written, for the noise component, as :

$$\begin{aligned}
\mathbf{s}^2(\nabla\Delta_{12}^{AB}n) &= \mathbf{s}_1^{A^2}(n) + \mathbf{s}_1^{B^2}(n) + \mathbf{s}_2^{A^2}(n) + \mathbf{s}_2^{B^2}(n) \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}n, \nabla\Delta_{13}^{AB}n) &= \mathbf{s}_1^{A^2}(n) + \mathbf{s}_1^{B^2}(n) \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}n, \nabla\Delta_{12}^{AC}n) &= \mathbf{s}_1^{A^2}(n) + \mathbf{s}_2^{A^2}(n) \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}n, \nabla\Delta_{13}^{AC}n) &= \mathbf{s}_1^{A^2}(n)
\end{aligned} \tag{6.12}$$

where the noise variances of each observation are calculated using Eq. (6.11). All the parameters involved in Eq. (6.11) are either calculable or available from calibration, thus forming the theoretical variance-covariance model.

### 6.2.1 -Signal Power and Theoretical Noise Variance Modelling

Eq. (6.11) indicates that receiver noise is dependent on the carrier to noise ratio (CNR) of the incoming signal. This is the CNR at the receiver front end, and as a results depends on the transmitted power of the signal, the transmitting beam pattern, atmospheric attenuation, free space loss, the antenna gain pattern, and the effects of the antenna preamplifier and line losses. In total, this can be expressed as (Ward, 1996b):

$$C / No = P_T - 10 \log(kT_0) - L_F - L_{ATM} + G_{ANT} + G_{LNA} - L \quad (\text{dB-Hz}) \tag{6.13}$$

with	$C/No$	... carrier to noise ratio at receiver front end (dB)
	$P_T$	... power radiated by antenna in direction of user (including transmitting antenna gain effects) (dB)
	$10\log(kT_0)$	... thermal noise density (-204 dB-Hz)
	$L_f, L_{ATM}$	... free space / atmospheric loss (dB)
	$G_{ANT}, G_{LNA}$	... receiver antenna gain and low noise amplifier gain (dB)
	$L$	... other system losses, including line and connector losses (dB)



The free space loss depends on the frequency of the carrier wave,  $f$ , and the radiation distance,  $R$ , and can be calculated by :

$$L_f = 20 \log \left( \frac{f}{4\pi R} \right) \quad (\text{dB}) \quad (6.14)$$

Typical values of  $L_f$  are 184.4 dB for L1 frequencies and 182.2 dB for L2 signals.

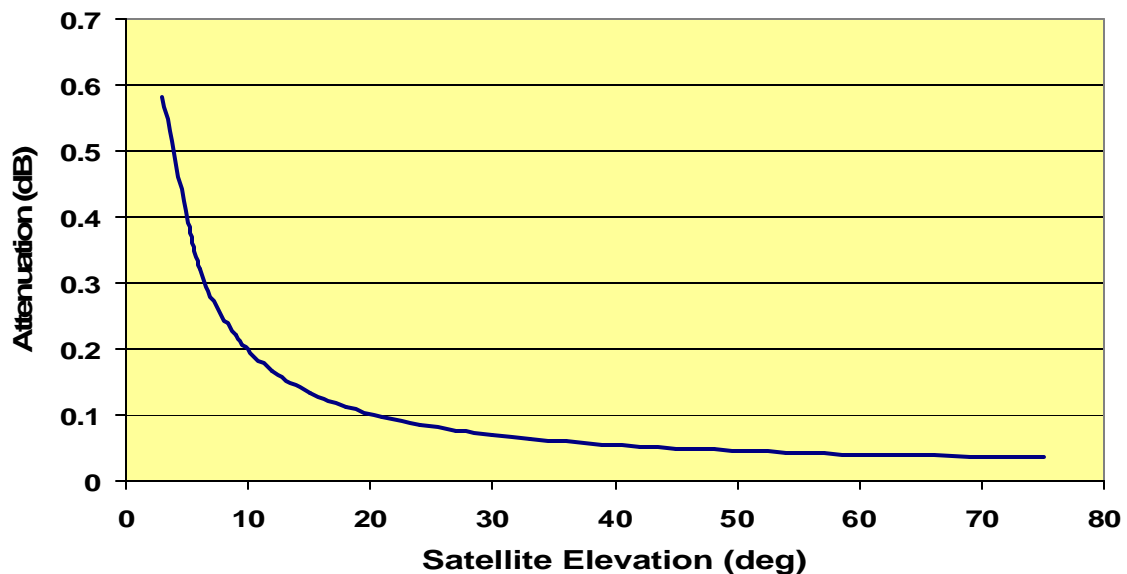
In addition, the atmospheric attenuation can be broken down into two parts – that of the troposphere and the ionosphere. Ionospheric scintillation is due to irregularities in the ionospheric electron content which can cause diffraction and refraction of the GPS signal. This results in fading of the signal up to 20dB, depending on solar activity and the latitude and longitude of the observer. (Klobuchar, 1996). In general, the severity of ionospheric scintillation is much worse around 21:00 local time and varies with the 11-year solar cycle (large sunspot numbers are highly correlated to active ionospheres). Typically, however, ionospheric scintillation causes losses of a few dB and is very difficult to predict. Thus we ignore it in the following analysis.

Tropospheric attenuation at GPS frequencies is mainly due to absorption by oxygen and can be expressed as (Spilker, 1996) :

$$L_{ATM} = \frac{2 \cdot 0.035 \cdot (1 + h_m / 2R_e)}{\sin E + \sqrt{\sin^2 E + 2 \cdot h_m / R_e + (h_m / R_e)^2}} \quad (\text{dB}) \quad (6.15)$$

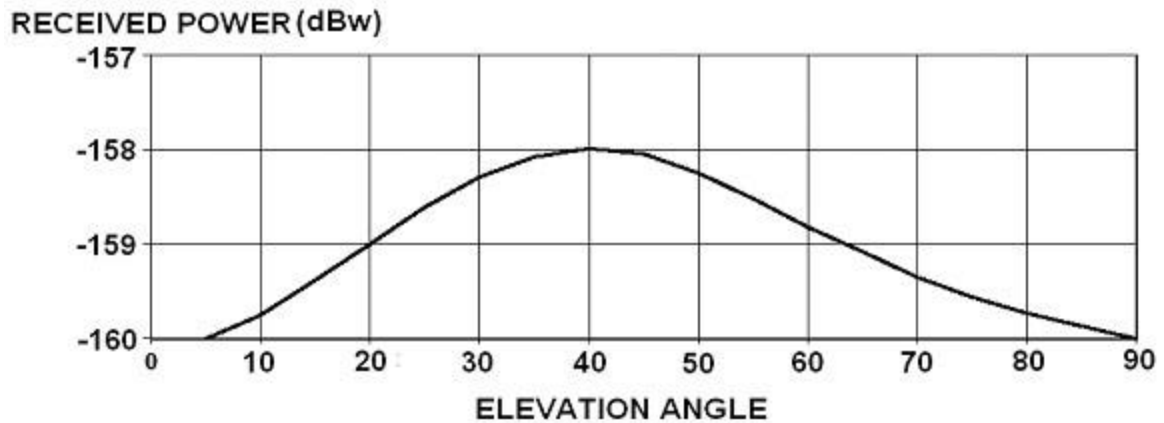
where  $E$  is the elevation angle of satellite,  $h_m$  is the equivalent height of the troposphere, usually set as 6 km, and  $R_e$  is the radius of the Earth, approximately 6378 km.

Figure 6.1 shows the variation in tropospheric attenuation with elevation angle. The attenuation can increase by a factor of 10 when elevation angles become low. This effect is partially offset by the gain pattern of the transmitting antenna, as discussed below.



**Figure 6.1.** Atmospheric Attenuation Versus Elevation Angle.

Unfortunately, it is difficult to predict the transmitted antenna power directly. Figure 6.2 shows the minimum signal power on the L1 carrier guaranteed by the GPS Standard Positioning Service Signal Specification (GPS-SPS, 1995) to be received by a user on the Earth. The power estimates assume i) a 3 dB antenna gain ii) a 2 dB atmospheric attenuation and iii) a minimum SV elevation of 5 degrees above the horizon and iv) that the SV attitude error is 0.5 degrees (toward reducing levels). Note that Figure 6.2 shows the *minimum* power received, and that in reality the signal strengths have been observed to be higher by up to 6 dB. Also, the maximum power received occurs when the satellite is at about 40° elevation. This is due to the shaping of the transmitting antenna. Also, the L2 power is guaranteed to be no more than 5 dB less than the L1 power.



**Figure 6.2.** Minimum Received L1 Signal Power (from GPS-SPS, 1995)

To predict the actual power at the receiver front end, the elevation of the observed satellite is first calculated and then the appropriate minimum power taken from Figure 6.2. 3 dB is then subtracted from the minimum power to strip the effect of the estimated antenna gain and 2 dB are added to remove the estimated atmospheric attenuation. The actual atmospheric attenuation is then subtracted using values calculated using Eq. (6.15) and the true antenna gain added using manufacturer specifications. The antenna gain is the most elevation dependant quantity, varying by as much as 20dB over the entire range of elevations and is shown for three varieties of NovAtel antennas in Figure 6.3. Finally, corrections for the line losses and preamplifier gain are made using manufacturer specifications. Since Figure 6.2 shows conservative power levels, the resulting predicted power should be a conservative estimate as well. Given the values of the carrier loop noise bandwidth and the preintegration time, all the double differenced noise variance required can be calculated.

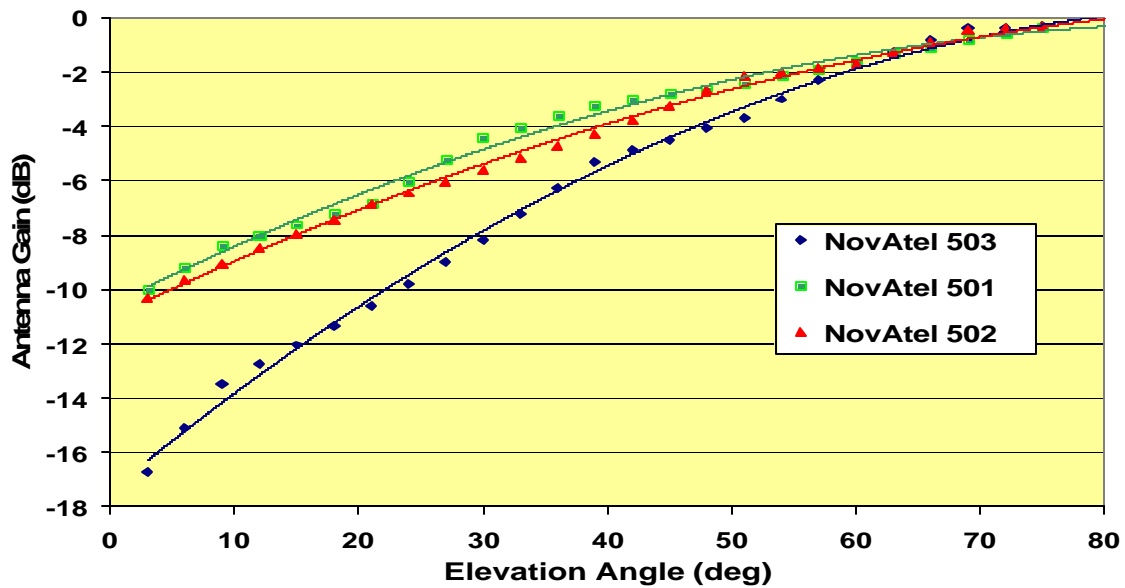


Figure 6.3. L1 Antenna Gain Patterns for Three Varieties of Antennas.

### 6.2.2 -Empirical Assessment of the Noise Variance Behaviour

A very useful way to analyse receiver noise is to perform a *zero-baseline test*. The test consists of two receivers processing a signal collected at a single antenna and split to each receiver. Upon double differencing the results, the effects of all error sources occurring prior to the signal passing the splitter are completely differenced out, leaving only receiver noise and the integer ambiguities. These ambiguities are easily solved for due to the small value of the carrier phase noise relative to the wavelength of the carrier (several millimetres versus 19 cm or 24 cm).

A zero-baseline test was conducted using three different GPS receivers with standard (not chokeringed) antennas. Two hour data spans were used with a zero degree elevation mask and 1Hz sampling interval. In all cases, the base satellite was above 40° elevation. Each double-difference sample was then binned according to satellite elevation in 3° bins and the standard deviation of samples in each bin computed. The results are shown in Figure 6.4 for all receivers using L1 data while Figure 6.5 shows the case of L2 data. The standard

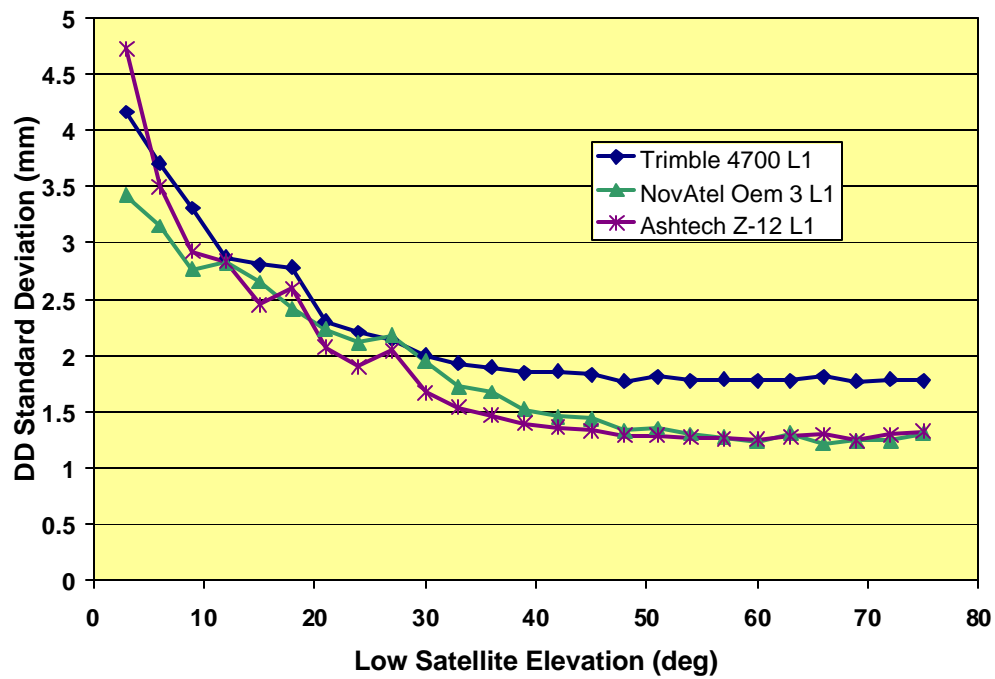


Figure 6.4. Comparison of double difference standard deviations for L1 signals.

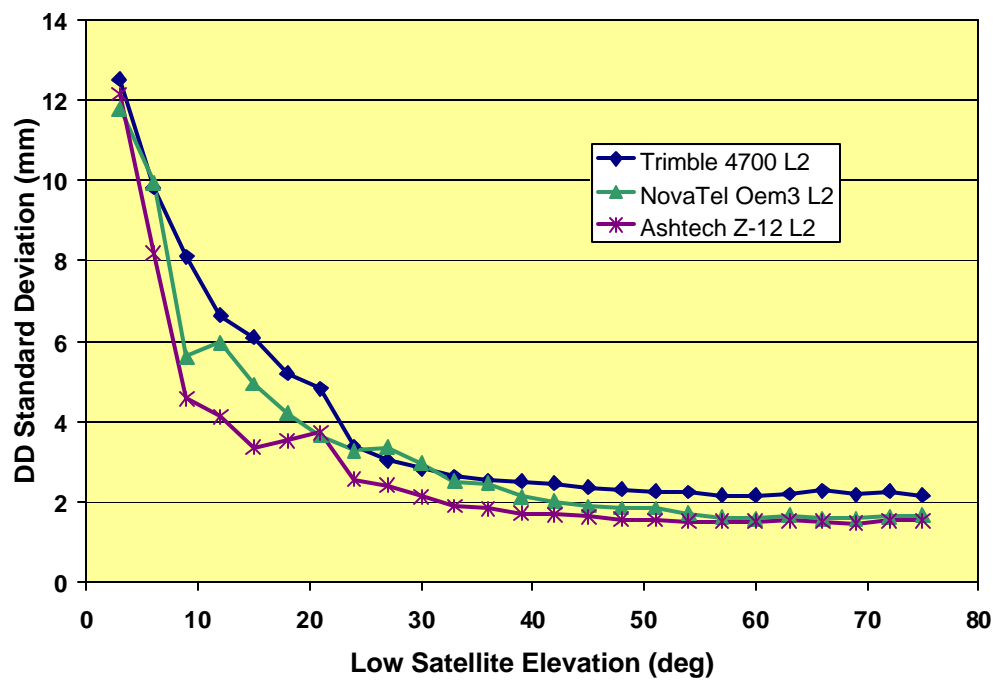


Figure 6.5. Comparison of double difference standard deviations for L2 signals.

deviations of the L2 data sets are higher due to tracking of the P codes in semi-codeless mode (resulting in a signal loss of 14 dB) and the longer wavelength of the L2 carrier.

As expected, the double differenced noise drops as the elevation of the lower satellite increases. This is mainly due to the effect of the antenna gain pattern, which is usually lower at low elevations to more effectively attenuate multipath signals. Above 40° elevation, the standard deviations become relatively constant for all three receivers.

If we assume that the double-difference standard deviation is constant when the low satellite elevation angle is above 40°, and that the performance of both receivers is equal, the variance of a double differenced observation above 40° can be calculated using Eq. (6.12) :

$$\mathbf{s}^2(\nabla\Delta_{HL}^{AB}n) = \mathbf{s}_n^{A^2}(\mathbf{e}_H) + \mathbf{s}_n^{A^2}(\mathbf{e}_L) + \mathbf{s}_n^{B^2}(\mathbf{e}_H) + \mathbf{s}_n^{B^2}(\mathbf{e}_L) = 4\mathbf{s}_n^2(\mathbf{e}_H) \quad (6.16)$$

where the notation implies that the noise standard deviations are dependant on the elevation angles of the low and high satellites,  $\mathbf{e}_L$  and  $\mathbf{e}_H$  and the receivers,  $A$  and  $B$ .

Once the variance of the high elevation observations has been calculated, the variance of *any* single observation can be calculated from its double differenced variance using :

$$\begin{aligned} \mathbf{s}^2(\nabla\Delta_{HL}^{AB}n) &= 2\mathbf{s}_n^2(\mathbf{e}_H) + 2\mathbf{s}_n^2(\mathbf{e}_L) \\ \therefore \mathbf{s}_n^2(\mathbf{e}_L) &= \frac{\mathbf{s}^2(\nabla\Delta_{HL}^{AB}n) - 2\mathbf{s}_n^2(\mathbf{e}_H)}{2} \end{aligned} \quad (6.17)$$

since both receivers are assumed to have identical noise profiles.

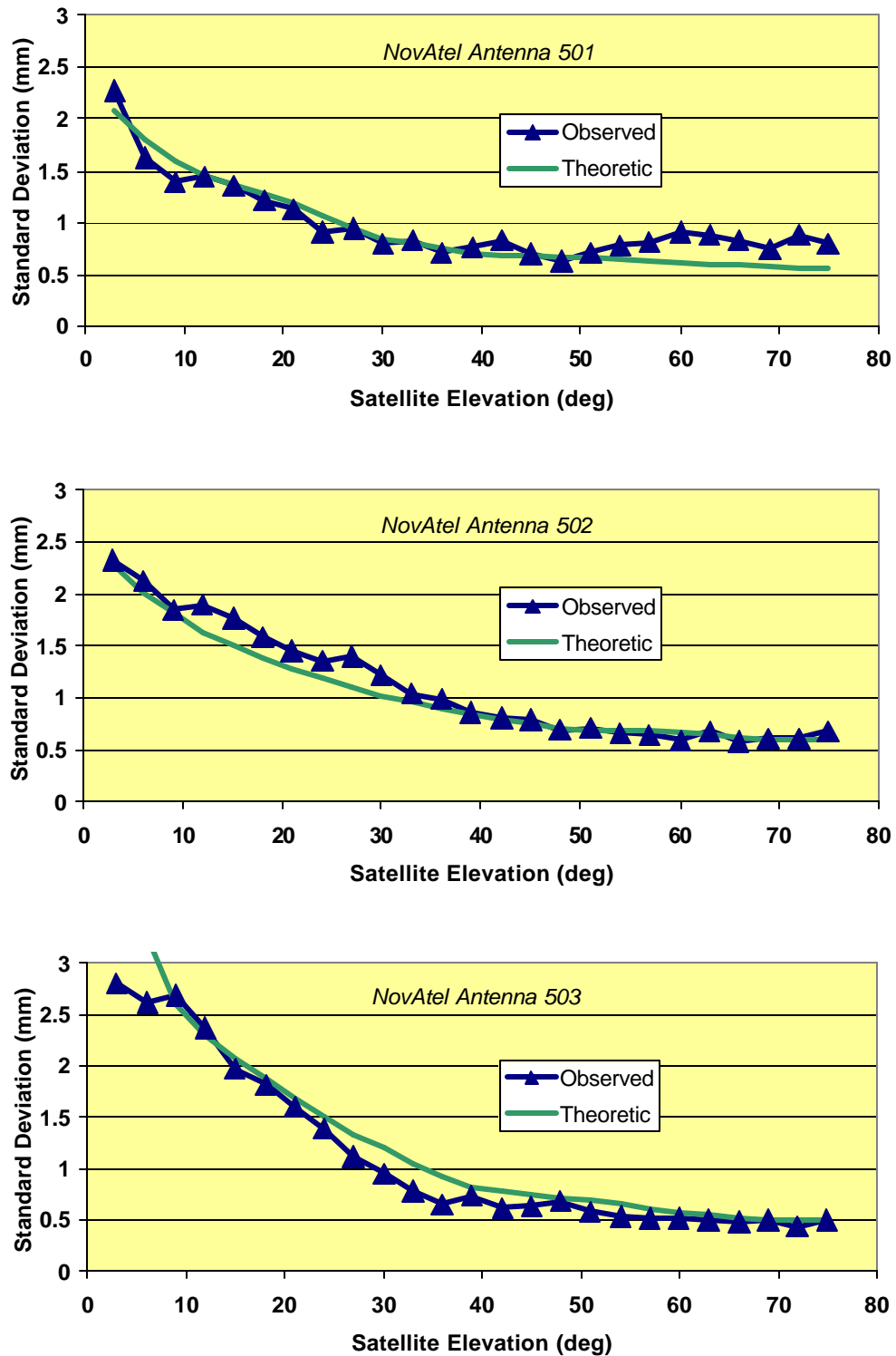
Another test was performed on a 2 hour data set collected using one receiver (NovAtel OEM 3) equipped with three types of antennas. Manufacturer's specifications for the receiver and the antennas gain patterns in Figure 6.3 were then used to predict the noise profile for each test using Eq. (6.9). The signal power was predicted as per Section 6.2.1, but an added 3dB loss was taken into consideration due to the splitting of the signal into two receivers and thus halving the available power. Finally, the true noise profile was then determined from the zero-baseline data and Eq. (6.15). Figure 6.6 shows the results for each antenna.

As seen in Figure 6.6, the theoretic noise levels closely match the noise measured in the zero-baseline test. Not only does this confirm the manufacturers' specifications and the validity of Eq. (6.11), it also illustrates the impact that antenna selection has on results. The NovAtel Antenna 503 is a chokeringed antenna designed to provide high multipath mitigation at low elevations. As a result, it has low gain at low elevation angles. Although this does mitigate multipath (discussed in the next section), it also means that the noise on low elevation satellites is increased.

