Department of Geomatics Engineering

Parameterization of DGPS Carrier Phase Errors Over a Regional Network of Reference Stations
(URL: http://www.geomatics.ucalgary.ca/GradTheses.html)

by

Georgia Fotopoulos

August 2000

UNIVERSITY OF
CALGARY

UNIVERISTY OF CALGARY

Parameterization of DGPS Carrier Phase Errors Over a Regional Network

of Reference Stations

by

Georgia Fotopoulos

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

AUGUST, 2000

ABSTRACT

As the distance between a GPS mobile user and a reference receiver increases, differential GPS (DGPS) positioning errors become more decorrelated. These distance-dependent errors include ionospheric, tropospheric, and satellite orbit errors that do not cancel or reduce as they do for short baseline cases. This limits the ability for carrier phase ambiguities to be resolved and results in a poorer positioning accuracy. Over the past few years, a significant amount of research has been conducted on the formulation of carrier phase corrections using multiple reference stations in order to enhance ambiguity resolution and to increase the distances over which precise positioning can be achieved. Recently the use of a network of multiple GPS reference stations for generating carrier phase based corrections has emerged with great promise for use in real-time environments. However, little research has been conducted on the distribution of these corrections to potential GPS users located within, and surrounding, the network coverage area. This is an integral part of real-time kinematic DGPS and it must be adequately addressed before a practical realization of the multi-reference station concept is implemented.

In order to assess the positioning performance, it is important to know how the corrections change over the network area. This involves examining issues such as the spatial and temporal parameterization of these corrections and the correction transmission rate. This thesis investigates two methods for parameterizing corrections, namely generating a correction grid for the desired coverage area on an epoch-by-epoch basis as

well as approximating the correction behavior by a low-order surface. Advantages and disadvantages associated with each method are also discussed and supported with network data results. From a real-time implementation perspective, the correction update rate is evaluated as a function of position accuracy. Results on the spatial and temporal behavior of the errors over a permanent array of reference stations located in Sweden are presented with an emphasis on the overall effect in the position domain.

PREFACE

This is an unaltered version of the author's Master of Science thesis of the same title. This thesis was accepted by the Faculty of Graduate Studies in August, 2000. The faculty supervisor of this work was Dr. M. Elizabeth Cannon, and the other members of the examining committee were Dr. G. Lachapelle, Dr. S. Skone and Dr. D. Irvine-Halliday.

ACKNOWLEDGEMENTS

I would like to express my appreciation and thank my supervisor, Dr. M.E. Cannon, for her support and guidance throughout my graduate studies. Her continuous encouragement and advice were greatly appreciated. The remaining professors on my supervisory committee, Dr. Lachapelle, Dr. Skone, and Dr. Irvine-Halliday are also thanked for their comments improving the final version of this thesis.

I would also like to thank all of my colleagues, especially the navigation group, who have helped me along the way with our 'scientific' discussions. My very good friend and colleague, Christopher Kotsakis, is also thanked for his insightful comments during the proofreading of my thesis.

Funding for my Master's studies was provided from numerous sources, including the Natural Sciences and Engineering Research Council (NSERC), the University of Calgary, the Department of Geomatics Engineering as well as my supervisor. The National Land Survey of Sweden and the Onsala Space Observatory are gratefully acknowledged for making the SWEPOS data available.

Finally, I would like to thank my dear brothers for providing a very noisy and impossible work environment at home, which allowed me to procrastinate the writing of this manuscript. Last, but not least I thank my parents for their endless support and encouragement. I could not have done it without their support.

Στους γονείς μου,

Αθανάσιο και Ειρήνη

και στα αδέλφια μου Κώστα, Αλέκο, και Χρήστο.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

L1   GPS signal transmitted at a frequency of 1575.42 MHz

L2   GPS signal transmitted at a frequency of 1227.60 MHz

p Pseudorange measurement

ρ Geometric range between the satellite and the receiver antenna

$d\rho$ Satellite orbit errors

d_{ion} Ionospheric delay

ϕ Carrier phase measurement

λ Carrier phase wavelength

N Carrier phase ambiguity

Δ Single difference operator

$\Delta\nabla$ Double differencing operator

db Baseline error

f Frequency

σ Standard deviation

$\delta\ell$ Corrections generated for the carrier phase observations

$C_{\delta\ell}$ Covariance matrix of network reference station carrier phase observations

d Baseline distance

$R(kT)$ Discrete autocorrelation function

P_h Power spectral density function

Δx Grid resolution in the longitudinal direction

Δy Grid resolution in the latitudinal direction

LIST OF ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
C/A	Coarse/Acquisition
CAA	Conditional Adjustment Algorithm
DGPS	Differential Global Positioning System
FFT	Fast Fourier Transform
GPS	Global Positioning System
IF	Ionospheric-Free
IS	Ionospheric Signal carrier phase combination
LIA	Linear Interpolation Algorithm
OTF	On-the-fly
PDA	Partial Derivative Algorithm
PRN	Pseudo-random Noise
PSD	Power Spectral Density
RINEX	Receiver Independent Exchange Format
RMS	Root Mean Square
RTK	Real-Time Kinematic
SA	Selective Availability
SSN	Southern Swedish Network
SV	Satellite Vehicle
UTC	Universal Time Coordinated
VRS	Virtual Reference Station
WAAS	Wide Area Augmentation System
WADGPS	Wide Area Differential GPS
WL	Widelane carrier phase combination

Chapter 1

INTRODUCTION

1.1   Background

Over the years, the Global Positioning System (GPS) has evolved into a significant tool

for civilian navigation and positioning. Today new and challenging applications are

emerging which demand that this tool is used to its maximum potential. In this view, the

key motivation of this thesis is to establish a means whereby users of GPS can obtain

very high accurate positions (at the centimetre-level) for a number of applications.

The use of a reference station generating corrections for various error sources, to be

applied to a rover/user station and known as differential GPS (DGPS), is a common

method for achieving higher accuracy (Leick, 1995). However, as the civilian

community of GPS users increases and the range in applications grows, the limits of the

accuracy envelope are being stretched even further. This has prompted the development

of a relatively novel concept whereby multiple reference stations are used to generate

corrections for error sources which routinely inhibit the achievable accuracy level.

To date, numerous regional networks consisting of an established array of permanent reference stations are scattered throughout the world. With these infrastructures in place, it was only a matter of time before multiple reference station algorithms would be developed which employ the network concept. Such algorithms include the partial derivative algorithm (Wübbena et al., 1996; Varner, 2000), the linear interpolation algorithm (Wanninger, 1995; Han and Rizos, 1996), and the condition adjustment algorithm (Raquet, 1998), to name a few. An overview of these methods is also provided in Fotopoulos and Cannon (2000b). Regardless of the method/algorithm employed, they are all subject to the disturbances caused by the atmosphere (labeled as ionosphere and troposphere for GPS purposes) and inaccuracies in the satellite orbit prediction, not to mention multipath and receiver noise errors. For high accuracy positioning applications such as post-glacial rebound studies, long-term deformation monitoring, and other geophysical surveying areas, these arrays have been instrumental for providing positioning information, since the errors can be modelled in post-mission using large volumes of archived data to derive observed positions. However, for many applications this post-mission information and analysis is insufficient. Users require data and position information 'as it happens', thus pushing the need for higher accuracy in (near) real-time. Real-time applications include navigating vehicles on land, navigating ships in marine environments, earthquake monitoring, and so on. Thus, the problem increases in complexity to include cm-level accuracy in real-time.

Real-time applications pose a challenging area of research for many users, as it is difficult to predict the behaviour of detrimental error sources at every epoch in time. Global indicators of the level of atmospheric activity available to GPS users have limited usefulness for a user who is only affected by very local phenomena. In this respect, the regional network, spanning a limited area over a few hundreds of kilometres in latitude and longitude, can benefit the user by providing valuable information regarding the local trends, behaviour and activity of various error sources, namely the distance-dependent or correlated error sources.

The major issues for study are not only focused on the formulation of network corrections for errors, but also on the communication and subsequent dissemination of these data to the user in an accurate, reliable and timely manner. It is in this latter area where the focus of this research lies. Practically speaking, a GPS user located within and/or surrounding a permanent array coverage area, must be able to obtain horizontal and vertical accuracies at least at the cm-level, in real-time. For such purposes, the more precise GPS observable, that is the carrier phase, is used. Code based (and carrier smoothed code) implementations of the network approach have been successfully developed for a number of years (Kee and Parkinson, 1992; Ashkenazi, et al., 1992; Abousalem and Bogle, 1997). However, the more precise applications bring to light issues which were often masked by the error in the code itself and therefore, they were not investigated. Also, as the number of civil GPS users increases, it is desirable to provide services which have a seamless integration, requiring few modifications to existing equipment and software.

Therefore, all proposed multiple reference station methodologies should be accompanied with practical implementation schemes. This will benefit GPS users by enhancing existing technologies and software, rather than making them obsolete. The objectives described in the following section have been identified as the most important aspects of this research.

1.2   Objectives

Given the eminent implementation of regional networks across the globe, it is imperative for operators and users of such systems to understand all aspects associated with multiple reference station approaches. The focus of this research work is to provide an understanding on the efficient spatial parameterization and dissemination of corrections and offer possible options for real-time use. Therefore, the following research objectives have been identified as the most important aspects of this thesis.

(i)   To investigate various options available for the spatial parameterization (or modelling) of corrections, formulated for reducing the effects of the correlated error sources, as a function of user position. This involves the identification of viable methodologies, which can be relatively seamlessly integrated into existing user equipment and software. Spatial aspects such as the dimensionality of the problem and any trends in the correlated error sources should also be defined.

Results supporting the spatial analysis in the position domain must be included in order to determine the achievable accuracy at the user level.

(ii)   To analyze the temporal characteristics of the spatially correlated error sources over a regional network. The relationship between the temporal aspects investigated and the various spatial parameterization techniques should also be assessed. Results in both the correction and the final position domain must be included in order to substantiate the variation in achievable accuracy over time.

(iii)   To identify the most viable data transmission options available for the implementation of the multi-reference approach in real-time. This involves the actual transmission information required for the various parameterization schemes as well as the required transmission rates derived from the temporal analysis.

It should be noted that each of the three objectives listed above are inter-related. Therefore, throughout this thesis results relating both spatial and temporal aspects will be presented. Detailed analysis of each area is conducted in separate arenas and brought together in various discussions of results in order to provide an integrated set of options for the parameterization and communication of corrections.

1.3   Thesis Outline

The statement of the problem and related issues have been used to define the objectives of this research in the previous section. Essentially, this thesis is structured in such a way as to address each of the objectives separately in detail in Chapters 2 through 5, as follows.

The discussion in Chapter 2 contains the background information on related areas of research as well as overview discussions based on fundamental GPS theory and multiple reference station methodologies. This includes an overview of the derived observables used for differential GPS processing and their error sources. For the purposes of this research, the error sources are separated in terms of their distance-dependency, with the uncorrelated errors discussed first followed by the correlated errors. A number of useful carrier phase combinations used for generating corrections are derived and detailed for future reference. It is assumed throughout this work that the reader is somewhat familiar with basic GPS concepts. However, as a supplement to the overview presented in this chapter, a number of useful references are included here for further reading on GPS theory.

The second part of this chapter introduces the multi-reference station network concept. Details on four specific categorized methodologies are included and brief inter-comparisons are made between each of the algorithms. Since a number of multi-reference correction formulation methodologies currently exist, an introduction to the

algorithm used for the research conducted herein will also be presented. In addition, the context of this work in the multi-reference arena is identified and defined.

In Chapter 3 the issues presented in objective (i) above, are addressed. A test network of regional network data is introduced and will be used for the majority of the analysis throughout this thesis. Specifically, the spatial parameterization options are described and analyzed providing results in both the correction and position domains. The two main categories of spatial parameterizations investigated are a (1) grid-based approach and (2) low-order surface modelling. The advantages and disadvantages of each approach are discussed and supported with results at the user level.

Chapter 4 continues the discussion on parameterizations, but from a temporal point of view. Here, the temporal characteristics of the correlated error sources are investigated over the coverage area. Issues such as data decimation, varying the time of day, and spatial parameterization scheme update rates, are analyzed in detail with supporting results in the correction and position domains. This chapter focuses on addressing objectives (ii) and (iii). Again, the analysis in the position domain brings the discussion to the user level where the achievable accuracy of the final output is observed.

The spatial and temporal analyses presented in the previous two chapters provide important information regarding the correction field behaviour over the regional network coverage area. However, there are limits to the analysis that could be conducted as the

results could only be characterized in the space domain. Although this provides a solid foundation for the analysis, further tests are conducted in the spectral domain in order to verify and/or identify any inconsistencies. In Chapter 5, both the spatial and temporal aspects are looked at from a spectral analysis point of view. The key motivation behind this analysis is to confirm, if possible, the results obtained in the previous chapters using an independent methodology, as well as extract any additional information on the spectral content of the correction field. Results using the same data sets as in the previous chapters are presented for consistency and comparison purposes.

Finally, in Chapter 6 conclusions based on the research work presented in the previous chapters will be outlined. Since this area is a topic of ongoing research and only a sub-set of the modules involved in the practical realization of a real-time multi-reference station approach were investigated, a number of recommendations for future work are also included in this chapter.

Chapter 2

OVERVIEW OF GPS THEORY AND REGIONAL NETWORK ALGORITHMS

The discussion presented in this thesis assumes that the reader is somewhat familiar with the basic theory of the Global Positioning System (GPS). An overview of the fundamental GPS observables and differencing processes will be given, followed by a description of the major error sources which inhibit the positioning accuracy achievable using differential GPS (DGPS).

Since its inception in the early 1970s (Parkinson, 1996) GPS has become a common civilian tool for a widespread base of applications focusing on, but not limited to, positioning and navigation. The advancement has also prompted numerous additions to GPS literature covering fundamental aspects in great detail. For a basic overview of GPS the reader is referred to Leick (1995) and Hofmann-Wellenhof et al. (1994). A more detailed and more advanced account of GPS fundamentals is provided in Parkinson and Spilker (1996a), which is complimented by an applications-oriented approach in a second volume, Parkinson and Spilker (1996b).

The topic presented herein deals with precise positioning, often sought after for geodetic applications. In such cases it is advantageous to study the positioning problems according to the inter-station baseline lengths, since the analysis and processing may vary according to the distance between stations (see Teunissen and Kleusberg, 1996, chapters 11-13 for more details). Another useful reference, which outlines both GPS fundamentals and issues related to precise geodetic applications, is Seeber (1993). The GPS signal specifications can be found in the *GPS Signal Specification Interface Control Document* (ICD-GPS-200C, 1993). Finally, an excellent additional reference for receiver operations and signal tracking aimed at GPS users can be found in Kaplan (1996).

It should also be noted that GPS is currently undergoing some major changes in signal availability and specifications, which will impact how it is used in the future. However, for the applications in this thesis, an understanding of basic GPS theory and its error sources provides the stepping stone for the spatial and temporal investigations presented in later chapters.

2.1   GPS Observables

The two fundamental GPS observables used to compute a position are the pseudorange p and the carrier phase ϕ. The pseudorange can be thought of as the propagation time required for the signal transmitted from the satellite to reach the user receiver antenna (Kaplan, 1996). If the satellite and receiver clocks were perfectly synchronized with each

other and the GPS system time, p would represent the true geometric range between the satellite and the receiver antenna. However, this is not the case and because of the bias between the system time and the receiver clock, and to a lesser extent between the system time and the satellite clocks, the range measurement is called a *pseudo*range. The integrated carrier phase measurement is made on the beat frequency, which is the difference between the receiver generated reference frequency and the actual received frequency.

The two main frequencies that GPS satellites currently transmit on are known as L1 and L2 and are at 1575.42 MHz and 1227.6 MHz, respectively. Currently the coarse acquisition (C/A) code is modulated only on L1 and the P-code is modulated on both frequencies, although the P-code is denied to civilian users by the United States Department of Defence, through encryption (called anti-spoofing or AS). However, as part of the *GPS Modernization* the addition of the second civil signal to GPS satellites is scheduled to begin in 2003, with both civil signals being broadcast with the same C/A code. Furthermore, it should be noted that a third civilian frequency has recently been approved by the United States Department of Defence and after much deliberation it is planned to be at 1176.45 MHz (Challstrom, 1999). This so called L5 frequency will become available to civilian users in the near future when the first set of satellites equipped with this capability are launched in 2005 (more information on this and other GPS modernization efforts can be found on the NAVSTAR GPS Joint Program Office webpage at *www.laafb.af.mil/SMC/CZ/homepage/*).

Due to the fact that the GPS signals do not travel through a vacuum, but rather complex media, all of the measurements are affected by several error sources. A mathematical representation of a pseudorange measurement p, including these error sources, is provided in the following basic equation:

$$p = \rho + d\rho + c(dt - dT) + d_{ion} + d_{trop} + \varepsilon_{m_p} + \varepsilon_p \qquad (2.1)$$

where

ρ is the geometric range between the satellite and the receiver antenna in metres,

$d\rho$ is the satellite orbit error in metres,

c is the speed of light (299792458.0 metres/second),

dt is the satellite clock error term with respect to GPS time in seconds,

dT is the receiver clock error term with respect to GPS time in seconds,

d_{ion} is the ionospheric delay error in metres,

d_{trop} is the tropospheric delay error in metres,

ε_{m_p} is the code range multipath error in metres, and

ε_p is the receiver code noise in metres.

The carrier phase measurement ϕ (in metres) can also be mathematically represented by a formulation parallel to eq. (2.1), with some modifications, as follows:

$$\phi = \rho + d\rho + c(dt - dT) + \lambda N - d_{ion} + d_{trop} + \varepsilon_{m\phi} + \varepsilon_{\phi} \qquad (2.2)$$

where

λ is the carrier wavelength in metres,

N is the integer ambiguity in cycles,

$\varepsilon_{m\phi}$ is the carrier phase multipath error in metres, and

ε_{ϕ} is the receiver carrier noise in metres.

A descriptive overview of the error sources present in eqs. (2.1) and (2.2) will be provided in Section 2.2. For now, it is interesting to note some of the differences between the code and carrier formulations. Firstly, the ionospheric delay error term d_{ion} has the opposite sign in both expressions. This is due to the fact that the code is delayed by the ionosphere while the phase is advanced (see Leick, 1995, pp. 294-298 for more details on code delays and phase advances). Another difference between the two equations is the inclusion of an ambiguity term N in eq. (2.2), which represents the number of whole carrier waves between the satellite and receiver. This term accounts for the ambiguous nature of carrier phase measurements as opposed to the absolute character of pseudoranges. The resolution of the ambiguity term is a topic of great research that has been advancing over the years. However, it continues to be one of the biggest obstacles posed for precise carrier phase based positioning over longer baseline distances. A discussion on ambiguity resolution related to the analysis presented in this thesis is

provided in Section 3.2.1. Other differences include the code noise and code multipath terms in eq. (2.1), which are replaced by the carrier noise and carrier multipath in eq. (2.2), respectively and are discussed below.

2.2   Differential GPS Carrier Phase Error Sources

The observables presented in the previous section are most useful when they are used in a derived difference form, which significantly reduces the effects of some of the error sources. The level of reduction depends on the distance between the receivers (known as the baseline). Throughout the discussion in this thesis GPS errors will be categorized according to their distance-based correlation. That is, some of the errors are correlated over space (or distance) and are therefore reduced when differencing between satellites and receivers as in differential GPS processing. As mentioned above, the correlation depends on the baseline length. For short baselines (i.e. hundreds of metres), the effects of the correlated errors, namely ionospheric, tropospheric and satellite orbit are reduced. Depending on the application and the required accuracy level, the effects of these errors when using DGPS may be considered negligible. Uncorrelated errors are those which do not depend on the baseline distance (may be referred to as receiver/site specific errors) and are generally multipath and receiver noise.

One such differencing option, known as between receiver single differencing, is represented by the following equation for the pseudorange observable:

$$\Delta p = \Delta\rho + \Delta d\rho + c\Delta dT + \Delta d_{ion} + \Delta d_{trop} + \Delta\varepsilon_{m_p} + \Delta\varepsilon_p \qquad (2.3)$$

where Δ is the differencing operator. Similarly, for the carrier phase observable, single differencing can be represented as follows:

$$\Delta\phi = \Delta\rho + \Delta d\rho + c\Delta dT + \lambda\Delta N - \Delta d_{ion} + \Delta d_{trop} + \Delta\varepsilon_{m_\phi} + \Delta\varepsilon_\phi \qquad (2.4)$$

From eqs. (2.3) and (2.4), it is observed that when performing single differencing, the satellite clock error term dt is eliminated. This process is illustrated in Figure 2.1. The receiver whose location is precisely known is called the reference station (also referred to as the base or monitor station) and the receiver whose position is sought after is called the user station/receiver (sometimes called the rover receiver).
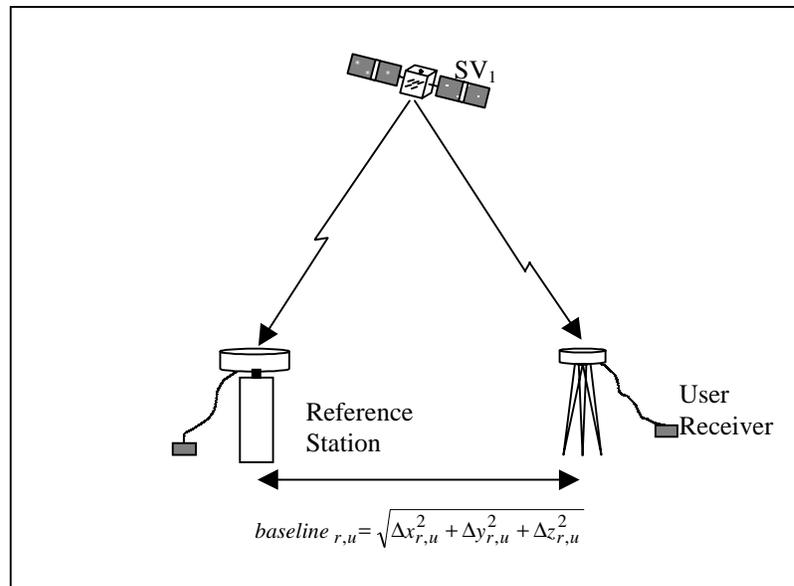
Figure 2.1: Between Receiver Single Differencing ($\Delta\rho$ and $\Delta\phi$)

By performing an additional difference between receivers and satellites, the following double difference equations are formed for the pseudorange and carrier phase observables, respectively:

$$\Delta\nabla p = \Delta\nabla\rho + \Delta\nabla d\rho + \Delta\nabla d_{ion} + \Delta\nabla d_{trop} + \Delta\nabla\varepsilon_{m_p} + \Delta\nabla\varepsilon_p \qquad (2.5)$$

$$\Delta\nabla\phi = \Delta\nabla\rho + \Delta\nabla d\rho + \lambda\Delta\nabla N - \Delta\nabla d_{ion} + \Delta\nabla d_{trop} + \Delta\nabla\varepsilon_{m_\phi} + \Delta\nabla\varepsilon_\phi \qquad (2.6)$$

where $\Delta\nabla$ is the double differencing operator.

Examining the differences between eqs. (2.3) and (2.4) with (2.5) and (2.6), it is evident that an additional error term is eliminated, namely the receiver clock error dT. The remaining correlated error sources are also reduced at some level, depending on the baseline distance. The elimination and reduction of the error sources, as shown above, is a major obvious advantage to using double difference equations in order to compute a receiver's unknown position. However, the disadvantage with differencing is that the noise is increased by a factor of $\sqrt{2}$ with each difference operation (i.e. for the case of double differencing the noise is increased by a factor of 2). Nevertheless, the advantages far outweigh the disadvantages and double differencing is commonly used in practice. This between receiver, between satellite double differencing process is illustrated in Figure 2.2.
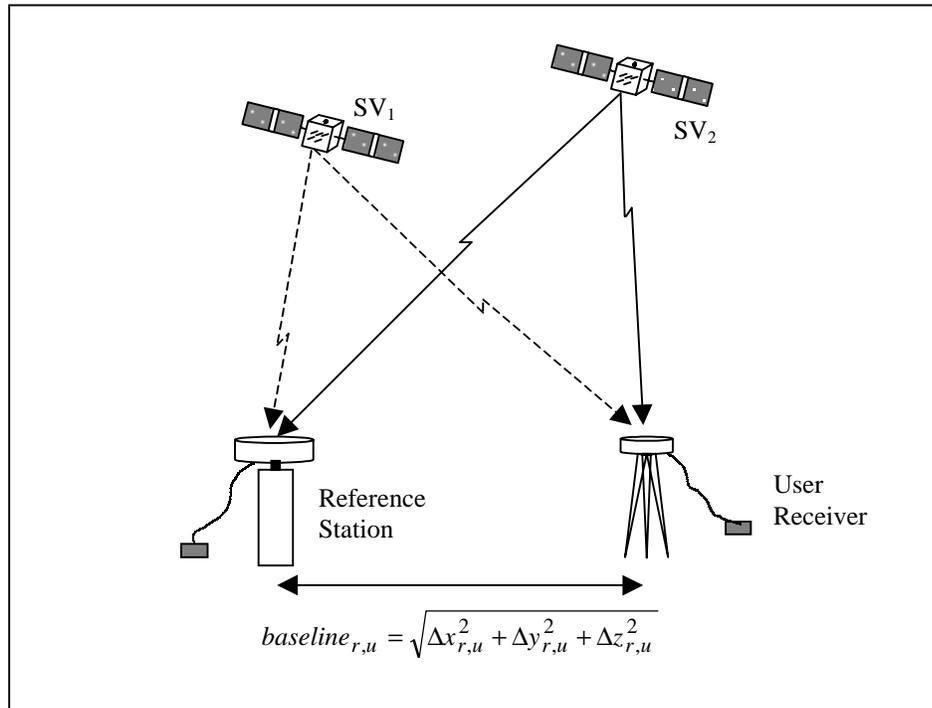
Figure 2.2: Between Receiver Between Satellite Double Differencing ($\Delta\nabla\rho$ and $\Delta\nabla\phi$)

The following few sections will describe the uncorrelated error sources, namely multipath and receiver noise, as well as the correlated error sources (ionospheric, tropospheric and satellite orbit), admissible after double differencing.

2.2.1 Multipath

The largest uncorrelated error source is multipath, a phenomena whereby a signal arrives at the receiver via multiple paths (Braasch, 1996). It is caused by the reflection and diffraction of the transmitted signal by objects in the surrounding area. These surfaces may include a variety of objects such as buildings, streets, vehicles, etc. as demonstrated in Figure 2.3.
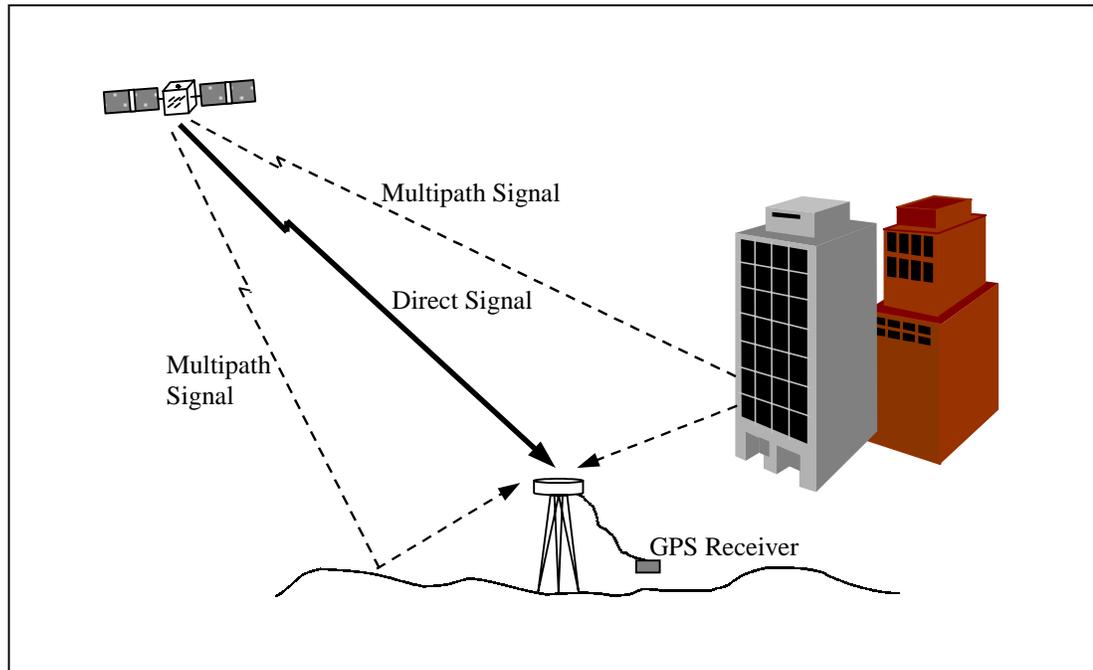
Figure 2.3: Direct and Multipath Signals

Multipath distorts the C/A and P-code modulations and the carrier phase at different levels. For carrier phases, the maximum amount of multipath can be derived as per Seeber (1993, pp. 307-309) and Leick (1995, pp. 311-316), and is summarized below. The direct carrier phase signal S_D can be represented as follows:

$$S_D = A \cos \varphi \tag{2.7}$$

where A is the signal amplitude and φ is the signal phase. In the simplest case, consider one reflected (or multipath) signal S_R, which can be represented as follows:

$$S_R = \alpha A \cos(\varphi + d\varphi) \tag{2.8}$$

where α is the damping factor (or amplitude reduction factor) and always satisfies $0 \le \alpha \le 1$. As α increases from a minimum value of zero to a maximum value of one, the reflected signal S_R strength also increases, equating the direct signal S_D strength at $\alpha = 1$. The phase shift caused by multipath is $d\varphi$. The superposition of the direct and reflected signals shown in (2.7) and (2.8) respectively, combine to form:

$$S_C = S_D + S_R = A \cos\varphi + \alpha A \cos(\varphi + d\varphi) = B \cos(\varphi + \vartheta) \tag{2.9}$$

The resultant signal amplitude B is computed as follows,

$$B = A\sqrt{1 + \alpha^2 + 2\alpha \cos d\varphi} \tag{2.10}$$

and the carrier phase multipath delay can be computed by

$$\vartheta = \tan^{-1}\left(\frac{\alpha \sin d\varphi}{1 + \alpha \cos d\varphi}\right) \tag{2.11}$$

By analyzing eqs. (2.10) and (2.11), it is evident that when $\alpha = 1$, the maximum value for ϑ occurs (namely, 90°). This corresponds to one quarter of the carrier phase signal. For

the case of the L1 phase observable, this value is approximately 4.8 cm, corresponding to the maximum theoretical level of carrier phase multipath.