A further impact of these results is that they imply that the noise variances at a given receiver may be modelled by a zenith term, corresponding to the noise encountered by high elevation satellites, multiplied by a *mapping function*, which allows the calculation of the actual noise variances at lower elevations. Thus the noise of any particular observation is calculated as

$$\mathbf{s}_1^{A^2}(n) = \mathbf{s}_A^2(n) \cdot m_n(\mathbf{e}_1) \quad (6.18)$$

where  $\mathbf{s}_A^2(n)$  is the zenith noise variance at receiver  $A$  and  $\mathbf{e}_1$  is the elevation angle of satellite 1 and  $m_n()$  is the noise mapping function. Collins and Langley (1999) suggest a  $1/\sin(\mathbf{e})$  model for the mapping function, but this does not always sufficiently simulate the



**Figure 6.6.** Observed versus Theoretic L1 Carrier Noise Standard Deviations for NovAtel Antennas 501, 502 and 503 and OEM 3 receivers.



shape of the theoretic noise profile at low elevations. Instead, an examination of Eq. (6.11) shows that the noise variance can be approximated, to the first order, by the expression :

$$\mathbf{s}_1^{A^2}(n) = \frac{c^A}{P_1^A} \quad (6.19a)$$

where  $c^A$  is a site-dependant constant and  $P_1^A$ , is the power of the observation from receiver A to satellite 1, given by Eq. (6.13) and equivalent to

$$P_1^A = 10^{b^A + f_A(e_1)} \quad (6.19b)$$

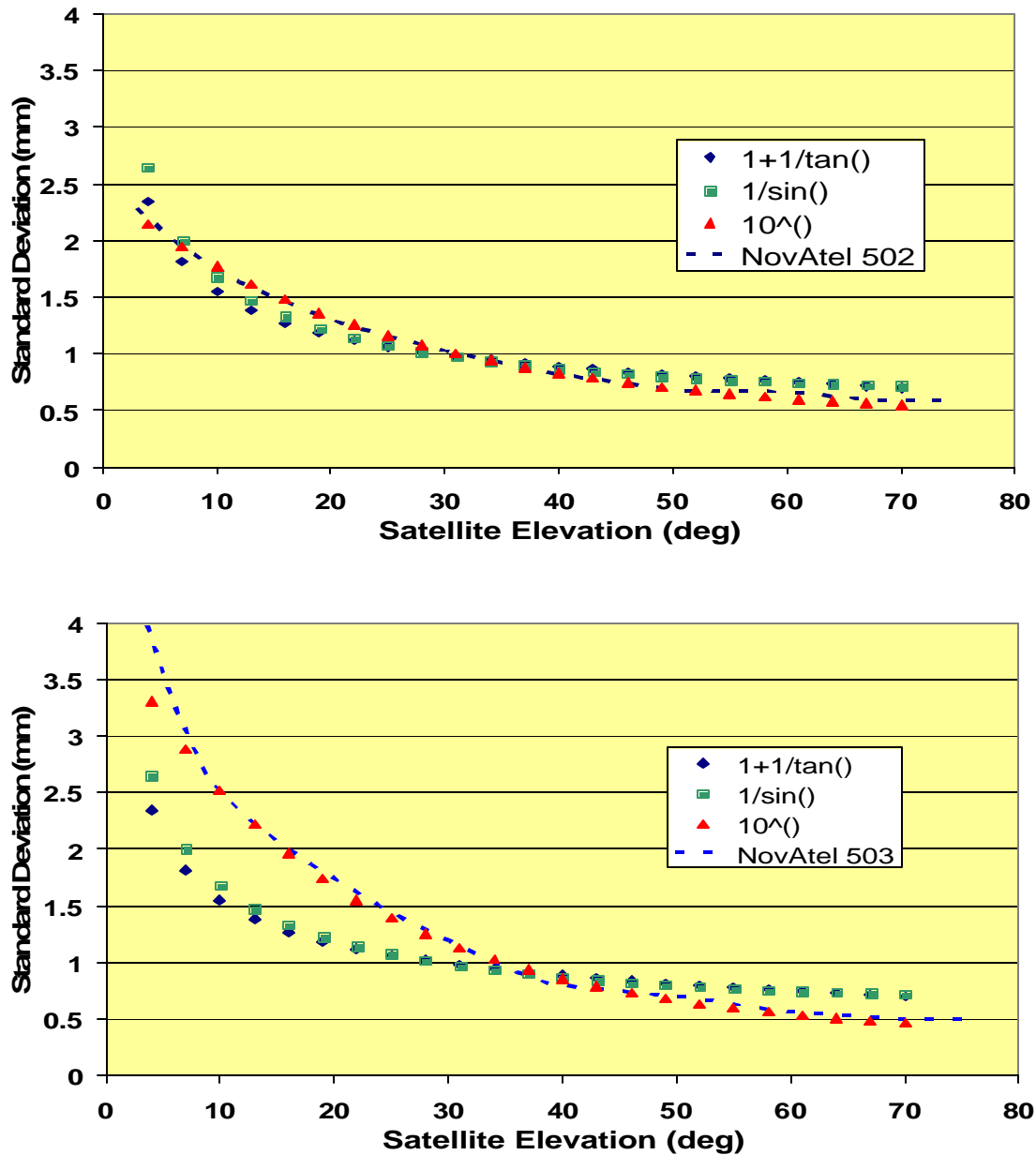
where  $b^A$  is another antenna dependant constant representing the average power received and  $f_A()$  is the power that varies with elevation angle. This is dominated by the variation in antenna gain. The gain patterns plotted in Figure 6.3 indicate that a quadratic relation adequately models the gain. This implies that the power can then be modelled by two additional parameters, since the offset is absorbed by the  $b^A$  term.

Putting these relations together and collecting constant terms results in an improved model for the receiver noise, namely

$$\mathbf{s}_1^{A^2}(n) = \mathbf{s}_A^2(n) \cdot 10^{c^A \cdot (90 - e_1)^2 + d^A (90 - e_1)} \quad (6.19c)$$

where  $c^A$  and  $d^A$  describes the variation of the antenna gain pattern with the elevation angle. Figure 6.7 shows a comparison of the theoretic noise profile with various popular weighting schemes and that of Eq. (6.19c) for two of the tested antennas. The latter gives the best fit over all elevation angles, especially for the chokeringed antenna and is simply

calculated from empirical data available. From a practical standpoint, use of Eq. (6.19c) to describe the variance of noise is preferred over that of Eq. (6.11) since manufacturer's data is often not available. Conversely, the three parameters required by Eq. (6.19c) can be easily determined by fitting the curve through data available from a zero-baseline test.



**Figure 6.7.** Comparison of Various Noise Mapping Functions for Two Antennas.

### 6.3 -Multipath Error

Multipath error is the error caused when a receiver receives a transmitted signal via multiple paths. Graphically, the effect is illustrated in Figure 6.8(a). The measured signal is a superposition of all the received signals and as such can be described as

$$s_c = s_d + \sum_i s_i^m = A \cdot G_D \cos \mathbf{f}_d + \sum_i \mathbf{b}_i \cdot A \cdot G_i \cos(\mathbf{f}_d + \Delta \mathbf{f}_i) \quad (6.20)$$

where  $s_c$  is the measured signal,  $s_d$  is the direct signal with amplitude  $A$  and phase  $\mathbf{Df}_i$ ,  $s_i^m$  is the  $i^{\text{th}}$  multipath signal with amplitude damped by the factor  $\mathbf{b}_i$  and phase delay  $\mathbf{Df}_i$ .  $G_D$  and  $G_i$  refer to the antenna gain patterns of the direct and reflected signals, and are amplitude factors between 0 and 1, rather than power levels as in the previous section.

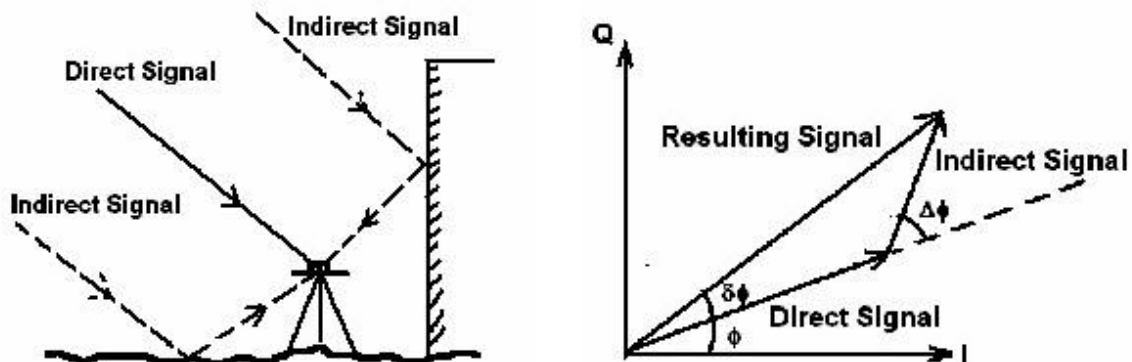


Figure 6.8. Multipath Effect (a) and Resulting Multipath Error(b).

The effect of multipath on carrier phase signals can be best seen in the form of a phasor diagram as in Figure 6.8(b). Each signal component is expressed as a complex number  $s = A \cdot e^{j\mathbf{q}}$ , where  $s$  is a complex number,  $A$  is the amplitude of the signal and  $\mathbf{q}$  is the phase of the signal. From Figure 6.8, it is apparent that the multipath phase error is the difference in

phase between the direct signal and the composite signal. The resulting multipath error,  $\mathbf{df}$ , is given by (Ray, 2000):

$$\mathbf{df} = \tan^{-1} \left[ \frac{\sum_i \mathbf{b}_i \cdot G_i \sin \Delta \mathbf{f}_i}{G_D + \sum_i \mathbf{b}_i \cdot G_i \cos \Delta \mathbf{f}_i} \right] \quad (\text{radians}) \quad (6.21)$$

For the receiver to be able to lock onto the direct signal, the sum of the reflective coefficients must be less than the antenna gain attenuating the direct signal, otherwise the reflections would overpower the direct signal. This implies that  $\mathbf{b}_i \ll 1$ . Using this assumption, Eq. (6.21) can be approximated by

$$\begin{aligned} \mathbf{df} &\approx \left[ \frac{\sum_i \mathbf{b}_i \cdot G_i \sin \Delta \mathbf{f}_i}{G_D + \sum_i \mathbf{b}_i \cdot G_i \cos \Delta \mathbf{f}_i} \right] \\ &\approx \left( \sum_i \mathbf{b}_i \cdot G_i \sin \Delta \mathbf{f}_i \right) \cdot \left( G_D - \sum_i \mathbf{b}_i \cdot G_i \cos \Delta \mathbf{f}_i \right) \\ &\quad \left[ G_D + \left( \sum_i \mathbf{b}_i \cdot G_i \cos \Delta \mathbf{f}_i \right)^2 \right] \cdot \left[ G_D + \left( \sum_i \mathbf{b}_i \cdot G_i \cos \Delta \mathbf{f}_i \right)^4 \right] \end{aligned} \quad (6.22)$$

using the expansion for  $1/(x+1)$  and the small angle approximation for the arctangent. The maximum value of the carrier phase multipath error using these assumptions is  $\frac{1}{4}$  of a wavelength. Again, the code multipath is not considered as it is too noisy for precise positioning.

### 6.3.1 - Theoretical Multipath Variance Modelling

The variance of multipath can be derived by expanding Eq. (6.22), and calculating the variance of each term, taking into account the orthogonality of the sinusoids involved. The resulting expression, neglecting higher order terms, is given in metres by :

$$\begin{aligned} \mathbf{s}^2(m) &= \left[ \frac{(\mathbf{b}_{mean}^2 + \mathbf{b}_{var}) \cdot n \cdot \mathbf{I}^2}{2} \right] \cdot \frac{1}{G_D^2} \\ &= \mathbf{s}_m^2 \cdot m_m(\mathbf{e}) \end{aligned} \quad (6.23)$$

where  $\mathbf{b}_{mean}$  is the mean reflectance of reflectors in the environment incorporating the effects of the antenna gain pattern,  $\mathbf{b}_{var}$  is the variance of the reflectivities of these reflectors, and  $n$  is the number of reflectors in the environment. Note that the terms in the brackets can be considered “site-specific” and constant over the observing period, while the antenna gain varies with the elevation of the direct signal. Thus the multipath variance can be expressed as the product of a zenith multipath variance multiplied by an elevation dependent mapping function, much like in the case of noise.

Furthermore, Eq. (6.23) essentially relates the total power of the reflected signals to the power of the direct signal, since the power of a signal is related to the square of its amplitude (Lathi, 1992). Thus, Eq. (6.23) can be re-written as

$$\mathbf{s}^2(m) = \frac{\mathbf{I}^2}{2} \frac{P_{ref}}{P_D} \quad (6.24)$$

where  $P_{ref}$  is the reflected power and  $P_D$  is the direct power in watts. According to Eq. (6.13), the power in decibels is equal to the sum of individual gains and losses along the transmission path. Since the reflect and direct differences are small with respect to the total transmission path, the only differences to be considered are the attenuation of the reflectors and the antenna gain pattern affecting the direct signal. As a result, Eq. (6.24) is equivalent to

$$\mathbf{s}^2(m) = \frac{\mathbf{I}^2}{2} \left[ \left( 10^{P_{Trans}} \cdot 10^{P_b} \right) / \left( 10^{P_{Trans}} \cdot 10^{P_G} \right) \right] \quad (6.25)$$

where  $P_{Trans}$  is the maximum received power,  $P_b$  is the power loss due to the reflections,  $P_G$  is the power loss on the direct signal due to the antenna gain pattern, in decibels. The form of the antenna gain term is equal to that encountered in Eq. (6.19b) when determining the noise mapping function. The conclusion is that the multipath variance for a single observation can be modelled via :

$$\mathbf{s}_1^{A^2}(m) = \mathbf{s}_A^2(m) \cdot 10^{c^A \cdot (90 - \epsilon_1)^2 + d^A (90 - \epsilon_1)} \quad (6.26)$$

where the  $c^A$  and  $d^A$  factors are identical to that for the noise model and  $\mathbf{s}_A^2(m)$  term is the site-dependant zenith multipath variance.

To complete the theoretical model, it can be shown that multipath is uncorrelated between satellites and receivers by taking the cross-variance of Eq. (6.22) for two different observations. Thus the double-differenced multipath variance-covariance model follows the same form as that of the noise, namely :

$$\begin{aligned} \mathbf{s}^2(\nabla\Delta_{12}^{AB}m) &= \mathbf{s}_1^{A^2}(m) + \mathbf{s}_1^{B^2}(m) + \mathbf{s}_2^{A^2}(m) + \mathbf{s}_2^{B^2}(m) \\ \mathbf{s}^2(\nabla\Delta_{12}^{AB}m, \nabla\Delta_{13}^{AB}m) &= \mathbf{s}_1^{A^2}(m) + \mathbf{s}_1^{B^2}(m) \\ \mathbf{s}^2(\nabla\Delta_{12}^{AB}m, \nabla\Delta_{12}^{AC}m) &= \mathbf{s}_1^{A^2}(m) + \mathbf{s}_2^{A^2}(m) \\ \mathbf{s}^2(\nabla\Delta_{12}^{AB}m, \nabla\Delta_{13}^{AC}m) &= \mathbf{s}_1^{A^2}(m) \end{aligned} \quad (6.27)$$

### 6.3.2 -Empirical Assessment of the Multipath Variance Behaviour

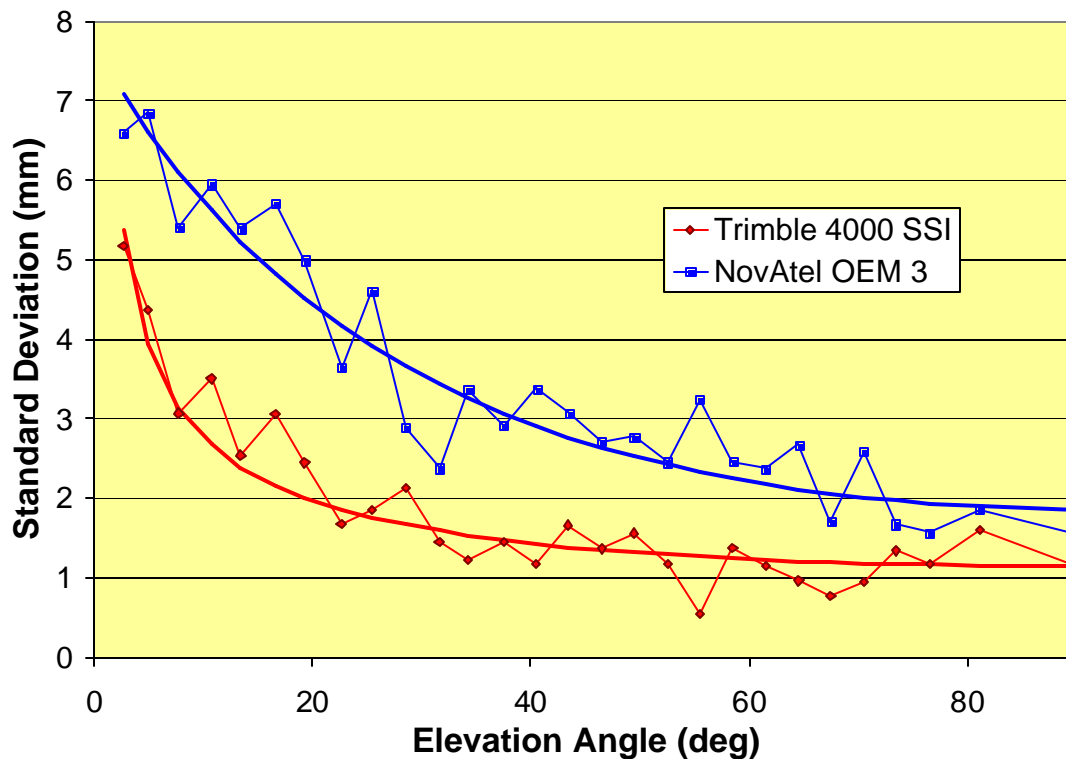
If a local-area network of receivers with known coordinates is observed over time, and the noise profiles of the receivers are known, it is possible to determine the double-differenced multipath variances for each baseline by removing the predicted noise variance from the observed variance, based on the principle of independence of error sources.

Such a test was performed using a short-baseline network consisting of two NovAtel OEM3 receivers and two Trimble 4000SSI receivers, located on the roof of the Engineering Building at the University of Calgary. Baseline lengths were less than 10 metres, ensuring

that tropospheric, orbital and ionospheric errors essentially differenced out completely and that the multipath environment would be the same for all receivers. Two hours of L1 carrier phase data was collected, resulting in 6000 observed double-differenced residuals. These residuals were binned according to elevation angle and the variance of each bin calculated, as in the noise variance study.

The predicted noise variances were then removed from these observed double-differenced variances using the noise model previously derived from a zero-baseline test. The remaining double differenced residuals for the NovAtel and Trimble baselines were then processed in the same manner as the double-differenced noise variances to produce undifferenced multipath variances for the Trimble and NovAtel receivers, using Eq. (6.17). Note that a key assumption in this procedure is that the zenith variance of both receivers in pair are equal, which was why the receiver pair were kept close together, ensuring identical multipath environments. The resulting multipath standard deviations for both receiver types are plotted in Figure 6.9. In addition, the theoretical multipath variance was calculated using the observed zenith multipath variance and the previously determined noise mapping function, which according to Eq. (6.26) is identical to the mapping function for multipath. The predicted multipath standard deviations are shown in Figure 6.9 as smooth curves

Two important observations can be made from Figure 6.9. Firstly, the shape and offset of the observed multipath variances is different for the two receiver types, despite being exposed to identical multipath environments. This is explained by the dependence of the multipath error on the antenna gain, which is different for the two receiver types. Also, the predicted and actual curves show good fit, which validates the model implied by Eq. (6.26). Residual scatter is due to changes in the particular reflectors affecting the signal at different elevation angles, which is not taken into account by the single zenith variance model. Modelling of the particular reflectors in an environment a difficult task, and requires specialized techniques and equipment (Ray, 2000). For the purposes of variance modelling, the single zenith model is sufficient.



**Figure 6.9.** Comparison of Observed and Theoretical L1 Multipath Variance for Two Receivers.

#### **6.4 -Tropospheric Error**

The tropospheric error was introduced in Chapter 2 as the delay a radio wave encounters due to its passage through the Earth's neutral atmosphere, or the troposphere. This region extends to an altitude of approximately 10 kilometres above the Earth's surface and is composed mainly of nitrogen, oxygen and argon. Water vapour is also found in this region. The presence of these gasses causes the index of refraction to become greater than one, which results in a retardation of the radio wave. At the frequencies of the GPS signals, the atmosphere is non-dispersive, implying the delay is equal for L1 and L2 frequencies (Hofmann-Wellenhof et al, 1994).

Mendes (1999) provides a very thorough overview of the physics behind the tropospheric delay. Essentially, tropospheric modelling consists of developing a model of the pressure,



temperature and water vapour variation throughout the atmosphere. A refractivity model is then developed around this atmospheric model. Early pioneers in this field include Hopfield (1963) and Saastamoinen (1973), who developed models still in use today.

Essentially all modern models separate the tropospheric delay problem into two parts – a wet and dry delay (Hopfield, 1969). More precisely, the dry delay is the component attributable to hydrostatic equilibrium, which includes the contribution of water vapour. The wet component is attributable to the water vapour not in equilibrium with the atmosphere. At zenith, the hydrostatic component causes approximately 2.3 metres, while the wet delay can account for up to 30 centimetres (Spilker, 1996).

A further common feature of models in use today is modelling of each component in the form

$$T_i = T_i^Z \cdot m_z(\mathbf{e}) \quad (6.28)$$

where  $T_i^Z$  is the zenith tropospheric delay for the wet or dry component and  $m_z(\mathbf{e})$  is a tropospheric mapping function relating the zenith delay to delays at lower elevation angles. Ifadis and Savvaidis (1999) provides a study of the performance of various wet and dry tropospheric mapping functions, as does Mendes (1999). Individual zenith delay models and mapping functions vary in terms of their accuracies and the input parameters required. For example, the NMF model (Niell, 1996) requires the latitude and height of the observing station, as well as the day of year. Others, such as the CfA-2.2 (Davis et al, 1985) require surface measurements of pressure, temperature and humidity. In general, it has been found that models depending on surface measurements have resulted in poorer performance, since surface measurements do not necessarily correlate with conditions even a few hundred metres above the surface (Brunner and Welsch, 1993). This is the justification for

use of the UNB2 model (Collins, 1999) which is parameterized solely in terms of the user's latitude and height.

#### 6.4.1 -Theoretical Tropospheric Variance Modelling

The basis of all tropospheric models is that the atmosphere can be completely described by a certain refractivity profile with regard to altitude, and that the atmosphere is homogenous from point-to-point. The real atmosphere, of course, is much more complex and as a result, a residual tropospheric delay error remains. This error is basically an integration of the point-to-point refractivity errors along the line-of-sight to the satellite. Assuming that the *variations* of refractivity are constant throughout the atmosphere, the delay error variance is then proportional to the path length of the observation ray. The path length is in turn proportional to the mapping function (Marini, 1971) for the observation and so the delay variance can be written as :

$$\mathbf{s}_1^{A^2}(T) = E(\mathbf{dT}_1^A \cdot \mathbf{dT}_1^A) = E(m_T(\mathbf{e}_1)\mathbf{dT}_Z^A \cdot m_T(\mathbf{e}_1)\mathbf{dT}_Z^A) = m_T(\mathbf{e}_1)^2 \mathbf{s}_Z^2(T) \quad (6.29)$$

where the  $\mathbf{dT}$  terms represents the residual tropospheric errors along the line of sight and in the zenith direction at site  $A$ , respectively.  $\mathbf{s}_Z^2(T)$  represents the zenith tropospheric variance and  $m_T()$  is the tropospheric mapping function used.  $E()$  is the expectation operator and it is assumed that the residual tropospheric delays are zero-meant, as they are due to variations about the average refractivity of the atmosphere. Also, note that the mapping function is assumed to be errorless, which is acceptable for the purposes of variance modelling, as the majority of the errors are due to the variation of water vapour in the atmosphere. In the same vein, although the hydrostatic component contributes 90% of the total tropospheric effect, it is the uncertainty in modelling the wet component that results in the majority of the error (Mendes, 1999). As a result, only the wet component is considered for variance-modelling.

In the case of observations made from a station to two satellites, the residual tropospheric errors are expected to be correlated if the line-of-sights to the two satellites pass through the same region of atmosphere, since the refractivity variations are assumed to be smoothly varying in space. From spherical trigonometry, the angle of separation between two observation line-of-sights,  $\mathbf{q}$ , can be written as :

$$\cos \mathbf{q} = \sin \mathbf{e}_1 \sin \mathbf{e}_2 + \cos \mathbf{e}_1 \cos \mathbf{e}_2 \cos(A_1 - A_2) \quad (6.30)$$

where  $A_1$  and  $A_2$  are the azimuths of line-of-sights to satellites 1 and 2 and  $\mathbf{e}_1, \mathbf{e}_2$  are the corresponding elevation angles.

For satellites in the same region of the sky ( $\mathbf{q}$  near zero), the line-of-sight vectors pass through similar portions of the troposphere and as a result the respective delay errors are the sum of the integration of similar point-to-point refractivity errors. Thus the covariance of the two delay errors should be nearly equal to the variance of a single observation (or the covariance of an error with itself). For satellites in very different parts of the sky ( $\mathbf{q}$  near  $180^\circ$ ), the point-to-point refractivity errors will be very uncorrelated and the covariance of the delays nearly equal to zero. Consequently, the covariance of the delay errors will be modeled in similar fashion as the variance of a single delay, with an exponential factor added to take into account decorrelation with increasing angular separation. This results in a model of the form :

$$\mathbf{s}_{12}^A(T) = E(\mathbf{dT}_1^A \cdot \mathbf{dT}_2^A) = E(m(\mathbf{e}_1) \mathbf{dT}_Z^A \cdot m(\mathbf{e}_2) \mathbf{dT}_Z^A) = m(\mathbf{e}_1) m(\mathbf{e}_2) \exp(-\mathbf{q} / \Omega) \cdot \mathbf{s}_Z^2(T) \quad (6.31)$$

where  $\mathbf{W}$  is a constant representing the angular decorrelation of the tropospheric errors.

A similar model is used for the correlation between observations made to the same satellite from different receivers. In this case, the tropospheric delay errors should cancel exactly if the receivers are collocated, and become completely uncorrelated if the receivers are separated by a great distance. Using an exponential decay model, the covariance is expressed as :

$$\mathbf{s}_1^{AB}(T) = E(\mathbf{dT}_1^A \cdot \mathbf{dT}_1^B) = E(m(\mathbf{e}_1^A) \mathbf{dT}_Z^A \cdot m(\mathbf{e}_1^B) \mathbf{dT}_Z^B) = m(\mathbf{e}_1)^2 \exp(-d/D) \cdot \mathbf{s}_Z^2(T) \quad (6.32)$$

where  $d$  is the distance between receivers and  $D$  is a constant representing the correlation distance for the zenith tropospheric errors. In addition, the mapping function to the satellite is assumed to be identical at both receivers, which is admissible considering that a 100 kilometre separation results in only a  $0.3^\circ$  change in elevation angles.

Note that Eq. (6.32) assumes that the zenith tropospheric variance is equal at both receivers, which is compatible with the assumption that the accuracy of the base refractivity model is the same everywhere. However, it is known that changing weather affects the performance of the tropospheric models through their unpredictable effects on temperature, pressure and humidity (Gregorius and Blewit, 1998). This in turn implies that for receivers separated at distances similar to the scale of weather systems (i.e. hundreds of kilometres) this assumption may not exactly hold. However, weighing the increase in variance modelling accuracy against the complexity of modelling such an effect, a constant tropospheric variance must be assumed.

The final covariance to be considered is that of two observations made between different satellites and receivers. A composite of Eq. (6.31) and (6.32) provides the necessary model, namely:

$$\begin{aligned} \mathbf{s}_{12}^{AB}(T) &= E(\mathbf{dT}_1^A \cdot \mathbf{dT}_2^B) = E(m(\mathbf{e}_1^A) \mathbf{dT}_{Z1}^A \cdot m(\mathbf{e}_2^B) \mathbf{dT}_{Z2}^B) \\ &= m(\mathbf{e}_1) m(\mathbf{e}_2) \exp(-\mathbf{q} / \Omega) \exp(-d / D) \cdot \mathbf{s}_Z^2(T) \end{aligned} \quad (6.33)$$

Taking Eqs. (6.31), (6.32) and (6.33) and substituting them into the double-differenced variance equations contained in Eq. (6.6) yields the theoretical tropospheric model which, after collecting terms, is :

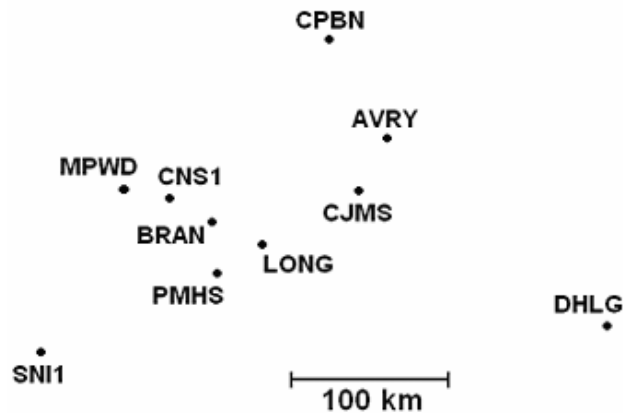
$$\begin{aligned} \mathbf{s}^2(\nabla\Delta_{12}^{AB}T) &= 2 \cdot (m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)^2 - 2m(\mathbf{e}_1)m(\mathbf{e}_2) \exp(-\mathbf{q}_{12} / \Omega)) \cdot (1 - \exp(-d_{AB} / D)) \cdot \mathbf{s}_Z^2(T) \\ \mathbf{s}^2(\nabla\Delta_{12}^{AB}T, \nabla\Delta_{13}^{AB}T) &= 2 \cdot \left( \begin{array}{l} m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)m(\mathbf{e}_3) \exp(-\mathbf{q}_{23} / \Omega) - m(\mathbf{e}_1)m(\mathbf{e}_2) \exp(-\mathbf{q}_{12} / \Omega) \\ - m(\mathbf{e}_1)m(\mathbf{e}_3) \exp(-\mathbf{q}_{13} / \Omega) \end{array} \right) \\ &\quad \cdot (1 - \exp(-d_{AB} / D)) \cdot \mathbf{s}_Z^2(T) \\ \mathbf{s}^2(\nabla\Delta_{12}^{AB}T, \nabla\Delta_{12}^{AC}T) &= (m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)^2 - 2m(\mathbf{e}_1)m(\mathbf{e}_2) \exp(-\mathbf{q}_{12} / \Omega)) \\ &\quad \cdot (1 + \exp(-d_{BC} / D) - \exp(-d_{AB} / D) - \exp(-d_{AC} / D)) \cdot \mathbf{s}_Z^2(T) \\ \mathbf{s}^2(\nabla\Delta_{12}^{AB}T, \nabla\Delta_{13}^{AC}T) &= \left( \begin{array}{l} m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)m(\mathbf{e}_3) \exp(-\mathbf{q}_{23} / \Omega) - m(\mathbf{e}_1)m(\mathbf{e}_2) \exp(-\mathbf{q}_{12} / \Omega) \\ - m(\mathbf{e}_1)m(\mathbf{e}_3) \exp(-\mathbf{q}_{13} / \Omega) \end{array} \right) \\ &\quad \cdot (1 + \exp(-d_{BC} / D) - \exp(-d_{AB} / D) - \exp(-d_{AC} / D)) \cdot \mathbf{s}_Z^2(T) \end{aligned} \quad (6.34)$$

where the only three unknowns required to completely describe the double-differenced tropospheric variance are  $\mathbf{W}$ ,  $D$  and  $\mathbf{s}_Z^2(T)$ .

Inspection of Eq. (6.34) shows that the variances all decrease towards zero as receiver separation goes to zero and the satellite line-of-sights become coincident. This is expected since the tropospheric errors become increasingly correlated and thus cancel out. At the other extreme, the double differenced variance simply becomes twice the accuracy of the tropospheric model itself, as no cancellation occurs. This is theoretically appealing, as the ultimate lower accuracy limit of the double-differenced observations must be that of the undifferenced observations themselves.

#### 6.4.2 -Empirical Assessment of the Tropospheric Variance Behaviour

To determine the feasibility of modeling the residual tropospheric covariance as described above, a network of ten reference stations forming part of the Southern California Integrated GPS Network (SCIGN) was used. The network configuration is shown in Figure 6.10. Station elevations ranged from -103 to 933 meters above the WGS84 ellipsoid and receiver separations varied from 30 to 464 kilometres. Data was collected at 30 second intervals using ASHTECH Z-XII3 receivers. Reference coordinates were generated by processing a 24hr session of data in a simultaneous adjustment of all observations available and a solution regularized by the centre of mass constraint discussed in Chapter 5. Processing resulted in 1203 L1 and L2 ambiguities, of which 143 were independent. 83 of these ambiguities could be successfully resolved, yielding a total of ~170 000 fixed L1/L2 data points for analysis since only observations with fixed L1 and L2 ambiguities were considered.



**Figure 6.10.** Network Configuration for Tropospheric Study.

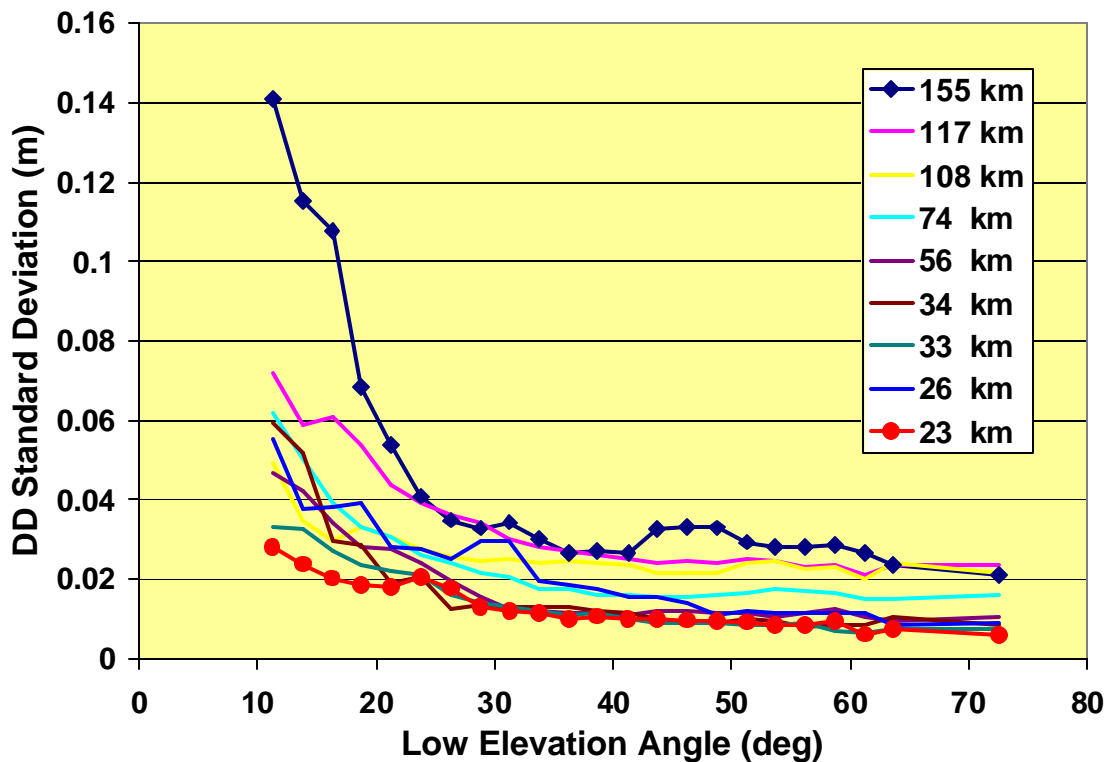
To extract the tropospheric error, the ionospheric-free combination (Hofmann-Wellenhof et al, 1994) is created via :

$$\mathbf{f}_{IF} = \mathbf{f}_{L1} - \frac{f_{L2}}{f_{L1}} \cdot \mathbf{f}_{L2} \quad (\text{cycles}) \quad (6.35)$$

where  $f_{L1}$  and  $f_{L2}$  are the carrier phase measurements in cycles on the L1 and L2 frequencies,  $f_{L1}$  and  $f_{L2}$ . As will be shown in Chapter 7, this combination removes the effect of the ionosphere. The tropospheric error in metres is the same as on the L1 and L2 frequencies, but a propagation of errors analysis shows that the multipath and noise variance in metres is increased by a factor of 3.2. The wavelength of this observable is 0.4844 metres. The fixed ambiguities were removed from the L1 and L2 observations prior to the application of Eq. (6.35) as well as the theoretic ranges, resulting in a direct estimate of the tropospheric error for each observation pair.

As in the case of the multipath and noise studies, the resulting double-differenced ionospheric-free residuals were binned into elevation ranges according to the low satellite of the double difference and the variance of each bin calculated. The result for a sample of baseline lengths is presented in Figure 6.11. A previous study of Radovanovic et al (2001) found that a similar model (UNB3) contained significant, receiver-height based biases, but no such biases were found in the UNB2 model. This is a finding consistent with that of Collins (1999). The elevation dependence of the tropospheric variance is clearly evident, in accordance with Eq. (6.34). In addition, the variance curve is scaled upwards as the baseline length increases, which also follows the behaviour predicted by Eq. 6.34.

The remaining problem is to determine a value of the three tropospheric model coefficients,  $W$ ,  $D$  and  $s_Z^2(T)$ . The process of calculating the variances shown above for the entire network resulted in 810 variance ‘‘observations’’, as the data from each of the 45 baselines was binned into 18 5° elevation bins. For a given set of initial approximates for the model parameters, the theoretical variance of each bin/baseline pair could be calculated using the first expression in Eq. (6.34). A least-squares adjustment then determined the values of the tropospheric model parameters that resulted in values of  $W$ ,  $D$  and  $s_Z^2(T)$  which minimised the overall error of the observed binned variances compared to the theoretical variance. An average multipath variance was also incorporated into the problem, as individual multipath variances for each station were unknown. The resulting parameter estimates are shown in Table 6.1.



**Figure 6.11.** Double-Differenced Residual Error Standard Deviations for Various Baseline Lengths

**Table 6.1.** Estimated Tropospheric Model Parameters.

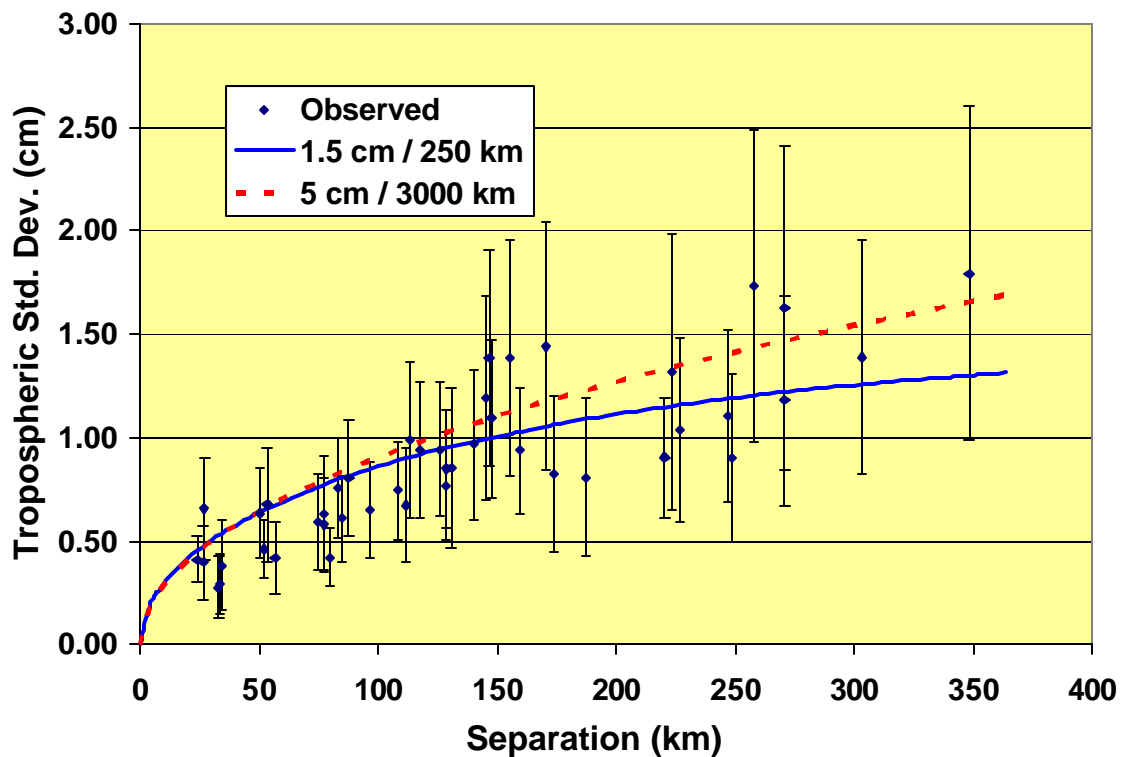
Epoch	Correlation Angle (degrees)	Correlation Distance (km)	Zenith Tropospheric Std. Dev. (cm)	Average Zenith Multipath Std. Dev. (mm)
<i>November 11, 2001</i>	14.3	253	1.51	4.2
<i>June 1, 2002</i>	12.2	232	1.52	3.8
<i>June 22, 2002</i>	12.0	240	1.48	3.9



The experiment described above was repeated for two other epochs spaced 6 and 7 months later. The results from these sessions are also contained in Table 6.1. Several conclusions can be drawn from these results. Firstly, there is little difference between the epochs separated by one month, indicating that the tropospheric variance parameters are stable over short time periods and as such do not need to be reevaluated frequently in practical applications. Surprisingly, these parameters are stable even over different seasons. Also, note that the multipath standard deviations listed are for the ionospheric-free observable and should be divided by a factor of 6.6 to arrive at the multipath standard deviation on the L1 signal. Comparison to the results in Section 6.3 indicate that these values are not unusual.

The zenith tropospheric standard deviation is significantly lower than the value of 5 centimetres reported by Collins and Langley (1998). However, this value was determined solely using radiosonde data and the resulting accuracy is reported to be at the centimetre level. Regardless, the above adjustment for the tropospheric variance models was repeated, holding the value of the zenith tropospheric standard deviation at 5 centimetres. This yielded the same correlation angle and multipath variance, but the corresponding correlation distance was extended to approximately 3000 kilometres. Figure 6.12 shows the observed mean zenith tropospheric variances for each baseline. The error bars correspond to the range of observed zenith tropospheric variances for each baseline. The two smooth lines correspond to the theoretical zenith variances calculated using zenith variance / correlation length combination of ( 1.5 cm / 250 km ) and ( 5 cm / 3000 km ), respectively.

The results immediately show that the two models do not significantly differ over the range of baseline lengths considered. On one hand, this is unfortunate since it implies that it is not possible to effectively separate the actual tropospheric variance and correlation distances using the experiment conducted. Rather, only the ratio of the two quantities is determinable. Conversely, the results also suggest that if receiver separations are limited



**Figure 6.12.** Comparison of Observed and Theoretical Zenith Tropospheric Variances for Various Separations.

to several hundred kilometres, it is not necessary to separate these quantities for accurate variance modelling and a solution can be made by constraining one of the parameters based on external knowledge.

### **6.5 -Ionospheric Error**

There exists above the Earth a region of positively and negatively charged ions, created by the interaction of the upper atmosphere with incoming ultraviolet light produced by the Sun. This region is known as the ionosphere, and its interaction with electromagnetic waves in the radio frequency is of great importance to satellite-based navigation systems.

The physics behind the effects of the ionosphere on radio waves is well documented by Komjathy (1997) and the detail provided here is only meant to be cursory. Essentially, unlike the troposphere, the refractive index of the ionosphere is dispersive at radio

frequencies, and as a result the index of refraction is not equal for different frequencies. Also, a further effect is that the velocity of the energy of a given radio wave (or group velocity) is not the same as the *apparent* velocity of its phase (or phase velocity). Thus the code information and carrier of a ranging signal are retarded by different amounts. Denoting the indices of refraction for the group and phase as  $n_{gr}$  and  $n_{ph}$ , the following relation holds (Hofmann-Wellenhof et al, 1994):

$$n_{gr} = n_{ph} + f \cdot \frac{dn_{ph}}{df} \quad (6.36)$$

where  $f$  is the frequency of the wave and the derivative is zero for non-dispersive media; hence the group and phase delays are identical for the troposphere.

A first-order approximation phase index of refraction is given by (Seeber, 1993) as

$$n_{ph} = 1 - \frac{40.3 \cdot N_e}{f^2} \quad (6.37)$$

where  $N_e$  is the electron density, and it always positive. Taking the derivative of Eq. (6.37) and substituting the result into Eq. (6.36), it can be shown that the group index of refraction is as much greater than one as the phase index of refraction is below one. Thus the existence of the ionosphere retards the group as much as it advances the phase. Note that this implies that the phase velocity is greater than the speed of light in vacuo, but that this limit only applies to the transfer of energy and that the phase velocity is only an apparent velocity.

Overall, the group range effect of the ionosphere is given by (Klobuchar, 1996)

$$I = \frac{40.3 \cdot TEC}{f^2} \quad (\text{metres}) \quad (6.38)$$

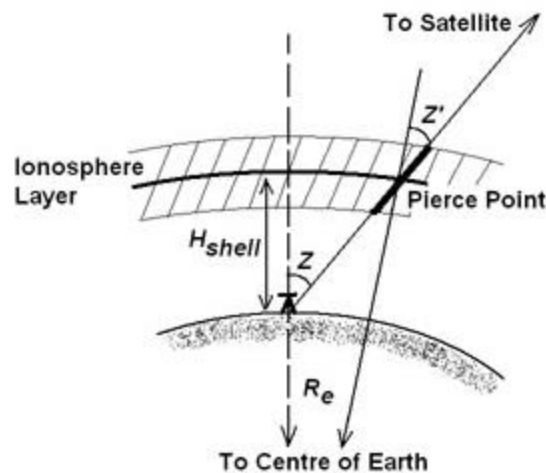
where  $TEC$  is the total electron content along the line of sight in units of  $10^{16}$  electrons per  $m^2$ . The corresponding phase ranges are shortened by this amount.

### 6.5.1 - Klobuchar Ionospheric Model

The usual way to model the ionospheric effect is to assume that the entire ionosphere, which ranges in altitude from 50 to 1500 kilometres, is concentrated in a thin shell at some height  $H_{shell}$  above the Earth's surface. The height used often varies from 300 to 500 kilometres (Schaer, 1999) but is typically selected at 350 kilometres. Referring to Figure 6.13, the resulting  $TEC$  affecting a given observation is

$$TEC = \frac{VTEC}{\cos z'} \quad (6.39)$$

where  $VTEC$  is the vertical total electron content at the ionospheric pierce-point compressed to the shell, and  $z'$  is the zenith angle at the ionospheric point.



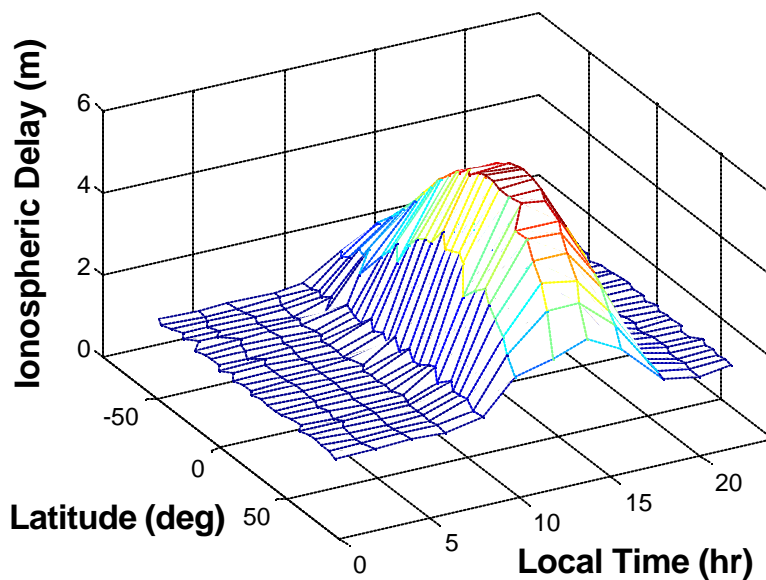
**Figure 6.13.** Relationship between  $TEC$  and  $VTEC$ .

The value of  $z'$  is related to the height of the shell and the zenith angle,  $z$ , at the station via the equation (Hofmann-Wellenhof et al, 1994) :

$$z' = \text{asin} \left[ \frac{R_E}{R_E + H_{shell}} \cdot \sin z \right] \quad (6.40)$$

where  $R_E$  is the nominal radius of the Earth.

Putting Eq. (6.39) into Eq. (6.38) shows that the ionosphere can be modelled by the familiar form of a zenith delay, multiplied by a corresponding, elevation dependant mapping function. The remaining problem is to determine the *VTEC* corresponding to a particular observation. The standard model used in GPS positioning is that of Klobuchar (1986), which parameterizes the *VTEC* with a half-cosine model, dependant on the local solar time, the latitude of the ionospheric point and several empirical constants. These constants are broadcast with the GPS navigation message and are usually updated every 10 days (Newby, 1992). Figure 6.14 shows the resulting vertical ionospheric delay estimates given by the Klobuchar model broadcast in June of 2002, and clearly shows that the ionospheric effect is greatest during the daytime at 14:00 hr local time, and near the equator.



**Figure 6.14.** Vertical Zenith Delay Estimates from Klobuchar Ionospheric Model (June, 2002)

The actual *VTEC* is strongly correlated to the number of sunspots, and as a result, varies according to an approximately 11-year cycle (Klobuchar, 1996). At the time of writing, the ionospheric effect was near the maximum of the cycle. In addition, the average *VTEC* changes over the year, peaking in the Northern Hemisphere in the winter months (Kleusberg, 1998). Furthermore, sudden ionospheric irregularities, such as those caused by solar storms, can drastically and suddenly change the *VTEC* over a particular region, an effect that sometimes manifests itself as the aurora borealis in the auroral zones (Skone, 1998).

### 6.5.2 -Theoretical Modelling of the Ionospheric Error

Since the distribution of *VTEC* is highly variable in space and time, it is difficult to make general statements regarding the accuracy of a particular ionospheric model (Schaer, 1999). As a result, the goal of this work in modelling the variance-covariance properties of the ionosphere is to provide a general specification on the stochastic proprieties of the ionospheric error on a large scale. Small scale disturbances caused by solar storms and scintillation are not considered.

The starting point of the model development will be the assumption that the majority of the ionospheric error results from a mis-modelling of the vertical total electron content. Therefore, the error in estimating the ionospheric delay (or advance),  $dI_1^A$ , is given by

$$dI_1^A = \frac{40.3}{f^2} \cdot \frac{dVTEC_1^A}{\cos z'} \quad (\text{metres}) \quad (6.41)$$

This implies that the variance of the observation can be calculated as

$$s_1^{A^2}(I) = \left( \frac{40.3}{f^2} \right)^2 \frac{s_A^2(VTEC)}{\cos^2 z'} \quad (6.42)$$

where  $\mathbf{s}_A^2(VTEC)$  is the variance of the *VTEC* estimation error at the ionospheric point for the observation.

Aside from localized disturbances, the physical distribution of the *VTEC* is smoothly varying (Schaer, 1999). Consequently, the errors in the *VTEC* estimation are smoothly varying in space as well and so the spatial correlation of the *VTEC* is modelled in the same way as the troposphere, using an exponential decay shaped by a correlation distance  $D$ . Thus, for two observations, the correlation between the *VTEC* errors affecting them is determined by calculating the distance between their ionospheric pierce-points. The covariance between any two observations is given by

$$\begin{aligned} \mathbf{s}_{12}^{AB}(I) &= \left( \frac{40.3}{f^2} \right) \frac{\mathbf{s}^2(VTEC)}{\cos z_1^{A'} \cdot \cos z_2^{B'}} \cdot e^{-d_{12}^{AB}/D} \\ &= \frac{\mathbf{s}_Z^2(I)}{\cos z_1^{A'} \cdot \cos z_2^{B'}} \cdot e^{-d_{12}^{AB}/D} \end{aligned} \quad (6.43)$$

where  $d_{12}^{AB}$  is the distance between the two pierce points, and  $z_A^{1'}$  and  $z_B^{2'}$  are their respective zenith angles. Eq. (6.43) can be modified to express the *VTEC* variance as a zenith ionospheric delay,  $\mathbf{s}_Z^2(I)$ , by multiplying the terms in brackets by the *VTEC* variance to arrive at a variance in  $m^2$ .

Incorporating Eq. (6.43) into Eq. (6.6) produces the resulting expressions for the double-differenced ionospheric error, after simplification and collecting like terms :

$$\begin{aligned}
\mathbf{s}^2(\nabla\Delta_{12}^{AB}I) &= 2 \cdot \left( m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)^2 - 2m(\mathbf{e}_1)m(\mathbf{e}_2)\exp(-d_{12}/D) \right) \cdot (1 - \exp(-d_{AB}/D)) \cdot \mathbf{s}_Z^2(I) \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}I, \nabla\Delta_{13}^{AB}I) &= 2 \cdot \left( \begin{aligned} &m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)m(\mathbf{e}_3)\exp(-d_{23}/D) - m(\mathbf{e}_1)m(\mathbf{e}_2)\exp(-d_{12}/D) \\ &- m(\mathbf{e}_1)m(\mathbf{e}_3)\exp(-d_{13}/D) \end{aligned} \right) \\
&\quad \cdot (1 - \exp(-d_{AB}/D)) \cdot \mathbf{s}_Z^2(I) \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}I, \nabla\Delta_{12}^{AC}I) &= \left( m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)^2 - 2m(\mathbf{e}_1)m(\mathbf{e}_2)\exp(-d_{12}/D) \right) \\
&\quad \cdot (1 + \exp(-d_{BC}/D) - \exp(-d_{AB}/D) - \exp(-d_{AC}/D)) \cdot \mathbf{s}_Z^2(I) \\
\mathbf{s}^2(\nabla\Delta_{12}^{AB}I, \nabla\Delta_{13}^{AC}I) &= \left( \begin{aligned} &m(\mathbf{e}_1)^2 + m(\mathbf{e}_2)m(\mathbf{e}_3)\exp(-d_{23}/D) - m(\mathbf{e}_1)m(\mathbf{e}_2)\exp(-d_{12}/D) \\ &- m(\mathbf{e}_1)m(\mathbf{e}_3)\exp(-d_{13}/D) \end{aligned} \right) \\
&\quad \cdot (1 + \exp(-d_{BC}/D) - \exp(-d_{AB}/D) - \exp(-d_{AC}/D)) \cdot \mathbf{s}_Z^2(I)
\end{aligned} \tag{6.44}$$

where  $m(e)$  and  $D$  now refer to the ionospheric mapping functions and correlation distances, respectively. Two distance quantities must be calculated, of the form  $d_{AB}$  and  $d_{12}$ . The first refers to the distance separation between the receivers and the second is the average separation of the pierce points corresponding to satellites  $I$  and  $2$ . Note that although the ionospheric model is similar in form to the tropospheric model, it contains one less term, as the  $D$  parameter controls both the inter-station and inter-satellite correlations and all pierce-points are referred to the ionospheric shell. Using different considerations, Yang and Goad (1997) also arrive at an exponential decorrelation model for the double-differenced ionospheric variance, but do not include the correlations between different double-differences, nor the mapping function dependant portion of the variance model.