Pseudorange multipath is similar to carrier phase multipath, with the exception that it is a few orders of magnitude larger. Previously it was shown that the maximum level of carrier phase multipath could be computed as a fraction of the wavelength. In this manner, the maximum level of code multipath is computed as a fraction of the chip length or rather equal to one chip length. Appropriately then, the effect of multipath on the P-code is 30 metres (i.e. chipping rate of 10.23 MHz) as opposed to the C/A code pseudoranges which is 300 metres (i.e. chipping rate of 1.023 MHz).

For more details on multipath and its effects on both carrier phase and code based positioning, refer to Braasch (1996) and Ray (2000). Over the past decade, there have been numerous improvements to receiver and antenna technology which aid in mitigating the effects of multipath such as the narrow correlator (Van Dierendonck et al., 1992) and using antennas equipped with choke rings or ground planes. Most recently, a promising new development for reducing carrier phase multipath was developed at the University of Calgary, which uses an array of antennas for mitigating multipath effects (see Ray, 2000 for more details).

Despite the improvements and advances of all of these technologies, the simplest method for most GPS users to mitigate multipath is simply to avoid it, by selecting receiver

station sites that are free of any reflective obstructions. For the case of multiple reference

or permanent regional arrays, this concept applies to all contributing reference stations.

Ideally, reference stations are set up at carefully selected sites, which are located in

relatively low multipath environments. However, often the location of the sites is

dictated by existing infrastructure and may or may not be ideally situated for GPS

purposes. The difficulty in finding a low multipath environment is enhanced for user

stations (especially in kinematic applications), where the user's environment is inherently

decreed by the application and most probably corrupted with multiple obstructions.

2.2.2 Receiver Noise

Another uncorrelated error source is receiver noise. This is any noise which is generated

by the receiver itself during the measurement process of both code and phase

observables. Compared to the other error sources, receiver noise is considered to be

small in magnitude, however, as the ability to model and mitigate other error sources

improves the effects of this error source becomes more prevalent.

For the code measurements, high frequency thermal noise jitter and the effect of dynamic

stress on the receiver's code tracking loop cause receiver noise. As in the case of

multipath, the magnitude of the error on the C/A-code and P-code varies depending on

the chip width. Thus, a C/A code receiver noise is usually one order of magnitude higher

compared to that of a P-code receiver. Receiver noise on the carrier phase measurements

can be attributed to the thermal noise, dynamic stress and oscillator phase noise and is on the order of a few millimetres, for most modern receivers. It should be noted that the magnitude of receiver noise is very much dependent on the receiver itself and can effectively be measured through a zero-baseline test. As a general rule, receiver noise tends to be less than 1% of the observable wavelength (Weisenburger, 1997).

2.2.3 Satellite Orbit Errors

For the majority of GPS users there are at least two options available for computing the satellite positions, that is, using broadcast ephemeris information provided in the broadcast navigation message, or using precise ephemeris. Although precise ephemeris information is much more accurate than the broadcast information, it has the limitation of only being available to the user in post-mission. Therefore, for real-time applications, the current source for satellite orbit information is what is broadcast in the navigation message. The satellite positions are computed from a set of Keplerian orbit and perturbation parameters and clock parameters, which are predicted states for the satellite orbit and are updated every two hours. The slant range RMS accuracy associated with broadcast orbits is approximately 2.1 metres, whereas for precise orbits this is reduced to approximately 6 cm (Rothacher, 1997).

For short baselines, the orbit error is cancelled when differential processing is performed. However, satellite orbit errors are spatially correlated and the baseline length establishes

the level of cancellation/reduction. A conservative and perhaps even pessimistic estimate of the decorrelation of satellite orbit errors based on the baseline length is provided by the following linear relationship:

$$\frac{db}{b} = \frac{dr}{\rho} \tag{2.12}$$

where db is the baseline error for a baseline length b. The satellite range is represented by ρ and used to compute the orbit error dr. Using a nominal satellite range of 22,000 km above a set of receivers located on the surface of the Earth, forming various baselines, the expected orbit error as computed by eq. (2.12) is summarized in Table 2.1. To illustrate this point, a maximum allowable baseline error of 1 cm is used with baseline lengths ranging from 100 metres to 500 km. As evidenced in the table, the allowable errors in satellite position computation decreases as the baseline length increases.

Table 2.1: Admissible Orbit Errors for Various Baselines when $db = 1\,cm$

Baseline Length, b (km)	Orbit Error, dr (m)
0.10	2200
10	22
100	2.2
250	0.88
500	0.44

Ultimately, our goal is to provide a cm-level positioning accuracy for (near) real-time applications, thus, the broadcast ephemeris provided in the navigation message was used

for all data processed in the following chapters (excluding the ambiguity resolution process as described in Section 3.2.1).

2.2.4 Tropospheric Delay Errors

Atmospheric errors account for the largest part of the correlated error sources. For GPS users, the atmosphere is separated into two main areas, the troposphere (or neutral atmosphere) and the ionosphere. Understanding the behaviour of both over regional multi-reference station networks is imperative in order to obtain results in the cm-level domain. In this section, a brief description of the troposphere and the concerns associated with GPS will be presented. The next section will focus on the ionospheric errors.

For the purposes of GPS, the troposphere is considered to be a neutral part of the atmosphere ranging from 0 km to 40 km above the surface of the Earth. When GPS signals are transmitted from the satellites and travel through the troposphere they suffer the effects of tropospheric attenuation, delay and short-term variations (scintillation). The magnitude of these effects are a function of the satellite elevation and conditions such as the temperature, pressure and relative humidity of the atmosphere during signal propagation. The troposphere is also a non-dispersive medium for GPS frequencies, which means that tropospheric range errors are not frequency dependent and therefore cannot be cancelled through the use of dual-frequency measurements (unlike the ionosphere, see Section 2.2.5).

In terms of the magnitude of the delay errors, it is useful to study the troposphere in two components often referred to as the wet and dry components. The dry part contributes 80% to 90% of the delay, however it can be modelled with an accuracy of less than 1%. The wet part, although contributing only 10% to 20% of the delay, is more unpredictable and therefore more difficult to model. The dry tropospheric zenith delay is approximately 2.3 metres (the majority of which can be modelled as stated above), while the wet tropospheric zenith delay ranges from approximately 1 cm to 80 cm (Spilker, 1996). For shorter baselines with receiver locations that do not have a large vertical height separation, the majority of these tropospheric effects cancel through differential processing because of its non-dispersive property. However, for longer baseline processing, the spatial correlation of the troposphere becomes more evident and the residual tropospheric delay remaining after double differencing may be quite significant.

Over the past few years, numerous models have been developed, which reduce the effects of the troposphere. Some of the most common models are the Hopfield model (Hopfield, 1969), the Saastamoinen model (Saastamoinen, 1972), and the modified Hopfield model (Goad and Goodman, 1974). The required input parameters for each of these models vary as does their performance. An excellent reference which provides a comprehensive description of these models and their differences can be found in Mendes (1999), (see also, Hofmann-Wellenhof et al., (1994), pp. 108-117, and Spilker (1996), pp. 517-544). In order to gain some insight into the behaviour and effect of the residual tropospheric effect over a network of multiple reference stations, it is also useful to analyze the

ionospheric-free carrier phase measurement combination (see Section 2.3), which does not contain the first order ionospheric effects and therefore it is indicative of the remaining correlated errors. This approach was used in order to estimate the residual tropospheric delays for GPS networks located in Norway and part of Sweden (see Zhang, 1999 for more details). Reported levels of residual double difference ionospheric free RMS values for data collected on April 30, 1998 on a five station network in southern Sweden were on the level of approximately 1.6 cm to 3 cm for baselines ranging from 70 km to 200 km respectively (Zhang, 1999).

2.2.5   Ionospheric Delay Errors

The ionosphere can be considered as one of the most detrimental factors affecting high accuracy positioning. It is a weakly ionized plasma extending from approximately 50 km up to 1000 km above the surface of the Earth and can affect radiowave propagation in various ways (Klobuchar, 1996). At GPS frequencies the ionosphere's presence is experienced through group delay, carrier phase advance, range-rate errors, and amplitude and phase scintillations, to name a few. However, unlike the troposphere, it is a dispersive medium, which means that dual-frequency users can correct for the first order ionospheric range (and range-rate) effects. These range errors vary from a few metres to tens of metres at the zenith (Klobuchar, 1996) depending on numerous factors such as, the time of day, time of year, place in solar cycle and the user's geomagnetic latitude

(higher effects are usually experienced in the Northern latitudes and the equatorial region).

In addition to using dual-frequency measurements, the remaining ionospheric error is further reduced through double differencing. Normally, over short baselines DGPS users can effectively remove the majority of the ionospheric errors. However, as positioning over larger baselines is performed, effectively resulting in a larger coverage area, the spatial decorrelation of the ionosphere increases and residual effects remain even after differencing. The effect of these residual effects can pose as a significant limitation for regional network users (see El-Arini et al., 1995; Skone, 1998). Recent investigations conducted in the area of the St. Lawrence Seaway showed maximum differential ionospheric effects of up to 10 ppm (1 m in 100 km) during a test period conducted in November of 1998 (Lachapelle et al., 1999). Further tests conducted in August of 1999 revealed the ionospheric activity in the same region to be relatively active with a RMS differential effect of 2 ppm to 3 ppm and a maximum differential ionospheric effect in excess of 7 ppm (Lachapelle et al., 2000).

As with the troposphere, numerous empirical and physical ionospheric models exist to combat the effects of the ionosphere. Models over an extensive network area usually involve the formulation of the ionospheric effects at discrete grid points on an ionospheric shell located at approximately 350 km in altitude. However, modelling ionospheric effects continues to be a very challenging area of research because of its

relatively unpredictable behaviour. The ionospheric-free linear phase combination can be derived which removes the first order ionospheric effects for dual-frequency users. A brief description and outline of the derivation can be found in the following section on carrier phase combinations.

2.3   Carrier Phase Combinations

For applications which use a network of multiple reference stations in order to position one or more users in real-time, it is expected that dual-frequency receivers will be used. As it was mentioned previously, over short distances, the impact of the ionosphere is negligible. However, as the baseline lengths increase, the spatial decorrelation of the ionosphere becomes more evident, and does not necessarily cancel. In an attempt to improve the attainable position accuracy and ambiguity resolution strategies, as well as reduce the effects of the ionosphere, several carrier phase combinations can be derived.

In general, carrier phase combinations involve measurements on two frequencies, L1 and L2, which are combined in the following manner:

$$\phi_{ij}^{kl}(\beta_1, \beta_2) = \beta_1 \phi_{ij}^{kl}(L1) + \beta_2 \phi_{ij}^{kl}(L2) \tag{2.13}$$

The equation above, represents a combination of dual-frequency measurements for an arbitrary set of two stations (denoted by the subscripts ij) and two satellites (denoted by

the superscripts kl) which are used to form double difference measurements at a particular epoch (see eq. 2.6). A number of combinations can be formed using eq. (2.13) depending on the choice of the coefficients, β_1 and β_2, which can take on any value ($\beta_1, \beta_2 \in \Re$). The resulting combined phase $\phi_{ij}^{kl}(\beta_1, \beta_2)$ has a frequency f_{β_1, β_2} of

$$f_{\beta_1, \beta_2} = \beta_1 f_{L1} + \beta_2 f_{L2} \qquad (2.14)$$

and a wavelength $\lambda_{\beta_1, \beta_2}$ of

$$\lambda_{\beta_1, \beta_2} = \frac{c}{\dfrac{\beta_1}{\lambda_{L1}} + \dfrac{\beta_2}{\lambda_{L2}}} \qquad (2.15)$$

One of the most commonly used carrier phase combinations is derived when $\beta_1 = 1$ and $\beta_2 = -1$. The aforementioned combination is referred to as the widelane because of the increase in the wavelength to approximately 86.2 cm, facilitating improved ambiguity resolution in terms of ease and time to fix. The narrow lane combination is formed when $\beta_1 = \beta_2 = 1$. The resulting wavelength is much shorter, only 10.7 cm, however the measurement noise is significantly reduced over other common phase combinations. In addition to this, the wide and narrow lanes are linked through an 'even-odd' relationship, enforcing ambiguity resolution strategies. Simply stated, the widelane ambiguities ($N_{WL} = N_{L1} - N_{L2}$) and the narrow lane ambiguities ($N_{NL} = N_{L1} + N_{L2}$) are always of

the same even or odd nature (i.e. they satisfy $N_{WL} \bmod 2 = N_{NL} \bmod 2$). Therefore, if one of the two ambiguity types can be resolved, the possibility of resolving the second ambiguity type is effectively enhanced.

It is appropriate at this time to recognize the fact that ambiguity resolution is an integral part of proper precise real-time positioning and a significant amount of research is continuously being conducted on efficient (and real-time) ambiguity resolution. The matter is further addressed in the overview discussion on multi-reference station methods in the latter parts of this chapter and references for additional readings are provided.

The ionospheric-free (IF) phase combination can be derived, which theoretically eliminates the first order effects of the ionosphere. The remaining higher order effects (approximately 0.1%) may be on the order of a few centimetres under high ionospheric conditions, which is significant for some precise applications (Klobuchar, 1996). To form an ionospheric-free combination, the following postulate must be satisfied:

$$\beta_1 d_{ion}(f_{L1}) + \beta_2 d_{ion}(f_{L2}) = 0 \tag{2.16}$$

where one of the two coefficients is arbitrarily chosen and the second is solved for as per eq. (2.16). Therefore, it is evident that there are a number of IF combinations which can be formed depending on the arbitrary choice for coefficients. One of the most common combinations, which is also used for the work presented in the following chapters, is

formed by $\beta_1 = 1$ and $\beta_2 = -\frac{f_{L2}}{f_{L1}}$. This derivation is straightforward and can be found in Hofmann-Wellenhof et al. (1993, pp. 105-108). The final IF phase combination can therefore be written as

$$\varphi_{L1,L2} = \beta_1 \varphi_{L1} + \beta_2 \varphi_{L2} = \varphi_{L1} - \frac{f_{L2}}{f_{L1}} \varphi_{L2} \tag{2.17}$$

Two interesting artifacts can be extracted from this derivation. Even though the first order ionospheric effect is removed, the measurement noise also increases. It can be shown, through simple error propagation of eq. (2.17), that the effective noise of the IF observable increases with respect to the L1 and L2 observables as follows:

$$\sigma_{IF} = \sqrt{(\beta_1 \sigma_1)^2 + (\beta_2 \sigma_2)^2} \tag{2.18}$$

A second *artifact* of the IF combination, is that the integer nature of the unknown ambiguities is not retained. Following the form of eq. (2.17), the IF combination ambiguities are computed as:

$$N_{IF} = \beta_1 N_{L1} + \beta_2 N_{L2} \tag{2.19}$$

which are real number values.

A summary of the carrier phase combinations and the associated wavelengths presented is provided in Table 2.2.

Table 2.2: Useful Carrier Phase Combinations as per eq. (2.13)

Combination Signal	β_1	β_2	Effective Wavelength (m)
L1 phase	1	0	0.190
L2 phase	0	1	0.244
Widelane	1	-1	0.862
Narrow Lane	1	1	0.107
Ionospheric Free	1	$-f_{L2}/f_{L1}$	0.484

One final combination that has not been mentioned and can be thought of as the converse of the IF, is the ionospheric signal (IS) combination. The IS is derived to contain only the differential ionospheric effect, as its name implies. Following the form of eq. (2.13) and substituting $\beta_1 = \lambda_{L1}$ and $\beta_2 = -\lambda_{L2}$ to get,

$$\varphi_{IS} = \lambda_{L1}\varphi_{L1} - \lambda_{L2}\varphi_{L2} \tag{2.20}$$

This signal is useful for the insight that it provides into the ionospheric behaviour after double differencing is performed.

2.4   A Note on Selective Availability

It should be noted that this is an exciting period for GPS because several changes in the signal availability have recently been announced, the effects of which remain to be seen.

Specifically, on May 1, 2000 at midnight GMT, the intentional degradation of the signal, called selective availability (SA), was officially turned down to zero following a statement made by the United States government in a White House press release (ION Newsletter, 2000). This has numerous implications for future and current GPS users. However, all of the data collected for the analysis presented in this thesis was corrupt by SA at the time and therefore this recent development does not impact the results. Also, the use of DGPS is an effective manner for eliminating the effects of SA, thus all of the corrections generated and analyzed in later chapters do not include the effects of SA. Throughout this dissertation, an attempt has been made to comment on the effect of SA being turned down to zero wherever it is applicable/significant. In most cases, the effects of SA are minimal for a user with differential capability, with the exception of reducing the required bandwidth capacity and perhaps improving the robustness of the system (Divis, 2000).

2.5   Overview of Multi-Reference Station Methods

All of the techniques discussed in the previous sections were presented using a single reference station. In this section, the concept is extended to include multiple reference stations. The use of several stations in a wide area differential GPS (WADGPS) network to improve code-based (or code with carrier smoothing) positions have been employed with great success, both in post-processing and real-time modes, see Kee and Parkinson (1992), Ashkenazi et al. (1992), and Mueller (1994). A natural extension of this concept

is to use the more precise carrier phase measurements in a similar network approach. However, the transition from code-based to carrier phase based schemes is not a trivial task, mainly due to the more stringent accuracy requirements in the latter case, which expose new and interesting problems. For instance, in most code-based differential GPS approaches the limiting error source is in fact the measurement noise and multipath of the code itself (ionospheric effects come into play only over longer baseline distances). Thus, errors such as atmospheric and satellite orbit, the effects of which were somewhat masked by the lower accuracy requirements in code-based DGPS, are now brought to the forefront in carrier techniques. Also, the accurate and reliable resolution of integer ambiguities is required for high accuracy carrier phase based positions. Finally, the extension of this concept to meet competitive real-time user requirements, uncovers the many issues related to the optimal correction parameterization schemes and their dissemination/communication to potential users located within (or surrounding) the network coverage area.

Recently, the use of multiple reference stations instead of standard single baseline carrier phase approaches has been receiving a significant amount of attention from the GPS community. As in any area of scientific research, this topic is dependent on several other major (and quite involved) sub-modules, which combine to create the final solution. Initially, each area must be investigated fully on its own and then the solutions must be integrated to arrive at a final formulation of a realizable cm-level positioning implementation (in real-time). To this end, it is not possible to discuss all steps related to

multi-reference station methodologies, such as network ambiguity resolution (over long baselines and in real-time), error correction generation schemes (using a multiple reference station network adjustment), the proper parameterization of corrections (i.e. grid-based or function approximations), the subsequent dissemination and communication of these corrections to users, not to mention the user implementation issues as well as the quality issues inherent in an accurate and reliable service system (see Figure 2.4), in the limited extent of this overview section.
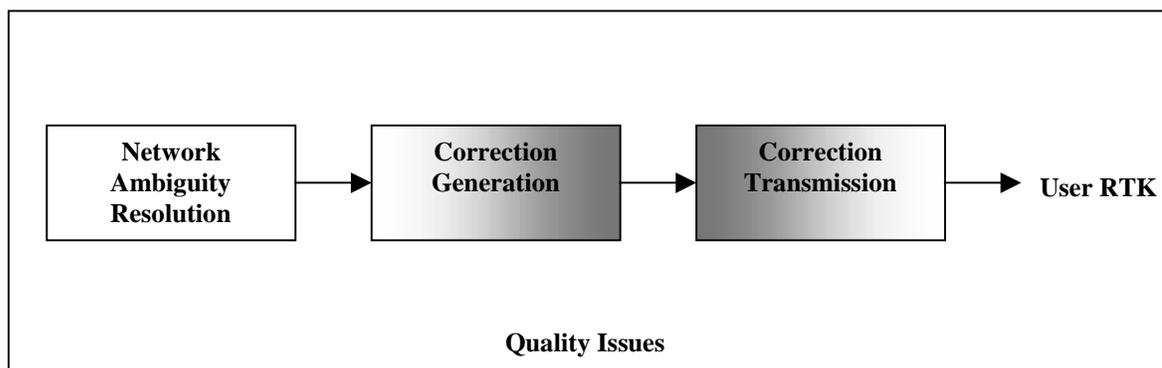
Figure 2.4: Multi-Reference Station Modules for RTK Positioning

In view of all of the aforementioned issues related to carrier phase based network positioning, the focus of this section has been narrowed down to presenting a comprehensive summary of *some* of the most common multi-reference station methods, with specific attention directed towards the correction generation and transmission processes (shaded region in Figure 2.4). More specifically, the various multi-reference station methodologies will be categorized according to their underlying correction generation framework, but will be discussed in terms of the correction dissemination

options (part of correction transmission) presented by the various authors. This approach does not take each topic (correction generation and correction transmission) on its own, rather it attempts to incorporate a comprehensive overview based on both areas.

To date, numerous methods have been developed to deal with the resolution of the carrier phase ambiguities in *near* real-time, known as on-the-fly (OTF) techniques (see, Frei and Beutler, 1990; Hatch, 1990; Chen and Lachapelle, 1995; de Jonge, 1997; Teunissen, 1998), as well as resolving ambiguities with longer inter-station baseline lengths (Blewitt, 1989). In fact, one of the key motivations behind the development of the multi-reference station methods presented in this section, is the improvement in the accuracy of the observations, which leads to improvements in ambiguity resolution (see Wanninger, 1995; Han and Rizos, 1997). The issue of ambiguity resolution will not be pursued any further here, however it should be noted that all of the multi-reference station techniques presented herein require *a-priori* knowledge of the double difference ambiguities $\nabla \Delta N_n$ between the network reference stations. These are generally determined as integers, although some work has been done when using float ambiguities between the network reference stations (see Jensen and Cannon, 2000 for more details).

The four main algorithms/methods to be summarized, are the (1) partial derivative algorithms, (2) linear interpolation algorithms, (3) conditional adjustment algorithms and (4) virtual reference station methods. Each of these methods is discussed in detail according to the general methodology and correction parameterization requirements.

Examples using a sample set of regional network data are included throughout to aid with the discussion.

2.5.1   Why Use a Multiple Reference Station Approach?

The implementation of multiple reference stations in a permanent array for performing carrier phase based DGPS offers several advantages over the standard single baseline approach, and are discussed in the following paragraphs.

One of the most important advantages for multi-reference station network users, as compared to single baseline users, is the increase in reliability and availability of the service. In a network approach, if one or two reference stations fail at the same time (for some reason, such as a power failure), their contribution can be eliminated from the solution and the remaining reference stations can 'take over' in order to provide the user with carrier phase corrections, thus retaining the availability of the service. Although, in such cases, the position accuracy may suffer slightly, it will not be as poor as in the standard single baseline DGPS approach when the only reference station fails, resulting in single point positioning results. In addition to this, the use of a network approach allows for the quality of the corrections generated from each reference station to be checked with the remaining corrections. Thus, if a particular station is generating erroneous corrections (i.e. a blunder enters into the solution), the network allows for the possible detection and elimination of this blunder from the final solution.

Another quite important aspect of the network approach is that it allows for the modelling of the distance-dependent or spatially correlated errors, such as ionospheric, tropospheric and satellite orbit effects (note that some methods may also reduce the effects of the non-spatially correlated or station-specific errors). By combining observations from a number of permanent reference stations (with known coordinates), the effect of the above mentioned distance-dependent errors can be reduced, through various parameterization/modelling techniques. Thus, corrections representing these errors can be generated and disseminated to the user in order to improve the overall positioning accuracy. A direct result of modelling the spatially correlated errors is the ability to improve the resolution of carrier phase ambiguities that are required for cm-level positioning results. In fact, the improvement in ambiguity resolution over longer baselines, leads us to another important advantage of the network approach, namely larger allowable inter-station distances between the network reference stations. This directly translates to a larger coverage area over which DGPS methods can be performed. For instance, in most cases, the inter-station baseline distances increase from a maximum of few tens of kilometres, using the standard single baseline approaches, to a few hundred kilometres, using the multi-reference station approaches.

Finally, the network approach permits the generation of observations for a *fictitious* or virtual reference station (VRS), which can be situated closer to the user station (than any of the other permanent physical reference stations) resulting in improved DGPS positioning. In other words, this approach does not require a receiver to be *physically*

located at the reference site. Also, the advent of the virtual reference station approach is more flexible in terms of permitting users to utilize their current receivers and software without involving any 'special' carrier phase software to deal with simultaneously receiving corrections from a series of reference stations.

As with any approach, there are some drawbacks associated with using a network of reference stations. More specifically, depending on the method used, there is a possible increase in the data transmission load (requirements) and complexity of user implementation over standard single baseline DGPS. Combining all of the observations from the network at a central station for processing, and then transmitting either corrections for a user or synthetic observations for a virtual reference station (or both), are the preferred methods for overcoming the data bandwidth limitations. Finally, in order for the method to be readily available to all users it is necessary that the appropriate reference stations, with very well-known coordinates, are in place and are disseminating corrections in a manner that is acceptable for users within the coverage area. To date, several regional networks are operational world wide, however we are by no means in a position to provide coverage to *all* users. This is mainly due to an additional drawback of the network approach, which is the cost of implementing and maintaining a quality service.

In general, however, the advantages provided by the permanent array approach combined with the multitude of effort that has been placed on formulating algorithms and

processing methodologies, take precedence in providing users with an efficient and reliable multi-reference station carrier phase based solution.

Several methods for formulating corrections from network station data have been developed over the past few years. The main methods investigated herein can be categorized as follows, according to the correction generation algorithms adopted and investigated by the various authors:

(a) Partial derivative algorithms (as per Wübbena et al., 1996; Varner et al., 1997),

(b) Linear interpolation algorithms (as per Wanninger, 1995, 1997; Gao et al., 1997; Han and Rizos, 1996; Odijk, et al., 2000),

(c) Condition adjustment algorithm (as per Raquet, 1998; Raquet et al., 1998), and

(d) Virtual reference station methodologies (described in Wanninger, 1995, 1997; Odijk et al., 2000; van der Marel, 1998).

The first three methods presented above, concentrate on the actual generation and subsequent parameterization of carrier phase based corrections (in some cases this strictly involves residual ionospheric effects), whereas the last method deals with a 'parameterization' in a virtual reference station framework. It is included here because of its plausibility in a real-time user environment. Each class of multiple reference station approaches is presented in more detail in the following sections.