### 6.5.3 -Empirical Assessment of Ionospheric Variance Behaviour

The same data set used in the previous section on tropospheric variance behaviour was used to investigate the feasibility of the ionospheric model presented. To extract the ionospheric error, the following linear combination was formed for all the fixed-ambiguity data available :

$$\mathbf{f}_{ION} = (\mathbf{I}_{L1} \cdot \mathbf{f}_{L1} - \mathbf{I}_{L2} \cdot \mathbf{f}_{L2}) \frac{f_{L2}^2}{f_{L2}^2 - f_{L1}^2} = \mathbf{dI} \quad (\text{metres}) \tag{6.45}$$



where the known ambiguities have been removed from the phase measurements  $f_{L1}$  and  $f_{L2}$ . The remaining quantities are a scaled value of the residual ionospheric error and an amplified noise and multipath term, whose variance in metres is approximately 3.8 times greater than the multipath and noise on the L1 carrier alone.

Figure 6.15 shows the residual *undifferenced* ionospheric errors calculated from code data collected by the network receivers during the June, 2002 campaign, plotted as a function of local time and scaled to the zenith using Eq. (6.39). Results before the Klobuchar model was applied and after are shown. As can be seen, the Klobuchar model does not significantly improve the scatter (and hence the variance) of the data, but it does remove the mean diurnal trend. Assuming an average zenith code noise and multipath standard deviation of 25 cm (Langley, 1998; Tiberius et al, 1999) the resulting ionospheric standard deviation was estimated to be 90 cm, independent of the time of day. These are results comparable to those obtained by Newby (1992), who performed a study during the previous solar cycle and found no significant difference in the residual accuracy between night time and day time conditions and estimated a model accuracy of 1 metre. In addition, the results for the November, 2001, data set are plotted for comparison. Interestingly, while the actual ionospheric effect is markedly different for the two sessions, the residual ionospheric error after applying the Klobuchar model has approximately the same variance in both seasons. However, this accuracy consistency is not expected to hold in periods of unsettled ionospheric activity, such as in the auroral region or during ionospheric storms.

Having established an estimate of the ionospheric variance, the remaining task is to evaluate the correlation length of the residual ionospheric error. A similar approach to that demonstrated in the tropospheric study was used. The available fixed ambiguity data was binned according to elevation angle and receiver separation, and the observed variance for each bin was calculated, along with the theoretical multipath variance, according to the values determined in the tropospheric study. Using the observed minus multipath variances as observations, a correlation length and zenith variance was found that minimized the error in the observed and predicted bin variances. A comparison between the observed and

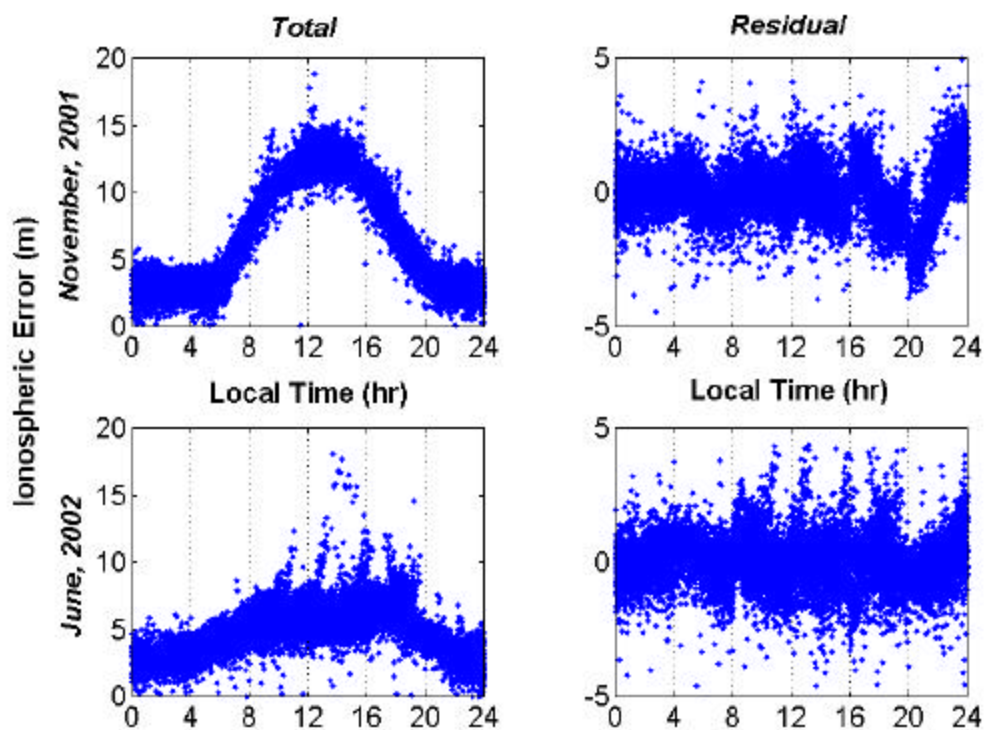


Figure 6.15. Total and Residual Ionospheric Error as a Function of Local Time (from code data).

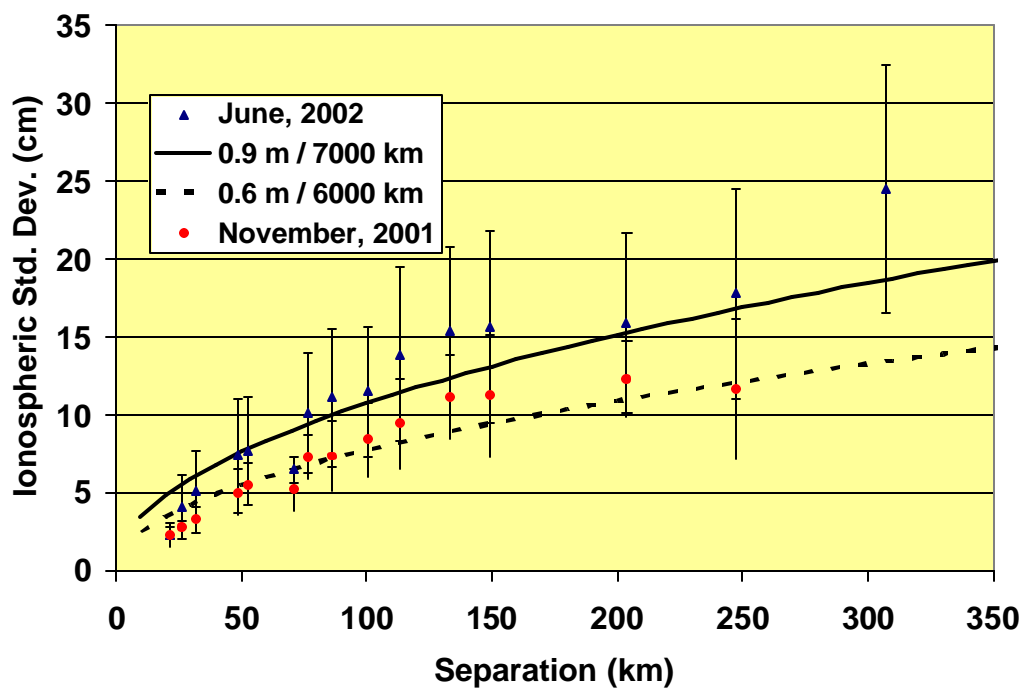


Figure 6.16. Comparison of Observed and Theoretical Zenith Ionospheric Variances for Various Separations in Winter and Summer.

theoretical zenith ionospheric variances is provided in Figure 6.16 for the November, 2001 and June, 2002 data sets. The error bars shown correspond to the spread of the observed variances at each separation. There exists a definite difference between the trends in the zenith variances for the two sessions, with the summer ionospheric errors being worse but more highly correlated with pierce-point separation. An analysis of a July, 2002 data set indicated an identical zenith ionospheric variance as the June, 2002 set, but with a decorrelation distance of 6000 km rather than 7000 km. This may imply that the zenith variance varies slowly over the course of the year, whereas the correlation distance is sensitive to the current state of the ionosphere. It should be noted that no ionospheric storm events were recorded for any of the sessions processed.

### 6.6 -Treatment of Temporal Correlations

One consequence of the high data rate capabilities of satellite-based navigation systems is the fact that the observations are *temporally correlated*. Essentially, the temporal correlation describes the correlation between two observations (typically from the same data stream) separated by some period of time. Given a data stream of zero-centred residuals of length  $L$ , the temporal correlation  $f(\mathbf{t})$ , of the series can be calculated as :

$$f(\mathbf{t}) = \frac{1}{\mathbf{s}^2} \frac{\int_{t=0}^{t=T-\mathbf{t}} r(t) \cdot r(t + \mathbf{t}) dt}{L - |\mathbf{t}|} \quad (6.46)$$

where  $\mathbf{t}$  is the time separation between the observations in the data stream and  $\sigma^2$  is the variance of the data stream (assumed to be constant over time). In the limit, as  $\mathbf{t}$  becomes zero the correlation function becomes equal to one, which is the maximum allowable correlation for any value of  $\mathbf{t}$ . A high value of  $f(\mathbf{t})$  indicates that samples separated in time by  $\mathbf{t}$  are highly correlated. In general, the temporal correlation between two different data streams,  $r(t)$  and  $g(t)$ , can be calculated by replacing one  $r(t)$  in Eq. (6.46) by  $g(t)$  and an appropriate value of the covariance between the data streams.

In the case of traditional terrestrial observations, observations taken at different epochs were either considered independent or contaminated by a constant bias. Unfortunately, this is not so in the case of GPS errors, which are often slowly varying over time. El-Rabbany (1994) was the first to study this property in depth, and proposed an exponential decay to model the temporal correlation of GPS errors :

$$f(t) = e^{-t/T} \quad (6.47)$$

where  $T$  is the temporal correlation length of the GPS errors, set at approximately 300s. This correlation length has been used in subsequent studies, including Han and Rizos (1995) and Howind et al (1999).

Given two GPS observations of equal variance separated by a time  $t$ , the covariance between the two observations is given by :

$$\mathbf{s}_{AB}(t) = f(t) \cdot \mathbf{s}^2 \quad (6.48a)$$

where  $\mathbf{s}^2$  is the variance of the data stream. The above can be expanded to the case of the covariance matrix for two sets of observations to yield

$$\begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{C} & f(t)\mathbf{C} \\ f(t)\mathbf{C} & \mathbf{C} \end{bmatrix} \quad (6.48b)$$

where  $\mathbf{C}_1$ ,  $\mathbf{C}_2$  are the variance-covariance matrix of the observations at epochs 1 and 2 and  $\mathbf{C}_{12}$  is the cross-covariance matrix for the observations. Note that since the double-

difference operator is linear, the full double differenced variance-covariance matrix can be expressed as

$$\mathbf{C}_{DD} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_2 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{B}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^T \end{bmatrix} = \begin{bmatrix} \mathbf{B} \cdot \mathbf{C} \cdot \mathbf{B}^T & f(t)\mathbf{B} \cdot \mathbf{C} \cdot \mathbf{B}^T \\ f(t)\mathbf{B} \cdot \mathbf{C} \cdot \mathbf{B}^T & \mathbf{B} \cdot \mathbf{C} \cdot \mathbf{B}^T \end{bmatrix} \quad (6.49)$$

where  $\mathbf{B}$  is the Jacobian with respect to the observations. The significance of the above expression is that the temporal correlation of a stream of double-differences is identical to the temporal correlation of the component observations.

El-Rabbanny (1994) shows that use of the exponential model has several advantages in the formation and inversion of the full variance-covariance matrix. For example, if the satellite geometry is constant over time and the variances vary smoothly with time, the inverse of the variance-covariance matrix becomes band-limited. This property was exploited by Han and Rizos (1995), who show that the temporal correlation of errors causes a scaling of the normal equations, with the factor

$$k = \frac{1 - f + 2f/n}{1 + f} \quad (6.50)$$

where  $f$  is equal to  $f(t)$ , with  $\tau$  equal to the sampling rate, and  $n$  is the number of samples used. The corrected normal equations  $\mathbf{N}$  and  $\mathbf{u}$  are simply multiplied by the above factor. Thus, the actual parameter estimates are not significantly affected by neglect of the temporal correlations, but the resulting variance-covariance matrix is scaled. Since  $k$  is usually less than one, the reported variances will be greater after correction for the temporal correlation. As most commercial processors neglect the temporal correlations, it

has become the norm for the statistical results from these packages to be considered grossly optimistic.

Eq. (6.50) further introduces the futility of dense data sampling in GPS. Assume that a 10 minutes segment of data has been collected and that the satellite visibility is constant. Since the satellite geometry varies slowly with time, it can be shown that the normal matrix formed via the standard summation of normal equations at every epoch is equivalent to a normal matrix formed by averaging the normal matrices formed by the first and last epochs and multiplying the resulting matrix by the number of observations. This can be expressed as

$$\mathbf{N}_{\text{Tot}} = \sum_i \mathbf{N}_i \cong n \cdot \frac{\mathbf{N}_1 + \mathbf{N}_2}{2} = n \cdot \mathbf{N}_{\text{Av}} \quad (6.51)$$

where  $\mathbf{N}_{\text{Tot}}$  is the total normal matrix, neglecting temporal correlations,  $\mathbf{N}_i$  is the normal matrix formed from each epoch and  $\mathbf{N}_{\text{av}}$  is the average single-epoch normal equation. Using Eq. (6.51) and (6.50), the normal matrix corrected for temporal correlations,  $\mathbf{N}_{\text{cor}}$ , will be :

$$\mathbf{N}_{\text{cor}} = \frac{1-f + 2f/n}{1+f} \cdot \mathbf{N}_{\text{Tot}} = \frac{(1-f) \cdot n + 2f}{1+f} \cdot \mathbf{N}_{\text{Av}} \quad (6.52)$$

The expected accuracy of the parameters is dependant on the inverse of the above matrix. Given two identical observing sessions, distinguished only by their sampling rates, the ratio of their expected overall standard deviations will be

$$\frac{s_1}{s_2} = \sqrt{\frac{1+f_1}{1+f_2} \cdot \frac{(1-f_2) \cdot n_2 + 2f_2}{(1-f_1) \cdot n_1 + 2f_1}} \quad (6.53a)$$

with

$$f_i = e^{-t_i/T} \quad (6.53b)$$

where  $n_1, n_2$  are the number of samples collected in each session,  $t_1, t_2$  are the sample spacings for the two sessions. The poor return on collecting higher data rates is quite shocking. For example, based on a 20 minute data span and a correlation period of 300 s, the improvement in collecting one sample every 5 seconds instead of 30 is only 2%! Truly, the most effective way to improve positioning accuracy is through lengthening the observation period, which not only allows errors to decorrelate, but also adds more variation in the satellite geometry observed.

### **6.7 -A Note on Inter-Frequency Correlation**

One aspect of correlation modelling that does not often receive direct treatment is the correlation of GPS errors between frequencies. For example, since the tropospheric error is due to the neutral atmosphere and not dependant on carrier frequency, it expected that the tropospheric error on two carrier phases made from the same receiver to the same satellite but on two different frequencies are identical. Conversely, errors such as noise are often considered to be independent between frequencies. This is not necessarily the case however, if L1 code is used to aid tracking of the L2 frequency, as in many receivers. Still, such effects are very difficult to quantify and as a result, the noise and multipath is considered uncorrelated between frequencies.

Perhaps the most significant inter-frequency correlation is due to the ionosphere. With reference to Eq. (6.41), the covariance between the ionospheric error on the L1 and L2 frequencies for a given observation can be written as :

$$\mathbf{s}_{L1L2}^2 = \frac{40.3^2}{f_{L1}^2 \cdot f_{L2}^2} \cdot \mathbf{s}_{TEC}^2 \quad (\text{metres}) \quad (6.54)$$

where  $\mathbf{s}_{TEC}^2$  is the variance of the *TEC* along the line of sight.

Given a set of dual frequency observations, the variance-covariance matrix for the set can be expressed by :

$$\mathbf{C}_l = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \quad (\text{cycles}) \quad (6.55a)$$

$$c_{11} = \frac{\mathbf{s}_T^2}{I_{L1}^2} + I_{L1}^2 \frac{40.3^2}{c^4} \mathbf{s}_i^2 + \mathbf{s}_n^2 + \mathbf{s}_m^2 \quad (6.55b)$$

$$c_{22} = \frac{\mathbf{s}_T^2}{I_{L2}^2} + I_{L2}^2 \frac{40.3^2}{c^4} \mathbf{s}_i^2 + \mathbf{s}_n^2 + \mathbf{s}_m^2 \quad (6.55c)$$

$$c_{12} = c_{21} = \frac{\mathbf{s}_T^2}{I_{L1} \cdot I_{L2}} + I_{L1} I_{L2} \frac{40.3^2}{c^4} \mathbf{s}_i^2 \quad (6.55d)$$

where  $\mathbf{s}_T^2$  is the tropospheric variance in metres<sup>2</sup>,  $\sigma_n^2$  and  $\sigma_m^2$  are the noise and multipath variances in cycles<sup>2</sup> and  $\mathbf{s}_i^2$  is the ionospheric variance in units of TEC<sup>2</sup>. The double-difference variance covariance matrix can be made via the appropriate substitution of the double-differenced variances into Eq. (6.55). Furthermore, to create the variance-covariance matrix for a set of dual-frequency observations, the inter-frequency and inter-spatial correlations between all lines-of-sights observed must be generated. Fortunately, these can be simply generated by first calculating the inter-spatial correlations of the lines-of-sights, and then dealing with all the inter-frequency correlations using Eq. (6.55).



The importance of treating inter-frequency correlations is two-fold. Firstly, proper treatment of the ionospheric correlation is necessary for proper dual-frequency ambiguity resolution. Traditional techniques of ambiguity resolution have relied on the formation of wide-lane and ionospheric free linear combinations (Raquet, 1998) to allow for ambiguity resolution on longer baselines due to the effects of the ionosphere. However, such techniques result in ad-hoc methods of resolving the ambiguities. Rather, by properly incorporating the ionospheric inter-frequency covariance into the variance-covariance matrix of the double-differences, the LAMBDA method is able to implicitly estimate the ionospheric effect affecting each observation and resolve the L1 and L2 ambiguities as integers. As a result, one finds that the ambiguity success rate calculated via Eq. (5.11) is drastically higher when the ionospheric inter-frequency correlations are included than if the ionospheric error is considered uncorrelated between frequencies.

The second application of inter-frequency correlations lies in the field of determining optimal linear combinations of carrier phases which allow the compression of two or more observations into one single pseudo-observation. This is discussed in Section 7.3.

### ***6.8 -Practical Determination of Variance-Covariance Models***

The preceding sections have established stochastic models for the major errors sources present in differential GPS and validated these models using empirical data collected. However, the key element of the modelling methodology presented is the solution for defining parameters (i.e. zenith variance, correlation distance, etc.) from the data collected. Some parameters, such as those of the tropospheric model, have been shown to be stable over time, whereas other, such as the ionospheric parameters, vary over the course of months. Also, it is expected that the defining parameters may change with location, the types of receivers used, and the environment surrounding the receivers. Thus a practical method of establishing the appropriate variance-covariance model for a network under consideration must be developed.

In engineering networks of extent less than a few kilometres, the dominant error effects will be noise and multipath. The noise variance model is simply determined via a zero-baseline test of the receivers to be used in the network. Since most engineering projects will use identical receivers at each station, only one pair of receivers need to be tested. However, since in practice it is difficult to accurately calculate the actual power received at each receiver, this test is not useful for calculating the zenith noise variance. Rather, its utility lies in establishing the *shape* of the noise mapping function.

Section 6.3 established that noise and multipath are both affected by the same mapping function – thus both effects may be grouped together into a single zenith variance value. Since the multipath effect depends on the reflective environment surrounding a receiver, multipath/noise zenith variances must be calculated for each receiver in the network. A practical method of doing this is to take one roving receiver to each site and place it a few metres away from the station to be analysed. Collecting thirty minutes of data creates a short-baseline on which the multipath variances at both receivers are expected to be identical, despite the fact the *actual* multipath quickly decorrelates spatially (Ray, 2000). Following the analysis of Section 6.3, combined with the now known noise mapping function allows the zenith multipath at the site to be established. This test can be repeated whenever it is felt that the reflective environment has significantly changed, such as when earthwork has been completed nearby. Otherwise, it is assumed that the multipath variance will remain quite stable, although a summer and winter value might be worth determining to account for the effects of snow cover, if appropriate.

If the network is of larger extent, the tropospheric and ionospheric stochastic models must be derived as well. The best method for accomplishing this is to use data collected at a larger reference receiver network that encompasses the site to be studied. For example, data from IGS stations may be used for this purpose, as well as data from a variety of scientific reference networks. The Scripps Orbit and Permanent Array Center (<http://sopac.ucsd.edu>) is an excellent source of data from more than 1000 sites located worldwide. The methods of Sections 6.4 and 6.5 can then be applied to establish the defining tropospheric and ionospheric parameters for the epoch under consideration. In

general, the tropospheric parameters are expected to be stable over time, whereas the ionospheric model should be updated once a week to take into account the variability of the ionosphere. Of course, if the engineering network is of small extent, then the actual values of the tropospheric and ionospheric models are not significant, and the results presented in this thesis may be used without loss of positioning accuracy.

## **7.0 - APPLICATION OF GLOBAL NAVIGATION SATELLITE SYSTEMS TO NETWORK POSITIONING**

Traditional methods of applying data collected with satellite-based navigation systems have been heavily biased by the heritage of these systems as navigation instruments and the high cost of available equipment. For example, most early GPS surveys and research were completed using pairs of receivers, resulting in *baseline processing* techniques that yielded the vector between two occupied stations. Networks were built up by moving the receivers to all points of interest and accumulating individual interstation vectors. A network “adjustment” of these vectors was then completed to calculate the positions of all stations in the network. Furthermore, due to limitations in computing power and a lack of knowledge of the stochastic properties of the errors affecting GPS, variance-covariance matrices used in the adjustment of these individual baselines were often simply diagonal, and at best only modelled the mathematical correlations of the formed double differences were considered. Finally, datum definition was typically only implicitly accomplished by assuming a “base station” as fixed, with known coordinates.

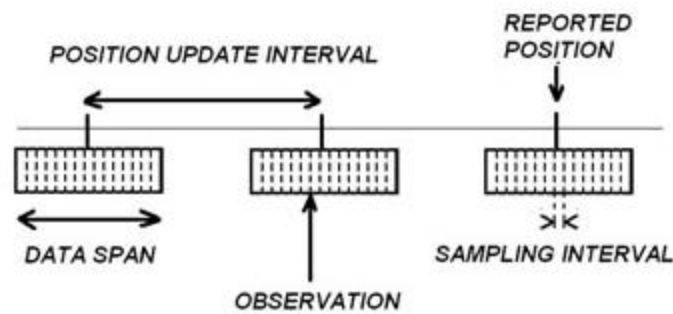
With a few exceptions, these processing methodologies remain today, primarily because despite improper treatment, the high accuracy GPS observations often yield results significantly better than those achievable by terrestrial methods, especially when long observation sessions are used and baseline lengths exceed a few kilometres. This chapter seeks to incorporate the datum definition and variance-modelling concepts developed in the previous chapters into a robust GPS processing methodology suitable for precise positioning. The capabilities of satellite-based systems for precise positioning on short and medium range static networks are analysed, and a novel application of variance-modelling for efficient network-based kinematic positioning is presented.

## ***7.1 - Application to Short Range Deformation Monitoring***

Short range deformation monitoring includes applications where the total network extent is less than a few kilometres. Such applications include dam and bridge deformation monitoring, the modelling of slope movements and building sway, among many others. The main appeal of applying GPS to these applications is the autonomous operation of the system, as the number of points to be monitored and the position update frequency can be quite high. For example, one application involving monitoring of the Diamond Valley Reservoir involves 250 points to be monitored every 12 hours (Duffy et al, 2001). Collecting such a great amount of data manually would be impossible.

### *7.1.1 - Principles of a GPS Based Deformation Monitoring System*

A hypothetical deformation monitoring system involving GPS would consist of a series of receivers, permanently mounted at monitoring points. Due to the temporal correlations of GPS observations, continuous data collection at a high data rate would be wasteful in terms of both memory and power, as closely spaced epochs do not provide independent information. Rather, an acceptable position update interval must be set by the project designer, possibly based on apriori assumptions of deformation rates. Within the constraint of the interval separation, the network designer can then chose an appropriate data sampling interval and data span to achieve a particular accuracy. The relationship between these concepts are shown in Figure 7.1. Note that increasing the sampling rate yields reducing returns as discussed in Section 6.6, due to the temporal correlation of the GPS errors, whereas increasing the data span can cause slow deformation trends to be absorbed into the individual static solutions, thus reducing the dynamic response of the system. For a typical application with an update interval of 1 hour, a suggested data span would be twenty minutes, with a sampling interval of 10 seconds.



**Figure 7.1.** Relationship Between Update Interval, Observation Span and Sampling Rate.

Data is then collected at the sites and transmitted to a central processing centre, possibly using radio, telephone or internet communication strategies. The data is then processed using a newly developed GPS processing package based on the results of this research, called PADRES- GPS (Position And Deformation Estimation System via GPS). At the central site, the data is first pre-processed to identify cycle-slips and obtain initial estimates for the receiver locations. At this stage, the station locations are compared to their locations in the previous epoch and any gross receiver movements are flagged (i.e. due to deliberate or inadvertent movement of the receivers to different locations by personnel). The observations are then processed first processed in float-mode using a full variance-covariance model and a suitable regularization scheme. The ambiguities are then resolved using the partial-fix method outlined in Chapter 5.

An important distinction at this point must be made between the *differencing base*, defined as the station used as the basis of all the double-differences formed and the concept of a *base station*, used in traditional processing modes as a datum definition. In the PADRES-GPS system, the differencing base is used to facilitate the ordering and formation of the normal equations by ensuring that all double-differences produced use one station in common. However, this station plays no special role in the regularization and its coordinates are treated as unknown in the same manner as all the other stations. Thus, the differencing base can be a station that is actively deforming, but should be located near the centre of the network to minimize the residual double-differenced errors, as discussed in

Section 7.2. It should also be a location where cycle-slips are a minimum and the multipath effect is low, thus providing a clean data stream to form the other double differences upon.

Several appropriate regularization strategies may be employed, depending on a priori knowledge of the network deformations. For example, if certain points are known to be stable with known coordinates, then the initial estimates of these points may be given very high weights. This will aid in ambiguity resolution by providing information regarding the relative geometry of the network. On the other hand, if no points may be considered stable or if their initial estimates are unknown, then the centre-of-mass regularization discussed in Chapter 5 should be used. Although no internal geometry information is provided by this method, the ambiguities will remain minimally biased. This is the default mode for processing, as it allows the greatest flexibility in the subsequent deformation monitoring stage.

Since the normal equations for the small-extent case are essentially rank deficient due to a poor control of translations, the results of a centre-of-mass regularization based on all the network points can be effectively transformed to those produced by a centre-of-mass regularization involving a subset of points via an S-transformation (Baarda, 1973). The S-transformation matrix can be defined as :

$$\mathbf{S}_{GD} = \mathbf{I} - \mathbf{G} \cdot (\mathbf{D}^T \cdot \mathbf{G})^{-1} \cdot \mathbf{D}^T \quad (7.1)$$

where  $\mathbf{G}$  and  $\mathbf{D}$  are the datum constraint matrices corresponding to different centre-of-mass regularization schemes and  $\mathbf{I}$  is an identity matrix. The results using the  $\mathbf{G}$  scheme can be transformed into the  $\mathbf{D}$  scheme using

$$\mathbf{d}_D = \mathbf{S}_{GD} \cdot \mathbf{d}_G \quad (7.2a)$$

$$\mathbf{C}_{xD} = \mathbf{S}_{GD} \cdot \mathbf{C}_{xG} \cdot \mathbf{S}_{GD}^T \quad (7.2b)$$

where  $\mathbf{d}_G$  and  $\mathbf{C}_{xG}$  are the correction vector and variance-covariance matrix calculated using the  $\mathbf{G}$  scheme. Note that it is assumed that both schemes use the same initial estimates for the parameters.

Eqs. (7.1) and (7.2) are particularly important if the position estimates of two epochs are to be compared. Obviously, the position results and the associated statistics are dependant on the regularization scheme used. As a result, for comparisons to be meaningful, two epochs must be compared using the same regularization scheme. This can be simply accomplished by using the  $\mathbf{S}$ -transform to convert the results of each epoch to the same regularization scheme. Again, it must be assumed that both epochs use the same initial estimates for station positions. The two epochs can then be compared for “significant” deformation using many different techniques – Yong-Qi (1983) gives an excellent review of prevailing theories.

In this thesis, a method outlined by Biacs (1989) for use in conventional networks is used, with modifications for the peculiarities involved in satellite-based systems. Given the position results of a two epochs, the displacement vector,  $\mathbf{d}$ , and associated variance-covariance matrix,  $\mathbf{C}_d$ , will be given by

$$\mathbf{d} = \mathbf{d}_2 - \mathbf{d}_1 \quad (7.3a)$$

$$\mathbf{C}_d = \mathbf{C}_{d2} + \mathbf{C}_{d1} - \mathbf{C}_{d2d1} - \mathbf{C}_{d1d2} \quad (7.3b)$$

where  $\mathbf{C}_{d2}$ ,  $\mathbf{C}_{d1}$  are the variances of the vector of corrections,  $\mathbf{d}_2$  and  $\mathbf{d}_1$ , for the two epochs and  $\mathbf{C}_{d2d1}$  and  $\mathbf{C}_{d1d2}$  are their cross-covariances. Note that the displacement variance-covariance matrix will contain the same near rank-deficiency in translation as the variance-covariance matrices of the original problem. Also, since both epochs have the same centre-of-mass, translations of the entire network are undetectable, although an overall rotation *is* detectable since the orientation is well defined by the satellite coordinates.



In general, two separate goals must be distinguished – the detection of a deformation, and its appropriate modelling. Deformation detection simply seeks to detect that a point or points in a network have been disturbed from their positions in a previous epoch. Deformation modelling, on the other hand, seeks to test the hypothesis that a particular deformation trend can be suitably described by a particular model. The latter is obviously the more difficult task, and is discussed by sources such as Teskey (1987) and Kuang (1996). In this thesis, the scope is limited to the detection of deformation of individual points, or conversely, the identification of stable points between epochs.

A method implemented by Grundig et al (1985), and Biacs (1989), determines stable points in a network via the following method. Given a set of network points, each point in the network is individually investigated for stability by partitioning the displacement vector as follows :

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_r \\ \mathbf{d}_i \end{bmatrix} \quad (7.4a)$$

with an associated weight matrix of

$$\mathbf{P}_d = \mathbf{C}_d^{-1} = \begin{bmatrix} \mathbf{P}_{rr} & \mathbf{P}_{ri} \\ \mathbf{P}_{ir} & \mathbf{P}_{ii} \end{bmatrix} \quad (7.4b)$$

where the  $i$  and  $r$  subscripts refer to the point under investigation and the remaining points, respectively, and the inverse of the displacement variance-covariance matrix is calculated as a reflexive generalized inverse (see Biacs, 1989 for implementation details). To effectively gauge whether the point under investigation has deformed or not, its solution must first be disengaged from its inclusion in the computational base. Thus an S-transform is used to convert the solution of the points from one where all points are used in the regularization to one where the point under investigation is excluded. This can be succinctly expressed as:

$$\mathbf{d}'_i = \mathbf{d}_i + \mathbf{P}_{ii}^{-1} \cdot \mathbf{P}_{ir} \cdot \mathbf{d}_r \quad (7.5)$$

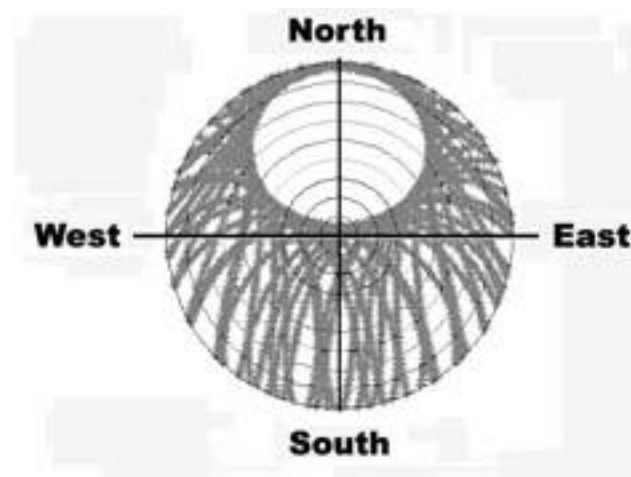
The test quantity used is then

$$T = \frac{\mathbf{d}'_i{}^T \cdot \mathbf{P}_{ii} \cdot \mathbf{d}'_i}{m \mathbf{s}_o^2} \quad (7.6)$$

where  $m$  is the dimension of the coordinate system (typically 3) and  $\mathbf{s}_o^2$  is the a posteriori variance factor, usually taken as one if appropriate variance-covariance modelling has been performed using the techniques of Chapter 6. The above test statistic follows the Fisher distribution with  $(m, \mathbf{Y})$  degrees of freedom. If the test quantity exceeds the allowable value of the Fisher distribution for the level-of-significance chosen, then the point cannot be assumed to be stable and so is flagged for possible removal from the regularization set. This procedure is repeated for the remaining points and the point with the largest significant test statistic removed from the list of stable points. The entire analysis is then repeated until no points can be considered displaced. Although computationally tedious, this procedure has been shown to be more sensitive than an inspection of individual error ellipsoids. Of course, the sensitivity of the test relies on the confidence level selected. However, appropriate selection of the confidence level and the associated problems of errors in hypothesis testing is beyond the scope of this thesis – interested readers are referred to Yong-Qi (1983).

An interesting peculiarity of GPS networks is that once displaced points have been identified, it is relatively simple to test the direction of the deformation. It is well known that the least accurate dimension observable in GPS networks is the zenith direction, due to the fact that all the satellites are above the horizon and so result in poor geometry. However, due to the configuration of the satellite orbits, there exists a polar “cap” that manifests itself as a lack of satellites visible in the north direction for stations in the northern hemisphere – the opposite effect occurs in the Southern. Figure 7.2 shows a

typical plot of the satellite orbits overhead of the network in question for a 24 hour period. As a result of this effect, when the variance-covariance matrix of a given position estimate is rotated into a local level frame, it was found that the principal axes of the error ellipsoid lies very nearly along the axes of the coordinate system – and thus the errors are largely decorrelated in these directions. The implication of this peculiarity is that if a station has been flagged as displaced, the components of the displacement vector in the north, east and zenith direction can be individually tested by dividing the movement in a particular direction by the standard deviation of the position estimate in that direction and comparing the resulting test statistic to the F-distribution with  $(1, \infty)$  degrees of freedom.



**Figure 7.2.** Typical Overhead Skyplot Showing Effect of Polar Caps.

### *7.1.2 - Test of the PADRES-GPS Software*

To field-test the capabilities of the PADRES-GPS software, a five receiver network was established in August, 2002, at the Province of Alberta Calibration Baseline, located east of the City of Edmonton. The network consisted of five concrete pillars with forced-centring plates, equipped with Trimble 4600LS GPS receivers. L1-only receivers were used since the additional cost associated with dual-frequency receivers precludes their use in large numbers. In addition, two stations were equipped with translation stages, which allowed deformations to be induced with an accuracy of 1/10th of a millimetre. The maximum station separation was 3 kilometres, and the average station spacing was 500 metres.

The first stage of this test was to establish an appropriate variance-covariance model for use in the subsequent processing. Due to the small extent of the network, the ionospheric and tropospheric variance model was identical to that presented in Chapter 6 for the July, 2002 SCIGN data set. The multipath variances for the stations were calculated by occupying each station for twenty minutes, with an additional (and identical model) receiver located 3 metres away. In this way, double-differenced observations were collected at various elevation angles, which could then be reduced to a single zenith multipath/noise variance for each station in the manner presented in Section 6.3. The resulting zenith multipath/noise variances for each station are shown in Table 7.1. A  $1/\sin(\epsilon)$  mapping function was found to suitably model the elevation variation in the variance. The results of Table 7.1 fall within the expected range of carrier phase multipath variance, based on prior experiences presented in Chapter 6, and indeed, the variations in the variances can be largely explained by the environment of the stations. Stations B, C and D, for example, lay inside fields, whereas stations A and E were located near paved roadways.

**Table 7.1.** Zenith Multipath/Noise Standard Deviations for Stations Involved in the Short Range Test.

Station	Multipath/Noise Zenith Standard Deviation (mm)
<i>A</i>	1.9
<i>B</i>	1.3
<i>C</i>	1.5
<i>D</i>	1.3
<i>E</i>	2.3

With the variance models established, the next stage was to collect data to test the deformation monitoring capabilities of the proposed system. For this purpose, three hours of data were collected on two different days. On the first day, the stations remained static, whereas on the second day known deformations were induced in stations B and C. These deformations were such that the station remained deformed at particular value for fifteen minutes. The data sampling rate was 5 seconds.

Reference coordinates were generated by processing the whole of the ninety minute static data for both days and averaging the results. Subsequently, the data was partitioned such that the position update rate was fifteen minutes with a ten minute data span. The spans were located such that they would encompass only one single deformation state. An adjustment of each epoch using PADRES-GPS was then compared to that of a commercial GPS processing package which processes data in a baseline-by-baseline fashion and relies on a float/fixed ambiguity resolution scheme. To allow meaningful comparison, the results of the PADRES-GPS software (which are not based on a fixed base station) were converted via S-transformation to a minimally constrained solution with station D considered fixed. The commercial processor used station D as a fixed base for all baselines processed.

Although there exist many varieties of GPS processors on the market, and the particular algorithms they apply vary widely, the author believes that the commercial processor chosen as a benchmark accurately reflects the state-of-the-art in general-purpose GPS processing. The processor used allows sophisticated adjustment of many processing parameters and has, in the author's experience, consistently provided superior position and accuracy estimates, in particular when considering problematic or short data sets. Of course, various processors can be optimized for particular applications, but the commercial processor used is representative of a high-quality general-purpose processor commonly used in industry.

### *7.1.3 -Deformation Detection Performance*

Figures 7.3-7.6 show the position estimates for each station over the epochs observed. The results are those stemming from a minimally constrained adjustment using the station D as a known station. The PADRES-GPS results are shown as solid symbols along with their associated one-sigma standard deviation error bars. The results from the commercial processor results (when available) are shown as open symbols. Finally, the theoretical position of the station at each epoch is shown by the lighter solid line. The overall RMS positioning error using the PADRES-GPS software are shown in Table 7.2, along with the average positioning accuracy predicted by the system, without any a posteriori scaling of the estimates.

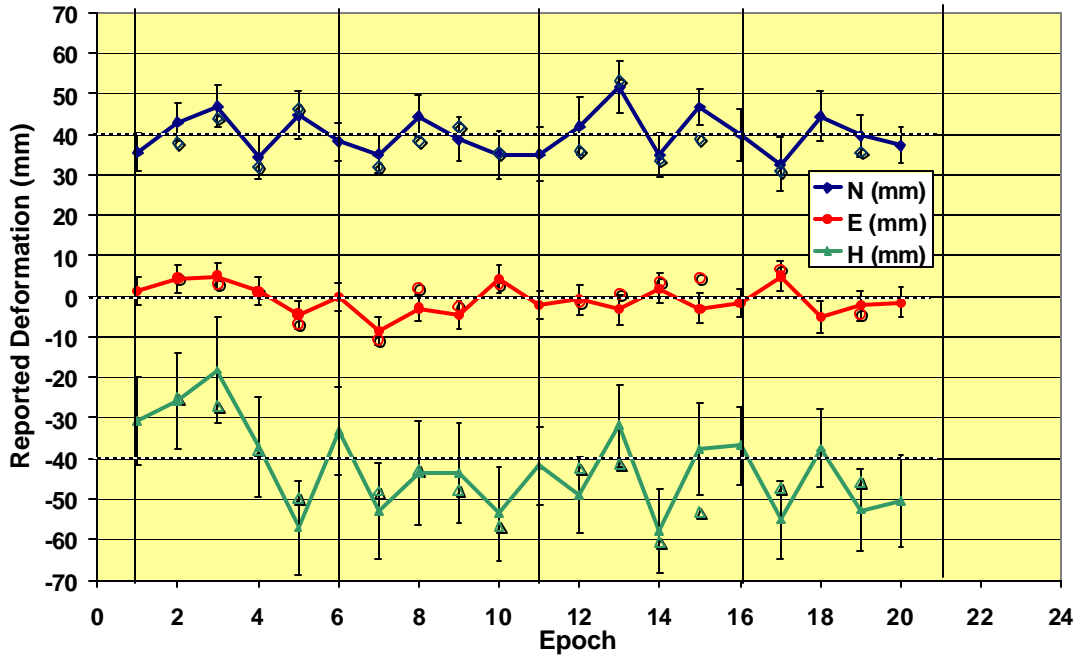


Figure 7.3. Position Estimates of Station A in Short Range Network Test.

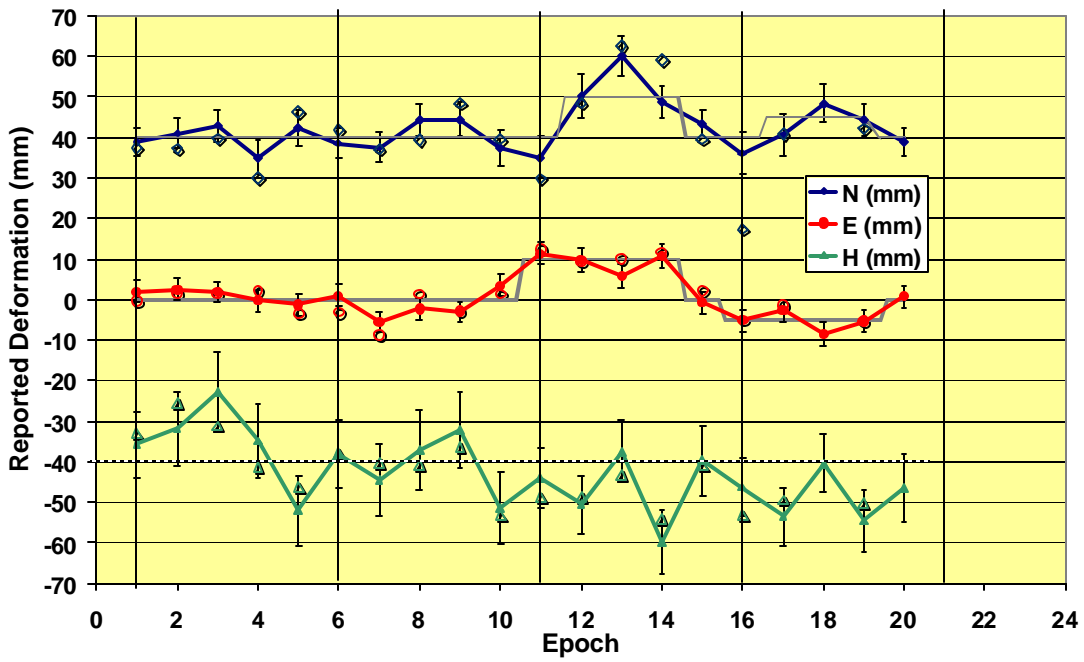


Figure 7.4. Position Estimates of Station B in Short Range Network Test.

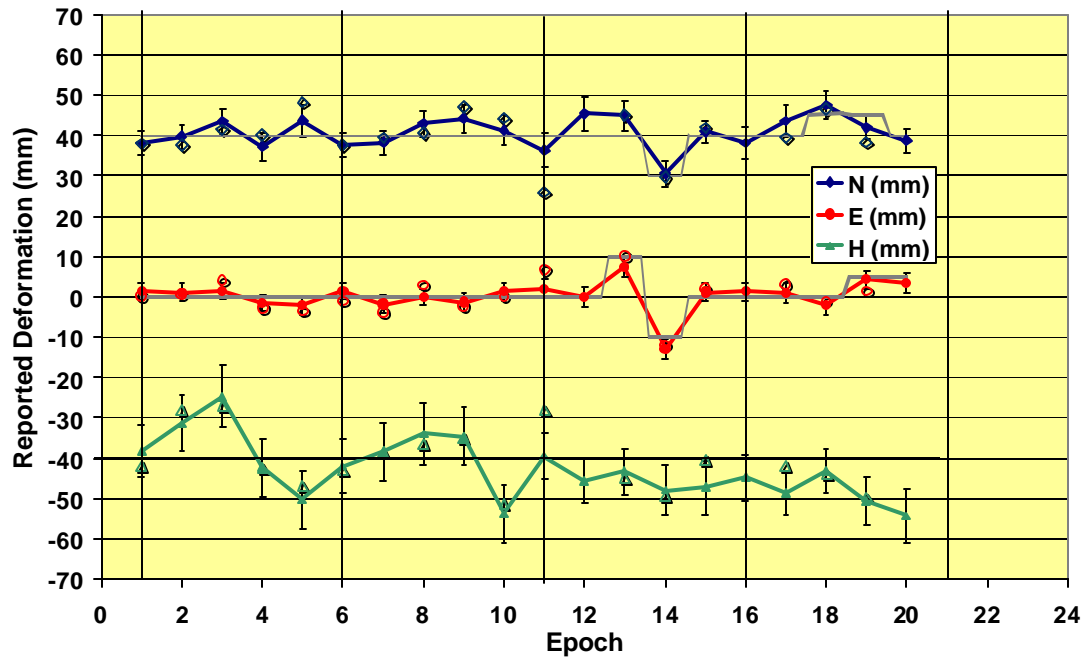


Figure 7.5. Position Estimates of Station C in Short Range Network Test.

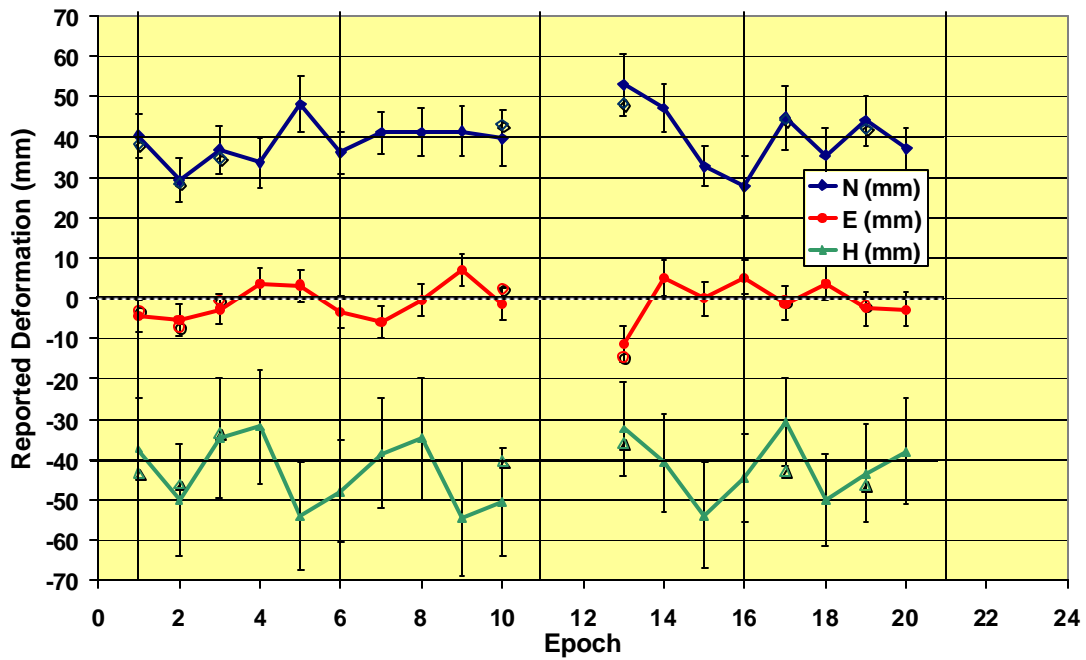


Figure 7.6. Position Estimates of Station E in Short Range Network Test.

**Table 7.2.** Observed and Predicted Positioning 1- $\sigma$  Accuracies of Stations During Short-Range Monitoring Test.

Station	Observed Accuracy (mm)			Predicted Accuracy (mm)		
	N	E	H	N	E	H
A	5.2	3.7	11.4	5.5	3.5	11.0
B	3.7	2.4	9.3	4.2	2.4	8.6
C	3.2	1.7	7.6	3.4	2.5	6.6
E	6.5	4.7	10.2	6.2	4.1	13.4

Note : Station D held fixed.

Several important findings are worth discussing. Firstly, in terms of position availability, the PADRES-GPS software was able to provide position estimates accurate to the centimetre level in 78 of 80 trials, whereas the commercial processor was able to provide centimetre level results in only 22 cases. Station E was particularly affected, possibly due to the severity of cycle-slipping due to the proximity of tree cover to the north of this station. This problem is aggravated by the relatively short time periods used. In the PADRES-GPS software, the partial-fix ambiguity strategy implies that the presence of unresolvable cycle-slips does not seriously impair the accuracy of the solution, since the fixed, high-accuracy carrier phase observation are weighted more heavily than their floating, low-accuracy counterparts. In the conventional processor, unresolvable cycle-slips can result in the processor reverting to float-only mode which provides position estimates accurate to the decimetre-level. The stability of the solution accuracy is of particular importance in deformation monitoring, as no reasonable assessment of the deformation can be made if two epochs are of grossly differing accuracies. In turn, this can pose a truly dangerous situation if continuity of monitoring is paramount, as in dam deformation monitoring. Note however, that when the conventional processor *is* capable of returning a valid solution, the position estimate is usually equivalent to that returned by PADRES-GPS within the accuracy reported by the processor developed.



A second key issue in deformation monitoring is the validity of the estimated accuracy of the position solutions. Position estimates are of little use in deformation monitoring if the analyst does not have accurate knowledge of their accuracy. A common shortcoming of most commercial processors is that the estimates of positioning accuracy returned are notoriously optimistic. This is largely due to neglecting temporal correlations (Section 6.6) and of the physical correlations between the observations. For example, in the test under study, the commercial processor steadfastly reported an accuracy of roughly 1mm laterally and 3 mm vertically for all stations. Inspection of the reported movements using this accuracy estimate would result in a plethora of false-deformations reported. On the other hand, as shown in Table 7.2, the actual and estimated accuracies are identical at any reasonable significance level.

However, this accurate assessment of the achievable positioning accuracy may come as a disappointment to those who believe that GPS is capable of achieving millimetre level precision in a semi-kinematic mode. This is quite simply not the case, mainly due to the effects of multipath and its temporal correlation. Even if an hour of data is used, the estimated one-sigma height accuracy remains limited to the 5 millimetre level – to achieve one millimetre accuracy in height at a 95% confidence level would require 10 such hour-long sessions combined together over the course of one day. Thus, if long-term deformation trends are to be studied, a permanent GPS array can provide excellent position estimates. However, if intra-day movements are to be studied, an accuracy of several millimetres *laterally* is expected to be the ultimate achievable accuracy.

Finally, an investigation of the results of Figure 7.3-7.6 brings to the fore the problem of robust deformation detection. Given that the error bars shown are at the one-sigma level, their length must be multiplied by 1.95 to form a 95% confidence interval. The simplest deformation monitoring system would be to take the observed displacements in the North, East and Zenith directions and divide them by their corresponding estimated accuracies. If this ratio exceeds 1.95 (the square root of  $F_{0.05}(1, \infty)$ ), a deformation is flagged as occurring at the 95% level. Several problems exist with this method. Firstly, the constraint of Station

D as a fixed-station implies that any deformations in D become mapped as deformations in the other stations, thus creating a false map of the actual deformation trends in the network. Also, while the position traces of stations B and C do follow the theoretic positions overall, it can be seen that upon scaling of the error bars, many of the smaller deformations will not be trapped at the required level of significance. A naive solution is to reduce the level of significance, and hence the size of the error bars, but this will result in greater false-deformations being reported during the static periods.