2.5.2 Partial Derivative Algorithm

One of the first algorithms employing multiple reference stations was presented by Wübbena et al. (1996). The fundamental concept underlying this work, from an error modelling perspective, is the use of a multi-station adjustment to derive parameters or 'network coefficients' for an appropriate geometrical model, which attempts to describe the behaviour of the distance-dependent errors. The modelling of spatially correlated errors is based on a first order partial derivative function, which can be interpolated to obtain the corresponding corrections of any user receiver within the network coverage area. This method requires the compilation of data from a minimum of three reference stations, which results in an inclined plane parameterization (see Figure 2.5).
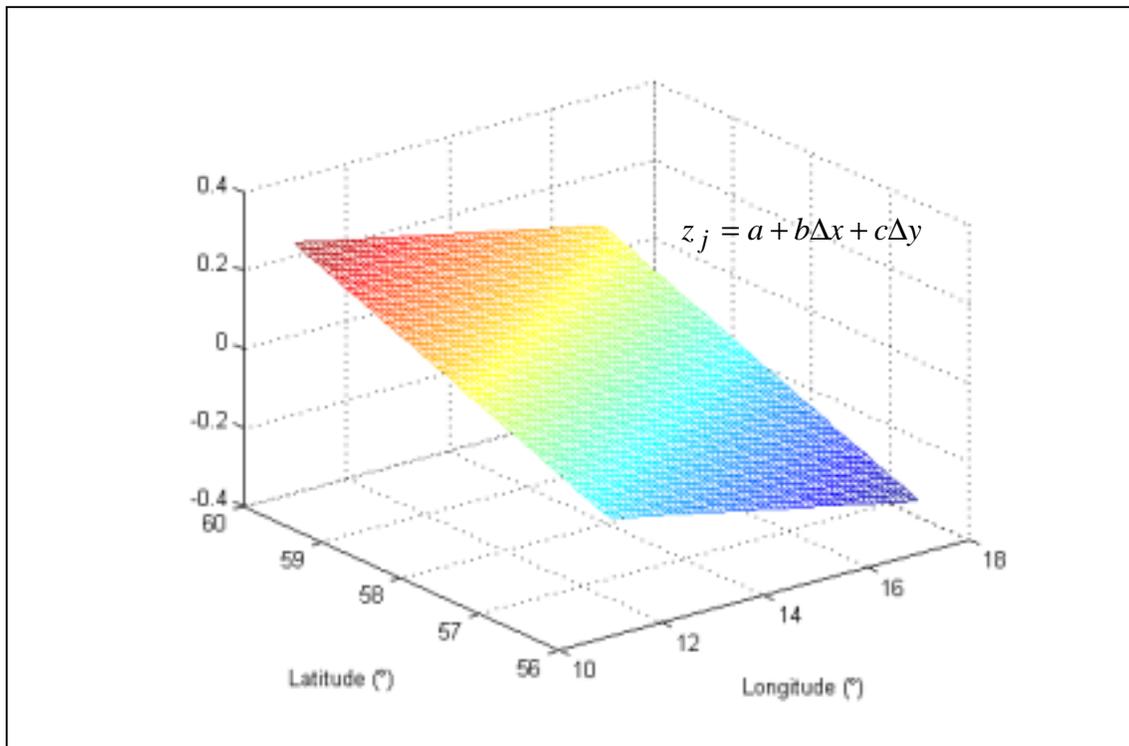
Figure 2.5: Inclined Plane Model (3 parameters) for Spatially Correlated Errors

For real-time implementation, this method offers the practicality of only having to disseminate the model parameters or network coefficients to the users, rather than the actual raw (or corrected) measurements. This implies that as the number of reference stations increases and the model becomes more complex (involving more coefficients), the increase in the transmission load is not as significant as if a new set of raw measurements for each reference station was required for transmission. In addition to this, initial results showed that the relatively low dynamic nature of the network coefficients implies that it is possible to periodically update the network coefficients, perhaps only every 30 to 60 seconds (Wübbena et al., 1996). Further tests conducted to verify the actual temporal correlation of the errors and corresponding network coefficient behaviour, show that the data transmission bandwidth requirements can certainly be reduced (see Chapter 4 for the detailed analysis and presentation of results).

An extension of the partial derivative algorithm discussed thus far, was presented by Varner and Cannon (1997), which in addition to the distance-dependent errors, investigated the use of this algorithm for station-specific errors (predominantly multipath). In this case, four different partial derivative functions were developed to estimate multipath and spatially correlated errors from network data. The unknown parameters are basically the partial derivative coefficients, which are estimated from the double difference carrier phase measurements and knowledge of the precise locations of the network reference stations. The general form of the PDA expressions can be derived

from a subset of the following expression of the GPS measurement error function, g, that is simply a truncated Taylor series expansion:

$$g(P) = \alpha + \beta \Delta x + \gamma \Delta y + \delta \Delta z + \varepsilon \Delta z^2 + \upsilon_g \qquad (2.21)$$

Each of the coefficients β, γ, δ represent the spatially correlated errors and are computed from the first order partial derivatives with respect to each of the horizontal axes x, y and the vertical axis z respectively, as follows:

$$\beta = \frac{\partial g}{\partial x}, \ \gamma = \frac{\partial g}{\partial y}, \ \delta = \frac{\partial g}{\partial z} \qquad (2.22)$$

The point about which the truncated Taylor series expansion or PDA is computed is known as the master station and all other network stations are referred to as secondary stations, used in computing the coordinate differences $(\Delta x, \Delta y, \Delta z)$ between the locations of the master (P_m) and the secondary (P_s) stations which are shown in eq. (2.21). The first coefficient α is a constant which represents the station-specific error at the master station, such that

$$\alpha = g(P_m) \qquad (2.23)$$

The fifth coefficient is computed from a second order partial derivative with respect to the vertical axis as follows:

$$\varepsilon = \frac{\partial^2 g}{2\partial z^2} \qquad (2.24)$$

which basically takes into account the non-linear effects evident in the vertical direction due to the ionosphere and troposphere. This model assumes that the non-linear effects in the horizontal directions are very insignificant (i.e. $\frac{\partial^2 g}{\partial x^2} = \frac{\partial^2 g}{\partial y^2} = \frac{\partial^2 g}{\partial x \partial y} = \frac{\partial^2 g}{\partial x \partial z}, etc. = 0$).

Finally υ_g is the model prediction error at the secondary station.

In summary, the partial derivative algorithm essentially estimates network field parameters for each satellite pair at a master station, which are then disseminated to the user receiver. In Varner (2000), the choice of the appropriate PDA algorithm is discussed, based on the spatial extent, the geometry of the network, and the number of reference stations, which define the level of complexity and accuracy of the PDA. For instance, one possible form of eq. (2.21) that would represent the inclined plane model (as defined by Wübbena et al., 1996) as well as taking into consideration the non-spatial errors, would be as follows:

$$g(P) = \alpha + \beta \Delta x + \gamma \Delta y + \upsilon_g \qquad (2.25)$$

with only three coefficients α, β, γ representing the non-spatial errors, the spatial decorrelation in the x direction, and the spatial decorrelation in the y direction, respectively. Of course, in all of these derivations it is assumed that the double difference ambiguities are correctly resolved in the network. This PDA also lends itself for real-time use because of the decreased transmission load associated with just sending the PDA coefficients instead of the raw measurements or actual correction values. However, one factor that has to be considered more closely, is the fact that each coefficient set is computed for a satellite pair. This implies that the master station and the user station must use the same base satellite, a condition that must be addressed for *practical* real-time use. Further details on the issues related to choosing an appropriate PDA model for a specified network configuration are given in Varner (2000).

2.5.3   Linear Interpolation Algorithm

A distance based linear interpolation algorithm for modelling the ionospheric delays at a user station, based on a network of reference stations, has been proposed by Gao et al. (1997), and a slightly modified version is presented in Gao and Li (1998). As in the case of the PDA, data is collected from all of the network reference stations and relayed to the master station, where ionospheric delay parameters are computed and broadcast to the user station (located somewhere within the coverage area) for interpolation, as follows:

$$\nabla\Delta\hat{I}_{user} = \sum_{j=1}^{m} \frac{s_j}{s} \nabla\Delta\hat{I}_j \qquad\qquad (2.26)$$

where the approximated residual ionospheric delay that remains after double-differencing for a specific reference station j is represented by $\nabla\Delta\hat{I}_j$ and is used to compute the corresponding value for the user station $\nabla\Delta\hat{I}_{user}$. In eq. (2.26), m is the number of reference stations minus one, and

$$s = \sum_{j=1}^{m} s_j \qquad\qquad (2.27a)$$

where,

$$s_j = \frac{1}{d_j} = \frac{1}{\sqrt{(x_j - x_u)^2 + (y_j - y_u)^2}} \qquad\qquad (2.27b)$$

This method is implemented by the user, based on the ionospheric delay parameters computed from the master station, which requires knowledge of the horizontal coordinates of the reference stations (x_j, y_j) and the estimated coordinates of the user station (x_u, y_u). The advantage of this method for real-time use over the PDA methods is that interpolations are performed on an epoch-by-epoch and satellite-by-satellite basis, which means that there is no *a-priori* requirement placed on the user station to use the same base satellite as the master station. A similar interpolator for ionospheric errors which is based on the choice of an appropriate covariance function and is linearly

dependent on the distance between reference stations, was proposed by Odijk et al. (2000).

In Wanninger (1995, 1997), a linear interpolation algorithm was presented for describing the differential effects of the distance-dependent errors (predominantly residual ionosphere). More specifically, the parameterization process involved creating a bilinear surface (or plane) defined by two parameters, one for the inclination of the plane in each of the north/south and east/west directions. A minimum of three reference stations is required for this process. In cases where more than three reference stations are available, the model parameters are computed via a simple least-squares algorithm, which also contains the benefit of mitigating the effects of multipath through averaging. The proposed implementation of this method involved the use of a virtual reference station, for facilitating a number of users within the network coverage area and is discussed in more detail in one of the following sections. Han and Rizos propose a similar linear interpolation approach for modelling the spatially correlated errors and mitigating other errors, such as multipath (Han and Rizos, 1996, 1997).

2.5.4   Conditional Adjustment Algorithm

The conditional adjustment methodology, developed by Raquet (1998), provides another alternative for computing corrections to the carrier phase measurements based on the estimated behaviour of the distance-dependent errors. Although the original derivation of

this method by Raquet was taken from optimal estimation theory, it is appropriate and advantageous to look at the problem from a classical conditional least squares approach. More specifically, the condition applied in the least squares adjustment is that the double differences of the adjusted measurements minus the calculated ranges are zero, which is valid in the absence of any errors. Essentially, the generated corrections are applied to the raw measurements from the reference and user receivers and then the double difference measurements are computed. It is important to understand that the corrections are computed and applied to the raw code and carrier phase measurements at the reference and user receivers (estimated by approximate positions).

A pre-requisite to implementing the CAA correction methodology as well as the PDA and LIA methods is the provision of accurate reference station coordinates. The responsible authority, in the case of a permanent regional reference network may provide these, or they may be obtained through a static survey of each station over long periods. Some results based on the use of the CAA under various conditions are provided in Fortes et al. (1999) and Townsend et al. (1999).

This method, although designed with real-time applications in mind, does not provide a practical parameterization scheme in order to model the corrections for several users within the network coverage area, since the corrections are computed directly for a specific user location. In order to overcome this limitation, Raquet (1997) proposed that the corrections be further parameterized according to either a grid-based or a function-

based model. These options are explored in detail in Chapter 3 and 4 of this thesis and are presented in a more compact form in Fotopoulos (2000) and Fotopoulos and Cannon (2000a).

It should be noted that the CAA was chosen for generating corrections required for the analysis in this thesis. Therefore, the methodology will be described in more detail in Section 3.1 and has only been included briefly here for the sake of completeness.

2.5.5 Virtual Reference Station Approach

To this point, all of the multiple reference station approaches have been presented in terms of the modelling techniques applied to reduce the effect of the spatially correlated errors. In most cases, it was possible to compute model parameters at a central processing facility (or master station) which communicates the corrections to the user located somewhere within the network. On a similar note, the use of a virtual reference station has been proposed by many, as a more feasible approach for relaying critical model information to network users (see, Wanninger, 1997; Odijk et al., 2000; van der Marel, 1998). This approach does not require that an actual *physical* reference station (with GPS receiver and data link) is available on site, rather, it allows for the user to access data of a non-existent *virtual* GPS reference station at any location within the network coverage area (van der Marel, 1998). Figure 2.6 shows a sample case of a virtual reference station and permanent network array.
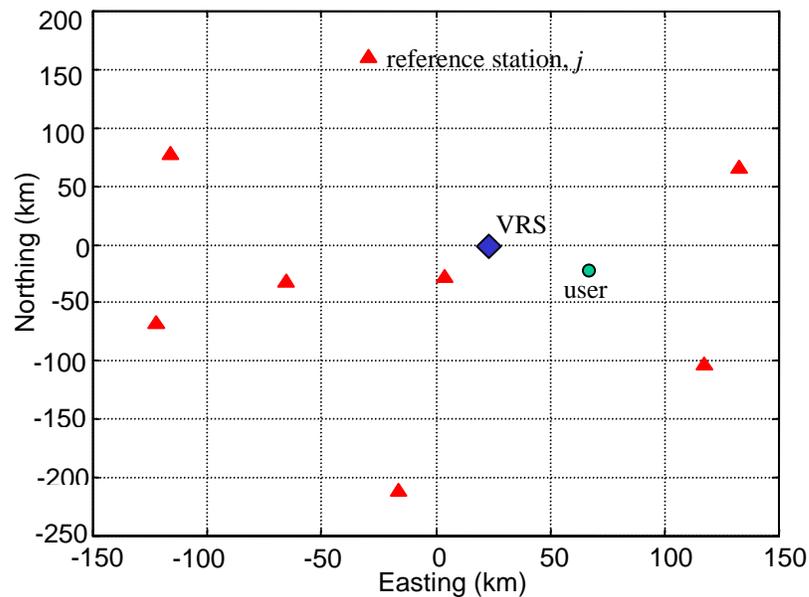
Figure 2.6: Sample of a Virtual Reference Station Network

The main objective in Wanninger (1997) was to use a network of reference stations and combine the data in such a way as to generate an *optimal* set of measurements for a virtual reference station, located approximately at the user's location, in order to determine the user's position using a standard single baseline approach. In this way, the user benefits from the reliability, availability and accuracy of a permanent network array, without having to invest in new processing software. Results using this approach have shown to be comparable to the multiple reference station approaches, where an actual physical reference station is used as the base station for processing (van der Marel, 1998; Odijk et al., 2000; Wanninger, 1995, 1997). It is also one of the most notable methods for real-time use due to the convenience and flexibility in virtual reference station locations as well as the reduced data bandwidth requirements.

2.6 Summary

In this chapter the basic foundations for differential GPS observables, residual error sources, and phase combinations were presented, followed by an overview of the multi-reference approach.   In the following chapters, results on the spatial and temporal characteristics of the corrections derived from the CAA will be analyzed. It is important to note that although the CAA is used, all correction formulation schemes, which are satellite-based and supply corrections for the raw measurements from the reference and the user stations can be employed with the modelling techniques presented.

Chapter 3

SPATIAL PARAMETERIZATIONS OF COMBINED CARRIER PHASE CORRECTIONS

The purpose of this chapter is to investigate various methods for parameterizing carrier phase corrections over a network of reference stations. The study focuses on the distribution of these corrections to potential GPS users located within, and surrounding, the network coverage area. This is an integral part of real-time kinematic DGPS and it must be adequately addressed before a practical realization of the multi-reference station concept is implemented. Specifically, numerous options for communicating these corrections to a user(s) in an efficient and accurate way are presented.

The discussion begins with a description of the correction generation methodology, followed by an introduction to the permanent operational array of GPS reference stations, located in southern Sweden, used for the analysis. The remaining parts of the chapter can be categorized into two main spatial parameterization techniques, namely (1) grid-based interpolation and (2) low-order surface modelling.

3.1   Computation of Corrections

As was shown in the previous chapter, differential GPS measurement errors can be categorized according to their spatial decorrelation. That is, the spatially correlated errors (ionospheric, tropospheric and satellite orbit) become more and more decorrelated as the receiver-to-receiver baseline distance increases. The remaining errors are said to be independent of the baseline distance (i.e. receiver or station specific) and include receiver noise and multipath. Here, combined corrections for the correlated errors were computed using an approach based on a linear minimum variance estimator, hereinafter referred to as the NetAdjust method and was briefly introduced as the CAA in Chapter 2. Although the original derivation of this method by Raquet (1998) was taken from optimal estimation theory, it is appropriate and advantageous to look at the problem from a classical conditional least squares approach. Essentially, this method imposes the condition that the double differences of the adjusted measurements minus the calculated ranges are zero, which is a valid condition in the absence of any errors. After the adjustment is performed the generated corrections are applied to the raw measurements from the reference and user receivers and then double difference measurements are computed (as in eq. 2.6).

This method can be summarized by two equations used for generating the corrections, which are applied to the carrier phase (and to the code with some minor modifications) observables as follows:

$$\hat{\delta}\ell_n = C_{\delta\ell_n} B_n^T \left(B_n C_{\delta\ell_n} B_n^T\right)^{-1} \left(B_n \ell_n - \lambda\Delta\nabla N_n\right) \tag{3.1}$$

$$\hat{\delta}\ell_u = C_{\delta\ell_u,\delta\ell_n} B_n^T \left(B_n C_{\delta\ell_n} B_n^T\right)^{-1} \left(B_n \ell_n - \lambda\Delta\nabla N_n\right) \tag{3.2}$$

where,

$\hat{\delta}\ell_n$ are the corrections generated for the carrier phase observations at the network reference stations (in metres),

$\hat{\delta}\ell_u$ are the corrections generated for the carrier phase observations at the user receiver (in metres),

ℓ_n are the measurement-minus-range carrier phase observations ($\Phi - \rho$), in metres, which assume known coordinates required for the geometrical range, ρ, computation,

B_n is the double difference Jacobian matrix for the network, $B_n = \partial\Delta\nabla\ell_n / \partial\ell_n$,

$C_{\delta\ell_n}$ is the covariance matrix of the network reference station carrier phase observations (in m^2),

$C_{\delta\ell_u,\delta\ell_n}$ is the cross-covariance matrix of the user receiver carrier phase observations (in m^2),

λ is the carrier phase wavelength (in metres), and

$\Delta\nabla N_n$ are the double difference integer ambiguities between the network reference stations (in cycles).

It should be noted that unless otherwise stated, all references to corrections correspond to the combined L1 phase corrections computed by the method given in equations (3.1) and (3.2) above (i.e. corrections include the ionospheric, tropospheric and satellite orbit effects). Although the majority of the discussion of results presented in this chapter will concentrate on the L1 phase corrections, the parameterization methodologies can be applied to a number of other carrier phase combinations (see Section 2.3).

A pre-requisite to implementing the NetAdjust correction methodology is the provision of accurate reference station coordinates. The responsible authority in the case of a permanent regional reference network may provide these, or they may be obtained through a static survey of each station over long periods. In addition, the correct double difference integer ambiguities between reference stations, $\Delta\nabla N_n$, must be known. These values are described in more detail in Section 3.2.1.

The computation of the covariance matrices, $C_{\delta\ell_n}$ and $C_{\delta\ell_r,\delta\ell_n}$, can be decomposed into two mathematical functions. First, a correlated variance function which maps the zenith variance of the correlated errors over the network area is computed and shown in the following equation:

$$\sigma^2_{c_z}(p_m, p_n) = k_1 d + k_2 d^2 \tag{3.3}$$

where, $\sigma^2_{c_z}(p_m, p_n)$, is the differential zenith variance of the correlated errors for points p_m and p_n in the network. This function is based on the two-dimensional distances between the reference stations, d, in kilometres. The importance of this dependence will be revisited in the spatial dimensionality discussion in Section 3.4.1. The values of the coefficients, k_1 and k_2, are provided in Table 3.1 for the case of the Swedish network (presented in Section 3.2). Second, a mapping function is needed to map the zenith correlated and uncorrelated errors to the elevation of the satellite at each epoch as follows:

$$\mu(\varepsilon) = \frac{1}{\sin \varepsilon} + k_\mu \left(0.53 - \frac{\varepsilon}{180°} \right)^3 \tag{3.4}$$

where, $\mu(\varepsilon)$, is a unitless scale factor, which when multiplied by the zenith variance obtained from eq. (3.3), provides the correlated variance for the specified satellite elevation, ε, and k_μ is a coefficient, also shown in Table 3.1 for the case of the Swedish network. The computation process of these coefficients and the mapping function is described in detail in Raquet (1998). The values shown in Table 3.1 and the uncorrelated zenith variances in Table 3.2, were originally derived for the Swedish network by L.P.S. Fortes at the University of Calgary and subsequently recomputed and verified by the author. Figure 3.1 illustrates the correlated error function computed for the southern Swedish network as well as the value of the mapping function for satellite elevations

ranging from 10° to 80°. Note that the correlated error variances are shown in cycles for the L1 phase and WL configurations (see Table 2.2 for the effective wavelengths). The computation of these coefficients and variance factors does not have to be performed on an epoch-by-epoch basis. In fact, the same values can be used for significantly long periods of time, depending on the network conditions (Raquet, 1998). For the results presented in the following sections, the values contained in Tables 3.1 and 3.2 were used for the entire data set.

Table 3.1: NetAdjust Coefficient Values for the Swedish Stations

Coefficients	L1 Code	L1 Phase	WL Phase
k_1	9.25737e-7 (m^2/km)	1.61782e-4 $(cycles^2/km)$	1.29975e-5 $(cycles^2/km)$
k_2	1.04204e-7 (m^2/km^2)	1.35091e-6 $(cycles^2/km^2)$	1.56561e-7 $(cycles^2/km^2)$
k_μ	10.322 (unitless)	8.307 (unitless)	12.349 (unitless)

Table 3.2: Covariance Function Elements of the Uncorrelated Errors at the Zenith for the Swedish Stations

Swedish Station	Uncorrelated Zenith Variance ($\sigma^2_{u_z}$)		
	L1 Code (m^2)	L1 Phase $(cycles^2)$	WL Phase $(cycles^2)$
KARL	6.63629e-2	4.4273e-5	3.0794e-5
VANE	3.41326e-2	4.4273e-5	3.0794e-5
BORA	4.83736e-2	4.000e-5	3.000e-5
ONSA	3.77836e-2	4.4273e-5	3.0794e-5
HASS	4.19348e-2	4.4273e-5	3.0794e-5
NORR	2.42581e-2	4.4273e-5	3.0794e-5
OSKA	4.79479e-2	4.4273e-5	3.0794e-5
JONK	4.83736e-2	4.000e-5	3.000e-5

Figure 3.1: Differential Zenith Correlated Variance Error Function (left) and Mapping Function (right) for the Southern Swedish Network in September 1998

Further details on these functions can also be found in Raquet (1998). For the purposes of the analysis that follows in this thesis, it is important to understand that the corrections are computed and applied to the raw code and carrier phase measurements at the reference and user receivers (at approximate positions).

3.1.1   Why Use the Condition Adjustment Approach?

As it was shown in Chapter 2, numerous techniques have been developed for computing carrier phase corrections based on a regional reference station network. For the spatial investigations conducted herein, the NetAdjust (or CAA) correction generation algorithm was chosen because it embodied a number of useful characteristics. Firstly, the CAA generates corrections on a per satellite basis. This is a desirable characteristic because it

does not place any onus on the user to apply the same base satellite for double differencing as in the network approach. This satellite-based correction scheme provides the user with the flexibility of using any satellite pair for double differencing, which may or may not coincide with the network approach. Also, the user and reference receivers do not have to have exactly the same set of visible satellites available, as long as there are sufficient satellites for computing a position.

From a user implementation point of view, the CAA is also beneficial because the corrections are applied to both the user and reference raw measurements and then double differencing is performed. In this way, the user can continue to use the same standard software available in differentially capable receivers, without making significant changes to accommodate any 'special' kind of observables. Finally, this scheme has been tested under various conditions on a number of different networks and shown to provide positive results. Overall positioning improvements using the NetAdjust method have been shown to be on the average of more than 25% (see Fortes et al., 1999 and Townsend et al., 1999). This indicates that the quality of corrections is high, which directly influences the quality of the results obtained after the spatial parameterizations are performed.

3.2   Description of Data

The Swedish Permanent GPS Network, commonly referred to as the SWEPOS network, is a national network of 21 permanent reference stations which span all of Sweden with an average separation of 200 km (see Figure 3.2). Table 3.3 contains a partial list of the SWEPOS reference stations and the abbreviated names by which they will be referred to throughout this thesis. The network was established in August of 1993 by the National Land Survey of Sweden and the Onsala Space Observatory. Each station is equipped with two Ashtech Z-XII dual frequency receivers and a Dorne-Margolin antenna. All GPS data is downloaded daily via telephone lines at the Network Control Center at the National Land Survey in Gävle (Hedling and Jonsson, 1996). Here the data is converted to *R*eceiver *IN*dependent *EX*change format (RINEX) and archived after performing numerous tests for data quality. The network is used for applications requiring metre to millimetre levels of accuracy. Some of the most common applications include tectonic deformation monitoring, post-glacial rebound studies and other geophysical phenomenon, surveying, mapping and research in several areas including precise navigation and positioning.

Table 3.3: SWEPOS Reference Station Abbreviations

Location	Station Name	Location	Station Name
Borås	BORA	Norrköping	NORR
Hässlehom	HASS	Onsala	ONSA
Jönköping	JONK	Oskarshamn	OSKA
Karlstad	KARL	Vänersborg	VANE
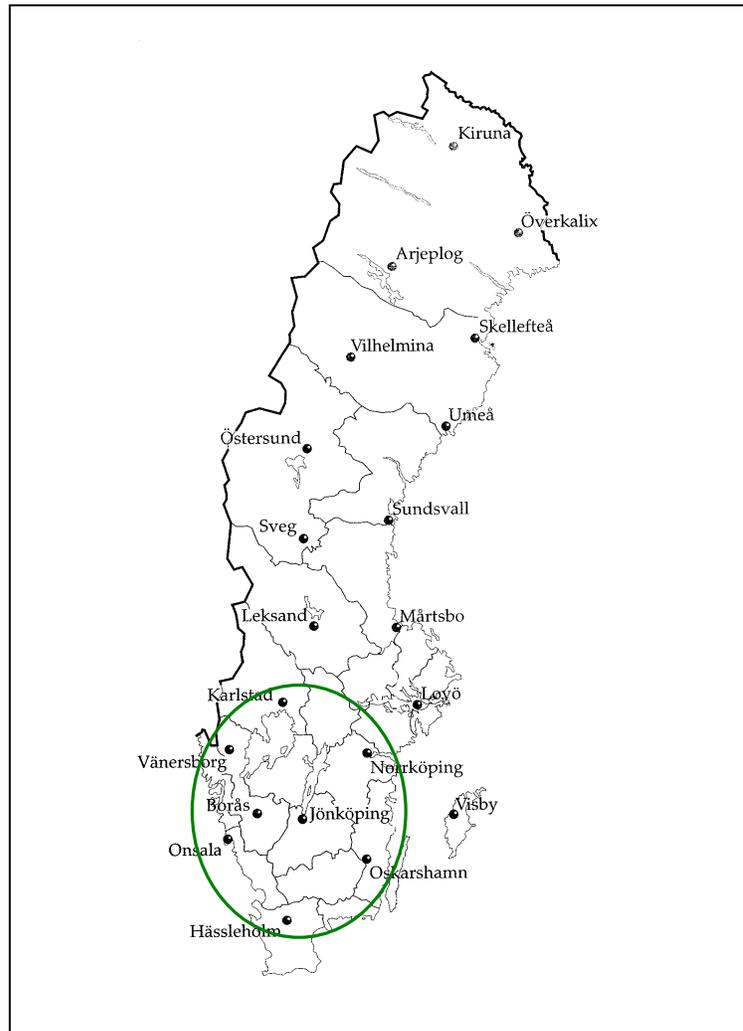
Figure 3.2: The SWEPOS Network

A 24-hour continuous data set collected on September 16, 1998 at 1 Hz is used for the results presented herein. A sub-network consisting of eight of the southern SWEPOS stations, refered to as SSN, was created, see Figure 3.3 (and the enclosed stations in Figure 3.2). Table 3.4 contains the precise position coordinates of the reference stations as obtained from the National Land Survey of Sweden (where heights include the L1 phase centre offset).