As a result, the more involved deformation detection approach outlined previously was implemented to identify deformed stations before using the simpler test to identify the most probable directions of the movements. An added benefit of this method is that it does not rely on the fixing of one station in particular, which not only improves the positioning accuracy of all the stations in general, but also frees the system from the requirement of a stable base station. At each epoch, the reported positions of the stations were compared to their reference positions and each station tested for stability at the 95% level using the rigorous test for localized deformation. Unstable stations were removed from the regularization base and the results correspondingly adjusted via appropriate S-transforms. The significant direction of movements for the unstable stations were then tested using the 95% error ellipsoid corresponding to the station. Unstable stations for which no significant direction can be established are termed “marginally stable.”

The resulting deformation report is shown in Table 7.3. Only significant epochs are shown, included if an actual deformation occurred during the epoch, or a deformation was reported. A summary of the performance is presented in Table 7.4. As can be seen, the PADRES-GPS software was very effective in detecting deformations of one-centimetre, even when the deformations were in several directions and at multiple stations. However, its performance was notably poorer when deformations at the 5 millimetre level were induced. Interestingly, at three epochs, Station D was reported to be either deformed or marginally stable. Reference to Figure 7.2 shows that the positions of all four other stations show apparent, but identical, movements in the height direction at these epochs, highlighting the problems encountered when fixing a base station.

**Table 7.3.** Deformation Report Produced During Short-Range Test and Actual Movements.

Epoch	Station	Reported	Actual
		<i>Overall</i>	<i>Overall</i>
		<i>Direction</i>	<i>Direction</i>
5	D	Movement	No Movement
		Z : 14.7 mm	
11	B	Movement	Movement
		E : 18.5 mm	E : 10 mm
12	B	Movement	Movement
		E : 11.6 mm	N : 10 mm E : 10 mm
	C	Marginally Stable	No Movement
	D	Marginally Stable	No Movement
13	B	Movement	Movement
		N : 9.2 mm E : 13.2 mm	N : 10 mm E : 10 mm
	C	Movement	Movement
E : 14.2 mm		E : 10 mm	
	D	Marginally Stable	No Movement
14	B	Movement	Movement
		N : 8.3 mm E : 8.1 mm	N : 10 mm E : 10 mm
	C	Movement	Movement
N : -9.5 mm E : -13.2 mm		N : -10 mm E : -10 mm	
16	B	Movement	Movement
		E : -6.5 mm	E : -5 mm
17	B	No Movement	Movement
			N : 5 mm E : -5 mm

**Table 7.3** (continued). Deformation Report Produced During Short-Range Test and Actual Movements.

18	<i>B</i>	Movement ----- N : 6.2 mm E : -7.6 mm	Movement ----- N : 5 mm E : -5 mm
	<i>C</i>	No Movement ----- N : 5 mm	Movement ----- N : 5 mm
19	<i>B</i>	Marginally Stable ----- N : 5 mm E : -5 mm	Movement ----- N : 5 mm E : -5 mm
	<i>C</i>	Movement ----- E : 5.5 mm	Movement ----- N : 5 mm E : 5 mm
20	<i>C</i>	Movement ----- E : 4.8 mm	Movement ----- E : 5 mm

**Table 7.4.** Summary of Deformation Monitoring Performance for Stable and Deformed Stations.

	<b>Total Instances</b>	<b>Incorrect Deformation Reported</b>	<b>Marginal Stability Reported</b>	<b>Missed Deformation</b>	<b>Correct Analysis</b>
<i>Stable</i>	<b>87</b>	1	3	n/a	<b>83</b>
<i>Deformed</i>	<b>13</b>	0	1	2	<b>10</b>

Finally, note that false deformation rate was roughly 1% – an attractive feature in an automated deformation monitoring system. The number of true deformations missed can be reduced by lowering the level of significance of the stability test, but a corresponding increase in the number of false alarms is an unfortunate consequence. To simultaneously improve both the false-alarm rate *and* the detection sensitivity, the actual positioning accuracy must be improved.

Improved accuracy of the deformation monitoring system can result from several changes. The author feels that the best method of improving the system accuracy is to incorporate dual-frequency receivers even on network extents of one kilometre. The reason for this is two-fold. Firstly, according to the results of Chapter 6, the ionospheric error quickly grows with receiver distance, and is responsible for approximately 5mm of error even on these short separations. Dual frequency processing removes this effect by modelling the correlation of L1 and L2 ionospheric errors, as discussed in depth in Section 7.3. Also, since multipath and noise are generally uncorrelated between the two frequencies, the L2 observations afford a truly independent observation. Increasing the data sampling does not add useful information, as the temporal correlations reduce the value of the added observations. However, due to the high cost of dual frequency receivers, it is anticipated that this is not a practical approach to increasing the deformation sensitivity. Instead, common approaches such as equipping stations with chokerings to mitigate multipath, and simply extending the data spans used are possibilities.

### ***7.2 - Application to Medium Range Static Positioning and Results***

Perhaps the most important advantage of satellite-based measurement systems is that direct measurements between stations are never required. As a result, there is no theoretic limit on the separation between stations. In addition, traditional terrestrial networks are developed through the propagation of “chains” of distance or angular measurements. Consequently, the positioning accuracies can vary significantly depending on the network geometries and the location of stations with known coordinates (Vanicek and Krakiswsky, 1986). Satellite-based networks, on the other hand, can be solved via a simultaneous adjustment of observations directly to common satellites and so show greater homogeneity in their positioning accuracies throughout the network.

Medium range positioning refers to applications where the network extents range from tens to several hundreds of kilometres. The justification for this extent-based classification lies in the fact that the dominant error source affecting networks of this size is the ionosphere, with the troposphere playing a secondary role. As a result, the use of dual frequency

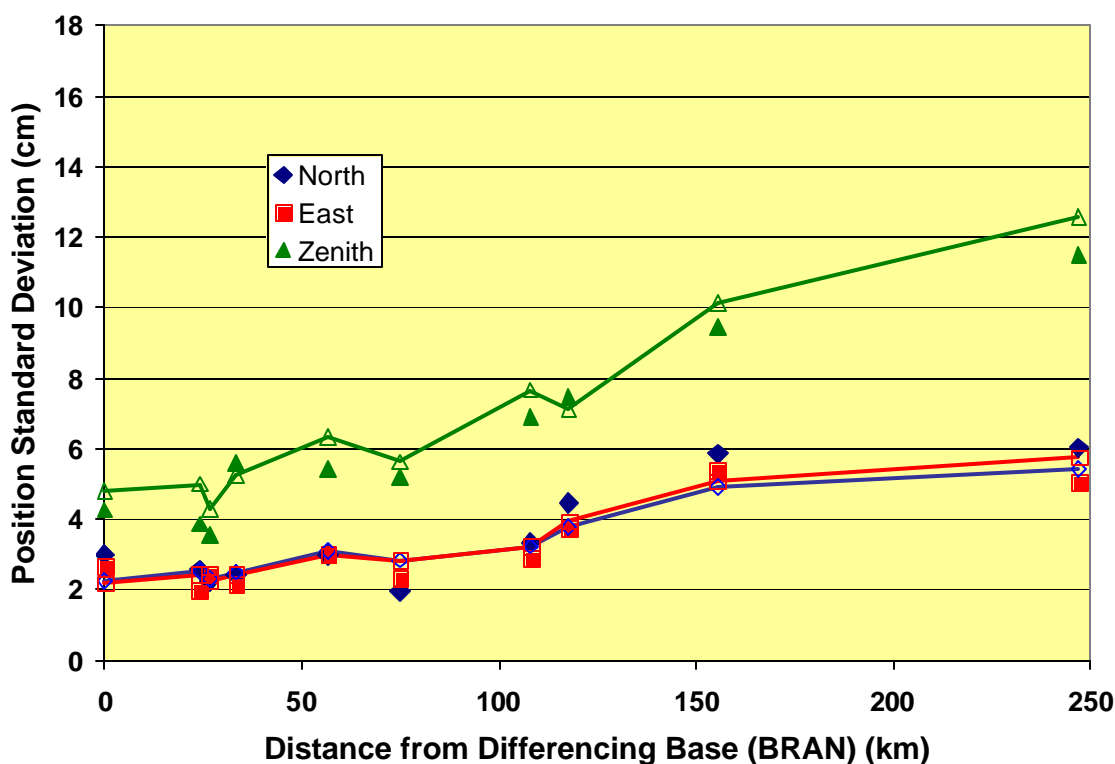
measurements is required for practical ambiguity resolution. In addition, the network extent is small enough to ensure that a suitable large set of common satellites is nominally visible from all receivers in the networks, allowing the establishment of a differencing base station at the centre of the network. As discussed previously, the differencing base is simply the station which forms part of all each double-difference and does not play any special role in the regularization scheme. At larger scales, orbital errors necessitate the use of precise ephemerides and specialized techniques to deal with the lack of common satellite visibility; topics which are beyond the scope of this thesis.

A key distinction between medium and short range positioning lies in the spatial correlation of errors. On short receiver separations, the ionosphere and troposphere are effectively cancelled out by the double-differencing operation, leaving only the multipath and noise effects. As a result, the double-differences are mainly only mathematically correlated, with little spatial correlation. However, as the receiver separations increase, the variances of the double-differenced observations grow as well, and the true variance-covariance matrix of the double-differences diverges from that created from only mathematical considerations. However, it was not well known what the effects of these spatial correlations were on both the positioning accuracy and on the estimates of the positioning accuracy.

Fortunately, the PADRES-GPS software allows processing of network GPS data using whatever variance-covariance model the analyst wishes to specify. As such, it affords a useful tool for the study of the effects of variance-covariance modelling on positioning accuracy and accuracy estimation. In particular, the effects of neglecting the spatial correlations between baselines entirely were studied, as well as the accuracies achievable in a typical medium range GPS network.

To test the capabilities of the PADRES-GPS software in a medium-range positioning mode, the July, 2002 data set collected from the SCIGN network for the purpose of tropospheric and ionospheric variance-modelling discussed in Chapter 6 was used. Fifteen forty minute data segments evenly spaced throughout the day were isolated and processed using the

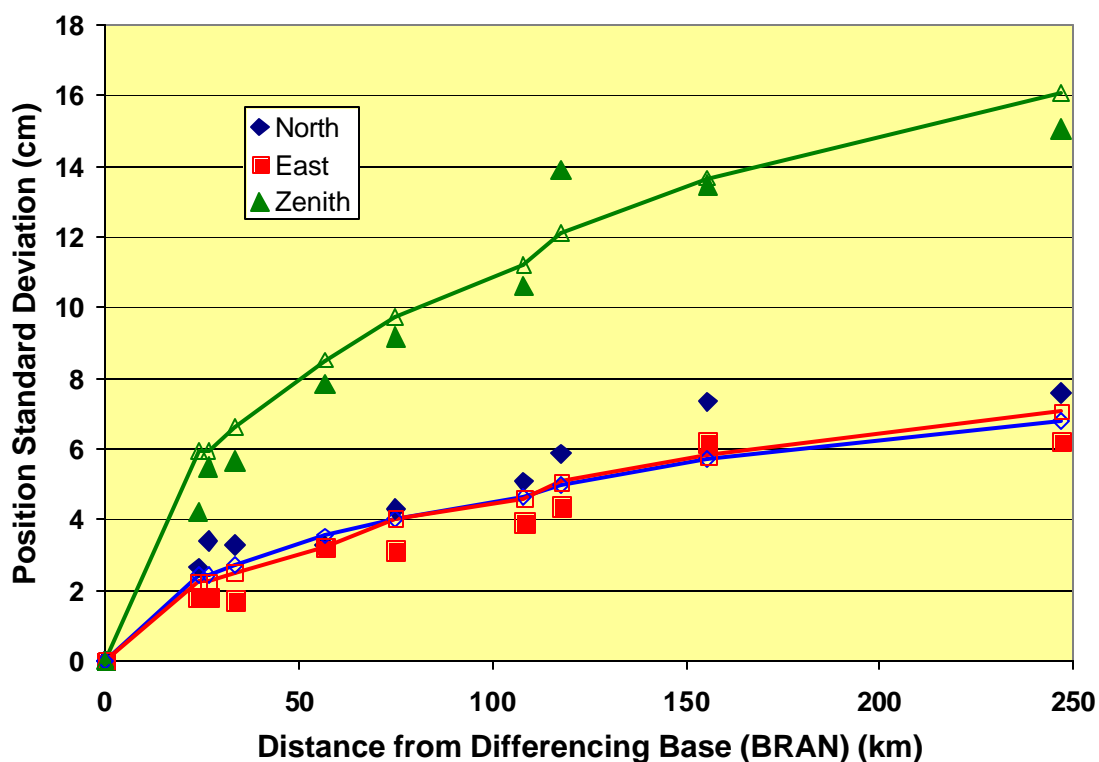
PADRES-GPS software. Dual frequency measurements were used, as was the centre-of-mass regularization scheme; incorporating the published site coordinates as initial estimates. Errors from the published coordinates for each station over the fifteen epochs were calculated and treated as formal errors. The resulting positioning accuracies for each station are plotted in Figure 7.7 as a function of station distance from the BRAN station, which was used as the differencing base. The average predicted accuracies are plotted as well, showing good agreement with the results actually observed.



**Figure 7.7.** Observed and Predicted Position Accuracies for Medium Extent Network using Centre-of-Mass Regularization. Estimated accuracies are shown as solid lines, observed accuracies as solid symbols.

Figure 7.8 shows the results achieved when the BRAN station is held fixed with known coordinates in a minimally constrained adjustment. The degradation in positioning accuracy is immediately apparent, mainly due to the propagation of errors at the BRAN station into the position results of the other stations. Whereas the positioning accuracy slowly degrades with receiver separation when the centre-of-mass regularization is used,

the error profile in the minimally constrained case follows a  $1-e^{-d/D}$  curve. The reason for this behaviour is most likely due to the increase in the tropospheric error with receiver separation, which follows a similar curve, as discussed in Section 6.4. In the centre-of-mass regularized case, these errors are averaged throughout the network to create a more homogenous accuracy behaviour. The residual degradation of positioning accuracy with receiver distance from the differencing base is mainly then due to the gradual increase in the proportion of ambiguities that cannot be resolved for that receiver.

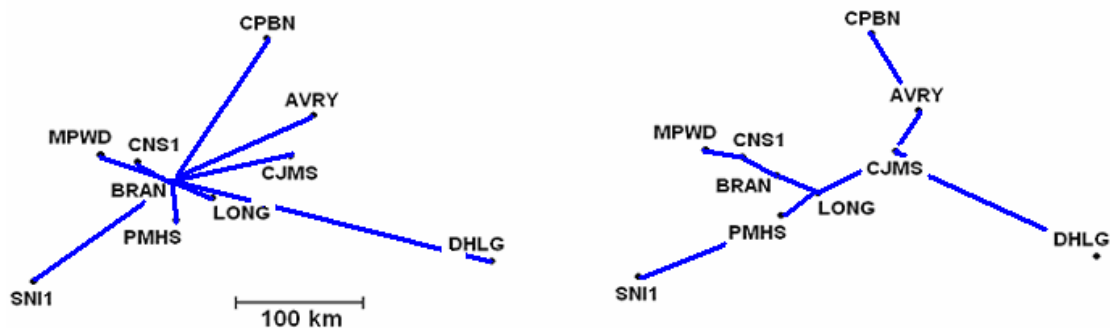


**Figure 7.8.** Observed and Predicted Position Accuracies for Medium Extent Network using Fixing of Station BRAN. Estimated accuracies are shown as solid lines, observed accuracies as solid symbols.

To further homogenize the positioning accuracies across the network and possibly improve the overall positioning accuracy, the notion of a single differencing base must be discarded. Figure 7.9(a) shows the relationship between receivers involved in producing double-differences stemming from a single differencing base. Essentially, the data is processed as individual baselines radiating from a central site, with the spatial correlation



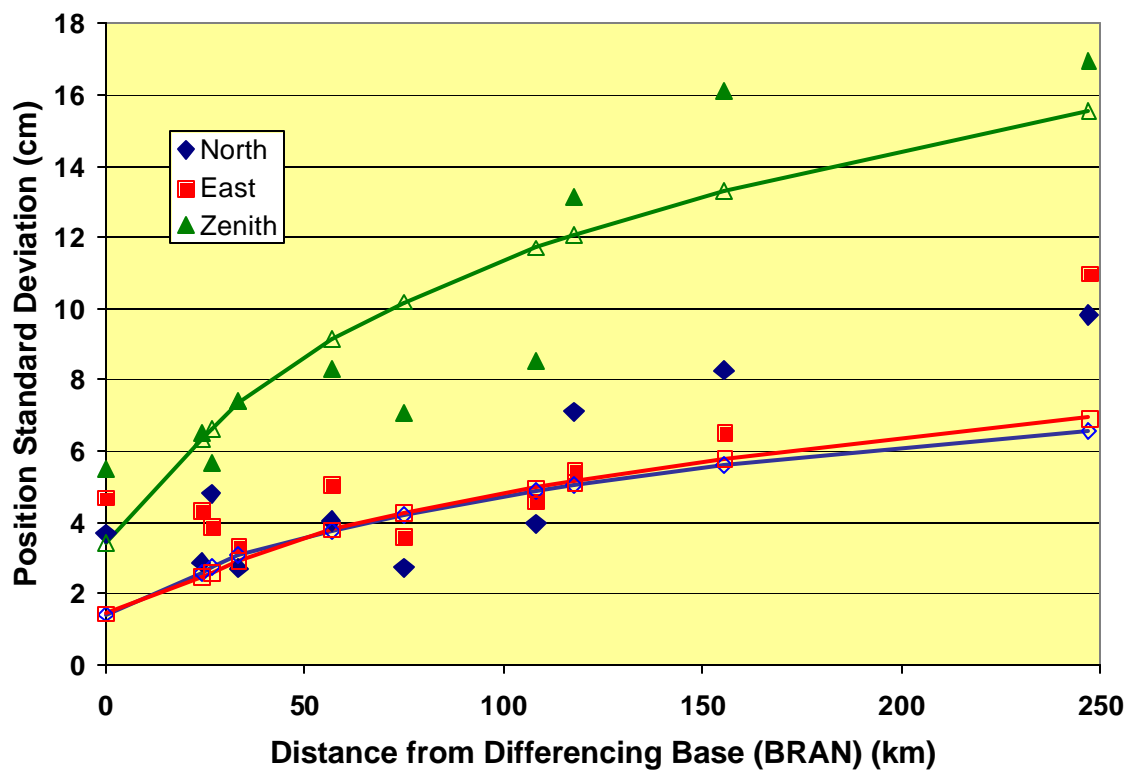
between the baselines rigorously modelled. As the baseline lengths increase, the differencing becomes less effective and as a result, ambiguities become less likely to be resolved as integers. This causes a corresponding degradation in positioning accuracy for the receivers located further away from the centre, although the modelling of the spatial correlations tempers this effect. A more effective scheme would be to form the double-differences by selecting pairs of receivers that would create the shortest baseline lengths, and hence minimize the effects of the undifferenced atmosphere. Of course, care must be taken to ensure that the number of baselines formed is equal to the total number of stations minus one, in order to ensure that the double-differences formed are all independent. As well, the algorithm must ensure that the topology of the baselines in the network is logical, such that there are no “hanging subnets” unconnected to the main network. Figure 7.9(b) shows the resulting optimized baseline selection.



**Figure 7.9.** Possible Baseline Configurations. (a) Differencing-base Scheme (b) Shortest Baseline Scheme

In theory, a system based on this method could maintain a consistent accuracy over a much larger extent, especially if the requirement for satellites to be visible at all stations is discarded. Unfortunately, the implementation of such a system is beyond the scope of this thesis, especially due to the logistics involved in properly generating the variance-covariance matrix for the double-differences formed in a network with such flexibility in its configuration. In addition, as the PADRES-GPS software was primarily designed for use on small-extent networks, the gains stemming from such flexibility are slight.

Aside from dependence on a fixed base station, the second most common feature of commercial GPS processors is a reliance on a baseline processing methodology where all baselines are formed separately and later combined, neglecting the correlations between baselines. To study the effects this has on positioning, the network under study was re-processed with PADRES-GPS, but the variance-model was changed to neglect the spatial correlations of the atmosphere. The centre-of-mass regularization was retained. The resulting accuracies are shown in Figure 7.10. Again, a significant degradation in positioning accuracy is apparent, especially for receivers located further away from the station BRAN. Overall, the effects of not considering the spatial correlation reduces the accuracy of positioning by an average of 22%. In addition, the agreement between the estimated and observed positioning accuracies is worse.



**Figure 7.10.** Observed and Estimated Position Accuracies for Medium Extent Network Neglecting Spatial Correlations. Estimated accuracies are shown as solid lines, observed accuracies as solid symbols.

When the correlations between the baselines are neglected, the network solution problem degenerates into a series of  $n$  independent and isolated least-squares problems, where  $n$  is the number of baselines formed. If the spatial correlations are properly accounted for, on the other hand, then the solutions for each baseline are strengthened by information provided by observations made at neighbouring stations. This behaviour is perhaps best explained through analogy with a two stage collocation process, where, in the first stage, the observed errors in the double-differences and the spatial correlation model are used to determine the “spatially varying error” (i.e. tropospheric and ionospheric error) across the network. The predicted error (equivalent to the signal in classical collocation) is then used to improve the observations in the subsequent position and ambiguity solution phase. Of course, in the actual adjustment under consideration, these two steps are implicitly and simultaneously completed via the inversion of the variance-covariance matrix of the double-differences, which are created using the spatial model, and the solution of the whole set of normal questions. However, the two-stage approach does find direct application when reference networks are used for kinematic positioning; a concept studied by Raquet (1998). Regardless of the implementation details of the network adjustment system, results show that proper accounting of the spatial correlations is critical in large network processing to ensure maximum accuracy.

### ***7.3 -Application to Kinematic Network Positioning and Optimal Linear Combination Determination***

A fast-developing area of application for GPS lies in the field of kinematic positioning using reference networks. Essentially, the goal is to improve the robustness and precision of positioning a moving platform by combining the data collected at the platform with observations made at a number of reference receivers with known coordinates in the project area. The current approaches to implementing such network positioning all rely on some level on the concept of the spatial correlation of GPS errors. For example, the method of Raquet (1998) uses empirical variance-covariance modelling and collocation, coupled with appropriate trend fitting (Fotopoulos, 1999) to generate corrections subsequently transmitted to the remote user.

At the same time, proposed improvements to GPS include the provision of extra frequencies available for civilian use (Lo verro, 2002). Given the addition of these extra observations, and the large number of reference receivers that may be available, it becomes evident that bandwidth limitations become a concern if data must be transmitted from one site to another. As a result, a method of compressing multi-frequency data has the potential to improve the efficiency of network based positioning. This section investigates how this can be done using variance-covariance analysis.

### 7.3.1 -Linear Phase Combinations

A linear carrier phase combination is simply a pseudo-observation formed from the scaled addition of two or more observed carrier phases. For the purposes of this thesis, a linear combination will be restricted to a combination of a particular carrier phase on the L1 and L2 GPS frequencies, although it is realized that there is no theoretical reason why combinations between multiple frequencies (when they become available) or even multiple systems (i.e. GLONASS or GALILEO) might not be possible.

The resulting linear phase combination,  $\mathbf{f}_*$ , can be expressed in cycles as

$$\mathbf{f}_* = [a \quad b] \cdot \begin{bmatrix} \mathbf{f}_{L1} \\ \mathbf{f}_{L2} \end{bmatrix} \quad (7.7)$$

where  $\mathbf{f}_1$  and  $\mathbf{f}_2$  are the parent carrier phase observations and  $a$  and  $b$  are the coefficients of the combination. Eq. (7.7) shows that the linear phase combination is actually the projection of a two dimensional observation space onto a single line. The effective wavelength,  $\mathbf{I}_*$ , of the new pseudo-observation is given by :

$$\mathbf{I}_* = \frac{\mathbf{I}_{L1} \cdot \mathbf{I}_{L2}}{a \cdot \mathbf{I}_{L2} + b \cdot \mathbf{I}_{L1}} \quad (7.8)$$

where  $I_{L1}$  and  $I_{L2}$  are the wavelengths of the L1 and L2 carrier.

In Section 6.7 it was shown that some of the errors affecting the two GPS frequencies are correlated between frequencies. Specifically, the tropospheric error, in metres, is identical between frequencies, whereas the multipath and noise was considered uncorrelated. Using these assumptions, and the expression for the ionospheric inter-frequency covariance from Eq. (6.55), the variance of the linear combination is calculated in cycles as :

$$\begin{aligned}
 \mathbf{s}_*^2 = & a^2 \left( \mathbf{s}_n^2 + \frac{\mathbf{s}_T^2}{I_1^2} + I_1^2 \frac{40.3^2}{c^4} \mathbf{s}_I^2 \right) \\
 & + 2ab \left( \frac{\mathbf{s}_T^2}{I_1 I_2} + I_1 I_2 \frac{40.3^2}{c^4} \mathbf{s}_I^2 \right) \\
 & + b^2 \left( \mathbf{s}_n^2 + \frac{\mathbf{s}_T^2}{I_2^2} + I_2^2 \frac{40.3^2}{c^4} \mathbf{s}_I^2 \right)
 \end{aligned} \tag{7.9}$$

where  $\mathbf{s}_n^2$  is the noise and multipath variance in cycles<sup>2</sup>,  $\mathbf{s}_T^2$  is the tropospheric variance in metres<sup>2</sup> and  $\mathbf{s}_I^2$  is the ionospheric variance in units of TEC<sup>2</sup>.

Both the process of double-differencing and linear-combination are linear processes, and as a result they are commutative. As a result, the above expressions hold for the double-differences formed at both frequencies, with the  $\nabla\Delta$  symbol, appropriately substituted into Eqs. (7.7) and (7.9). In addition, although the original and double-differenced ambiguities on both frequencies are integer quantities, the integer nature is not necessarily preserved in the linear combination if either coefficient is real valued. For the integer ambiguity resolution techniques of Chapter 6 to hold, the values of  $a$  and  $b$  must be held to integers as well.