The SSN covers an area from approximately 56º N to 59.5º N latitude and 11.9º E to 16.2º E longitude, which translates to approximately a 373 km × 262 km area. The reference station ellipsoidal heights range from approximately 41 metres to 261 metres. The baseline lengths for the independent baselines between reference stations within the network are depicted in Figure 3.3 and range from about 67 km to 177 km.
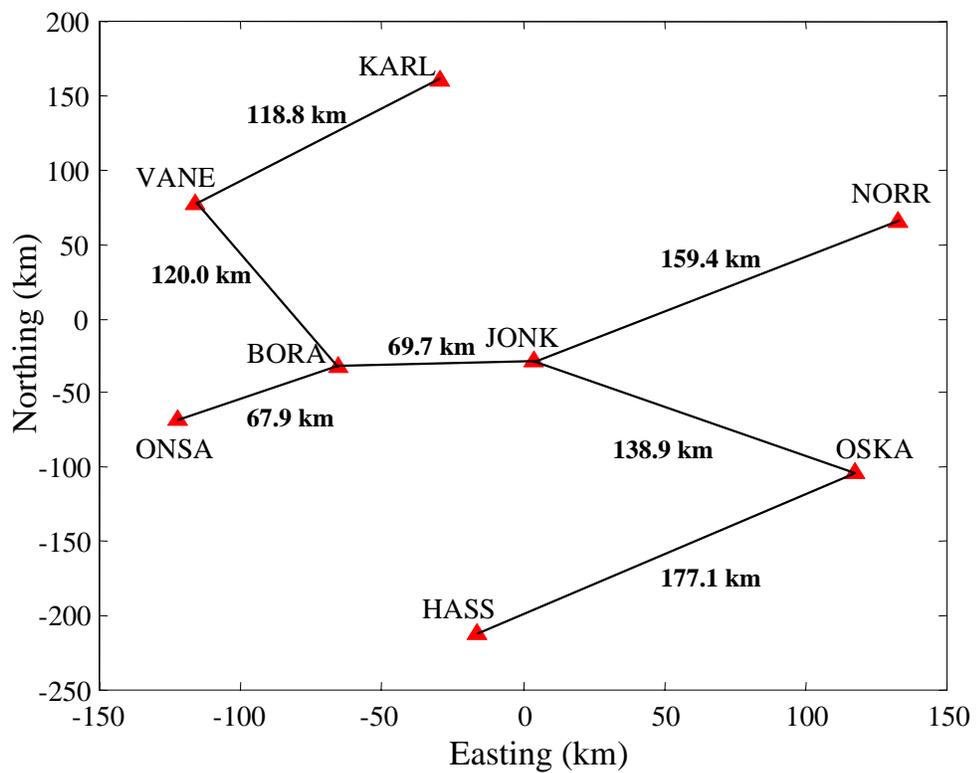
Figure 3.3: The Southern Swedish Network (SSN) and Independent Baseline Solutions

Table 3.4: SSN Reference Station Position Coordinates

Reference Station	Latitude (DMS)	Longitude (DMS)	Height (m)
ONSA	57º 23' 43".06645	11º 55' 31".85035	46.643
VANE	58º 41' 35".24963	12º 02' 06".00082	169.855
BORA	57º 42' 53".84161	12º 53' 28".84445	220.100
HASS	56º 05' 31".97408	13º 43' 05".06479	114.226
JONK	57º 44' 43".69653	14º 03' 34".58168	260.543
KARL	59º 26' 38".46721	13º 30' 20".24052	114.449
OSKA	57º 03' 56".29211	15º 59' 48".50412	149.950
NORR	58º 35' 24".82469	16º 14' 46".96527	41.088

The quality of the data was assessed by using the quality check program, TEQC (UNAVCO, 1994). In general the data was found to be of high quality with few cycle slips identified. In terms of atmospheric conditions, the maximum double difference (DD) ionsopheric delay for L1 was computed and found to be on the order of 4.6 ppm for a 69.7 km baseline located near the centre of the SSN (from Bora to Jonk). The maximum double difference misclosure value computed for the ionospheric free observable (eq. 2.17) was approximately 2.6 ppm. This latter value provides an indication into the residual tropospheric and satellite orbit effects for the data set. The double difference RMS values corresponding to the residual L1 ionospheric delay and the ionospheric free measurements for a number of baselines in the SSN are provided in Table 3.5. These values reveal moderate residual ionospheric effects ranging from 5 cm to 25 cm and tropospheric effects ranging from 2 cm to 8 cm (depending on the baseline length). Figure 3.4 provides an illustrative summary of the atmospheric levels described above, for the data set.

Table 3.5: Double Difference RMS Values for Selected Baselines

Baseline	Length (km)	DD L1 Ionospheric Delay (m)	DD Ionospheric Free (m)
Vane - Karl	118.8	0.080	0.037
Bora - Vane	120.0	0.069	0.032
Bora - Onsa	67.9	0.050	0.024
Jonk - Bora	69.7	0.055	0.026
Jonk - Norr	159.4	0.106	0.053
Jonk - Oska	138.9	0.102	0.042
Oska - Hass	177.1	0.163	0.066
Vane - Norr	244.2	0.155	0.065
Oska - Onsa	248.0	0.172	0.061
Oska - Karl	302.3	0.169	0.063
Norr - Hass	316.8	0.207	0.078
Karl - Hass	373.3	0.245	0.083
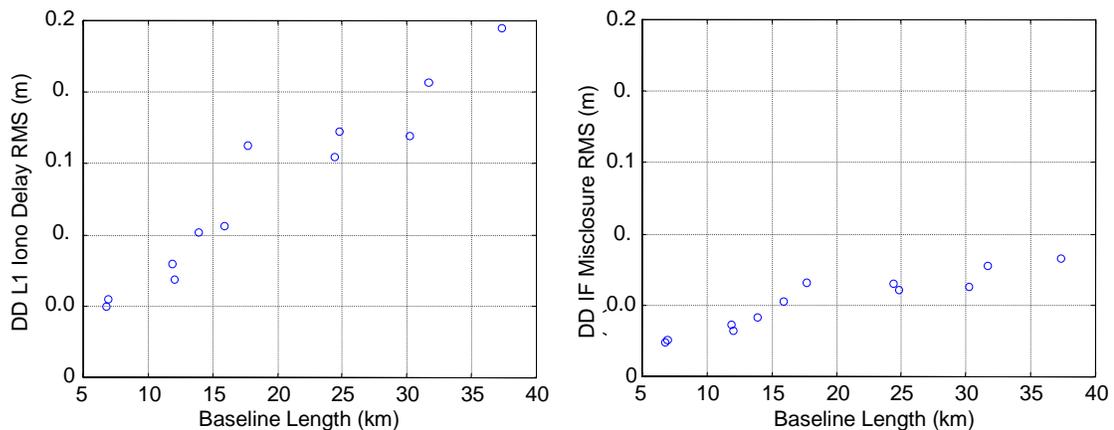
Figure 3.4: Double Difference RMS L1 Ionospheric Delay and Ionospheric Free Values for SSN Data

3.2.1 Resolution of Double Difference Ambiguities

The most difficult task in real-time kinematic (RTK) positioning over long baselines is correctly resolving the integer ambiguities. In some cases, this cannot be achieved

because of the long baselines and the corresponding spatial decorrelation of the atmospheric and satellite orbit errors. Although the average baseline length in the SSN was approximately 200 km, ambiguity resolution was possible using several methods. The majority of the L1 and L2 carrier phase integer ambiguities for the SSN were generated using GPSurvey™ (Trimble Navigation Limited, 1996). This is a commercial software package developed by Trimble Navigation Limited and is used for static and kinematic GPS data processing. The double differenced ambiguities were formed from the ambiguity output. The seven independent baselines were resolved at a 15º elevation mask and using precise ephemerides provided by the International GPS service (IGS). The well known Saastamoinen tropospheric model was also used to aid in the baseline resolution (Saastamoinen, 1972). The entire 24-hour data set, decimated to every 15 seconds, was used to process each independent baseline in the network separately (see Trimble Navigation Limited, 1996 for details on the processing methodology). It should be noted that the exact processing methodology is proprietary, however from the output generated in the log files, the major steps can be deduced. A basic flowchart outlining these major carrier phase data processing steps is provided in Appendix A.

GPSurvey™ was suitable for resolving the ambiguities for the independent baselines in the SSN, however in order to implement the multi-reference station approach, the ambiguities must be resolved for *all* of the possible baselines in the network configuration. In general, if there are *n* simultaneously operating receivers, the number of independent baselines is computed as follows:

$$bl_{ind} = (n-1) \tag{3.5a}$$

and the total number of possible baselines is computed by:

$$bl_{tot} = n(n-1)/2 \tag{3.5b}$$

For the case of the SSN, there were eight reference stations, and therefore 28 possible baselines. These baselines and the corresponding lengths are identified in Table 3.6.

Table 3.6: All Possible Baselines for the SSN

Number	FROM	TO	Baseline Length (km)
1	BORA	HASS	187.5
2	BORA	KARL	195.8
3	BORA	NORR	219.8
4	BORA	ONSA	67.8
5	BORA	OSKA	199.8
6	BORA	VANE	119.9
7	JONK	BORA	69.5
8	JONK	HASS	185.2
9	JONK	KARL	191.8
10	JONK	NORR	159.1
11	JONK	ONSA	133.2
12	JONK	OSKA	138.6
13	JONK	VANE	158.8
14	KARL	HASS	373.3
15	KARL	ONSA	246.0
16	NORR	HASS	316.8
17	NORR	KARL	183.6
18	ONSA	HASS	181.7
19	ONSA	NORR	287.6
20	ONSA	VANE	144.6
21	OSKA	HASS	176.8
22	OSKA	KARL	302.3
23	OSKA	NORR	170.4
24	OSKA	ONSA	248.0
25	VANE	HASS	306.6
26	VANE	KARL	118.6
27	VANE	NORR	244.2
28	VANE	OSKA	296.3

The ambiguities for the remaining baselines were created using the reference station chaining constraint method, which applies the well-known concept that the sum of the integer ambiguities in a closed loop is equal to zero. To illustrate this concept, consider a small network consisting of four points, as shown in Figure 3.5.
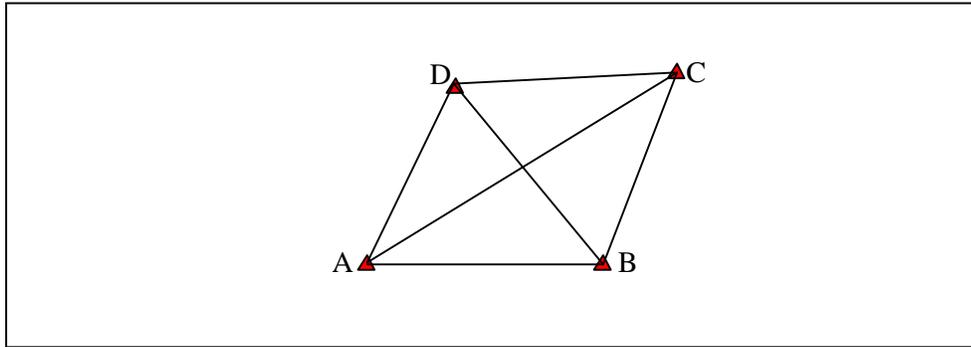
Figure 3.5: Sample Network for Closed Loop Ambiguity Constraint Concept

Using this network, five different ambiguity constraint equations can be written, one for each of the closed loops (four triangles and one quadrilateral). For triangle *BCD*, the following double difference integer ambiguity constraint equation applies:

$$\nabla\Delta N_{BC} + \nabla\Delta N_{CD} + \nabla\Delta N_{DB} = 0 \qquad (3.6)$$

where, $\nabla\Delta$ is the double difference operator and N is the integer ambiguity. The most trivial application of eq. (3.6) is when all but one of the baselines are resolved and the equation is solved for the missing baseline. In eq. (3.7), the constraint is used in order to solve for the ambiguity for the *BC* baseline.

$$\nabla\Delta N_{BC} = -(\nabla\Delta N_{CD_{resolved}} + \nabla\Delta N_{DB_{resolved}}) \qquad (3.7)$$

In the case of the SSN, the *resolved* ambiguities for all independent baselines were obtained from GPSurvey™ and this method was only used to generate the remaining (dependent) baselines. Using the constraint equation to directly calculate the third

ambiguity is a straightforward process and can easily be performed for any number of baselines. This method can also be used before the search process begins in order to reduce the search space of ambiguity candidates (Sun et al., 1999). Finally, reference station chaining can be used to verify the resolved ambiguities in a network. If all of the ambiguities satisfy the constraint equations then it is very likely that the resolved ambiguities are correct. However, there is also a possibility that all of the ambiguities are affected by the same systematic error and therefore incorrect by the same bias amount. Therefore, it is important to augment the reference station chaining approach with other ambiguity verification methods (i.e. sum of squared residuals check).

For the real-time implementation of the multi-reference station technique, the ambiguity resolution process must occur on-the-fly (OTF) and not in post-mission, as conducted in this case. However, since the focus of this research is on the spatial and temporal characteristics of the corrections, it was vital to work with correctly resolved integer values. It is recognized, however, that accurate and reliable OTF ambiguity resolution is a topic of extensive and ongoing research with tremendous potential for real-time applications and it is beyond the scope of this research.

3.3   Spatial Parameterizations

In essence, all of the spatial parameterization schemes investigated can be classified under the general area of interpolation. Interpolation can be defined as the approximation

of a representative function from its discrete values at a finite number of points (Dermanis, 1988). In this case a number of discrete points are provided which form a regularly spaced rectangular grid (which conform to some unknown function) of correction values and we are asked to compute the corresponding correction value at some arbitrary point, which most likely does not coincide with a grid node. In this situation, there exist two main cases of interpolation, namely *exact* and *approximate* methods (also called weighted average and fitted function method, numerical and mathematical surfaces, and value and coefficient problems).

The first case, termed exact interpolation, assumes the data values to be errorless quantities (or their error is accepted and permitted to propagate into the solution) and thus the resulting surface passes *exactly* through all of the data points. A simple one-dimensional situation that conveys this is shown in Figure 3.6 (analogous in every way to multi-dimensional problems, i.e. 2D or 3D). In such a case a function (1D) or a surface (two or more dimensions) is generated which exemplifies the data and each data point is reproducible (Watson, 1992). The second case is a smoother approximation, which does not readily reproduce the data values and tends to smooth the high frequency content of the signal, resulting in a function or surface which models (not necessarily exactly) the data. This is also shown in Figure 3.6.

Both cases of interpolation discussed will form the basis of the various spatial models presented. The exact interpolation method is employed in the grid-based schemes

(Section 3.4), while the approximate interpolation method is presented in the context of low-order surface modelling (Section 3.5).
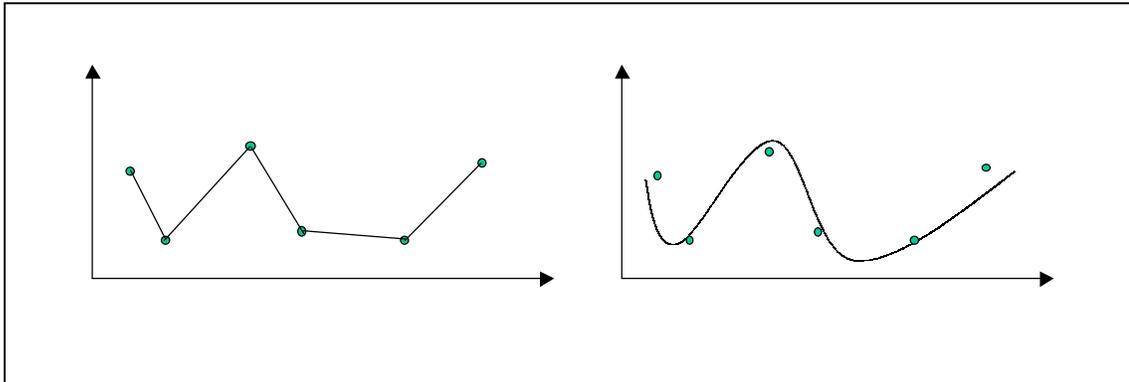
Figure 3.6: One-Dimensional Exact (Left) and Approximate (Right) Interpolation

Regardless of the parameterization method chosen, the overall objective remains, namely the fabrication of a surface that represents the spatial variation of the data, thus providing a solid foundation of the correction behaviour field over the area of interest.

3.4   Grid-Based Parameterization

The first parameterization scheme presented is a grid-based model. This concept has been implemented in varying degrees for code-based wide area differential GPS services (WADGPS) as well as more precise applications such as wide area ionospheric modelling (as in the wide area augmentation system, WAAS), see El-Arini et al. (1995) for more details. The key task is to apply this concept to formulate combined corrections at discrete locations over the regional network and then interpolate between grid points for any user within the coverage area.

Corrections were generated using the NetAdjust algorithm for the 24-hour data set at 15 second intervals, for various grid spacings over the SSN. The horizontal grid point locations were determined from the limits of the reference station network, which were 56°N to 60°N in latitude and 11°E to 17°E in longitude. However, the height of the grid points (prediction points) required some further investigation.

3.4.1   Spatial Dimensionality of the Problem

Before any work can be done on correction modelling, the spatial dimensionality of the problem must be addressed. A ground-based GPS network will always vary in latitude and longitude and depending on the locations of the permanent reference stations, the inter-station height differences may vary from metres to kilometres. The dimensionality of the problem affects the number of parameters required to model the correction surface accurately, therefore it must be determined for each data set. In this case, both the spatial dimensionality of the correction generation method and the network ground station geometry were evaluated.

The correlated error term used in generating the combined correction values was investigated and found to be a function based on two-dimensional distances between reference stations (see eq. 3.3), thus showing that it does not take the height of the reference stations into account. The second issue to examine for dimensionality is the network geometry. Given the possibility that significant height differences between

reference stations may imply residual tropospheric effects, several investigations were conducted to specifically target the sensitivity of the SSN to the vertical component.

In order to test if the model corrections generated for a grid of prediction points were sensitive to the vertical component, several tests were performed by generating combined corrections for three different surface lattice heights, namely, (a) 40 metres, (b) 160 metres and (c) 260 metres. Recall that for the case of the SSN, the reference station heights range from approximately 41 metres to 260 metres, with a mean height of about 140 metres. Corrections were generated for a 55 km $\times$ 55 km grid, over the coverage area for these three heights and each of the values computed were compared with the corresponding values at other surface lattice heights. Figure 3.7 shows a sample of the surface lattices computed at the various heights.
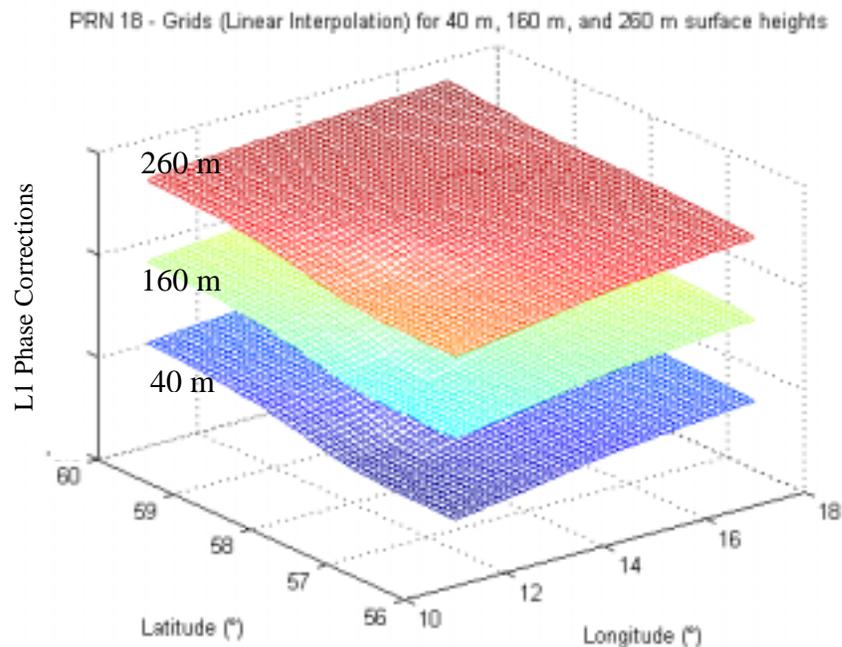
Figure 3.7: Surface Lattices at Various Heights

The comparison results showed that for approximately 97% of the epochs, there were no differences between the correction surfaces of varying heights. The maximum effect seen for a few satellites scattered throughout the afternoon hours was on the order of 0.001 cycles on L1, which translates to less than a millimetre and can therefore be neglected. Figure 3.8 is an example of a case where these minimal differences were noted, represented by the downward spikes (note the scale on the vertical axis).
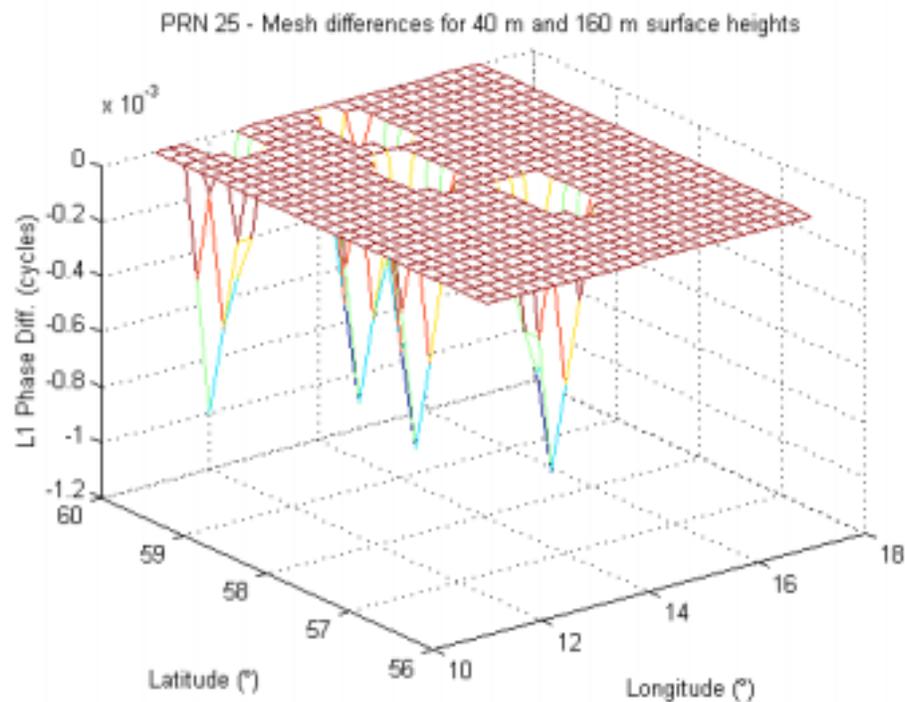
Figure 3.8: Sample of Differences in Corrections Generated at h = 40 m and h = 160 m

Given the results of these investigations, all prediction point grids presented herein will be computed at the mean height of the SSN (140 m). Other examples of networks with greater height differences also support this conclusion, such as in Schaer et al. (1999) where numerous investigations on solution types for modelling zenith path delay over a

network were performed.   The results showed small differences between two-dimensional and three-dimensional functions even for a regional network with more than 3,000 metres of inter-station height differences.

3.4.2   Interpolation Schemes

Three main interpolation schemes were investigated for possible use in the grid-based (exact interpolation) method namely, (1) nearest-neighbour interpolation (NNI), (2) bilinear interpolation (BI), and (3) bicubic interpolation (CI).   The level of complexity involved in the implementation of each method varies from nearest-neighbour being the simplest to bicubic interpolation being the most complex.   A brief description of each of the aforementioned interpolation schemes will be discussed.   For detailed mathematical formulations and discussions of the three interpolation schemes, see Watson (1992), Press et al. (1988), and Landcaster and Salkauskas (1986).

3.4.2.1 Nearest-Neighbour Interpolation

The simplest and crudest form of interpolation of the discrete rectangular grid correction values c_i incorporates the nearest-neighbour interpolation algorithm.   In the present case, the two-dimensional (could be 3D if the dimensionality of the problem called for it, resulting in a surface lattice at variable heights) distance from the closest grid nodes in a specified radius of interest are computed from the computational point z_{cp}.   The

computational point is simply assigned the correction value corresponding to the nearest point as follows:

$$z_{cp}^n = c(x_i, y_i)$$
(3.8)

where, $c(x_i, y_i)$ is the correction value of the closest point. The simplicity of this algorithm results in a discontinuous representation of the correction surface by a set of level surfaces of different values (i.e. a 'step-like' appearance). However, the algorithm offers several advantages in terms of data transmission requirements since, in theory, only one point is required for transmission. Therefore, it has not been dismissed as a viable method and is included here for the purpose of inter-comparing with the following, more complex and hence more accurate methods.

3.4.2.2 Bilinear Interpolation

A natural progression from the simple NNI approach is a distance-based bilinear interpolation scheme, which can be decomposed into three steps:

1) Compute the two-dimensional (three-dimensional distances can be computed if the correction surface is not located at a constant height throughout the

coverage area) distances between computational point (x_c, y_c) and each of the

grid points, $d_i = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2}$,

2)   Find the four closest points to the computational point(s), and

3)   Interpolate with the four closest points in order to obtain the value(s) at the

point(s) of interest $z(x_c, y_c)$, using the following formulas:

$$z(x_c, y_c) = (1 - t)(1 - u)z_1 + t(1 - u)z_2 + tuz_3 + (1 - t)uz_4 \qquad (3.8a)$$

where,

$$t = (x_c - x_1)/(x_2 - x_1) \qquad (3.8b)$$

and

$$u = (y_c - y_2)/(y_3 - y_2) . \qquad (3.8c)$$

For convenience, the grid points are numbered starting at the southwest corner and continuing in a counter-clockwise direction, as shown in Figure 3.9.

A grid of corrections at varying densities can be computed using NetAdjust within the desired coverage area. However, in order to test the results in the position domain, there must be raw observations available for the mobile user. In practice this will not always be the case, i.e. a receiver will not be stationed at every grid node. Therefore, a bilinear interpolation algorithm that generates corrections from the closest four points is used for obtaining correction values for the observations at the mobile user's location.
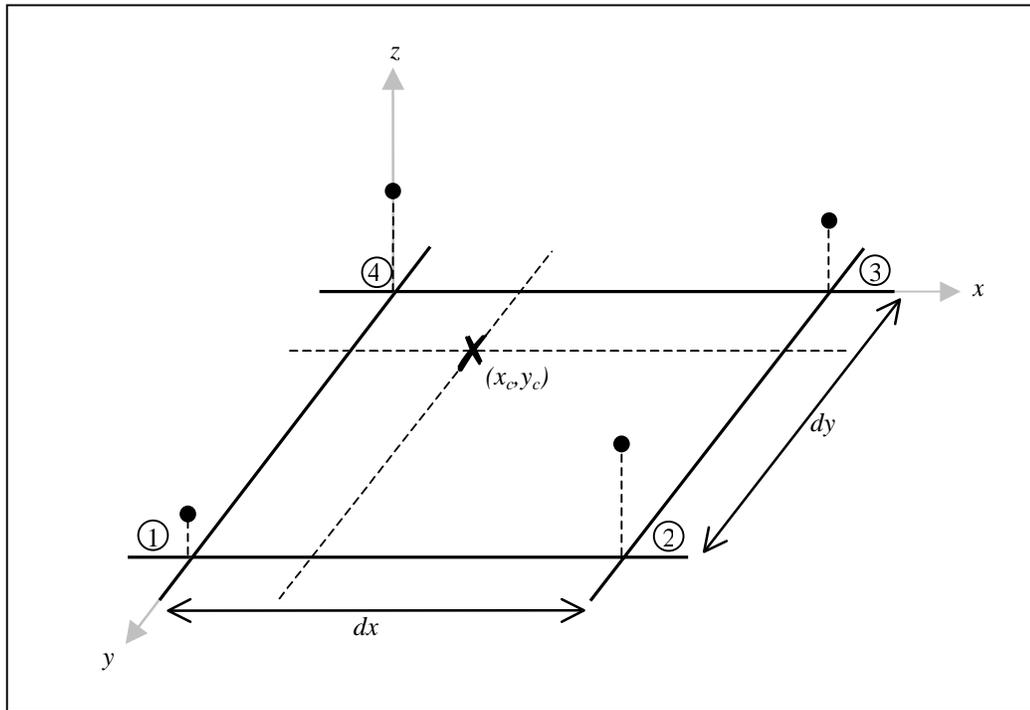
Figure 3.9: General Bilinear Interpolation Point Numbering Scheme

3.4.2.3 Bicubic Interpolation

Bicubic interpolation is the lowest order 2D interpolation procedure that maintains the continuity of the function and its first derivatives (both normal and tangential) across all boundaries (Russell, 1995). Unlike the previous two methods, it requires not only the values for the closest grid nodes c_{ij} (four points in total), but also three additional derivatives at each point, namely $\dfrac{\partial z}{\partial x}$ which provides the slope in the x-direction, $\dfrac{\partial z}{\partial y}$ which provides the slope in the y-direction and $\dfrac{\partial^2 z}{\partial xy}$ which is the cross-derivative. This

translates to a total of 16 values required for transmission. More details on the algorithm are provided in Russell (1995) and Press et al. (1988).

3.4.2.4 Choosing an Interpolation Scheme

The interpolation algorithms investigated represent a limited sub-set of a plethora of methodologies available. For instance, cubic spline interpolation is a commonly used algorithm for constructing a smooth surface from a number of data values. The goal in this case, is not to produce a surface which is pleasing in appearance (as in many image analysis applications), rather it is important to recover as much information as possible from the 'true' correction field, without over-sampling. This task is successfully accomplished using any one of the three interpolation schemes discussed above, depending on the grid resolution. One advantage, however, that other gridding methods such as kriging, may provide is an indication into the quality or accuracy of the interpolation algorithm itself (Stein, 1999). For the grid-based investigations, this can be accomplished through error propagation, provided covariance information for the NetAdjust corrections is available. This offers us an indication of the quality of the corrections *after* applying the spatial parameterization schemes (see Appendix B).

For RTK applications, the choice of an interpolation scheme is important because it is directly related to the number of correction values that are required for transmission (usually via a data link) to the user receivers. Data compression is important for the

efficient dissemination and communication of spatial data. Table 3.7 shows the minimum number of grid points required to perform each of the interpolation schemes as well as the number of bits to be transmitted for the NetAdjust corrections for a nominal case of nine satellites. The bit calculation was performed for L1 phase correction values at the specified number of grid points only, and may be further reduced using various compression/scaling techniques. A detailed analysis of the data transmission issues is included in Chapter 4.

Table 3.7: Grid Points Required for Interpolation Schemes

Interpolation Scheme	# of Grid Points	# of Bits (9 satellites)
Nearest-Neighbour	1	81
Bilinear Interpolation	4	405
Cubic Interpolation	16	1377

From Table 3.7 it is evident that as the number of grid points used for the interpolation scheme increases, so does the required number of bits. From a data transmission point of view, whenever more than one grid point is transmitted, the corresponding grid boundaries (north-south and east-west) and definition must also be specified.

A sample surface generated using each of the three interpolation methods is shown in Figure 3.10. It is evident that there are some visible differences between the correction surfaces of each type, with nearest-neighbour being the 'coarsest', and the bilinear and cubic being the 'smoothest'. The key to choosing an interpolation scheme is to balance the required amount of data and the achievable accuracy, based on the information

content. Also, because these distance-based interpolation schemes gather information from a specified number of surrounding grid points, it is important to maximize the contribution of the closest grid point values and minimize those that are farthest away. For instance, in using the cubic interpolation method, 16 values are employed. Depending on the grid spacing, the farthest corrections may introduce effects that are not indicative of the present computational point.

Figure 3.11 shows the differences between the interpolated surfaces (ignoring the vertical axis as the grid surfaces have been superimposed for visual analysis). From here it is noted that the differences between the linear and cubic interpolation schemes is almost negligible, on the order of less than a millimetre, while the greatest differences are noted between the nearest-neighbour and bilinear/cubic interpolations.
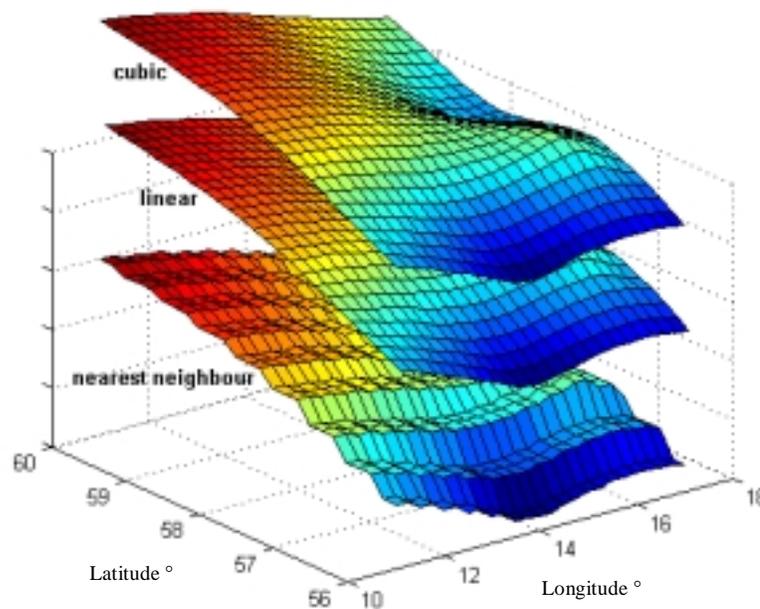
Figure 3.10: Three Interpolation Schemes used for L1 Phase Corrections for PRN 4
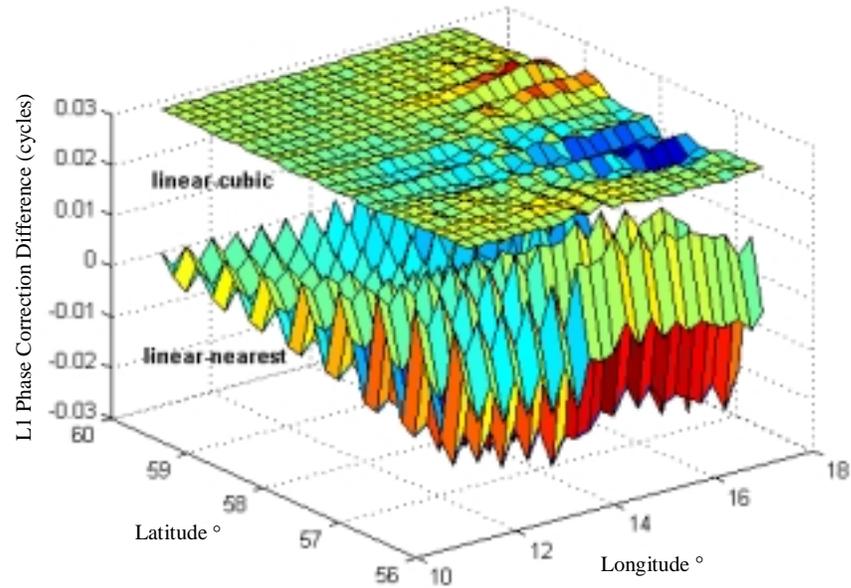
Figure 3.11: Differences in Correction Surfaces for PRN 4 at Noon Local Time

As a result of these investigations, bilinear interpolation was chosen to generate values for the computation points, since, it is less complex, requires fewer bits to be transmitted, and performs at approximately the same level as cubic interpolation. In Appendix B, the derivation of error variances for the parameterized corrections is shown. Given a covariance matrix of the computed corrections, the covariance information associated with the 'parameterized' values can be computed as shown.

3.4.3 Grid Resolution

The spacing between the grid points is directly related to the accuracy of the interpolated user correction. For real-time data transmission requirements it is desirable to use the highest resolution grid possible, without exceeding the data transmission bandwidth requirements (i.e. 2400 bits per second for 2400 baud rate, see Talbot, 1996 for more details). With this in mind, four different grid spacings were tested using bilinear interpolation to obtain the correction values for a user receiver, namely, (a) $0.5° \times 0.5°$, (b) $1° \times 1°$, (c) $1.5° \times 1.5°$, and (d) $2° \times 2°$, which corresponds to 117, 35, 24, and 12 prediction points respectively for the SSN. The user receiver was chosen to be near the centre of each grid (Jonk), and is shown along with the four grid spacings in Figure 3.12.
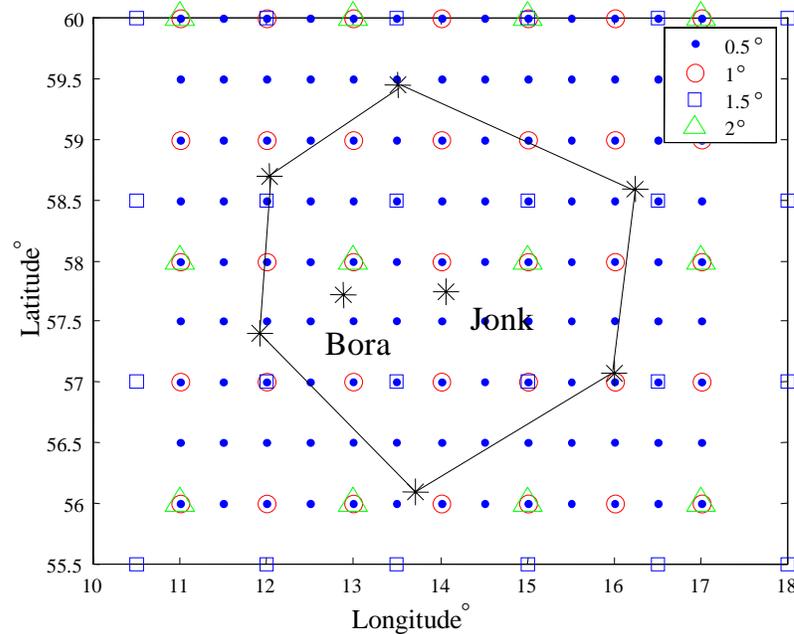
Figure 3.12: Overlay of the Southern Swedish Network and the Four Different Grid Spacings

Table 3.8 shows the statistics compiled over 24-hours for the differences between the combined corrections obtained from the various grid spacings for a common computation point, Jonk. All of the visible satellites were common among each grid. Differences are shown for each correction type, where code refers to code corrections on L1, WL is the widelane observable, IF is the ionospheric-free and IS is the differential ionospheric delay between L1 and L2 (see Section 2.3). The smallest discrepancies are between the two densest grids (0.5° and 1°), which average differences of less than a millimetre for L1 phase corrections and are depicted over the entire 24-hour period in Figure 3.13. As expected, the largest discrepancy occurs between the densest and the sparsest grids (0.5° and 2°), averaging just over three millimetres on L1 phase, which are also shown over the entire period in Figure 3.14. The significance of these differences in the correction domain will become apparent in the following discussion of the position domain results.
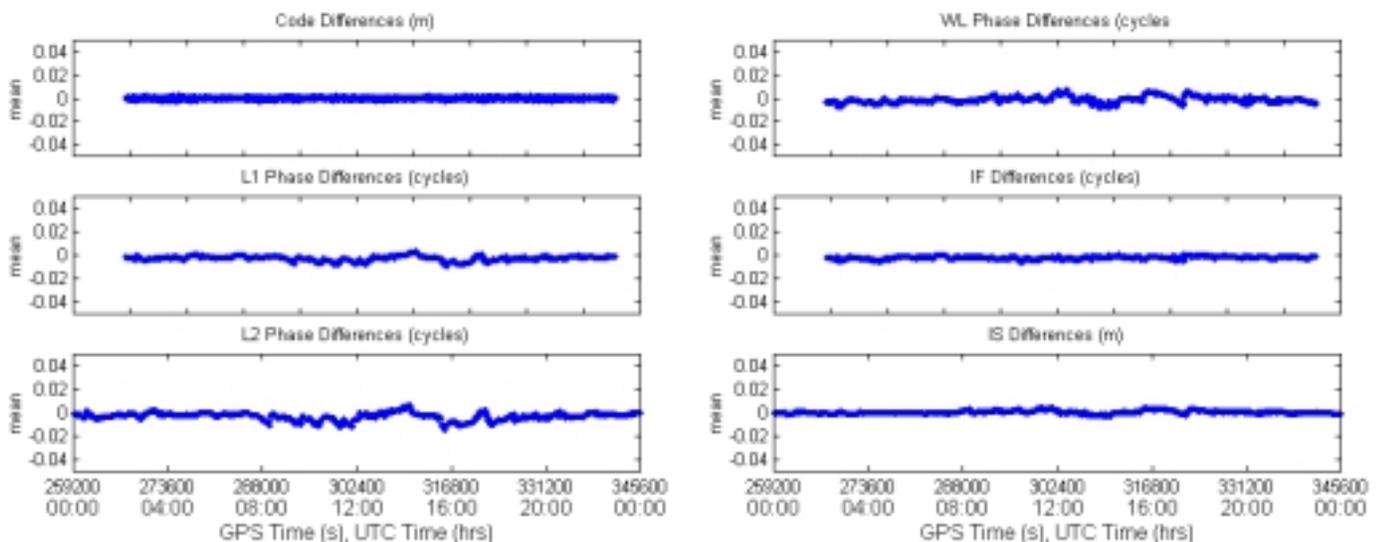
Figure 3.13: Mean Differences Between 0.5° and 1° Derived Grid Correction Values for Various Observable Combinations
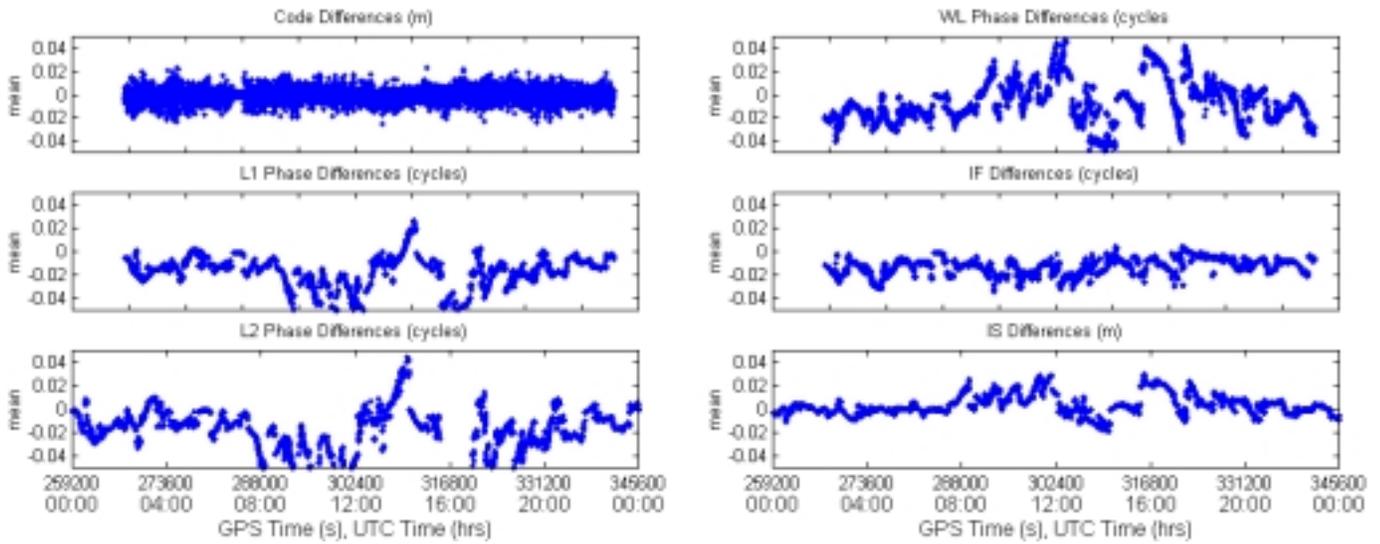
Figure 3.14: Mean Differences Between 0.5° and 2° Derived Grid Correction Values for Various Observable Combinations

From Figures 3.13 and 3.14, it is evident that the sparser grid resolutions have difficulty in maintaining the correct values for areas with higher ionospheric activity (i.e. in the early afternoon hours). This indicates the requirement for higher spatial sampling intervals in order to correctly recover the atmospheric effects over the network. Thus, as the grid resolution decreases, some error activity is not accounted for, as expected.

Table 3.8: Statistics for the Differences Between Corrections for Jonk (over 24 hours)

Statistic	Code (m)	L1 (cycles)	L2 (cycles)	WL (cycles)	IF (cycles)	IS (m)
0.5° × 0.5° Grid minus 1° × 1° Grid						
Minimum	-0.019	-0.048	-0.073	-0.044	-0.018	-0.017
Maximum	0.000	0.019	0.034	0.048	0.009	0.025
Mean	0.000	-0.003	-0.003	-0.001	-0.002	0.000
0.5°× 0.5° Grid minus 1.5° × 1.5° Grid						
Minimum	-0.004	-0.054	-0.090	-0.067	-0.032	-0.028
Maximum	0.000	0.032	0.059	0.074	0.013	0.036
Mean	0.000	-0.003	-0.003	-0.001	-0.002	0.000
0.5°× 0.5° Grid minus 2° × 2° Grid						
Minimum	-0.149	-0.331	-0.505	-0.296	-0.146	-0.114
Maximum	-0.001	0.166	0.256	0.315	0.050	0.182
Mean	-0.001	-0.017	-0.021	-0.006	-0.013	0.004
1° × 1° Grid minus 1.5° × 1.5° Grid						
Minimum	-0.018	-0.038	-0.059	-0.051	-0.032	-0.028
Maximum	-0.000	0.033	0.055	0.047	0.013	0.036
Mean	0.000	0.000	0.000	0.000	-0.002	0.000
1° × 1° Grid minus 2° × 2° Grid						
Minimum	-0.126	-0.148	-0.229	-0.267	-0.042	-0.157
Maximum	-0.001	0.284	0.433	0.263	0.127	0.097
Mean	0.001	0.015	0.018	0.005	0.011	-0.003

More intuitive results are presented in the position domain where the corrections are applied to the raw carrier phase measurements and then processed to compute epoch-by-epoch positions using Flykin™ (Lu et al., 1994), which is an on-the-fly ambiguity resolution software package developed at the University of Calgary. The software was modified to accept an input file of fixed integer ambiguities as opposed to allowing the software to compute them on-the-fly. This was done in order to maintain a consistent set of data between the correction generation and position computation processes. The position results are based on an L1-only solution for one of the shortest baselines in the network, 69 km, between the reference station, Bora, and the user, Jonk (see Figure 3.12).

Although Jonk was stationary, it was processed in 'kinematic' mode on an epoch-by-epoch basis, which simulated real-time conditions.

Table 3.9 shows the statistics for the position differences computed for the 69 km baseline. The data in the *no corrections* section corresponds to the case where raw measurements are used without applying any network corrections (i.e. a single baseline approach). The following section in the table corresponds to the case where the exact user station coordinates are known and network corrections can be directly computed via NetAdjust, without applying any interpolation scheme. This case essentially provides the improvement obtained when using the NetAdjust method. The last four sections correspond to each of the four grid-based cases where the corrections for the user are generated through bilinear interpolation of a correction grid at the specified spacing (eq. 3.8).

The results show that by using the various grid-based parameterizations, as much as 20%, 54%, and 27% improvement in RMS for latitude, longitude and height respectively, over the single baseline approach, can be achieved. In general an improvement is seen for all cases compared to not applying any corrections, except the very sparse $2° \times 2°$ grid, where an actual degradation in position is experienced. This shows that for the SSN conditions and the corresponding 24-hour data set, the grid-based method is valid up to $1.5°$ in each horizontal component. Beyond this grid spacing, there is no benefit to

applying the corrections. The percentage of improvement in RMS over no corrections for

each of the grids tested, are summarized in detail in Table 3.10.

Table 3.9: Statistics for Position Results using Single Baseline and Grid-Based Methods (Bora$_r$-Jonk$_u$)

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
No Corrections					
Latitude	-0.134	0.200	0.011	0.037	0.039
Longitude	-0.192	0.104	-0.030	0.050	0.058
Height	-0.370	0.341	-0.025	0.103	0.106
Directly Computed Corrections					
Latitude	-0.136	0.152	-0.001	0.031	0.031
Longitude	-0.091	0.100	0.002	0.026	0.026
Height	-0.293	0.343	-0.015	0.075	0.076
0.5°×0.5° Interpolated Grid					
Latitude	-0.138	0.165	-0.002	0.031	0.031
Longitude	-0.091	0.104	0.002	0.027	0.027
Height	-0.291	0.380	-0.013	0.076	0.077
1°×1° Interpolated Grid					
Latitude	-0.158	0.248	-0.008	0.034	0.035
Longitude	-0.091	0.117	0.004	0.029	0.029
Height	-0.283	0.612	0.001	0.080	0.080
1.5°×1.5° Interpolated Grid					
Latitude	-0.152	0.262	-0.003	0.036	0.036
Longitude	-0.091	0.131	0.006	0.031	0.032
Height	-0.288	0.636	0.003	0.086	0.086
2°×2° Interpolated Grid					
Latitude	-0.287	0.680	-0.053	0.066	0.084
Longitude	-0.120	0.190	0.017	0.047	0.050
Height	-0.317	1.860	0.075	0.128	0.148

Table 3.10: Percentage of RMS Improvement from Single Baseline Approach

Method	Latitude	Longitude	Height
Direct	21%	54%	28%
0.5° × 0.5°	20%	54%	27%
1° × 1°	10%	50%	24%
1.5° × 1.5°	7%	46%	18%
2° × 2°	-118%	14%	-15%

These results are quite significant, because they identify the limitations of the grid-based parameterizations. For RTK, if the requirements are such that only sparse grid data can be sent to the user, such as the 2° case, the user may achieve better results by opting to ignore the corrections and instead perform positioning based on raw data from the closest reference station.

Further investigations were conducted by assuming an even denser grid (i.e. resolutions greater than 0.5°). The results showed that spacings of approximately 44 km (or 0.4°) produced essentially the same information as the 0.5° grid. One should be cautious not to over sample the area which may lead to irregular oscillations in grid node values generating artifacts which are not indicative of the correction behaviour over the area. Therefore, the 0.5° grid was found to be the best alternative in terms of representing the spatial behaviour of the corrections. Figures showing the deviations in the L1 phase correction grid node values for the 0.5° grid as a function of baseline distance were generated for two different case satellites, representing the maximum decorrelation over distance (PRN 23) and a more moderate scenario (PRN 21) which are shown in Figure 3.15. The correction differences are computed for baseline lengths up to approximately 600 km, where the maximum difference is 0.33 cycles for PRN 23. More moderate values are obtained for PRN 21 where the average correction differences are approximately 0.08 cycles.
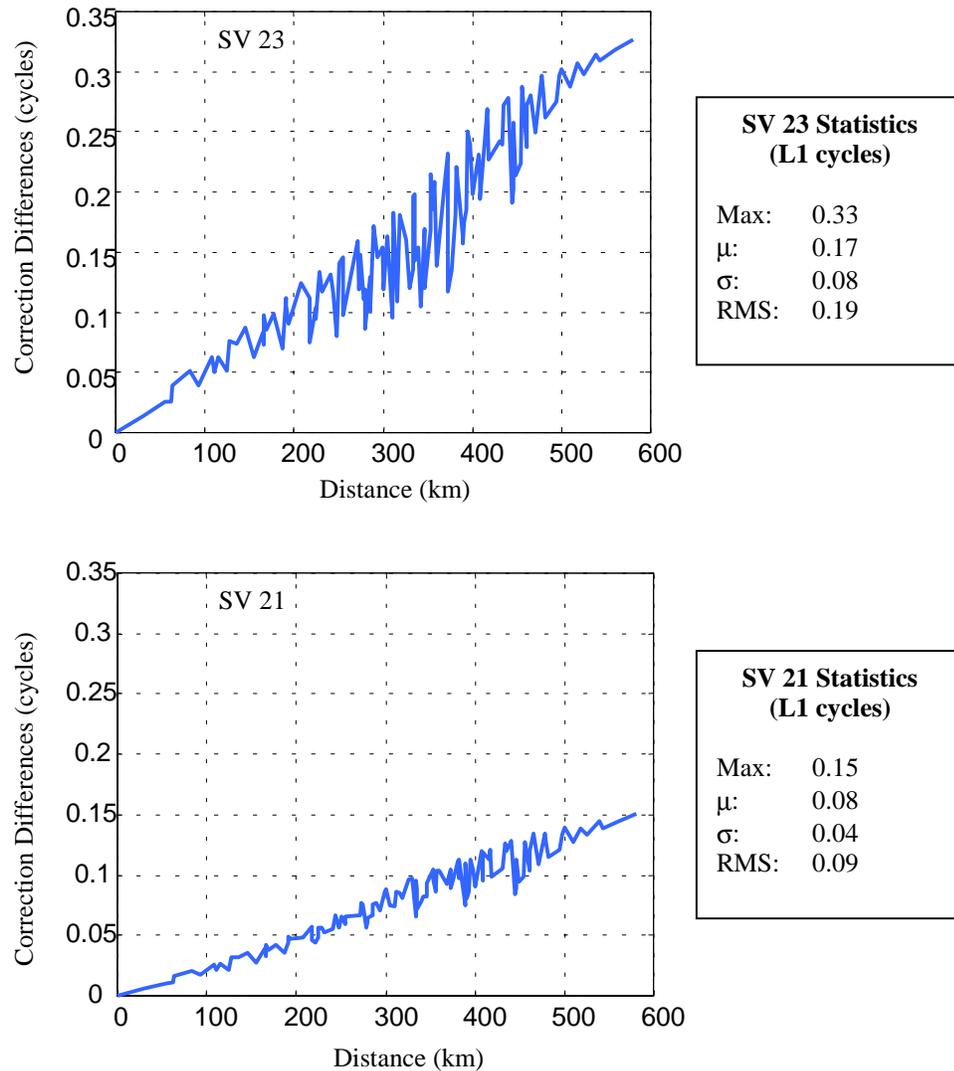
Figure 3.15: Correction Differences as a Function of Baseline Length

For the SWEPOS network, where the average station separation is around 200 km, the correction differences reach a maximum of 0.1 cycles (PRN 23) and an average of approximately 0.05 cycles for PRN 21. The latter translates to less than a centimetre difference. For periods of higher ionospheric activity, the decorrelating nature of the corrections (and hence the errors) will become evident over even shorter distances.

3.5   Low-Order Surface Modelling

In the previous section, the correction field was represented by interpolating a number of discrete gridded data points. This method worked well, especially with the higher resolution grids, such as $0.5° \times 0.5°$, where results are comparable to corrections computed directly for the user's location. However, for real-time applications it is desirable and necessary to limit the amount of data transmitted to the user receiver(s). With this concept in mind, a second method for parameterizing the correction field, described as low-order surface modelling, was investigated.

In this case, the role played by polynomials in curve fitting is extended to surface fitting by using polynomials of two variables, or bivariate polynomials. Such polynomial surfaces can be defined as a linear combination of a set of basis functions. These power basis functions and the corresponding polynomial class are shown in Table 3.11. In general P_n is a class of polynomials containing all functions of the form $x^i y^i$, where $0 \le i + j \le n$ and $i \ge 0$, $j \ge 0$ (Lancaster and Salkauskas, 1986). Essentially, each basis function can be considered as a surface on its own, where x and y are the horizontal coordinate values.

Table 3.11: Classes of Polynomial Functions

Class	Basis Functions
P_0	1
P_1	1 x y
P_2	1 x y x^2 xy y^2
P_3	1 x y x^2 xy y^2 x^3 x^2y xy^2 y^3

In its simplest form a polynomial surface is a plane. For RTK applications, this parameterization technique is favoured because of the *apparent* reduction in transmission bandwidth requirements. An apparent reduction is noted here, because the number of bits increases as the resolution of the values increases. Thus, in the case of polynomial surface coefficients, the number of bits per value may exceed those for lower resolution grid point values. This is an area of concern for RTK applications, which has secured the attention of many receiver manufacturers and researchers alike (Neumann et al., 1997; Hegarty, 1993) and is discussed further in the following chapter.

3.5.1   Plane Fit

In order to choose the best combination of power basis functions to model the satellite-based correction fields over the SSN coverage area, the 24-hour data set was decimated at one hour intervals. Samples of surfaces for each visible satellite were plotted to investigate their spatial behaviours. At this point, it was desired to isolate the spatial characteristics from the temporal aspects inherent in the corrections. Although, this cannot be accomplished in an absolute sense, sufficient samples were taken at all periods to obtain a representative data set. It became obvious from this analysis that a general trend from satellite-to-satellite or epoch-to-epoch, did not exist. As an example, Figure 3.16 shows the correction surfaces for six visible satellites at 2:00 pm local time over the coverage area. Modelling of irregular surfaces, as shown in Figure 3.16 is a common problem encountered in various applications, from topographic surface modelling to the

display of spatial image data. This analysis has taken the methods employed in these areas and applied them to a relatively new area, namely parameterizing corrections for DGPS positioning. Further details on surface fitting and modelling irregular surfaces, can be found in Landcaster and Salkauskas (1986), Watson (1992) and Junkins et al. (1973).

The criterion applied in the surface fitting approach was to limit the 'misfit' of the surface by imposing a minimum sum of squared differences between the approximated surface values and the gridded data. The resulting surfaces are known as trend or regression surfaces and they exhibit the major trend implied by each satellite at a particular epoch in time. Specifically, the coefficients for a first order surface (i.e. plane) were computed via a least squares adjustment with observation equations (Mikhail, 1976). The general form of the model is shown as follows,

$$p(x_c, y_c) = z = ax + by + c \qquad (3.9)$$

where $d^u = \begin{bmatrix} a & b & d \end{bmatrix}^T$ is the vector of unknown parameters and x, y, and z are the latitude, longitude and corrections of a pre-defined discrete set of grid crossover points, respectively. By minimizing the sum of the squared residuals ($v^T P v$ = minimum) and assigning equal weights to all observations, the adjustment gives the best-fit plane surface to the corrections. Therefore, the polynomial surface was not forced to pass through the discrete grid points used for establishing the plane coefficients. An example of a grid surface with a plane fit overlay for a single satellite at one epoch in time is shown in

Figure 3.17. It should be noted here, that large grids may lead to very large matrices in the least squares adjustment, however closed formulas can be derived which reduce the size of the matrix inversions (see Appendix C for derivations of first and second order polynomial coefficients).