### 7.3.2 - "Optimal" Linear Phase Combinations

In Chapter 2 it was shown that the variance-covariance matrix of the estimated parameters,  $\mathbf{C}_{x^*}$ , depends on the variance-covariance matrix of the observations used and their linearised mathematical model. The relationship is summarized again here, simplified as

$$\mathbf{C}_{x^*} = (\mathbf{A}^T (\mathbf{B} \mathbf{C}_l \mathbf{B}^T)^{-1} \mathbf{A})^{-1} \quad (7.10)$$

where again  $\mathbf{A}$  is the Jacobian matrix with respect to the unknown parameters,  $\mathbf{B}$  is the Jacobian matrix with respect to the observations and  $\mathbf{C}_l$  is the variance-covariance matrix of the observations. If a linear combination of two carrier phases is used, the number of effective observations is reduced by half, and thus the number of rows in the  $\mathbf{A}$  and  $\mathbf{B}$  matrices. The  $\mathbf{C}_l$  matrix is also transformed in accordance with the principle of the propagation of errors.

Inspection of Eq. (7.10) shows that the matrices  $\mathbf{A}$  and  $\mathbf{B}$  are invariant under various choices for the linear combination coefficients,  $a$  and  $b$ , assuming that the observation equations are expressed in meters. However, the  $\mathbf{C}_l$  matrix is directly affected by the choice of these coefficients, as it contains the variances of the new pseudo-observations, which are given by Eq. (7.9). Thus, by minimizing the variances of the pseudo-observations through an appropriate choice of  $a$  and  $b$ , the accuracy of the resulting position estimates will improve. The Optimal Linear Phase Combination (OLPC) is the combination that optimizes some aspect of the positioning problem through an appropriate choice of  $a$  and  $b$ .

Furthermore, due to the physical independence of error sources, one can derive OLPC's that minimize the variance of a particular error source. For example, inspection of Eq. (7.9) shows that an OLPC that eliminates the ionospheric error is given by  $(a,b) = (1, -f_{L2} / f_{L1})$ , which is the well known ionospheric free combination with  $f_{L1}$  and  $f_{L2}$  as the frequencies of the L1 and L2 carriers. Unfortunately, this is not an admissible OLPC since the  $b$  term is not integer. In addition, it can be shown that regardless of the OLPC chosen, the tropospheric

variance will remain the same in units of distance, although the wavelength of the OLPC will change. The only exception is the linear combination  $(a,b) = (1, -f_{L1} / f_{L2})$ , which eliminates the tropospheric error in cycles, but yields an observable with infinite wavelength, as well as not containing integer coefficients. Finally, as expected by the principle of the propagation of random errors, the noise variance in cycles goes up as  $a$  and  $b$  increase, but the variance in metres depends on the wavelength of the particular OLPC.

This illustrates one of the key difficulties in defining “optimality” for linear combinations. While it may be possible to minimize the cyclic error variance of a combination, it is not guaranteed that this combination will minimize the error in units of meters. Similarly, while short wavelengths generally reduce the error variance in metres<sup>2</sup> (which is desirable for positioning accuracy) it is more difficult to resolve ambiguities reliably under such circumstances. Thus, the definition of optimality is somewhat arbitrary.

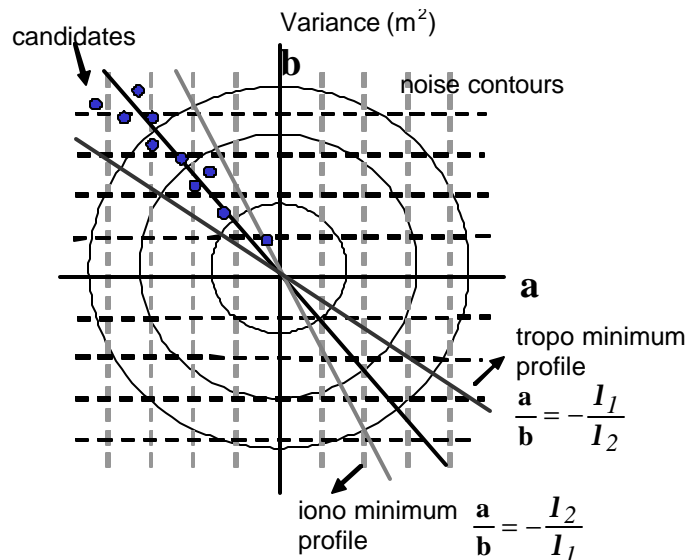
In this work, an optimal combination is defined as a combination that minimizes the total error variance in metres<sup>2</sup>, while maintaining an error variance in cycles<sup>2</sup> that is at least equal to that of the L1 observable. In this way, the positioning accuracy is maximized, but ambiguity resolution does not become more difficult than in the L1-only case. This is an effective strategy when baseline lengths are less than 50 kilometres, as the ionospheric error remains limited to less than ½ of the L1 wavelength. Hence the float ambiguities, although biased by the undifferenced ionosphere, are still likely to converge to their true integers. Once the ambiguities are resolved as integer, the key factor is a reduction in error variance in metres to achieve high positioning accuracy. On larger network extents, this optimality criteria may not be appropriate. Regardless, once the definition of optimality is set, the techniques presented herein can be used to find an OLPC to satisfy the definition.

Figure 7.11 shows the relationship between the error components considered for various values of  $a$  and  $b$ . In cycles, the noise error variance function forms a paraboloid with global minimum at  $(a,b) = (0,0)$ . Thus the variance contour lines form concentric circles about the origin. The tropospheric and ionospheric error variance functions, however, form parabolic cylinders, whose global minima are described by lines passing through the

origin. These relationships are shown graphically in Figure 7.11. The total error variance function is simply the sum of these surfaces and thus forms a saddle-shaped surface, whose *ridge lies* is described by the line passing through the origin satisfying the relation :

$$\begin{pmatrix} a \\ b \end{pmatrix}^* = \frac{-\frac{I_1}{I_2} \nabla \Delta \mathbf{s}_T^2 - \frac{I_2}{I_1} \nabla \Delta \mathbf{s}_I^2}{\nabla \Delta \mathbf{s}_T^2 + \nabla \Delta \mathbf{s}_I^2} \quad (7.11)$$

where the variances are all considered in cycles. This line is simply the weighted average of the tropospheric and ionospheric variance minimization lines.



**Figure 7.11.** Selection of an Optimal Linear Combination.

To determine an optimal integer combination, integer pairs lying on either side of the line described by Eq. (7.11) are chosen, and the combinations with total error variances less than that of the L1 double-differenced observations are retained. The remaining combination that minimizes the total error variance in metres<sup>2</sup> is then chosen as the optimal candidate. Thus, the errors are initially analysed in terms of cycles to derive a subset of potential candidates, and then in terms of metres to chose the optimal candidate.



In an operational sense, it is desirable to choose a single OLPC to apply to all the dual-frequency observations observed in a reference network, rather than calculating an OLPC for each observation pair. To do this, one simply has to calculate the average values of the double differenced noise, tropospheric and ionospheric variances observed in a given reference network. Fortunately, since the stochastic models describing the noise, tropospheric and ionospheric errors are fairly stable over time, the determination of the OLPC can be done in advance by using an almanac to predict the satellite orbits and an appropriate stochastic model to derive the predicted error variances

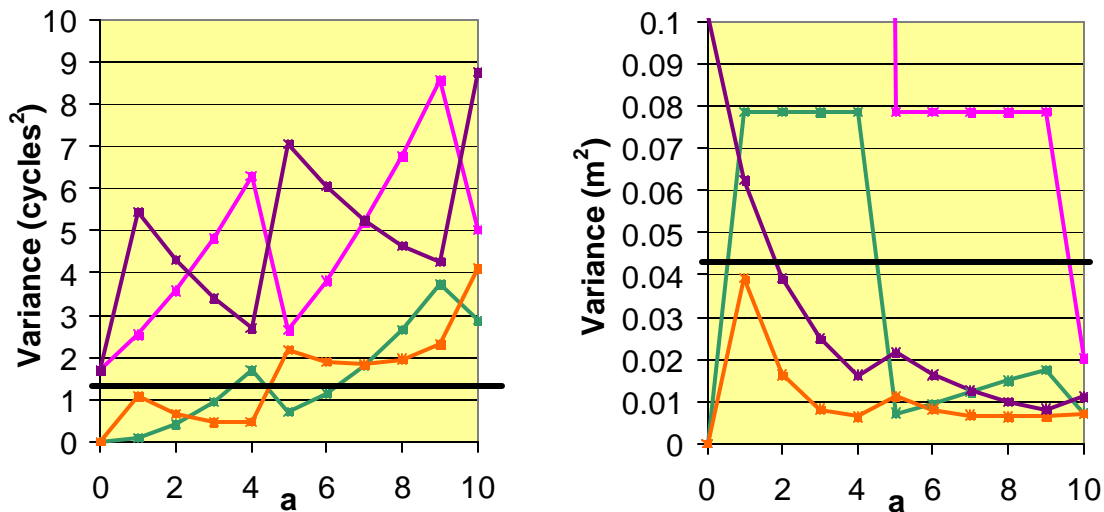
### *7.3.3 -Determination of Optimal Linear Phase Combinations From Reference Network Data*

To illustrate the procedure of determining an OLPC from data available from a reference network a test was conducted using the July 2002 data set analysed in the medium-extent network of the previous section.

Using the ephemerides available for the 24 hour period of data collected, the theoretical observations made during the observation period was generated. The variance-covariance models for the tropospheric, ionospheric and multipath effects developed in Chapter 6 were then used to generate average values of the variances of these error sources. Model values were those calculated for the July 2002 data set from the reference network. Based on these average double-differenced variances for both frequencies, the average variance for a linear combination based on a particular  $(a,b)$  selection can be calculated using Eq. (7.9). The resulting ionospheric/tropospheric error variance minimization line is given by the relation  $a/b = -1.232$ , which is very close to the pure ionospheric minimization line of  $a/b = -I_{L2}/I_{L1} = -1.283$ . This is due to the dominance of the ionospheric error over that of the troposphere, as discovered in Chapter 6.

Figure 7.12(a) shows the error in cycles<sup>2</sup> for a subset of potential integer combination candidates selected. The x-axis shows values of the a parameter, and the different lines represent different choices of the b parameter - namely selecting integer values on either side of the line defined by  $a/b = -1.232$ . The heavy black line indicates the error in cycles<sup>2</sup>

of the L1 observation. From Figure 7.12 it can be seen that the error of the combination in cycles increases as the values for  $a$  and  $b$  increase. This is primarily due to the amplification of noise.



**Figure 7.12.** Candidate combination variances in (a) cycles<sup>2</sup> and (b) m<sup>2</sup>.

According to the “optimality” criterion defined, the final step for determining the optimal combination is to choose the remaining integer candidate that minimizes the total variance in metres<sup>2</sup>. The results for each combination candidate are also shown in Figure 7.12. Once again, the heavy black line indicates the variance of the L1 observation. Interestingly, as the coefficient value  $a$  increases, the total variance converges to the tropospheric variance in metres<sup>2</sup>. After performing the above mentioned steps, the overall ‘optimal’ integer combination  $(a,b)$  is given by the pair  $(4, -3)$ . This resulting combination has an associated wavelength of 13.4 cm and an error variance (in m<sup>2</sup>) of  $\sim 5$  times smaller than the L1 observable (or roughly twice as small in units of cycles<sup>2</sup>).

#### 7.3.4 -Positioning with an Optimal Linear Phase Combination

To gauge the utility of OLPC's in kinematic network position, the following test was employed. A five-station subset of the SCIGN network was selected, along with a station approximately 1500m above the other five reference receivers. This station simulated the

conditions encountered by an aircraft in flight during a medium-scale photogrammetric survey. Any variation in the epoch-to-epoch position estimates of this static point can then be interpreted as an indication of positioning accuracy. Resulting baseline lengths from the reference receivers to the high platform ranged from 20 to 60 kilometres. Coordinates for the reference receivers and the test station were generated using a full 24 hours of dual frequency data processed using the PADRES-GPS software.

The first processing mode used was a rigorous adjustment using both L1 and L2 data simultaneously. The observations were processed using a full variance-covariance matrix based on the stochastic model described above. As this mode utilized the most information available, it was assumed that this would result in the best (most consistent) results. On the other hand, this mode requires the most data handling. Secondly, the data was processed using the (4, -3) OLPC. The data was then also processed using only L1 and using the widelane observable which corresponds to a (1,-1) linear combination. In all cases, the full variance-covariance between observations made at different receivers and satellites was propagated.

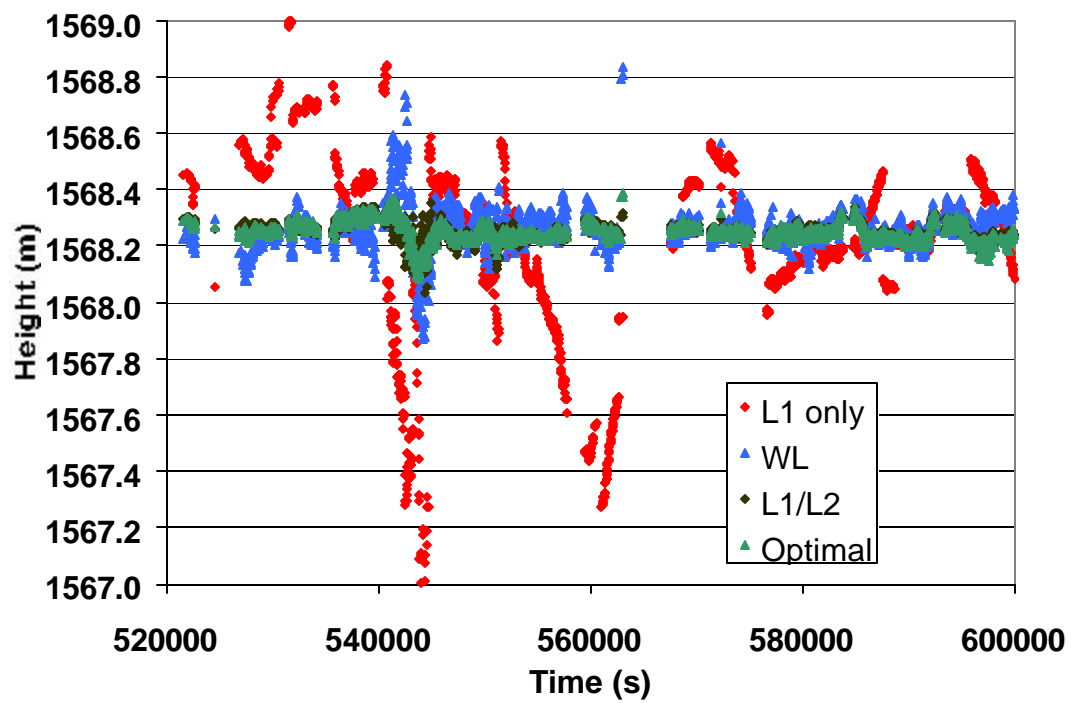
Table 7.5 shows the position variation standard deviation for each of the processing modes along with the average bias from the known coordinates. Figure 7.13 graphically shows the variation in the height coordinates the four processing modes. As expected, the L1/L2 mode performs the most consistently. Note that the centimetre-level epoch-to-epoch positioning accuracy has been achieved despite the relatively long baseline lengths, due to the rigorous incorporation of dual frequency data. However, the OLPC performs equally well and in fact, returns almost identical results every epoch. The L1 mode returns the worst variation, and the results are heavily affected by the variations in the ionosphere.

An explanation for the results lies in the domain of principal component analysis. As previously mentioned, for any dual-frequency measurement, the variance-covariance matrix takes the form of Eq. (6.55). Thus one can visualize an error ellipse around the two dimensional dual-frequency measurement pair. When the errors are contaminated only by

**Table 7.5.** Position Standard Deviations of Various Processing Modes.

	<b>L1</b>	<b>L1/L2</b>	<b>WL</b>	<b>Optimal</b>
<i>N</i>	14.5 (1.7)	1.8 (0.2)	3.5 (0.4)	2.0 (0.3)
<i>E</i>	12.4 (0.3)	0.9 (0.0)	2.2 (0.2)	1.0 (0.0)
<i>H</i>	32.6 (6.0)	3.3 (0.7)	7.8 (1.0)	3.5 (0.6)

Average bias shown in parenthesis. All values in cm.

**Figure 7.13.** Height Variation for Various Processing Modes

uncorrelated noise, the error ellipse has its principal axes aligned along the L1 and L2 "directions." However, if tropospheric or ionospheric effects exist, the variance-covariance matrix contains off-diagonal elements, and the error ellipse rotates accordingly.

A linear phase combination is simply a projection of the 2 dimensional dual-frequency set onto a single observation in the direction  $\langle a, b \rangle$ . As a result, the variance-covariance matrix is reduced to a single variance. The process of choosing an OLPC is essentially rotating the  $\langle a, b \rangle$  vector until the projected variance is a minimum. Thus when the dual-frequency data is pre-processed by applying the OLPC, the data is effectively filtered of the major correlated errors.

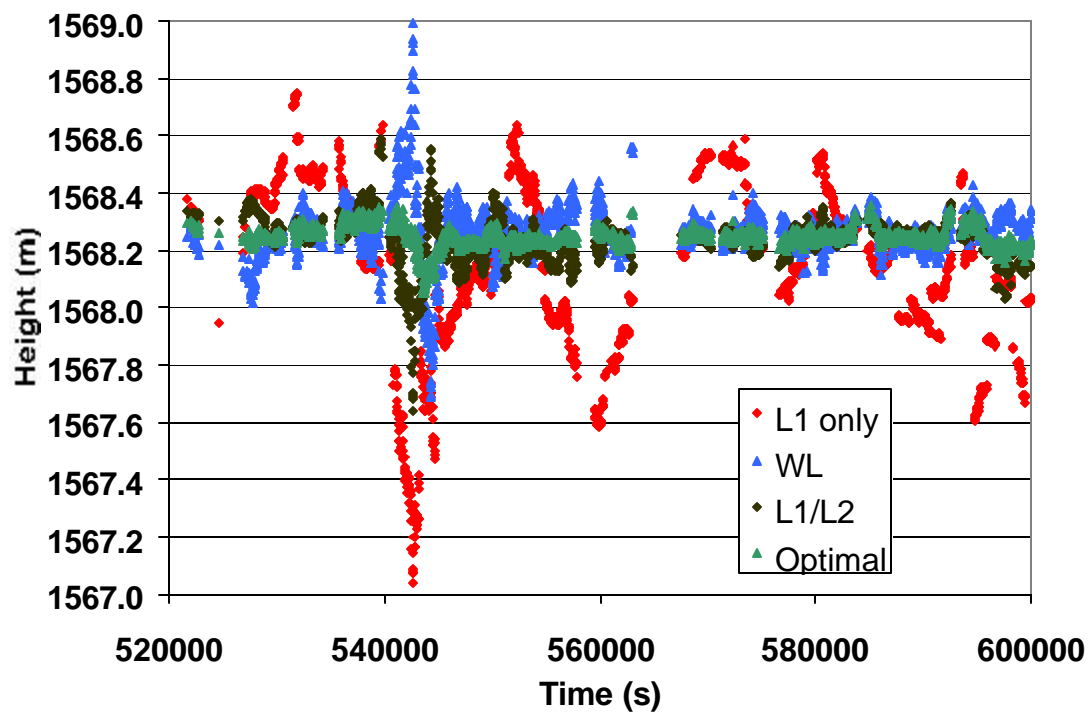
Since the ionospheric error significantly dominates the error budget in most medium to long range networks, the corresponding dual-frequency variance-covariance matrix is highly elongated. Therefore, the variance of the OLPC will be very small in relation to the semi-major axis of the error ellipse and the positioning accuracy will be high. In the widelane case, the corresponding  $\langle 1, -1 \rangle$  vector is rather arbitrary and thus cannot be expected to minimize the error variance in any fashion. Similarly, the L1 only mode is a projection in the  $\langle 1, 0 \rangle$  direction and so bears the brunt of the full ionospheric error.

The above test relied on the use of an accurate stochastic model to achieve correct relative weighting of the available observations. However, it is often difficult to estimate the parameters of the stochastic model, in particular in terms of the tropospheric and ionospheric correlation lengths. As a result, a second test was conducted in which the data was reprocessed under the assumption that there were no tropospheric or ionospheric error - i.e. only an elevation dependent weighting was applied and no physical cross-correlations exist. Table 7.6 contains the results of this test and Figure 7.14 shows the corresponding variation in the height coordinates.

**Table 7.6.** Position Standard Deviations of Various Processing Modes with an Improper Stochastic Model Assumed.

	<b>L1</b>	<b>L1/L2</b>	<b>WL</b>	<b>Optimal</b>
<i>N</i>	12.5 (3.4)	4.0 (0.2)	4.9 (0.8)	2.1 (0.3)
<i>E</i>	9.1 (1.0)	2.7 (0.0)	2.7 (0.2)	1.0 (0.0)
<i>H</i>	27.6 (11.0)	8.2 (2.2)	10.1 (1.0)	3.6 (0.6)

Average bias shown in parenthesis. All values in cm.



**Figure 7.14.** Height Variation for Various Processing Modes using an Incorrect Stochastic Model

As expected, the L1/L2 processing mode yielded better results than the L1 and widelane modes, mainly because twice as many observations were used and thus the ionospheric errors were able to cancel out somewhat, despite neglecting the correlation between the errors on both frequencies. Interestingly, the overall standard deviations of the L1 solutions improved, but the biases became significantly worse, as the effects of the uncorrected ionosphere propagated undetected into the solution. However, it is very surprising to note that the OLPC mode performed better than the L1/L2 mode. In fact, the OLPC mode performed almost as well as in the case of using a full-stochastic model.

The reason for the convenient behaviour of the OLPC lies again in the concept of the error ellipse. In the L1/L2 case, using the full stochastic model gave information to the least-squares adjustment regarding the correlation between the L1 and L2 frequencies. Implicitly, the adjustment was able to use this information to calculate the ionospheric error affecting each observation and remove this error effect. However, once the stochastic information was dropped, the adjustment could only treat the ionospheric effect as noise and the subsequent positioning accuracy was degraded. In the L1 case, this shows up as the large bias in the position estimates – the adjustment simply cannot “see” the ionospheric error, and consequently the ionospheric error causes a scaling of all the satellite-receiver ranges. This is also why the height component is the most affected.

In the case of the OLPC, a pseudo-observation that was relatively free of ionospheric effects was generated. Consequently, it is to be expected that the spatial correlations of the pseudo-observations would be significantly reduced and that the variance-covariance matrix for the set of observations becomes more diagonal as a consequence. In addition, the variance-covariance matrix of a set of uncorrelated observations (i.e. only affected by noise) is diagonal as well, and so the variance-covariance matrix of the OLPC observations becomes more like a scaled version of the noise only assumption (i.e. no stochastic information regarding spatially correlated errors). This is key, since it is well known from adjustment theory (Mikhail, 1971) that the results of a least-squares adjustment are invariant upon scaling of the variance-covariance matrix. An added benefit of the noise-like variance-covariance matrix lies in the fact that the inversion of such a matrix is much

more efficient than the inversion of a matrix where spatial-correlations cause significant off-diagonal terms.

In this work, the generation of an OLPC has been limited to the compression of two carrier phases into a single pseudo-observation. In the future it is anticipated that observations at multiple frequencies will be available. Fortunately, the concept of the single OLPC is still useful in such a scenario, and in fact, is even more attractive due to the restriction of transmission and processing loads to a single pseudo-observation, rather than a set of multiple frequency observations. In the case of observations at several frequencies, the variance-covariance information relating the observations can be visualized as a hyper-ellipsoid. The OLPC is then defined as the projection of the data collected at different frequencies into a single pseudo-observation, the “direction” of which is defined as the direction of the semi-minor axis of the error ellipsoid. Thus observations at additional frequencies are expected to allow for better extraction of the ionospheric error (as it is frequency dependent) and allow for the averaging out of multipath and noise.



## **8.0 - CONCLUSIONS AND FUTURE WORK**

The objectives of this research were satisfied. By eliminating traditional shortcuts and assumptions regarding GPS, an improved method of adjusting observations made with satellite-based navigation systems was developed. The key developments of this dissertation included :

- Study of the least-squares adjustment problem as applied to satellite-based positioning, in particular in terms of adjustment under constraints, proper propagation of parameter statistics and linearization of the positioning mathematical model.
- Investigation of the datum definition problem and the development of a regularization scheme to solve the problem of poor datum visibility.
- Development of a partial-fix ambiguity resolution scheme which depends on appropriate a priori variance models and minimally biased float ambiguity estimates, achievable through a centre-of-mass regularization scheme.
- Establishment of variance-covariance models for GPS error sources and the development of practical techniques to determine estimates of model parameters.

The above findings were then combined to form the basis of a very robust and flexible GPS processing package. The performance of this package was tested under small and medium network extents.

This chapter summarizes the key findings of each chapter individually, discusses other possible application areas, and makes recommendations for future research.

## **8.1 -Key Findings by Chapter**

### *Chapter 2. Basics of Space-Based Radio Ranging.*

Determination of satellite coordinates requires accounting of the frame rotation and satellite motion during transmission time to ensure accuracy of satellite orbits remains at the several metre level (as determined from the broadcast ephemeris). Otherwise orbital errors of up to 500 metres can result.

### *Chapter 3. General Solution to the Space-Based Positioning Problem.*

The GPS solution space is relatively flat with respect to the position unknowns, and completely linear with respect to unknown clock offsets and ambiguities. Position initial estimates accurate to the 300 metre level are required to preclude the need for iteration; initial estimates are not required for clock and ambiguity unknowns.

Double-differencing serves to eliminate both satellite and receiver clock offset unknowns from the positioning problem, significantly improving the efficiency of the solution for the receiver positions. Double-differencing maintains the integer nature of the ambiguities, but causes the full set of (originally independent) ambiguities inestimable, necessitating the solution for a reduced set of mathematically correlated double-differenced ambiguities.

The least-squares solution is complicated by the addition of constraints, in particular in terms of error propagation. Furthermore, modifications to the standard equations for variance propagation must be made if the regularization for the parameters used does not correspond to the a priori variance-covariance of the initial estimates.

### *Chapter 4. Datum Issues in Satellite-Based Navigation Systems.*

The GPS datum is overconstrained by the assumption of satellite coordinates as known quantities. However, assuming them as unknowns or as observations causes unacceptable inefficiencies in the solution for the receiver positions due to the size of the matrices involved as the satellites move over time.

Without differencing, the network solution is limited in accuracy to the accuracy of the satellite orbits. The resulting error imparted into the solution for the receiver positions is highly correlated between receivers. The overall accuracy of the network improves as the network extent increases, due to decorrelation of the orbital errors, but at the same time, the relative accuracy degrades for the same reason.