As in the case of the grid-based parameterizations, results were obtained in the position domain by applying the best-fit plane corrections to the carrier phase measurements and processing the data using Flykin^{TM} for the Bora-Jonk baseline. The results showed a 0.04 metre RMS in latitude and longitude and a 0.1 metre RMS for height. By comparing these values to the grid-based position results, it is evident that the grid technique is more accurate, as long as the grid spacing is less than or equal to 1.5°. With sparser grids, such as the 2° case, the results using the best-fit plane are marginally better, although they may require less data to be transmitted to the user.
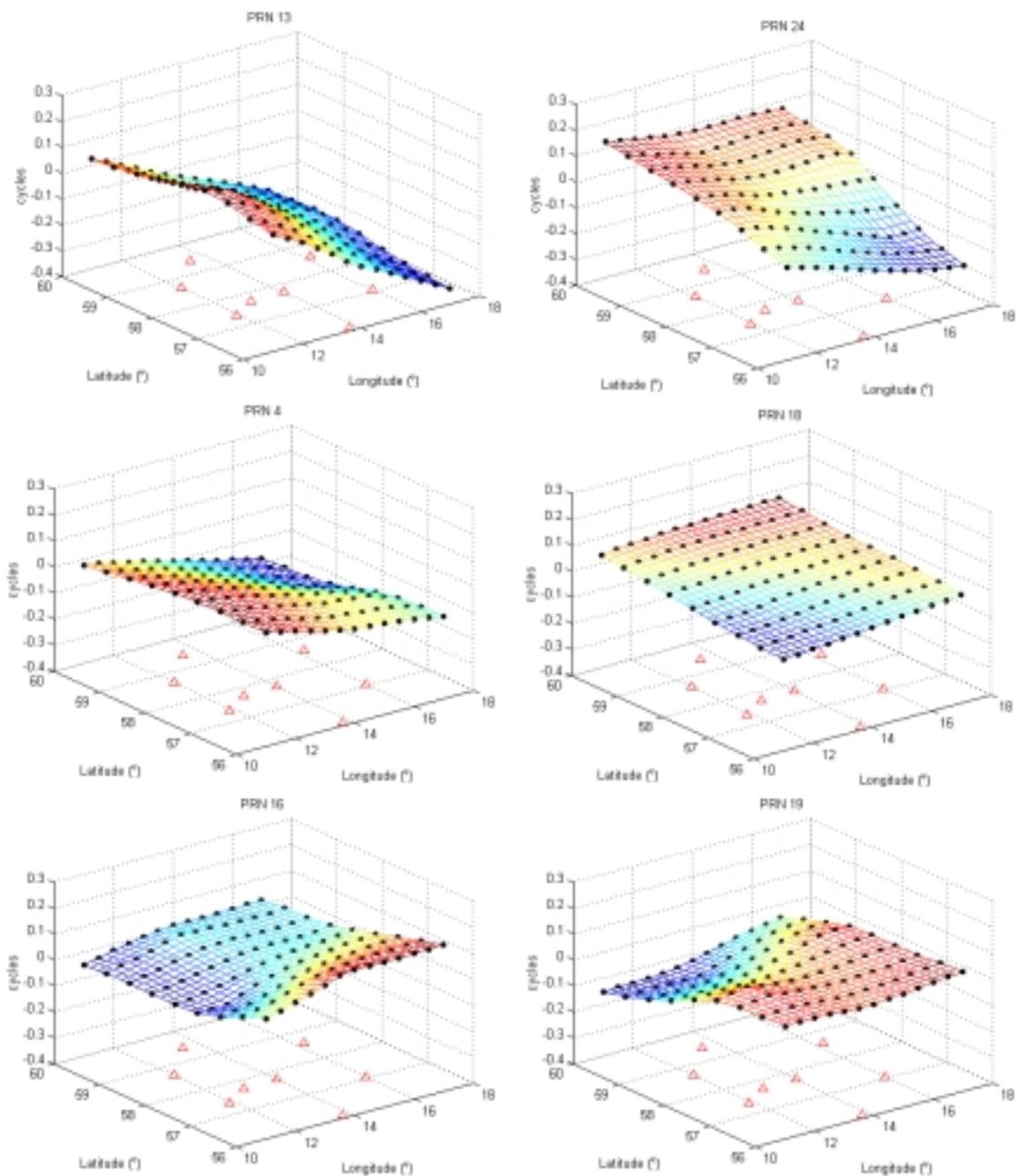
Figure 3.16: Samples of Combined L1 Phase Correction Surfaces for All Visible Satellites at 2:00 pm Local Time
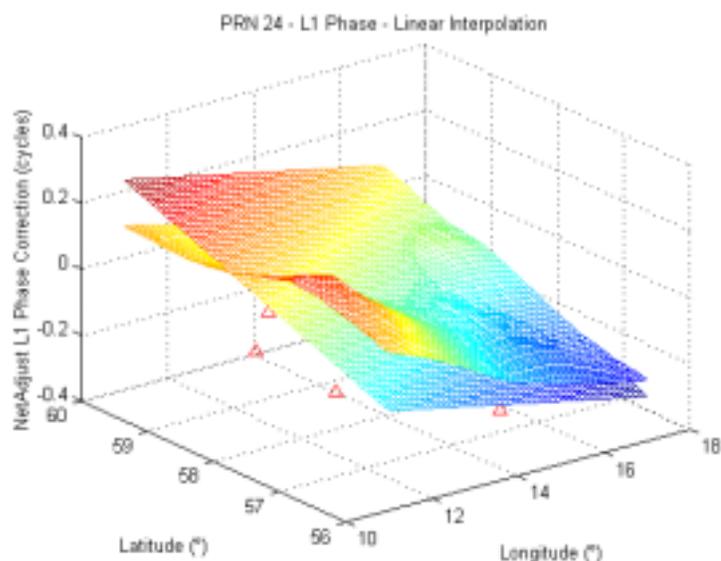
Figure 3.17: Example of a Correction Grid Surface and a Plane Fit Overlay

3.5.2 More Surface Fits

In addition to a simple three-parameter plane surface, the following higher coefficient polynomial surfaces were investigated,

a) $\quad z = ax + by + cxy + d$ (3.10a)

b) $\quad z = ax + by + cxy + dx^2 + e$ (3.10b)

c) $\quad z = ax + by + cxy + dx^2 + ey^2 + f$. (3.10c)

An example of a best fit polynomial surface of the format in eq. (3.10c) is shown in Figure 3.18. All of the above low-order polynomial surfaces, including the plane, are essentially subsets of the greater polynomial, *c*, shown above. Therefore, it is expected

that all surfaces will perform at approximately the same level. In other words, those coefficients which do not contribute to producing an optimal best-fit surface, will contain very small values output from the least squares adjustment. To test this hypothesis, position results were computed, using the method described above, for each of the second order polynomial surface fits.
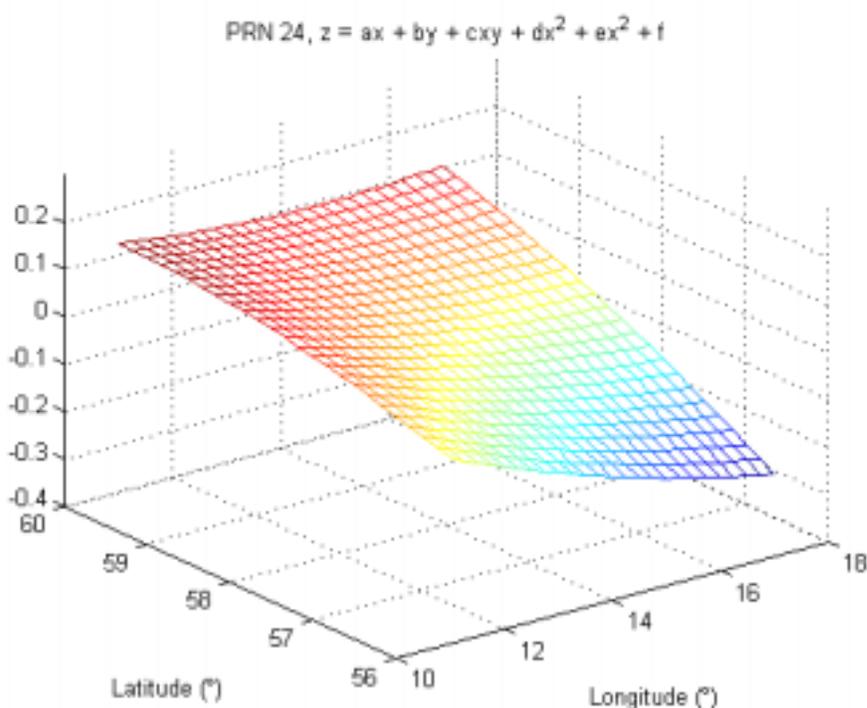
Figure 3.18: Example of the Six-Coefficient Correction Surface Fit

Table 3.12 summarizes the position domain results for the low-order polynomial surfaces. An interesting result is that the plane surface and polynomial surface (eq. 3.10a) performed at the same level, for latitude, longitude and height, which was parallel to the single baseline approach. This indicates that the contribution of the fourth coefficient in eq. (3.10a) is minimal for this case. However, the results in Table 3.12 also

show that by using the five and six coefficient second order models, RMS improvements over a single baseline approach, in latitude, longitude and height can be as much as 5%, 36%, and 18%, respectively. These values can be computed by comparing the values corresponding to the *no corrections* case in Table 3.9 with those of Table 3.12 for all three low-order surface fits.

A brief comparison of these position domain results with the grid-based parameterizations presented earlier, reveals that the achievable range of accuracy is approximately the same as the 1.5° grid spacing results. This is a tremendous improvement over the sparser grid-based results and has the potential of using fewer bits for data transmission. The exact number will depend on the method of data transmission as well as the actual magnitudes of the individual coefficient values.

Table 3.12: Statistics for Position Results using Low-Order Surface Fits (Bora$_r$-Jonk$_u$)

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
$z = ax + by + c$, $z = ax + by + cxy + d$					
Latitude	-0.158	0.187	0.021	0.037	0.043
Longitude	-0.135	0.096	-0.020	0.036	0.041
Height	-0.538	0.261	-0.064	0.083	0.105
$z = ax + by + cxy + dx^2 + e$					
Latitude	-0.078	0.181	0.020	0.032	0.038
Longitude	-0.119	0.095	-0.014	0.032	0.034
Height	-0.383	0.0257	-0.043	0.075	0.087
$z = ax + by + cxy + dx^2 + ey^2 + f$					
Latitude	-0.088	0.164	0.018	0.032	0.037
Longitude	-0.120	0.086	-0.016	0.033	0.037
Height	-0.383	0.259	-0.043	0.076	0.087

As it was mentioned previously, the key motivation for investigating the low order surface models was because of the potential for decreasing the data transmission load. Following this reasoning, it is not advantageous to generate complex higher order polynomial surfaces for representing the correction area. Not only will this defeat the purpose of reducing the required transmission capacity, but the possibilities of 'over' and 'under' shooting also increase. That is, with higher order fits, oscillations may be produced which do not accurately represent the error features in the network coverage area. Therefore, this is a situation where more is not necessarily better.

3.6   Remarks on Spatial Characterizations

Various methods for parameterizing carrier phase corrections for the correlated errors (ionospheric, tropospheric and satellite orbit) over a regional network of reference stations were presented. The grid-based method employed the formulation of corrections at discrete points, with spacings ranging from 0.5° to 2°. The results showed as much as 20%, 54% and 27% improvement in RMS for latitude, longitude and height respectively, over the single baseline approach. The achievable accuracies in position decreased with lower grid resolution, to the point where a 2° spacing was found to be worse than a single baseline approach. These findings are quite significant because they identify the limitations of the grid-based model for RTK applications.

In the interest of meeting data transmission bandwidth requirements, a second parameterization approach, namely low-order surface modelling, was investigated. In this case four different first and second order polynomial surfaces were used to model the corrections over the regional network. The results in the position domain showed up to 5%, 36% and 18% improvements in RMS for latitude, longitude and height respectively, over the single baseline approach, using a second order polynomial fit.

Comparing the two parameterization schemes, it is evident that both models are capable of producing results better than the single baseline solution. The following chapter contains results based on the temporal correlation of these errors, since it is the combination of the spatial and temporal aspects that will result in a practical realization of the multi-reference station concept for RTK applications.

Chapter 4

TEMPORAL CHARACTERISTICS OF CORRELATED ERRORS

In this chapter the temporal behaviour of the combined corrections is investigated. This is accomplished by establishing how the corrections change over time through studies on the effects of data decimation, analyzing the statistics for groups of data at varying times throughout the day and night, generating correction surface snapshots for numerous satellites and computing the correction rates of change. Based on the spatial parameterization techniques discussed in Chapter 3, the effect of correction data latency on the final positioning accuracy will also be determined for both grid-based and low-order surface models. This will provide valuable insight into the relationship between the correction update rates and the parameterization schemes.

Also, investigations are conducted on the required correction information for a message generated at a network/reference station and transmitted to all multi-reference station users. Although, current standards for correction messages exist, receiver manufacturers often adopt their own proprietary message format to serve the needs of GPS users. Regardless of the message format, specific information exists, which must be transmitted

and that will be identified in the latter parts of this chapter. Finally a summary of the information presented in this chapter as well as some contributions from Chapter 3 are included to aid with the identification of the best option for parameterization and dissemination.

4.1   Behaviour of Corrections Over Time

The correlated residual errors, which increase as the baseline distances increase, exhibit both spatial and temporal correlations. Spatial correlations involve observations from one station to different satellites or between two or more stations and the same satellite at one epoch in time. The latter was studied in Chapter 3. This involves the behaviour of the corrections over space, or more specifically, over the network coverage area. On the other hand, temporal correlations exist between observations at one station to one satellite over different epochs. Numerous studies have been performed on the implementation of a correlation function for modelling temporal correlations to aid with GPS positioning (El-Rabbany et al., 1992; Wang, 1998; Howind et al., 1999). Throughout these studies it was shown that physical temporal correlations do exist between observations at different epochs and they may impact precise positioning results. It is the purpose of this chapter to identify the behaviour of the NetAdjust corrections over time for a test area, in order to provide a better understanding of the temporal characteristics.

4.1.1   Data Decimation

One of the most difficult practical aspects encountered when implementing a real-time multiple reference station network approach is dealing with the large amounts of data coming from various streams. With this in mind, and in the interests of data transmission bandwidth limitations, it is common practice to decimate (thin) the available data at a rate which will not significantly reduce the amount of information provided to the network reference stations or the user receivers.

For the SSN 24-hour data set, the data was provided at 1 Hz. This non-decimated data was used as the base set for comparison. It is well known that the ionospheric activity usually peaks on the day side, between 12:00 and 14:00 hours local time. Investigations on the behaviour of the corrections showed that during this period the correction surfaces exhibited higher frequency variations than those during the night side. Therefore, for testing the data decimation issue, a four hour period representing the most variable period in the data set was chosen for the analysis, between 11:00 and 15:00 hours local time. Since the majority of the variations will occur during this period, any lost information due to data decimation will become evident. Figure 4.1 shows a plot of the L1 phase corrections for all visible satellites during this four-hour test period for the Bora-Jonk baseline.
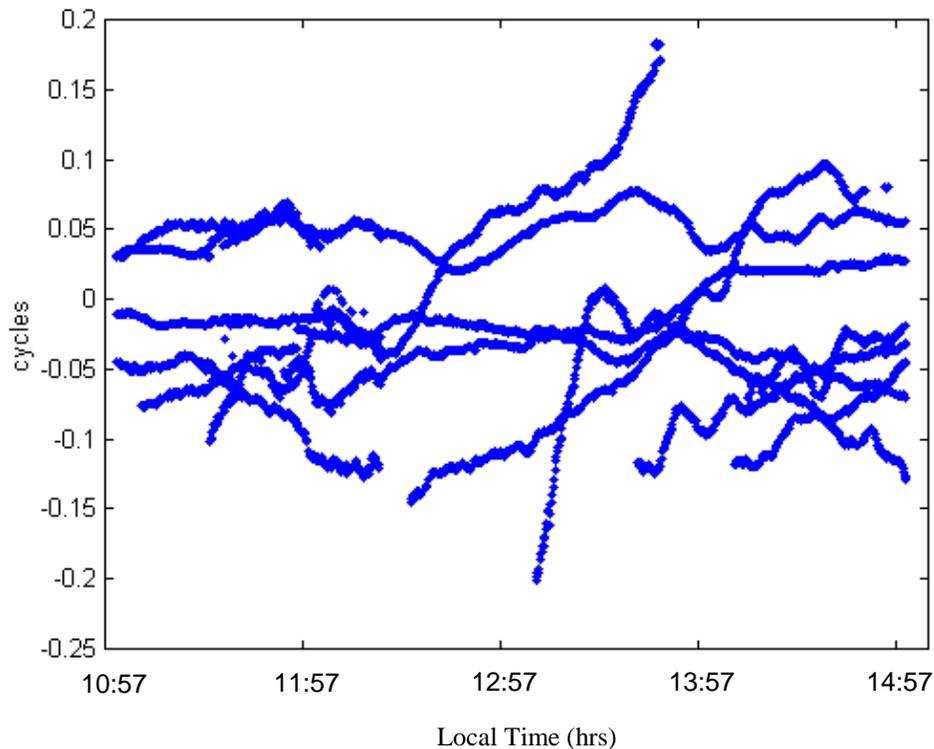
Figure 4.1: L1 Phase Corrections for All Visible Satellites of the Bora-Jonk Baseline

As it can be seen from the figure there are numerous threads beginning and ending at various times, representing the different satellites, which come in and out of visibility throughout the period. Figure 4.2 provides a time series of the satellites visible for the prediction point (Jonk) and the reference station (Bora). Only the satellites visible at both stations have generated corrections and are used in the combined solution for the final position.

Figure 4.2 is for the 1 Hz data set. Similar figures were generated for 5, 10, 15, 30 and 45 second data intervals and were found to be essentially the same. Figure 4.3 summarizes the number of visible satellites for both the user and reference stations. It should be noted that this was for a ~70 km baseline. As the distance between the user

and reference stations increases, the possibility of having the same number of common satellites will decrease, as will the corresponding set of corrections, resulting in fewer observations for computing the user position.
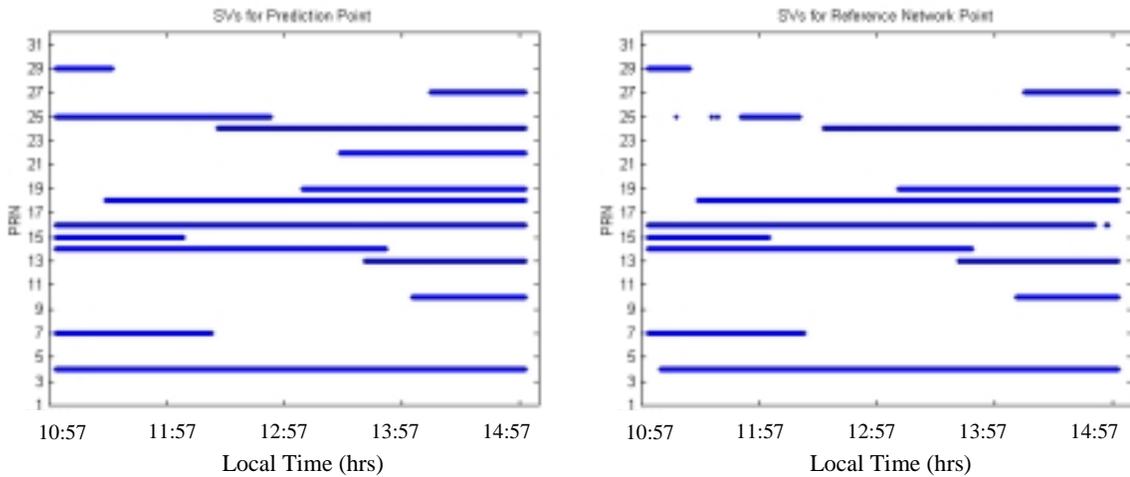
Figure 4.2: Satellite Visibility at the User (left) and Reference Station (right)
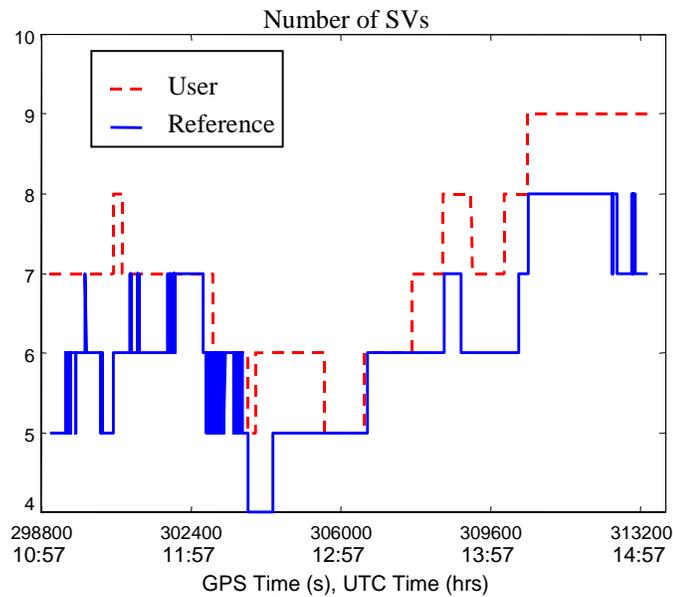
Figure 4.3: Number of Visible Satellites at the User and Reference Stations

From Figure 4.1 it is evident that the corrections for each satellite are different, varying between ± 0.2 cycles (translating to ± 3.8 cm for L1 phase) and they do not appear to have any predictable behaviour, with the exception of the greatest variation, occurring between 13:00 and 14:00 hours for all satellites. This is expected behaviour because of the peak in ionospheric activity during this time.

Figure 4.4 shows the widelane corrections for the same set of satellites during this day side period. These corrections exhibit higher frequency fluctuations most probably due to the amplification of the carrier phase multipath and noise inherent in the widelane phase combination, which is also seen in the relatively larger magnitudes of the WL corrections compared to the L1-only corrections.
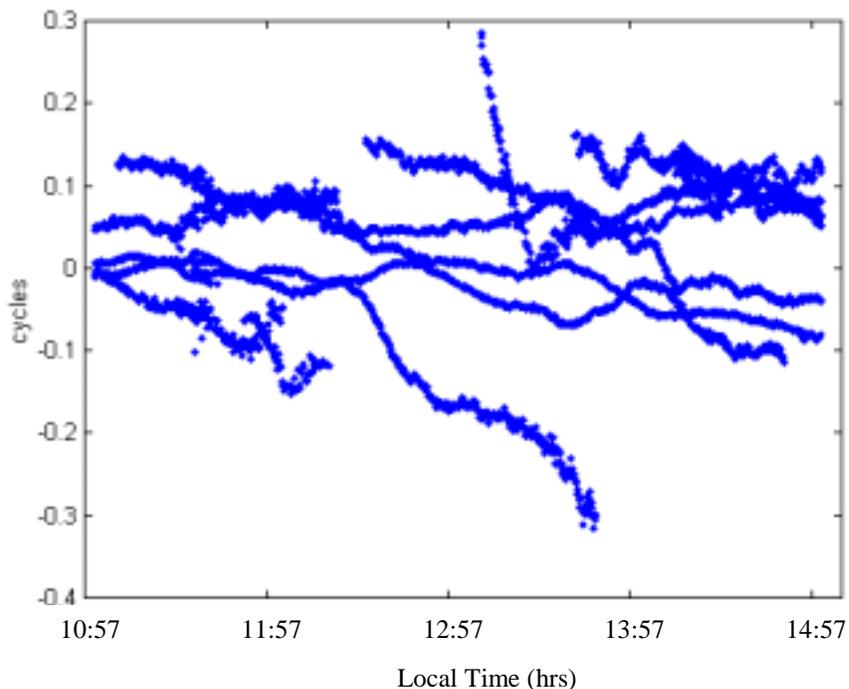
Figure 4.4: WL Phase Corrections for All Visible Satellites of the Bora-Jonk Baseline

Data decimation acts as a low pass filter, which limits the reduction of high frequency

components of the corrections. For real-time applications, only the corrections ($\hat{\delta\varphi}$)

present at a specific epoch in time are of interest, which do not take into account any

memory of the past corrections. Therefore, during data decimation any changes in the

corrections at a current epoch from the previous epochs will be lost. Investigations on the

rate of change of the corrections (*RCC*) were conducted using the following equation:

$$RCC = \frac{\hat{\delta\varphi}_t - \hat{\delta\varphi}_{t-\Delta t}}{\Delta t} \qquad (4.1)$$

The *RCC* was evaluated for the one second data interval, Δt. Statistics were computed

for all satellites that were visible for at least 2.5 hours, in order to provide an adequate

sample size, and are provided in Table 4.1.

Table 4.1: L1 Phase Correction Rates

SV	Min (cycles/sec)	Max (cycles/sec)	μ (cycles/sec)	σ (cycles/sec)	RMS (cycles/sec)
4	-0.006	0.013	-0.000004	0.000742	0.000742
14	-0.012	0.016	0.000018	0.000911	0.000911
16	-0.009	0.010	0.000003	0.000928	0.000928
18	-0.056	0.018	0.000006	0.001202	0.001202
24	-0.003	0.003	0.000008	0.000605	0.000605

The correction rates vary from satellite to satellite, with RMS values ranging from 0.0006

cycles/second to 0.001 cycles/second. This range can be considered as a worst case

estimate, since the test period was taken during the highest variability in the corrections.

Nonetheless, these RMS values were used and extrapolated in order to determine the expected data decimation error for data intervals as high as one minute and plotted in Figure 4.5. In most cases, the responsible authorities for distributing data for a permanent array of reference stations usually make the data available every 15 or 30 seconds. For the purposes of the majority of the position results presented in this thesis, a 15 second data decimation interval was chosen. For L1 phase corrections this translates to approximately 0.17 cm to at most 0.29 cm RMS (see Figure 4.5), which is on the level of the noise and thus an admissible error.
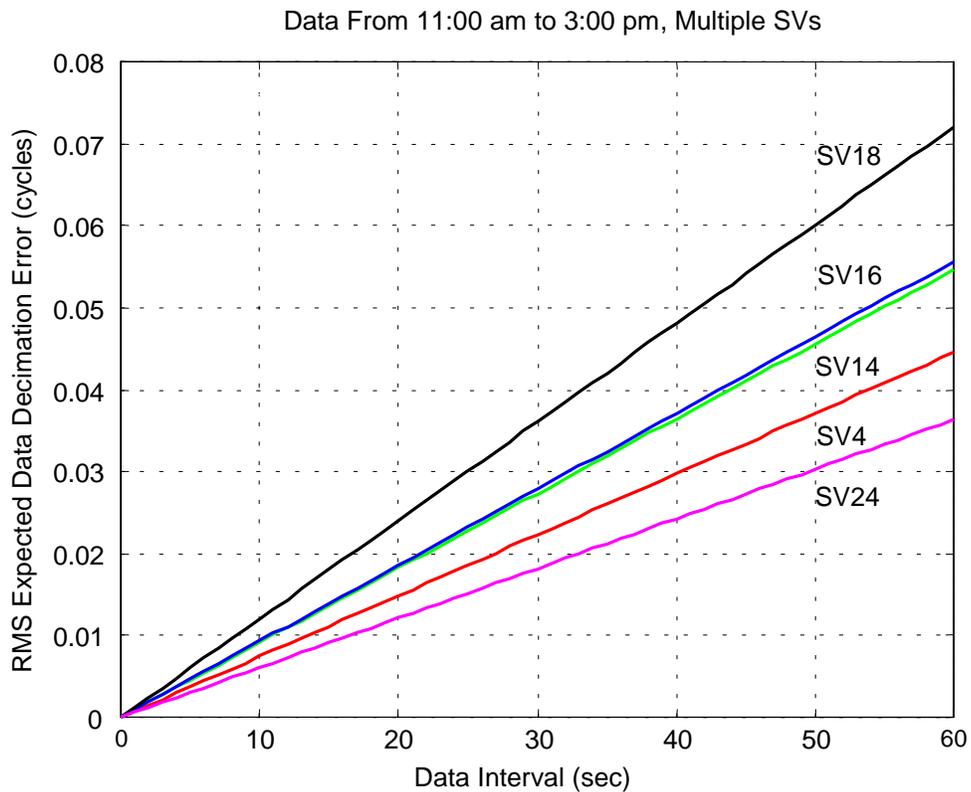
Figure 4.5: RMS Expected Data Decimation Error (L1 phase) for Various Data Intervals

4.1.2   Varying the Time of Day

To test the effects of the time of day on the shape and values of the correction surfaces, a series of 'snapshots' were taken for numerous satellites throughout their entire pass over the network coverage area and at intermittent periods throughout the day and night. Two specific periods were identified to represent the a) high variability and b) low variability of the correction values. The high variability period, also called the day side, occurred during 11:00 to 15:00 hours local time with a computed 4.6 ppm double difference ionospheric delay in terms of L1. The low variability period, referred to as the night side, occurred during 24:00 to 4:00 hours local time with an associated 3 ppm double difference ionospheric delay in terms of L1. Statistics of the range in correction values were computed for these two periods and are summarized in Table 4.2. The minimum and maximum values are included for five different phase combinations, including L1 phase, L2 phase, WL phase, IF phase and IS phase combinations. Note that the statistics in Table 4.2 are for the user corrections only, as opposed to the combined reference station and user corrections, which were used for generating surfaces in Chapter 3.

It is evident from Table 4.2 that the range in correction values remains approximately the same during the two periods, while the maximum (absolute value) phase corrections occur during the day side. This is expected due to the prominent higher ionospheric activity during this latter period (4.6 ppm as opposed to 3 ppm for period (b)). It is also interesting to note that the correction values are not mapped to zenith, therefore they are the actual corrections received at the user locations. Investigations based on the satellite

elevation also revealed that the lower elevation satellites produce surfaces with greater variability and higher correction values, indicating the higher magnitude of the residual errors that remain unmodelled over lower elevations. It should also be noted that a visible change in correction surfaces is observed as the satellite continues along its path and the azimuth varies from snapshot to snapshot.

Table 4.2: Range of Corrections During the Day side and Night side Periods

Correction Type	Day side (11 am to 3 pm)		Night side (12 am to 4 am)	
	Minimum	**Maximum**	**Minimum**	**Maximum**
L1 Phase (cycles)	-0.044	0.048	-0.029	0.015
L2 Phase (cycles)	-0.078	0.077	-0.045	0.021
WL Phase (cycles)	-0.055	0.086	-0.024	0.033
IF Phase (cycles)	-0.008	0.026	-0.013	0.013
IS (metres)	-0.029	0.035	-0.009	0.017

To illustrate these findings, several snapshots at one-hour intervals were taken for a single pass of a satellite and its corresponding correction surfaces were plotted (using the dense $0.5° \times 0.5°$ grid). The satellite number and elevation are indicated along with the epoch in Figures 4.6 and 4.7. Here again, the higher variability in the surface lattice during the day period is emphasized. The correction surfaces generated in these figures include the total values from combining the reference station and user corrections.

Table 4.2 also provides information on the contribution of the ionosphere to the combined correction values by looking at the various phase combinations, namely the ionospheric-free versus the L1 phase values. This indicates the significance in studying the effects of *separated* or error-specific corrections. This may involve the separation of

all three components engaged in the combined corrections to a more simple representation where correction values are assigned for the ionosphere separately and a combined troposphere and orbit effect. The implications of separating the corrections may lead to perhaps a lower update rate for the less variable/significant errors (i.e. tropospheric and orbit) and a higher update rate for the ionospheric contribution. However, further studies on this matter must be conducted in order to decipher any practical advantages.
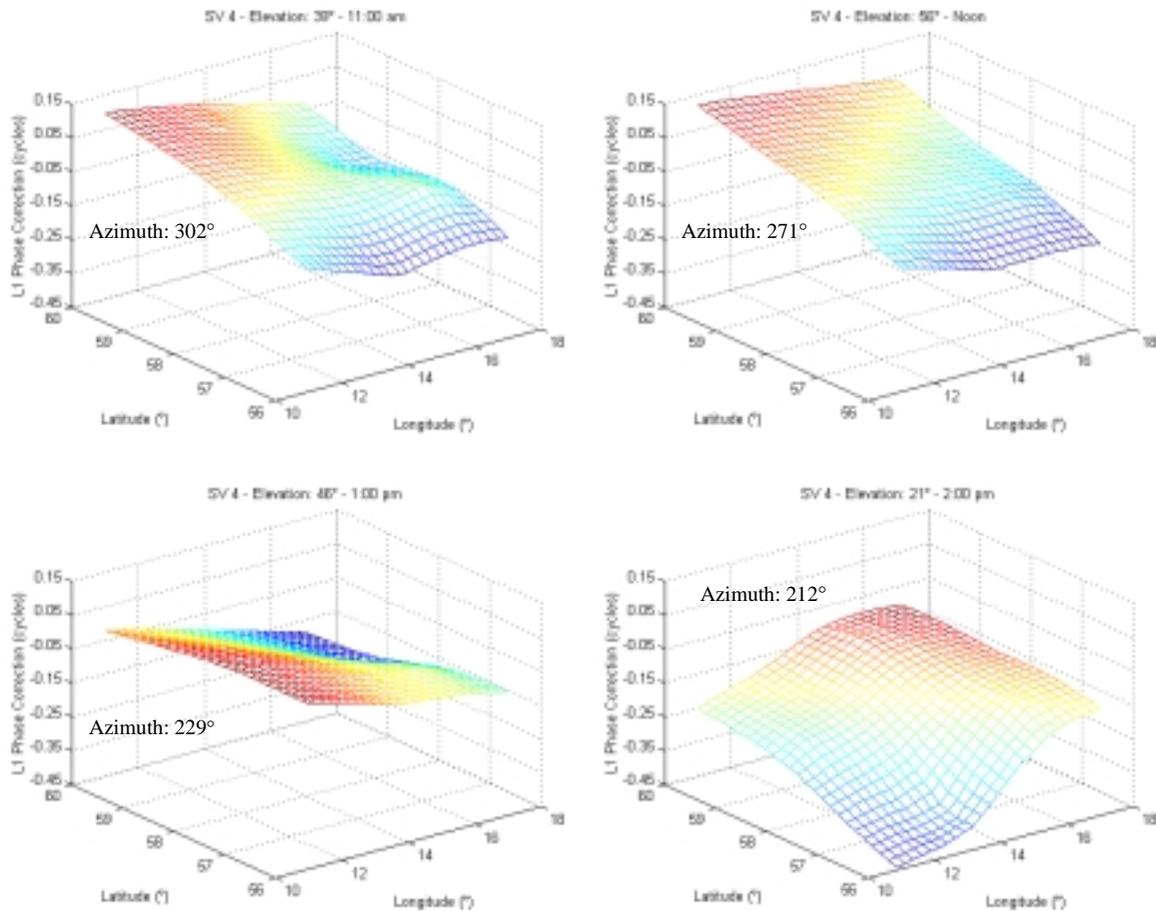
Figure 4.6: Snapshots of Satellite-Based Correction Surfaces for Day Time Period
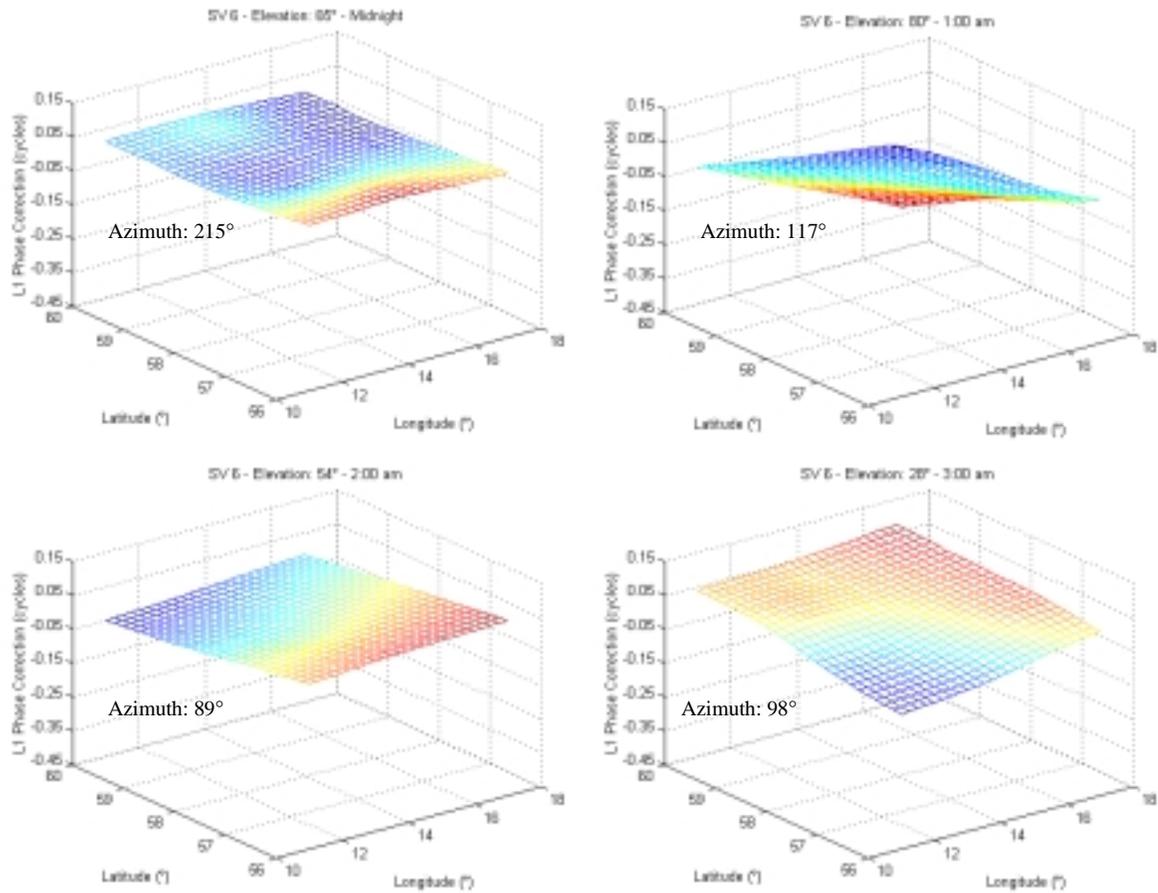
Figure 4.7: Snapshots of Satellite-Based Correction Surfaces for Night Time Period

4.2   Parameterization Scheme Update Rates

In this section, the results from numerous processing scenarios are presented in order to test the effects of the correction data update rate (also called latency) on the user's final position accuracy. Four different parameterization schemes were used to evaluate the position accuracy at eleven different correction ages, as described below.

In each case, the corrections were generated as described in Chapter 3 for the Jonk station (see Figure 3.3). In the interests of consistency and comparison purposes, positions were computed using the standard Bora-Jonk baseline. Under *no latency* conditions, the L1 carrier phase corrections are synchronously applied to the reference and user(s) measurements. This implies that the time tags associated with the measurements are 'matched' with the generated corrections. This is shown in general terms as:

$$\varphi'(t) = \varphi(t) + \hat{\delta}\varphi(t) \tag{4.2}$$

where $\varphi'(t)$ is the corrected carrier phase measurement (for the reference or rover stations), $\varphi(t)$ is the current carrier phase measurement at time t, and $\hat{\delta}\varphi(t)$ is the corresponding carrier phase correction. To test the effect of applying a correction κ_i seconds old, *new* carrier phase measurements are formed where,

$$\varphi'(t - \kappa_i) = \varphi(t) + \hat{\delta}\varphi(t - \kappa_i) \tag{4.3}$$

and $\varphi(t - \kappa_i)$ is the correction that is κ_i seconds old. For the results presented below a range of values from zero to 300 seconds were used, as follows:

$$\kappa = \begin{bmatrix} 0 & 10 & 15 & 20 & 25 & 30 & 60 & 100 & 150 & 200 & 250 & 300 \end{bmatrix} \text{seconds} \tag{4.4}$$

For the cases of the grid parameterizations ($0.5° \times 0.5°$ and $1.5° \times 1.5°$) the κ can be considered as the period between sequential grid node updates. Similarly, for the function parameterizations (plane and eq. 3.10c) κ is the polynomial coefficient update period.

The results of the position accuracy given a specified update rate are provided in Tables 4.3 to 4.7 for various schemes. Table 4.3 can be considered as the basis for comparison as it includes the statistics for a corrected single baseline approach. That is, the L1 phase combined corrections are directly computed and applied to the user's measurements and a single reference station. Tables 4.4 and 4.5 show the statistics for the same update rates as in the single baseline approach, with respect to the plane fit parameterization and the six-coefficient fit (eq. 3.10c). Finally, Tables 4.6 and 4.7 give the results for the grid parameterizations.

These tables show that in general (i.e. update rates less than approximately 250 seconds) the corrected single baseline approach is the best, followed by the $0.5° \times 0.5°$ and $1.5° \times 1.5°$ grid models, then the six-coefficient fit and plane fits.

Table 4.3: Statistics for Position Errors using Single Baseline with Various Update Rates

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
30 seconds Latency					
Latitude	-0.136	0.159	-0.001	0.031	0.031
Longitude	-0.093	0.102	0.002	0.026	0.026
Height	-0.289	0.324	-0.015	0.075	0.077
100 seconds Latency					
Latitude	-0.135	0.163	-0.001	0.032	0.032
Longitude	-0.098	0.109	0.001	0.027	0.027
Height	-0.302	0.273	-0.015	0.077	0.078
200 seconds Latency					
Latitude	-0.145	0.161	-0.001	0.034	0.034
Longitude	-0.105	0.122	0.001	0.028	0.028
Height	-0.306	0.312	-0.015	0.080	0.082
300 seconds Latency					
Latitude	-5.260	1.183	-0.001	0.085	0.085
Longitude	-1.505	1.127	0.001	0.036	0.036
Height	-6.321	0.464	-0.014	0.123	0.124

Table 4.4: Statistics for Position Errors using Plane Fit with Various Update Rates

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
30 seconds Latency					
Latitude	-0.148	0.197	0.021	0.037	0.043
Longitude	-0.135	0.105	-0.020	0.036	0.041
Height	-0.516	0.284	-0.064	0.083	0.105
100 seconds Latency					
Latitude	-0.154	0.236	0.021	0.038	0.043
Longitude	-0.134	0.108	-0.021	0.036	0.042
Height	-0.516	0.348	-0.064	0.084	0.106
200 seconds Latency					
Latitude	-0.157	0.270	0.021	0.039	0.044
Longitude	-0.131	0.119	-0.022	0.036	0.042
Height	-0.549	0.410	-0.064	0.086	0.107
300 seconds Latency					
Latitude	-0.409	2.120	0.022	0.056	0.060
Longitude	-0.218	0.162	-0.022	0.037	0.043
Height	-0.564	0.720	-0.061	0.092	0.111

Table 4.5: Statistics for Position Errors Using $z = ax + by + cxy + dx^2 + ey^2 + f$ with Various Update Rates

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
30 seconds Latency					
Latitude	-0.097	0.178	0.018	0.032	0.037
Longitude	-0.120	0.089	-0.016	0.033	0.037
Height	-0.378	0.260	-0.043	0.076	0.087
100 seconds Latency					
Latitude	-0.136	0.216	0.018	0.033	0.038
Longitude	-0.123	0.092	-0.017	0.033	0.037
Height	-0.376	0.318	-0.043	0.077	0.089
200 seconds Latency					
Latitude	-0.137	0.247	0.019	0.034	0.039
Longitude	-0.126	0.101	-0.017	0.033	0.037
Height	-0.374	0.376	-0.043	0.080	0.091
300 seconds Latency					
Latitude	-3.027	1.869	0.019	0.065	0.068
Longitude	-0.937	0.142	-0.018	0.036	0.040
Height	-3.679	0.648	-0.042	0.100	0.108

Table 4.6: Statistics for Position Errors using Grid (0.5°) with Various Update Rates

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
30 seconds Latency					
Latitude	-0.138	0.161	-0.001	0.031	0.031
Longitude	-0.094	0.106	0.002	0.027	0.027
Height	-0.287	0.355	-0.013	0.076	0.077
100 seconds Latency					
Latitude	-0.137	0.166	-0.001	0.033	0.033
Longitude	-0.098	0.113	0.002	0.027	0.027
Height	-0.296	0.277	-0.013	0.078	0.079
200 seconds Latency					
Latitude	-0.151	0.156	-0.001	0.035	0.035
Longitude	-0.105	0.127	0.002	0.028	0.028
Height	-0.310	0.312	-0.012	0.081	0.082
300 seconds Latency					
Latitude	-5.599	1.137	-0.001	0.089	0.089
Longitude	-1.599	0.133	0.001	0.037	0.037
Height	-6.726	0.451	-0.011	0.128	0.128

Table 4.7: Statistics for Position Errors Using Grid (1.5°) with Various Update Rates

	Min (m)	Max (m)	μ (m)	σ (m)	RMS (m)
30 seconds Latency					
Latitude	-.0156	0.249	-0.003	0.036	0.036
Longitude	-0.095	0.134	0.006	0.031	0.032
Height	-0.291	0.616	0.003	0.087	0.087
100 seconds Latency					
Latitude	-0.157	0.209	-0.003	0.037	0.037
Longitude	-0.100	0.143	0.006	0.032	0.032
Height	-0.308	0.516	0.003	0.088	0.089
200 seconds Latency					
Latitude	-0.185	0.162	-0.003	0.040	0.040
Longitude	-0.109	0.159	0.006	0.033	0.033
Height	-0.349	0.401	0.003	0.092	0.092
300 seconds Latency					
Latitude	-8.352	0.862	-0.003	0.125	0.125
Longitude	-2.360	0.167	0.006	0.047	0.048
Height	-10.018	0.408	0.003	0.170	0.170

To show the relative accuracy, graphs of the update rate versus the RMS error in latitude, longitude, and height are provided in Figures 4.8, 4.9, and 4.10, respectively. These graphs clearly reveal the most significant relationships between update rate and position error with respect to the parameterization scheme.   For the sake of discussion, concentration will be placed on describing the latitude RMS error behaviour first (Figure 4.8). The solid line constant at approximately 0.31 metres RMS provides the value for zero latency in corrections (as in eq. 4.2). Furthermore, for several seconds at the beginning of each plot, the parameterization scheme's optimal performance is given. As the period between updates increases, the achievable accuracy slowly degrades.

It is interesting to note that the corrected single baseline approach and the $0.5° \times 0.5°$ grid perform at essentially the same level throughout. Up to approximately 45 seconds, $0.5° \times 0.5°$ and the single baseline do not show any degradation in accuracy over the zero latency case. However, even after approximately 210 seconds, the RMS accuracy of these two schemes is better relative to any of the other parameterizations. At approximately 250 seconds latency, the function-based models, reveal better accuracy than the others, indicating the lower frequency behaviour of the polynomial function coefficients and the smoothing effect of the models. Also, up to approximately 215 seconds the plane fit is the worst in performance, which is replaced by the sparser $1.5° \times 1.5°$ degree grid parameterization after that.

The longitude and height RMS error plots show similar trends as the latitude. As expected the accuracy in height is worse than the horizontal components' accuracies. The horizontal component plots are plotted on the same scales for relative inter-comparison purposes, with only slightly different scales for the vertical component to allow for the greater range in values. Figures 4.8 through 4.10 highlight the fact that the parameters for all schemes do change slowly, allowing the user to apply the corrections at intervals greater than one second. This is a significant finding in terms of data transmission requirements. It implies that accommodations can be made for updating one satellite each second and consequently all of the visible satellites over a corresponding equivalent number of seconds (see the next section).
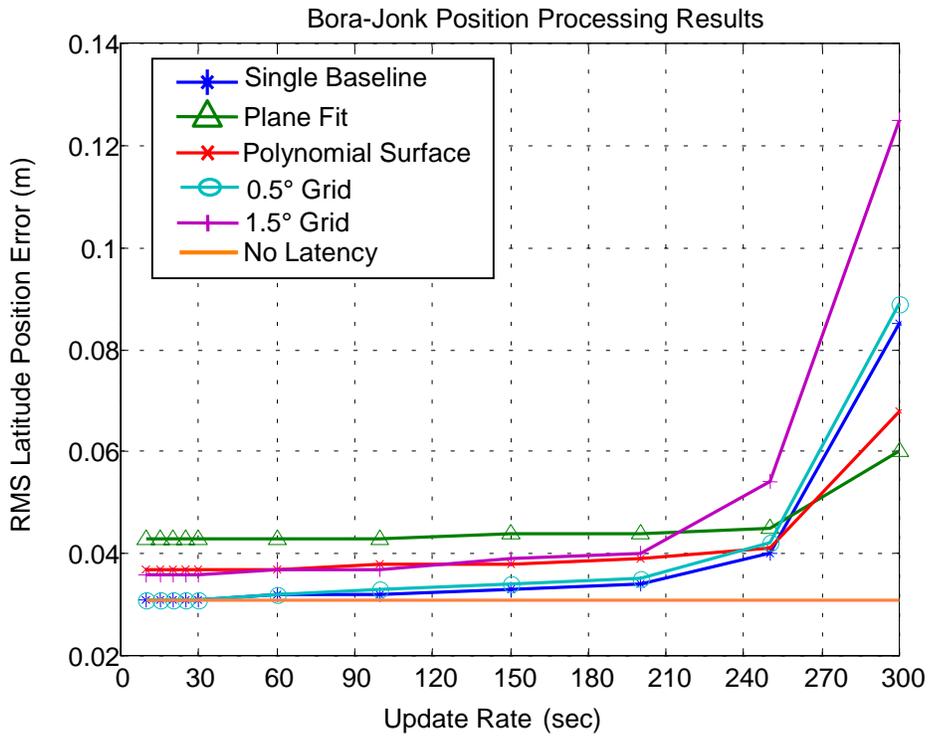
Figure 4.8: RMS Latitude Errors for Various Update Rates
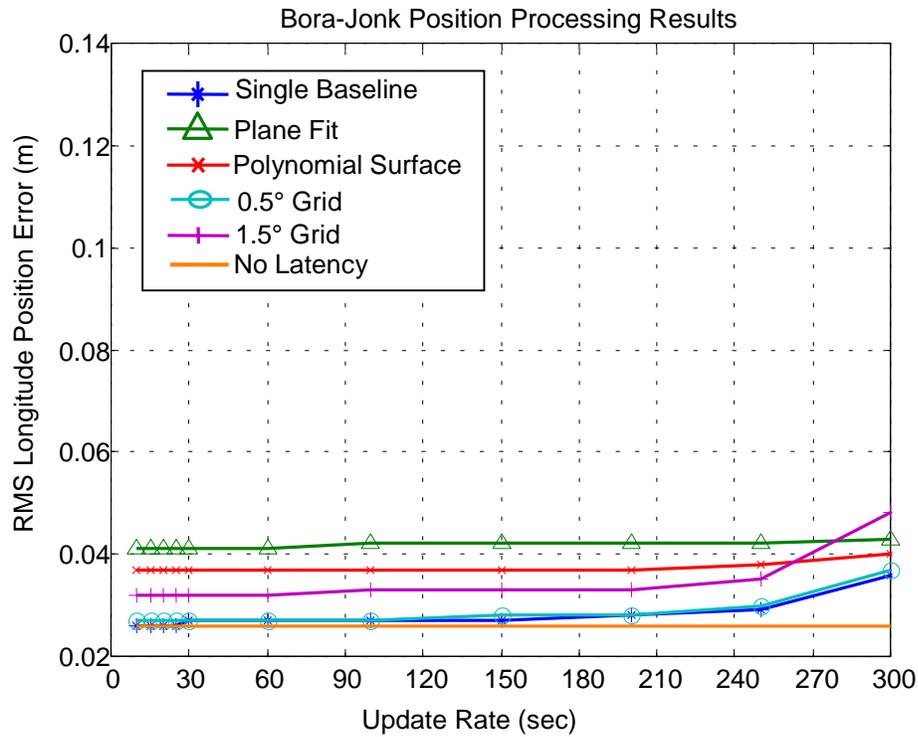
Figure 4.9: RMS Longitude Errors for Various Update Rates
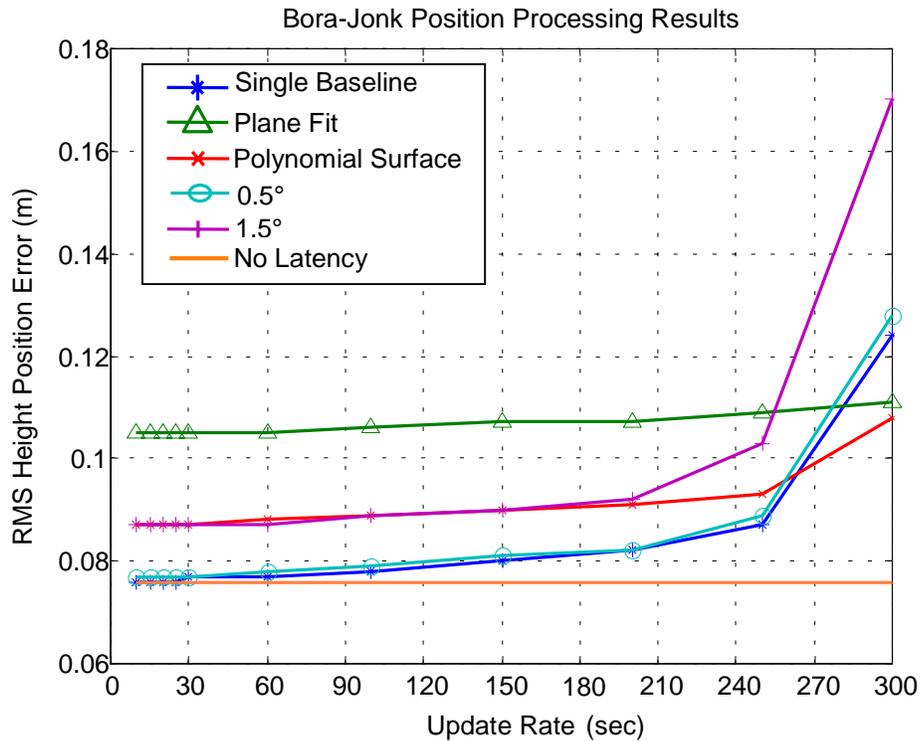
Figure 4.10: RMS Height Errors for Various Update Rates

Similar trends as those observed in this temporal analysis are found in the spectral analysis included in Chapter 5, where the update periods for each satellite are linked to the corresponding correlation length.

4.3 Correction Message Transmission Information

Without a proper communication process, the dissemination of the corrections is incomplete. The format chosen is based on many factors, such as the required accuracy level, the behaviour of the errors over the network coverage area, as well as the data transmission capacity. In previous discussions, the behaviour of the errors over space

and time were investigated and the analysis was carried through to the position domain. This provided insight into the achievable accuracy level associated with the various methods. Two main options for carrier phase correction parameterizations were proposed, thus providing two transmission options, namely:

1) transmit grid-based format, or

2) transmit low-order surface model (function-based) format.

Now, the issue of data transmission capacity for both options stated above, will be discussed.

Several properties of the NetAdjust corrections are common for both the grid-based and function-based models presented thus far. Firstly, the generated corrections are combined to contain all the effects of the correlated error sources of a given satellite. Secondly, the corrections are satellite-based, rather than an amalgamation of corrections for all useable satellites. Therefore, for any parameterization case, one set of reference station measurements and the carrier phase corrections must be transmitted. In this way, the user receiver can apply the corrections to its measurements and then use standard double-difference techniques already implemented in user software.

For transmitting the existing reference station measurements, the standard RTCM types 18/19 are available (RTCM, 1994) to be used or alternatively a proprietary 'special'

format can be implemented (see Talbot, 1996 and Neumann et al., 1997). Either way, the framework for this transmission process has been considered in previous discussions and is not repeated here. On the other hand, the transmission of carrier phase corrections is more complicated and has not been addressed in such depth. To date, there is no widely used standard message available for applying carrier phase corrections. RTCM message types 20/21 have been developed for this purpose, however many receiver manufacturers tend to gravitate towards using their own proprietary message format. Currently, there is a sub-committee of the RTCM Special Committee, RTCM SC-104, called the 'Network RTK Working Group' which is addressing the issue of network RTK messages.

To this end, it is not the purpose of this section to create another 'proprietary' message format. Instead, a discussion on the essential information for each parameterization scheme will be included. This will lay the foundation for a comparative analysis of models based on the data transmission load. Then based on the desired combination of accuracy, error behaviour and data transmission, the most practical alternative(s) can be pursued further for implementation in real-time.