Under differencing, the satellite-based datum becomes poorly visible, resulting in instabilities in the normal matrices required to be inverted. This problem can be alleviated by Tikhonov regularization. As the regularization is increased, the solution becomes more stable, but increasingly biased. As the network extent increases, the visibility of the datum improves, and the amount of regularization required is reduced.

For deformation monitoring, high regularization is desirable. Despite the highly biased results, the bias is correlated between epochs, and thus largely cancels out upon creation of the displacement vector.

#### *Chapter 5. Ambiguity Resolution for Precise Positioning.*

Regular least-squares return real-valued estimates for the integer ambiguities. The process of ambiguity resolution determines the most-likely integer values for the ambiguities based on their float solutions, and by applying these integer values as constraints, the solution for the receiver positions can be considerably improved.

Double-differenced ambiguities depend on the relative geometry of the receiver network and thus are largely insensitive to overall translations of the network. Heavy Tikhonov regularization results in highly biased ambiguities due to distortion of the relative geometry of the network. Instead, regularization using a constraint of the network centre-of-mass is preferred, although loss of relative geometrical information results in poorer accuracy of the float ambiguities. Nonetheless, this is preferable to a biased estimate if the integer values of the ambiguities are to be determined. Heavy regularization based on the centre-of-mass constraint can be applied without adversely affecting the ambiguity solutions, which is desirable from a deformation monitoring standpoint.

Integer values of the ambiguities can be efficiently determined via the LAMBDA method. A modification of this method was developed where the total number of ambiguities that can be resolved at the 99% level are chosen as a subset of the total ambiguities. This leads to a partial-fix methodology, where only some of the ambiguities are resolved as integer, and the remaining ambiguities are improved due to their correlation to the resolved ambiguities. The end result is more stable positioning accuracy between observation sessions.

*Chapter 6. Analysis of Error Sources in Global Navigation Satellite Systems and Stochastic Modelling.*

Proper stochastic modelling of the GPS observations is crucial for both positioning accuracy and the accuracy of the statistical estimates returned by the least-squares adjustment.

The stochastic modelling method employed herein combined a theoretical and empirical approach, relying on mathematical models developed from theoretical considerations of the error processes coupled to an empirical determination of the models' parameters. Once these models for the undifferenced errors are developed, they can be mathematically propagated depending on the differencing and network configuration scheme used.

All error sources were modelled using a zenith variance term multiplied by a mapping function. For noise and multipath, this mapping function is dependant on the antenna gain pattern, whereas for tropospheric and ionospheric errors this mapping function is dependant on the mapping functions used in the original error models.

The noise and multipath zenith variance are determinable using a zero-baseline and a short-baseline test, respectively. It was shown that the noise and multipath variances can be combined in practice, since their variance mapping functions are identical. Typical zenith variances of carrier phase noise and multipath are at the few millimetre level, and are site dependant.

The tropospheric and ionospheric variances were shown to be spatially correlated. This correlation was shown to decrease exponentially with increasing receiver distance and increasing angular separation of the receiver-satellite line-of-sight. The parameters of the resulting double-differenced variance models can be determined from a reference network. Although the parameters of the tropospheric model were fairly stable, it was shown that it is difficult to separate the zenith variance from the correlation length if the network extent is not larger than 500 kilometres. However, on networks of smaller extent, this also implies that only the ratio of the two quantities is of importance, and values of 1.5 cm / 250 km are suggested for the zenith standard deviation and correlation length respectively, assuming use of the UNB2 model. Typical model values for the ionosphere, using the Klobuchar model, are 0.9 m / 7000 km, but vary slightly over the course of the year and should be re-evaluated at 4 month intervals.

A simple method of adjusting the normal equations for the temporal correlations of the GPS errors was presented, based on the correlation period of the errors (assumed as approximately 300s), the sampling rate and the length of the observation session. It was shown that increasing the sampling rate has very limited gains in terms of improving the solution accuracy.

#### *Chapter 7. Application of Global Navigation Satellite Systems to Network Positioning.*

A processor adjusting GPS observations was developed for the primary purpose of high-precision short-range deformation monitoring. The PADRES-GPS software combines robust variance modelling, variable regularization schemes and rigorous least-squares processing into a package capable of returning stable, precise results with reliable statistical estimates.

The success of the system in short-range deformation was shown using a real-data set with known induced deformations. The PADRES-GPS software consistently provided more reliable results than a commercial GPS processing package used for comparison, mainly due to the short time spans used. This success was attributed to the partial-fix processing mode with regards to the ambiguities. Also, it was shown that the possibility for centre-of-

mass regularization allows more sophisticated deformation detection techniques to be used instead of a simple comparison of observed movements to their standard deviations. Application of the method was very successful in identifying lateral deformations on the level of 10 millimetres, and marginally successfully in identifying deformations on the 5 millimetre level. Observation session lengths were limited to 10 minutes, and it is believed that better positioning accuracy is possible if observation sessions are significantly extended (i.e. to the hour level), or if dual-frequency data is available.

PADRES-GPS was also shown to provide decimetre-level results when network extents of several hundred kilometres are involved and dual-frequency data is available, based on a forty minute observation span. The significant positive effects of centre-of-mass regularization and spatial correlation modelling were demonstrated.

Based on the variance-covariance modelling developed through this research, a novel method of combining dual-frequency data into a single optimal-linear phase combination (OLPC) was developed. Processing using this combination was shown to have advantages in terms of processing and data transmission efficiencies in kinematic network positioning. In addition, use of the OLPC yields results almost equivalent to that obtained when dual-frequency data is processed together with a proper stochastic model. When no such model is available, results using the OLPC remain largely unaffected, due to the implicit encapsulation of the stochastic model into the OLPC during its determination.

## ***8.2 -Thesis Contributions***

The author believes that the major contribution of this dissertation is the re-evaluation of the traditional dogma surrounding the processing of measurements collected using satellite-based navigation systems. Traditionally, satellite-based ranges have been treated as “special” observations, completely removed from traditional geodetic observations such as distances, angles and astronomic azimuths. In this thesis, every attempt has been made to treat GNSS measurements purely as very long range spatial distances and their similarities with those collected using terrestrial methods have been emphasised.

This change in perception directly led to the investigation of variance-covariance properties of the satellite-based measurements, and on datum definition and observation adjustment. At each stage, investigations were *led* by theoretical considerations and then *supported* by empirical data, rather than vice versa. The development of the multipath variance mapping function is an excellent example of this process. Although this thesis provided several contributions in terms of new variance-covariance models, ambiguity resolution techniques and datum regularization methods, the author feels the greatest contribution of this thesis is as an example of how a rigorous application of the hypothesis/testing/validation/refinement sequence can uncover previously unseen issues in an otherwise established problem.

### ***8.3 -Recommendations for Future Work***

Although these considerations led to an improved processing strategy, they also unveiled many new areas of research. Included below are some of the more pertinent topics :

- Proper variance-covariance modelling of the GPS observations should allow improved integration of these observations with terrestrial measurements. To date, integration of GPS and terrestrial measurements is done after processing of the GPS data to provide interstation vectors – these vectors are then used in a combined adjustment with the terrestrial data. Large problems exist in integrating the two sets of data due to improper weighting of the GPS interstation vectors. However, combining the normal equations corresponding to the terrestrial measurements and the GPS measurements directly, with a correct relative weighting, will allow the terrestrial data to directly improve ambiguity resolution by affording unambiguous information regarding the relative station positions – this in turn should result in overall improved accuracies. Similar research could involve the integration of photogrammetric flight data with GPS observations at the normal equations level.

- Further investigation of the covariance models developed should be performed. Specifically, the following areas are of interest
  - More comprehensive evaluation of the time varying nature of the variance model parameters for the multipath, tropospheric and ionospheric errors. In particular, it would be interesting to obtain better information regarding the variation in the ionospheric variance model under storm conditions and during different phases of the 11-year sunspot cycle. Similarly, it would be instructive to analyse the variation in the tropospheric model in different climate zones and under changing weather situations.
  - Further study of the temporal correlations is required. This dissertation limited itself to accounting for the temporal correlations based on the method of El-Rabbany (1994). However, initial research indicates that the correlation lengths for multipath and atmospheric effects may be different. Methods to determine and account for these differences are required, as well as an assessment on their effects on the adjustment of observations.
  
- The PADRES-GPS software was designed for use on small-extent networks. To simplify the formation of double-differences and the creation of their corresponding variance-covariance matrix, the concept of the differencing base was relied upon, although it played no part in datum definition. On larger network extents, ambiguities become more difficult to resolve for stations further away from the differencing base due to the decorrelation of errors and thus positioning accuracy suffers. Work should be undertaken to modify the PADRES-GPS to eliminate this formation constraint, thus forming double-differences between pairs of closest receivers.
  
- As observations on more frequencies become available, the applicability of forming Optimal Linear Phase Combinations using multiple frequencies should be investigated. Also, the applicability of a particular Optimal Linear Phase



Combination over time should be studied, as it may be discovered that one particular combination is optimal over a range of ionospheric conditions.

The author feels that the above recommended areas of study are only some of the topics that may arise from the work presented herein. Overall, the methods and concepts developed of dealing with spatial distances and the correlations of their errors will be widely applicable as various Global Navigation Satellite Systems are developed and put to use for precise positioning.

## 9.0 - REFERENCES

Ashjaee, J., and R. Lorenz (1992). "Precision GPS Surveying after Y-code." In : *Proceedings of the 5<sup>th</sup> International Technical Meeting of the Satellite Division of the Institute of Navigation*. September 16-18, 1992. Albuquerque, New Mexico. 657-659.

Baarda, W. (1973). *S-Transformations and Criterion Matrices*. Netherlands Geodetic Commission Publications on Geodesy. Volume 5, Number 1.

Beutler, G., W. Gurtner, I. Bauersima, M. Rothacher. (1986). "Efficient Computation of the Inverse of the Covariance Matrix for Simultaneous GPS Carrier Phase Difference Observations." *Manuscripta Geodaetica*. No. 11. 249-255.

Beutler, G., R. Weber, U. Hugentobler, M. Rothacher, and A. Verdun. (1998). "GPS Satellite Orbits". In : *GPS for Geodesy*. Teunissen, P. and A. Kleusberg, eds. Springer-Verlag. 43-109.

Biacs, Z. (1989). *Estimation and Hypothesis Testing*. Ph.D. Dissertation. Department of Surveying Engineering. The University of Calgary.

Bisnath, S. (2000). "Efficient, Automated Cycle-Slip Correction of Dual-Frequency Kinematic GPS Data". *Proceedings of the 13<sup>th</sup> International Technical Meeting of the Satellite Division of the Institute of Navigation*. Salt Lake City.. September 19-22, 2000. 145-154.

Blaaha, G. (1971). *Inner Adjustment Constraints with Emphasis on Range Observations*. Department of Geodetic Science Report No. 148. Ohio State University.

Braasch, M. (1992). *On the Characterization of Multipath in Satellite-Based Precision Approach and Landing Systems*. Ph.D. Dissertation. Department of Electrical and Computer Engineering, Ohio University.

Brunner, F. and W. Welsch (1993). "Effect of the Troposphere on GPS Measurements." *GPS World*. January 1993.

Collins, J. (1999). *Assessment and Development of a Tropospheric Delay Model for Aircraft Users of the Global Positioning System*. M.Sc. Dissertation. Technical Report No. 203. Department of Geodesy and Geomatics Engineering. The University of New Brunswick.

Collins, J. (1989). "The Satellite Solution to Surveying." *Professional Surveyor*. No. 2, Vol. 6. 12-17.

Collins, J. and R. Langley (1998). "The Residual Tropospheric Propagation Delay : How bad can it get?" *Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Nashville, Tennessee, 15-18 September, 1998.

Collins, J. and Langley, R. (1999) *Possible Weighting Schemes For GPS Carrier Phase Observations in The Presence of Multipath*. Technical Report. Contract No. DAAH04-96-C-0086 TCN 98151. Scientific Services Program. Department of Geodesy and Geomatics Engineering. University of New Brunswick.

Counselman, C., I. Shapiro, R. Greenspan and D. Cox (1979). "Astronomical application of differential interferometry." *Science*, No. 178. 607-608.

Davis, J., T. Herring, I. Shapiro, A. Rogers, G. Elgered (1985). "Geodesy by Radio Interferometer : Effects of Atmospheric Modelling Errors on Estimates of Baseline Length." *Radio Science*. Vol. 20, No. 6. 1593-1607.

De Jonge, P. and C. Tiberius. (1996). *The LAMBDA Method for Integer Ambiguity Estimation : Implementation Aspects*. Delft Geodetic Computing Centre LGR Series No. 12.

Duffy, M., C. Hill, C. Whitaker A. Chrzanowski, J. Lutes, G. Bastin (2001). "An Automated and Integrated Monitoring Program for Diamond Valley Lake." *Keynote Presentation at the 10th FIG International Symposium on Deformation Measurements* March 19-22, 2001, Orange California.

Fontana, R., W. Cheung, and T. Stanfell. (2001). "The Modernized L2 Civil Signal; Leaping Forward in the 21<sup>st</sup> Century." *GPS World*. September 2001. 28-34.

Fotopoulos, G. (2000). *Parameterization of DGPS Carrier Phase Errors Over a Regional Network of Reference Stations*. M.Sc. Dissertation. Department of Geomatics Engineering. The University of Calgary.

Francisco, S.G. (1996). "GPS Operational Control Segment". In : *Global Positioning System : Theory and Applications : Volume 1*. Parkinson, B. and Spilker, J., eds. American Institute of Aeronautics and Astronautics. 435-465.

Gerdan, G.P. (1995). "A Comparison of Four Methods of Weighting Double-Difference Pseudorange Measurements." *TransTasmanian Surveyor*. Vol. 1, No. 1. 60-66.

Goad, C., M. Yang (1997). "A New Approach to Precision airborne GPS Positioning for Photogrammetry." *Photogrammetric Engineering and Remote Sensing*. Vol. 63, No. 9. September 1997. 1067-1077.

GPS-SPS (1995). *Navstar GPS Standard Positioning Service Signal Specification*. 2<sup>nd</sup> Edition. GPS NAVSTAR JPO. June 2, 1995.

Gregorius, T. and G. Blewitt (1998). "The Effect of Weather Fronts on GPS Measurements." *GPS World*. May 1998. 52-60.

Grodecki, J. (1997). *Estimation of Variance Covariance Components for Geodetic Observations and Implications on Deformation Trend Analysis*. Ph.D. Dissertation. Technical Report No. 186. Department of Geodesy and Geomatics Engineering. The University of New Brunswick.

Grundig, L. (1985). "Special Adjustment Tools for Surveying Networks." *Papers for the Precise Engineering and Deformation Surveys Workshop*. Department of Surveying Engineering Publication No. 60004. The University of Calgary.

Guo, J. and S. Ge (1997). "Research of Displacement and Frequency of Tall Building under Wind Load Using GPS." In : *Proceedings of the 10<sup>th</sup> International Technical Meeting of the Satellite Division of the Institute of Navigation*. 1385-1388.

Han, S. (1995). "Ambiguity Recovery for GPS for Long Range Kinematic Positioning". *Proceedings of the 8<sup>th</sup> International Technical Meeting of the Satellite Division of the Institute of Navigation*. September 12-15, 1995. The Institute of Navigation Palm Springs. 349-360.

Hansen, P. C. (1998). *Rank Deficient and Discrete Ill-Posed Problems : Numerical Aspects of Linear Inversion*. Society for Industrial and Applied Mathematics.

Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins. (1994). *GPS : Theory And Practice*. Third Edition. Springer-Verlag.

Hopfield, H. (1963). "The Effect of Tropospheric Refraction on the Doppler Shift of a Satellite Signal." *Journal of Geophysical Research*. Vol. 68, No. 18. 5157-5167.

Hopfield, H. (1969). "Two-Quartic Tropospheric Refractivity Profile for Correcting Satellite Data." *Radio Science*. Vol. 74, No. 18. 4487-4499.

ICD-GPS-200C (1993). *Navstar GPS Space Segment / Navigation User Interfaces*., Revision C. GPS NAVSTAR JPO. April 12, 2000.

Ifadis, I., and P. Savvaidis (1999). "Review of Mapping Functions Proposed to Predict the Delay at Very Low Elevation Angles."

Joosten, P. and C. Tiberius. (2000). "Fixing the Ambiguities : Are You Sure They're Right?" *GPS World*. May, 2000. 46-51.

Kim, D., and R. Langley. (2001). "Instantaneous Real-time Cycle-slip Correction of Dual-frequency GPS Data." *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*. June 5-8, 2001. Banff. The University of Calgary. 255-264.

Kleusberg, A. (1998). "Atmospheric Effects in GPS." In : *GPS for Geodesy*. Teunissen, P. and A. Kleusberg, eds. Springer-Verlag. 599-625.

Klobuchar, J. (1986). "Design and Characteristics of the GPS Ionospheric Time-delay Algorithm for Single-Frequency Users." *Proceedings of the IEEE Position Location and Navigation Symposium*. Las Vegas, November 4-7, 1986. 280-286

Klobuchar, J. (1996). "Ionospheric Effects on GPS." In : *Global Positioning System : Theory and Applications : Volume 1*. Parkinson, B. and Spilker, J., eds. American Institute of Aeronautics and Astronautics. 485-514

Koch, K.R. (1988). *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer-Verlag.

Komjathy, A. (1997). *Global Ionospheric Total Electron Content Mapping Using the Global Positioning System*. Ph.D. Dissertation. Technical Report No. 188. Department of Geodesy and Geomatics Engineering. The University of New Brunswick.

Krakiwsky, E. (1975). *The Method of Least Squares : A Synthesis of Advances*. Lecture Notes No. 42. Department of Surveying Engineering. The University of New Brunswick.

Kuang, S. (1996). *Geodetic Network Analysis and Optimal Design : Concepts and Applications*. Ann Arbor Press, Inc.

Kujawa, L. (1998). "Comparison of Different Types of GPS Receivers and Antennas." *Politechnika Warszawska, Reports of Geodesy No. 9 (39)*. Warsaw University of Technology.

Langley, R. (1997) . "GPS Receiver System Noise." *GPS World*. June 1997. 40-45.

Langley, R. (1999). "GPS Receivers and the Observables." In : *GPS for Geodesy*. Teunissen, P. and A. Kleusberg, eds. Springer-Verlag. 151-186.

Lathi, B. (1992). *Linear Systems and Signals*. Berkeley-Cambridge Press.

Leick, A. (1990). *GPS Satellite Surveying*. John Wiley & Sons.

Loverro, D. (2002). "GPS Modernization." *Proceedings of the Institute of Navigation 57th Annual Meeting*. Plenary Session. June 11-13, 2001. Albuquerque, New Mexico. 1-32.

Marini, J. (1972). "Correction of Satellite Tracking Data for an Arbitrary Tropospheric Profile." *Radio Science*. Vol 7, No. 2. 223-231.

Mendes, V. (1999). *Modeling the Neutral Atmosphere Propagation Delay In Radiometric Space Techniques*. Ph.D. Dissertation. Technical Report No. 199. Department of Geodesy and Geomatics Engineering. University of New Brunswick.

Mikhail, E. (1971). *Observations and Least-Squares*. John Wiley & Sons.

Neill, A. (1996). "Global Mapping Functions for the Atmosphere Delay at Radio Wavelengths." *Journal of Geophysical Research*. Vol. 101, No. B2. 51-59.

Newby, S. (1992). *An Assessment of Empirical Models for the Prediction of the Transionospheric Propagation of Radio Signals*. M.Sc. Dissertation. Technical Report No. 160. Department of Surveying Engineering. The University of New Brunswick.

Nickerson, B. (1978). *A Priori Estimation of Variance For Surveying Observables*. Technical Report No. 57. Department of Surveying Engineering. University of New Brunswick.

Parkinson, B. (1996). "Introduction and Heritage of NAVSTAR, the Global Positioning System." In : *Global Positioning System : Theory and Applications : Volume 1*. Parkinson, B. and Spilker, J., eds. American Institute of Aeronautics and Astronautics. 585-599

Parrot, D. (1989). *Short-Arc Orbit Improvement for GPS Satellites*. Ph.D. Dissertation. Technical Report No. 143. Department of Surveying Engineering. University of New Brunswick.

Pope, A. (1972). *Some Pitfalls to be Avoided in the Iterative Adjustment of Nonlinear Problems*. American Society of Photogrammetry, March 12-17, 1972. 449-477.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. (1992). *Numerical Recipes in C. The Art of Scientific Computing*. 2<sup>nd</sup> Ed. Cambridge University Press.



Radovanovic, R., N. El-Sheimy, W. Teskey (2001). "Rigorous Combination of GPS Data From Multiple Base Stations for Mobile Platform Positioning." *3rd International Symposium on Mobile Mapping Technology*. Cairo, January 3-5, 2001.

Rao. C. (1971). "Estimation of Variance and Covariance Components – MINQUE." *Journal of Multivariate Analysis*. No. 1. 257-275.

Raquet, J. (1998) *Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers*. Ph.D. Dissertation. The University of Calgary.

Ray, J. (2000). *Mitigation of GPS Code and Carrier Phase Multipath Effects Using a Multi-Antenna System*. Ph.D. Dissertation. Department of Geomatics Engineering. The University of Calgary.

Remondi, B. (1984). *Using the Global Positioning System Phase Observable for Relative Geodesy : Modelling, Processing and Results*. Ph.D. Dissertation. Austin : Centre for Space Research, University of Texas at Austin.

Roulston, A., N. Talbot and K. Zhang. (2000). "An Evaluation of Various GPS Satellite Ephemerides." *Proceedings of the 13<sup>th</sup> International Technical Meeting of the Satellite Division of the Institute of Navigation*. September 19-22, 2000. The Institute of Navigation. Salt Lake City. 45-54.

Russell, S., and J. Schaibly. (1980). "Control Segment and User Performance". In : *Global Positioning System : Papers Published in Navigation : Volume 1*. Janiczek, P. ed. Institute of Navigation. 74-86.

Saastamoinen, J. (1973). "Contributions to the theory of atmospheric refraction." In three parts. *Bulletin Geodesique*, No. 105, 279-298; No. 106, 383-397; No. 107, 13-34.

Schaer, S. (1999). *Mapping and Predicting the Earth's Ionosphere Using the Global Positioning System*. Ph.D. Dissertation. Astronomisches Institut der Universität Bern. Universität Bern.

Schüler, T. (2001). *On Ground Based GPS Tropospheric Delay Estimation*. Ph.D. Dissertation. Studiengang Geodäsie und Geoinformation. Universität Der Bundeswehr München.

Schnadelbach, K. and T. Falkenberg (1987). "The Interaction Between Error Ellipses and Characteristic Oscillations of Geodetic Networks." *XIX-IGG / IUGG General Assembly*. August 9-22, 1987. Vancouver, Canada.

Seeber, G. (1993). *Satellite Geodesy : Foundations, Methods, and Applications*. Walter de Gruyter.

Skone, S. (1998). *Wide Area Ionosphere Grid Modelling in the Auroral Region*. Ph.D. Dissertation. Department of Geomatics Engineering. The University of Calgary.

Spilker, J. (1980). "GPS Signal Structure and Performance Characteristics". In : *Global Positioning System : Papers Published in Navigation : Volume 1*. Janiczek, P. ed. Institute of Navigation. 29-54.

Spilker, J. (1996). "Tropospheric Effects on GPS." In : *Global Positioning System : Theory and Applications : Volume 1*. Parkinson, B. and Spilker, J., eds. American Institute of Aeronautics and Astronautics. 517-547.

Teunissen, P. (1993). *Least-Squares Estimation of the Integer GPS Ambiguities*. Delft Geodetic Computing Centre LGR Series. No. 6.

Teunissen, P. (1997) "A Canonical Theory for Short GPS Baselines." In four parts. *Journal of Geodesy*. No.71, 320-336; No. 71, 389-401; No. 71, 486-501; No. 71, 513-525.

Teunissen, P. (2000). "The Success Rate and Precision of GPS Ambiguities." *Journal of Geodesy*. No. 74. 321-326.

Teunissen, P. (2002). "The Parameter Distributions of the Integer GPS Model." *Journal of Geodesy*. No. 76. 41-48.

Teunissen, P., P. de Jonge, C. Tiberius. (1994). "On the Spectrum of the GPS DD-Ambiguities." *Proceedings of the 7<sup>th</sup> International Technical Meeting of the Satellite Division of the Institute of Navigation*. September 20-23, 2000. The Institute of Navigation Salt Lake City. 115-124..

Teskey, W. (1987). *Integrated Analysis of Geodetic, Geotechnical, and Physical Model Data to Describe the Actual Deformation Behaviour of Earthfill Dams under Static Loading*. Ph. Dissertation. Institut fuer Anwendungen der Geodaesie im Bauwesen. Universitaet Stuttgart.

Tiberius, C., N. Jonkman, F. Kenselaar (1999). "The Stochastics of GPS Observables." *GPS World*. February, 1999. 49-54.

Tsuji, H., Y. Hatanaka, S. Miyazaki (1996). "Tremors : Monitoring Crustal Deformation in Japan." *GPS World*. April, 1996. 18-30.

United States of America, Office of Science and Technology Policy (2000). "Statement by the President Regarding The United States' Decision to Stop Degrading Global Positioning System Accuracy." Press Release. Office of the Press Secretary, United States of America. May 1, 2000.

Vanicek, P., Krakiswsky, E. (1986). *Geodesy : The Concepts*. Elsevier.

Wang, J., M. Stewart, M. Tsakiri. (1998). "Stochastic Modeling for Static GPS Baseline Data Processing." *Journal of Surveying Engineering*. November 1998. 171-181.

Ward, P. (1996a). "Satellite Signal Acquisition and Tracking" In : *Understanding GPS Principles and Applications*. Kaplan, E., ed. Artech House. 19-207.

Ward, P. (1996b). "Effects of RF Interference on GPS Satellite Signal Receiver Tracking" In : *Understanding GPS Principles and Applications*. Kaplan, E., ed. Artech House. 209-236.

Yong-Qi, C. (1983). *Analysis of Deformation Surveys – A Generalized Method*. Ph.D. Dissertation. Technical Report No. 94. Department Surveying Engineering. University of New Brunswick.

Zumberge, J., and W. Bertiger. (1994). "Ephemeris and Clock Navigation Message Accuracy". In : *Global Positioning System : Theory and Applications : Volume 1*. Parkinson, B. and Spilker, J., eds. American Institute of Aeronautics and Astronautics. 585-599