4.3.1   Correction Transmission Options

In terms of the actual transmission methods, there are several plausible alternatives depending on the reference network geometry, more specifically the inter-station distances, and cost. Common data link methods include radio/modem, satellite link,

cellular phone, FM sub-carrier (Talbot, 1996). The alternatives for operations involving physical ground-based data links are:

a) *Transmit one set of corrections (grid or function based) for the entire coverage area.* This requires only one data transmitter, usually located close to the centre of the coverage area. Multiple transmitters may be used if their extent is limited to less than the entire network coverage area, however they would each broadcast the same correction information. An example of this option is illustrated in Figure 4.11, where a broadcast tower represents the data transmitter.
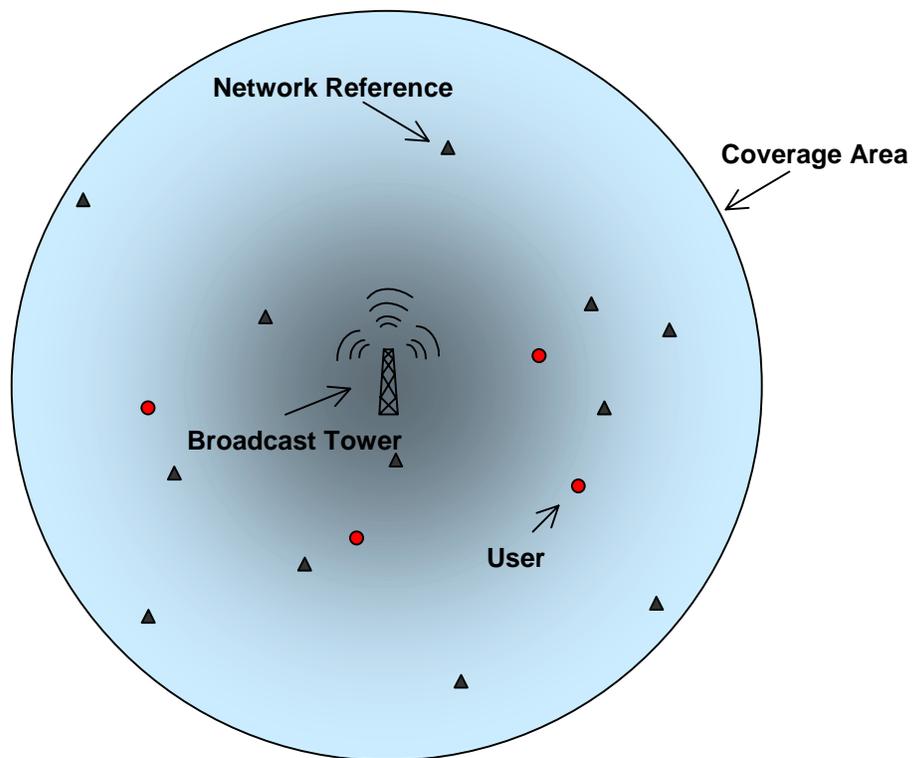
Figure 4.11: Example of Single Correction Data Transmitter Configuration

b) *Transmit multiple sets of corrections from multiple data transmitters sparsely situated throughout the network coverage area, which may or may not be co-located with physical reference stations.* Each transmitter communicates correction information for a sub-area of the entire coverage region, with some overlap between adjacent regions as shown in Figure 4.12. Essentially, a user would receive correction data from the closest available transmitter.
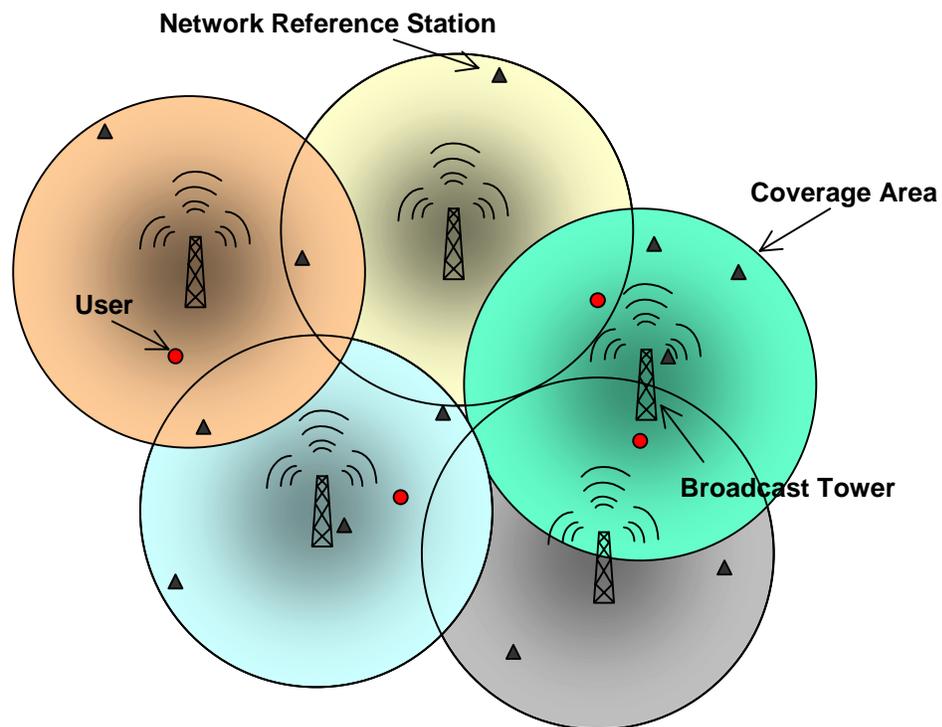
Figure 4.12: Example of Multiple Correction Data Transmitters Configuration

c) *Transmit both a set of corrections for the entire network (as in (a) above) and multiple sub-sets of corrections (as in (b) above).* The idea in this

method would be to exploit the ability of the smaller regions to model more local changes (i.e. higher frequency errors), while the larger set would model the more global, slowly changing (low frequency errors). This is similar to the *fast* and *slow* correction concept proposed by WAAS (FAA, 1999). See Figure 4.13 for an illustration of the combined approach.
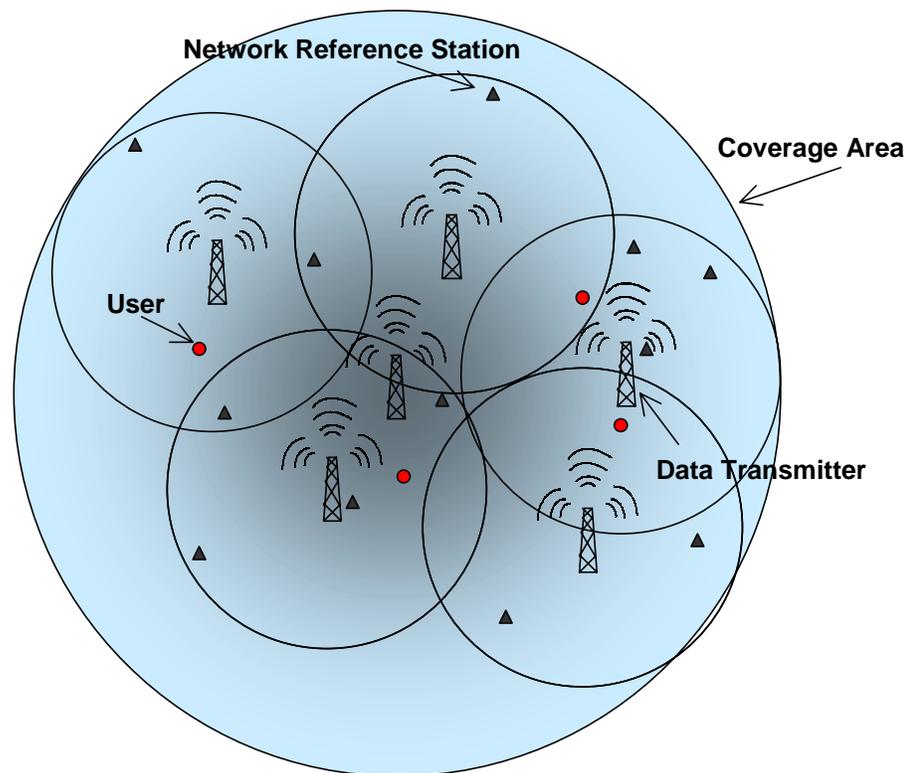
Figure 4.13: Example Combined Correction Data Transmitter Configuration

The advantage of having sub-sets is the ability to model more localized error sources. However, this depends on inter-station distances. If the network geometry and distances

are such that a denser grid or a more complex function model do not extract any additional information, then the use of densified broadcast sub-sets cannot be justified.

4.3.2   Grid-Based Transmission Message

For the grid-based parameterization there are several options available for correction transmission.   To illustrate these methods, correction transmission option (a), as described above, will be assumed.   Thus, in this case, a single set of corrections for the entire network is generated.   The most efficient method for this is to provide two types of messages, one for the grid definition and one containing the correction values on a satellite basis.   In such a manner, the grid definition does not have to be repeated at every epoch, or for every satellite since it will be defined based on the ground network boundaries.

For the case of a rectangular grid structure, the following basic, yet critical information is required to define the grid:

- latitude of one corner point (i.e. the southwest corner)
- longitude of the same corner point
- grid spacing in latitudinal direction
- grid spacing in longitudinal direction
- number of grid points in latitudinal direction
- number of grid points in longitudinal direction

In the tests conducted for the SSN, the grid spacing in latitude and longitude were the same, but the number of grid points in each direction were not necessarily equal. Table 4.10 summarizes this grid definition data and the number of bits required for transmission. These values are based on a certain range indicated in the final column, which also require the application of a scale factor. The total number of bits for this critical grid defining data is 79. Depending on the grid point resolution, size and compression methodologies, the corresponding number of bits will be altered.

Table 4.10: Critical Information for Grid Definition Message

Transmitted Information	# of Bits	Scale Factor	Units	Effective Range
Latitude of corner point	18	0.0005°	degrees	$\pm$ 90.000°
Longitude of corner point	19	0.0005°	degrees	$\pm$ 180.000°
Grid spacing in latitudinal direction	13	0.001°	degrees	0 - 5.000° (up to 8.191°)
Grid spacing in longitudinal direction	13	0.001°	degrees	0 - 5.000° (up to 8.191°)
Number of grid points in latitudinal direction	8	1	unitless	0 - 255
Number of grid points in longitudinal direction	8	1	unitless	0 - 255

The second message type required for transmission is the actual grid node corrections. Previously, it was found that an update rate of 1 Hz was not necessary. In fact, update rates of up to 60 seconds for the 0.5° grid resulted in virtually the same results as the single baseline approach. Therefore, it is possible to send one correction message each second per satellite. For instance if there are a maximum of 12 satellites with valid

corrections for a regional network area, corrections can be transmitted over 12 seconds while the grid definition message can be transmitted at intermittent periods (every 10-15 seconds), which will reduce the data transmission load.

The required information to be transmitted to the user for the grid-based corrections are:

- satellite ID
- correction type
- correction value
- correction quality indicator

The satellite ID is required because of the per satellite basis of the corrections. The correction type would involve information such as whether the correction is a combined correction consisting of atmospheric and orbital errors as in the case thus far, or separated corrections, which were also briefly mentioned. Then the actual correction value accompanied with a correction quality indicator would be sent. This quality indicator has not been addressed in this thesis, however it is equally as important as the corrections themselves. Ideally, this value would provide the user with some information regarding the quality of the corrections, based on the covariance function in the generation process and the corresponding parameterization (see Appendix B). Table 4.11 summarizes the grid-based correction message information and associated transmission parameters, in a manner similar to the previous table.

Table 4.11: Critical Information for Grid-Based Corrections

Transmitted Information	# of Bits	Scale Factor	Units	Effective Range
Satellite ID	6	1	unitless	1 - 32 (up to 63)
Correction Type	4	1	unitless	0 – 15
Correction Value	12	0.001	cycles	± 4.000 (up to 4.095)
Correction Quality Indicator	4	1	unitless	0 - 15

It is also important to note that other standard information must be sent such as message identification number, epoch, and parity bits. However, this information will not vary according to the parameterization scheme and was therefore not included in this discussion.

4.3.3   Function-Based Transmission Message

In the function-based parameterizations, only one type of message has to be transmitted since a single indicator can define the function. Therefore, the required information is:

- satellite ID
- function type
- correction type
- first coefficient
- second coefficient … up to n coefficients
- overall function quality indicator

It is evident that the total number of bits required for this message is heavily dependent on the complexity of the function parameterization. Also, the resolution of the coefficients must be, in general, higher than the grid node values (see Table 4.12) because the approximate horizontal coordinates of the user stations will scale the function coefficients in order to obtain the correction values.

In the case of the function-based parameterization, the accuracy of the approximate horizontal coordinates obtained from the user receiver will improve since SA was turned down to zero (see Section 2.4). Initial positions will be on the order of tens of metres as opposed to ~150 metres with SA on. This will produce initial function evaluations that are 'closer' to the direct evaluation for the user receiver.

Table 4.12: Critical Information for Function-Based Corrections

Transmitted Information	# of Bits	Scale Factor	Units	Effective Range
Satellite ID	6	1	unitless	1 – 32 (up to 63)
Correction Type	4	1	unitless	0 - 15
Function Type	4	1	unitless	0 - 15
Coefficient Value	20×6	0.00001	unitless	$\pm$ 8.00000 (up to 10.48575)
Overall Function Quality Indicator	4	1	unitless	0 - 15

4.4   Remarks

As it was shown in Chapter 3, the function-based parameterizations generally do not perform as well as the denser grid-based schemes. These function models do not represent the local anomalies in the area as well as a grid model does, acting as a low-pass filter and consequently resulting in a 'smoother' correction surface. However, if a user can afford to have positions at the level of accuracy available with the function models, there are advantages in terms of data transmission bandwidth. This is a direct result of the need to only transmit the function coefficients instead of both a grid definition type message and a number of grid node values. It can be argued that as the complexity of the function model increases, the number of coefficients and their corresponding resolution also increases, which may theoretically reach and exceed the number of bits for a grid-based scheme. An additional concern is that users are more likely to comprehend the implementation of a grid-based scheme due to its prevalence in representing other surfaces (i.e. ionospheric grid, digital terrain model, geoid model, etc.), rather than a polynomial function. With proper documentation and reducing the computational load on the user receiver, this latter problem can easily be overcome.

Chapter 5

CORRELATION AND SPECTRAL ANALYSIS OF DISTANCE-DEPENDENT ERRORS

In previous chapters the spatial (Chapter 3) and temporal (Chapter 4) characteristics of the carrier phase corrections generated per satellite over a regional network were investigated. The majority of the analysis was conducted in the space, time, correction and position domains, which provided insight to the spatial and temporal resolution of the corrections required to achieve highly accurate results.

In this chapter, additional analysis is performed on the data sets in the frequency domain. Here, results on the spatial and temporal characteristics of the correction fields are presented through spectral analysis techniques. The motivation behind this analysis is twofold. Firstly, it is often useful to study the behaviour of a signal using spectral analysis in order to extract additional information on small-scale variations (in space and time) which may not be readily available with other techniques. Secondly, spectral analysis techniques offer an alternative process to those already investigated. This is useful for verifying results found previously (in Chapters 3 and 4) in addition to

supplementing existing analysis tools which provide an opportunity to view the data from a different perspective.

As it was mentioned previously, it is difficult to completely isolate the spatial and temporal aspects for analysis. A satellite's time series of corrections reflects a number of factors including the satellite's position with respect to each of the network reference stations and user stations (i.e. satellite azimuth and elevation profile). Indeed each generated satellite-based correction field is unique and varies from satellite to receiver line-of-sight pair. An attempt has been made in this chapter to build the discussion and analysis about a representative set of real data sets. Specifically, the selection of sample fields covers various periods throughout the day and night for a number of satellites over different elevation profiles. This chapter continues the analysis conducted thus far, by using the results obtained from previous chapters in order to identify distinguishable samples of satellite series (in time and space) for evaluating the generated correction fields.

5.1   Temporal Correlation Analysis

One of the key concerns for providers and users of regional networks is the provision of timely corrections which adhere to transmission bandwidth limitations. A detailed analysis on the effects of various update rates for possible parameterization schemes was evaluated. The results showed that depending on the level of accuracy required, it is

possible to update or refresh the correction parameters from several tens of seconds to a few minutes (see discussion in Section 4.2). These values were obtained and analyzed in the correction and more intuitively in the position domains.

An additional method, which is useful for establishing the dependence of the correction values over time is the correlation function, defined as follows:

$$\phi(\tau) = \int\limits_{-\infty}^{\infty} g(\tau)h(t+\tau)d\tau \tag{5.1}$$

where, $g(\tau)$ and $h(\tau)$ are continuous data functions, whose 'correlation' is investigated by comparing both directly superposed and by shifting one by a lag of t (Press et al., 1988). Large correlation values indicate that the two data functions are close shifted copies of one another.

In practice, we do not generally have access to true continuous data functions, rather a discrete representation of the data is provided. In such a case the discrete correlation of two data functions (also termed lagged products) with period N is computed as follows,

$$\phi(kT) = \sum_{i=0}^{N-1} g(iT)\, h[(i+k)T] \tag{5.2}$$

where $\phi(kT)$, $g(kT)$, and $h(kT)$ are periodic functions. In the special case where both data functions are the same, we can compute the autocorrelation. Since the discrete autocorrelation of a sampled data function, $g(kT)$, is just the discrete correlation of the function with itself, it can easily be represented in the form of eq. (5.3) as:

$$R(kT) = \sum_{k=0}^{N-1} g(iT)g[(i+k)T] \tag{5.3}$$

A number of useful properties are important to take into consideration when interpreting an autocorrelation function (Borre and Strang, 1997). First an autocorrelation function takes its maximum value at zero shift (i.e. $R(0)$). Also, such functions are always symmetric with respect to the positive and negative lags (i.e. $R(kT)$ is an even function where $R(kT) = R(-kT)$). Lastly, the autocorrelation has a periodic component if the original data function has one. These properties will be observed in the results included in Section 5.1.3.
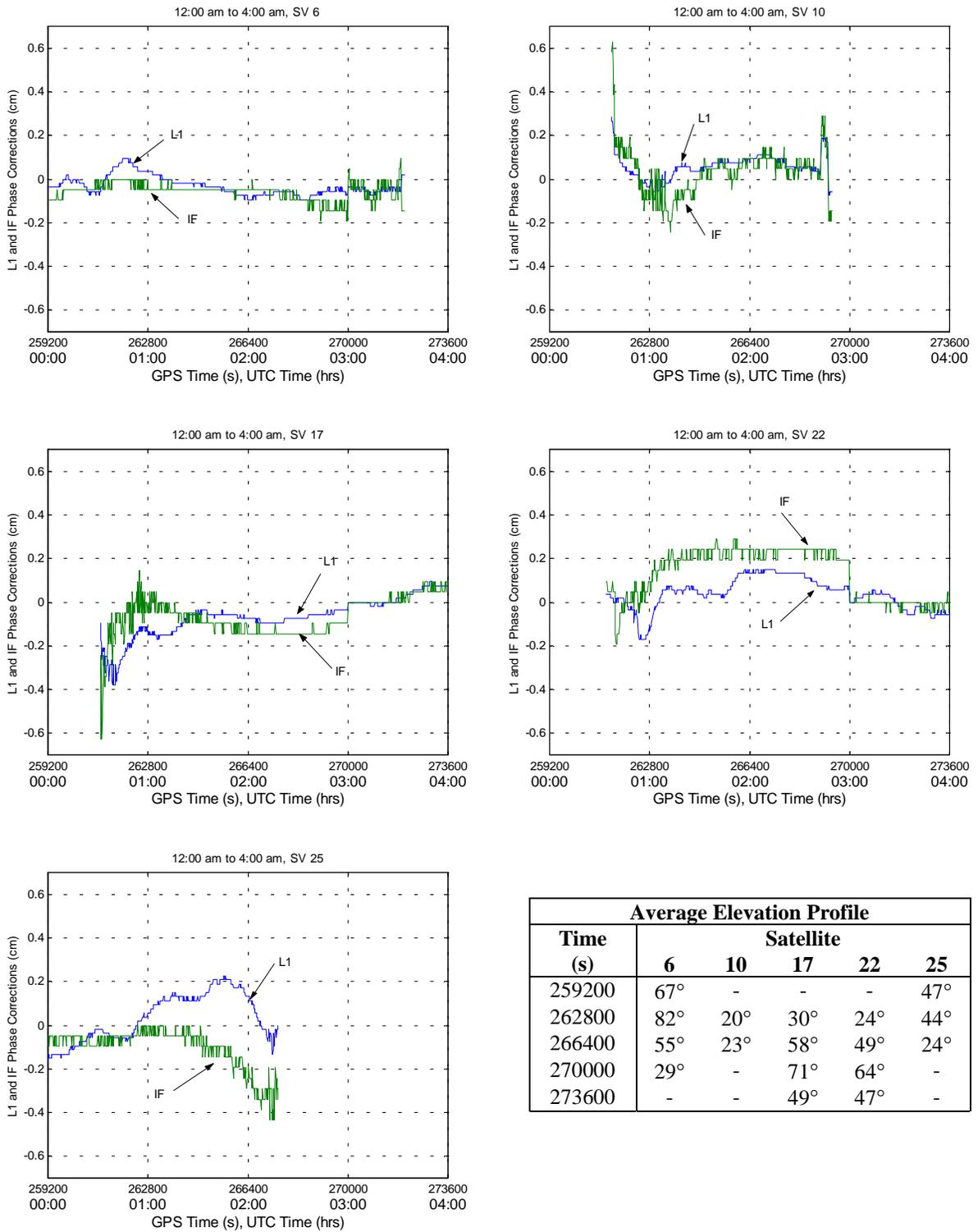
5.1.1   Representative Data Sets

To compute the autocorrelation functions of the correlated errors over time, representative samples of correction time series were evaluated. The corrections were generated on a per satellite basis for two main periods during the 24-hour data set. These periods were identified and described in Section 4.1.2 and are (a) 12:00 am to 4:00 am

and (b) 11:00 am to 3:00 pm. Spectral analysis of ten satellites in total, five for each time period, was performed. The time series of corrections is shown in Figures 5.1 and 5.2, corresponding to time periods (a) and (b), respectively. Each figure contains the combined L1 phase corrections and IF phase corrections.

It should also be noted that all corrections shown in Figures 5.1 and 5.2 are *user* values only. That is, a reference station from the network was not chosen as the monitor station in order to formulate combined corrections, as was done in part of the spatial analysis in Chapter 3. By including a reference station, an additional bias to the magnitude of the correction values would be observed. However, the overall 'shape' of the correction series would not significantly change. The magnitude of the bias would also change depending on the monitor station chosen (see *data encapsulation* effect discussion in Raquet, 1998). Therefore, the analysis for the user station presented here is equally applicable to any reference station within the network.

The appearance of the correction time series observed in Figures 5.1 and 5.2 can be described as discrete, or 'step-like'. Firstly, it should be noted that the corrections were generated for a 15 seconds decimated data set, with all successive values in between two data nodes linearly approximated. Also, the resolution of the corrections used for the plots was limited to 0.001 cycles, resulting in the discrete 'step-like' appearance, since in general the variations in corrections over several seconds is higher than the resolution of the plots.

Figure 5.1: L1 and IF Phase Corrections Over 12:00 am to 4:00 am Period
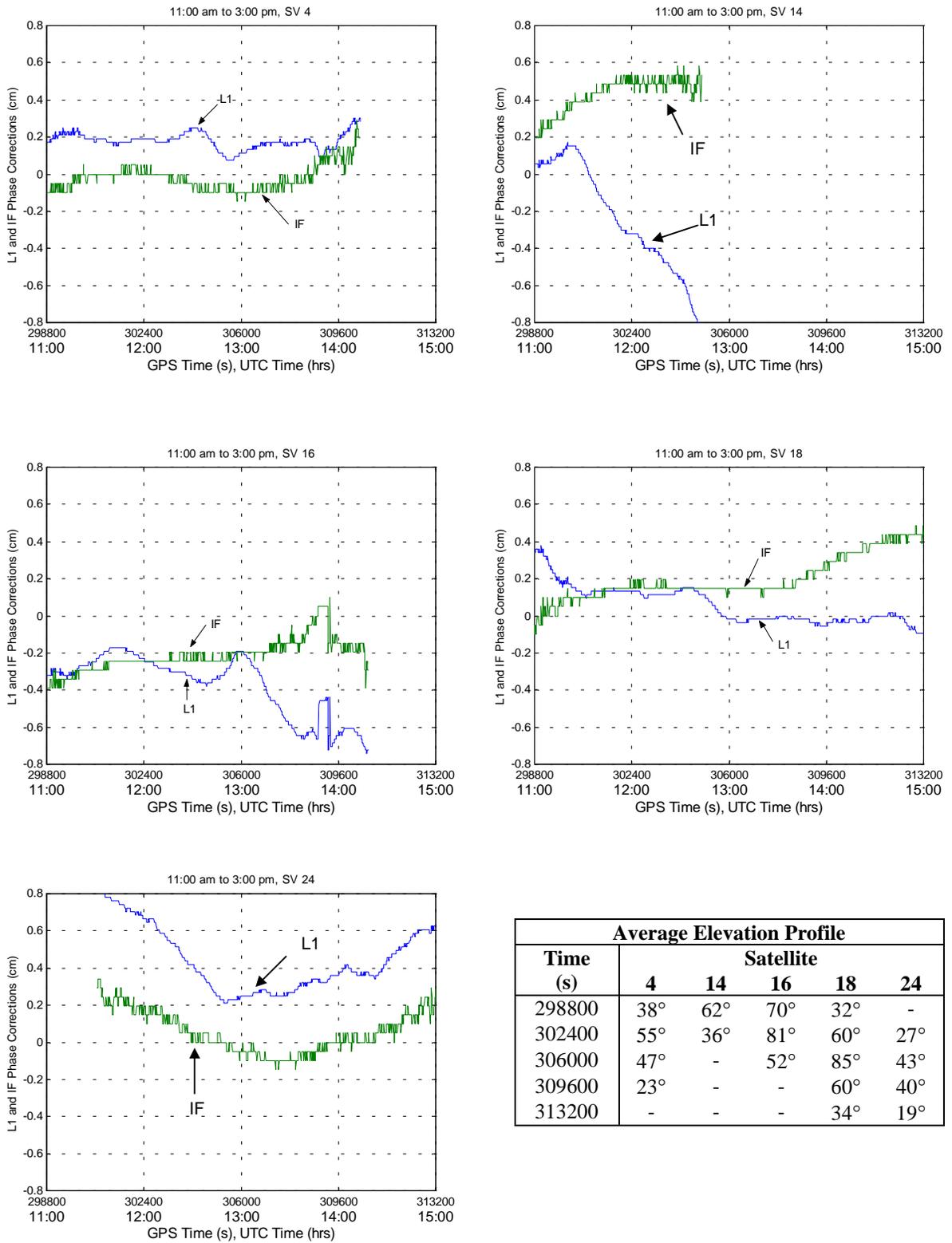
Figure 5.2: L1 and IF Phase Corrections Over 11:00 am to 3:00 pm Period

Average Elevation Profile					
Time	**Satellite**				
(s)	**4**	**14**	**16**	**18**	**24**
298800	38°	62°	70°	32°	-
302400	55°	36°	81°	60°	27°
306000	47°	-	52°	85°	43°
309600	23°	-	-	60°	40°
313200	-	-	-	34°	19°

In general, the range in combined L1 phase and IF corrections appear to be similar for the 12:00 am to 4:00 am period. This is expected, as the ionosphere is relatively quiet during this period, contributing less to the combined corrections than during more active periods. For instance, a comparison of the L1 and IF time series during the 11:00 am to 3:00 pm period reveal a larger difference between the two sets, due to the more active ionosphere. Also, the IF data is, on average, closer to zero mean than the combined L1 correction time series. This indicates that some residual systematic effect in the corrections due to the ionosphere is removed when first order effects are modelled.

The average elevation of the satellites over the network area are included in tables in the bottom right corner of Figures 5.1 and 5.2. Although in general the corrections exhibit relatively random behaviour, during the 11:00 am to 3:00 pm period, lower elevation satellites do contain higher correction values compensating for the predominant atmospheric effects.

5.1.2   Link to Parameterization Parameters

The random nature of the corrections observed in Figures 5.1 and 5.2, is also directly exhibited in the behaviour of the parameterization parameters. For example, Table 5.1 contains the values of the coefficients of a six-parameter polynomial fit (eq. 3.10c) over a one-hour period. The formulated coefficient values were recorded at five minute intervals. The randomness of the coefficients reflects the observed behaviour of the

corrections. It is also worth mentioning the relatively smaller contribution of parameters

c, d, and e, compared to the other coefficients (even excluding the intercept value f).

The smallest contribution is seen in coefficient e, which is the y_i^2 term, where y_i

corresponds to the latitudinal direction. From this sample, it is apparent that all six

coefficients are not necessary to model the correction surface. For data transmission

purposes, five coefficients sufficiently represent the correction field.

Table 5.1: Polynomial Coefficients at Various Times

GPS Time (s)	Coefficients					
	a	b	c	d	e	f
259200	-0.590	0.048	-0.001	0.005	-0.001	17.013
259500	-0.289	-0.105	0.002	0.002	0.001	9.123
259800	-0.200	-0.060	0.001	0.002	0.001	5.903
260100	0.120	-0.231	0.004	-0.002	0.001	-1.784
260400	0.023	-0.173	0.003	-0.001	0.001	1.054
260700	-0.399	-0.188	0.003	0.003	0.000	13.421
261000	-0.410	-0.049	0.001	0.003	-0.000	12.717
261300	-0.680	0.025	0.001	0.006	-0.002	19.880
261600	-0.287	-0.000	0.001	0.002	-0.003	9.049
261900	0.213	-0.039	0.001	-0.002	-0.001	-4.926
262200	0.263	-0.096	0.002	-0.003	-0.001	-5.826
262500	-0.154	-0.171	0.003	0.001	0.000	6.757
262800	-0.269	-0.192	0.003	0.002	0.000	10.285

5.1.3 Autocorrelation Functions

The discrete autocorrelation functions for each satellite time series described in the

previous section were evaluated. It should be noted that a scaled version of eq. (5.3) was

used where the function values were scaled by $1/M$, M being the number of samples.

This provides consistent correlation functions for evaluation, as the time lags range from 0 to 15000 seconds (half of the total sample length). The satellite visibilities for both periods are summarized in Table 5.2.

Samples of the autocorrelation functions computed for the satellites for both combined L1 and IF corrections are shown in Figures 5.3 and 5.4 for the two time periods, respectively.

Table 5.2: Satellite Visibility for Two Time Periods

12:00 am to 4:00 am Period		11:00 am to 3:00 pm Period	
SV 6	3.5 hours	SV 4	3 hours
SV 10	2 hours	SV 14	~ 2 hours
SV 17	3.5 hours	SV 16	3.3 hours
SV 22	3.5 hours	SV 18	4 hours
SV 25	2.3 hours	SV 24	3.5 hours

Some of the autocorrelation functions show a strong correlation over a larger time, and later the function decreases rapidly. In a few cases, as evidenced by Figures 5.3 and 5.4, the correlation functions closely resemble an exponential trend. Here, the functions decrease strongly for small time differences, as shown in satellite 6, 10 and 17 in Figure 5.3b. This also indicates a rougher signal with higher frequency variations, which are generall1y more difficult to model than smoother, slowly changing behaviour.
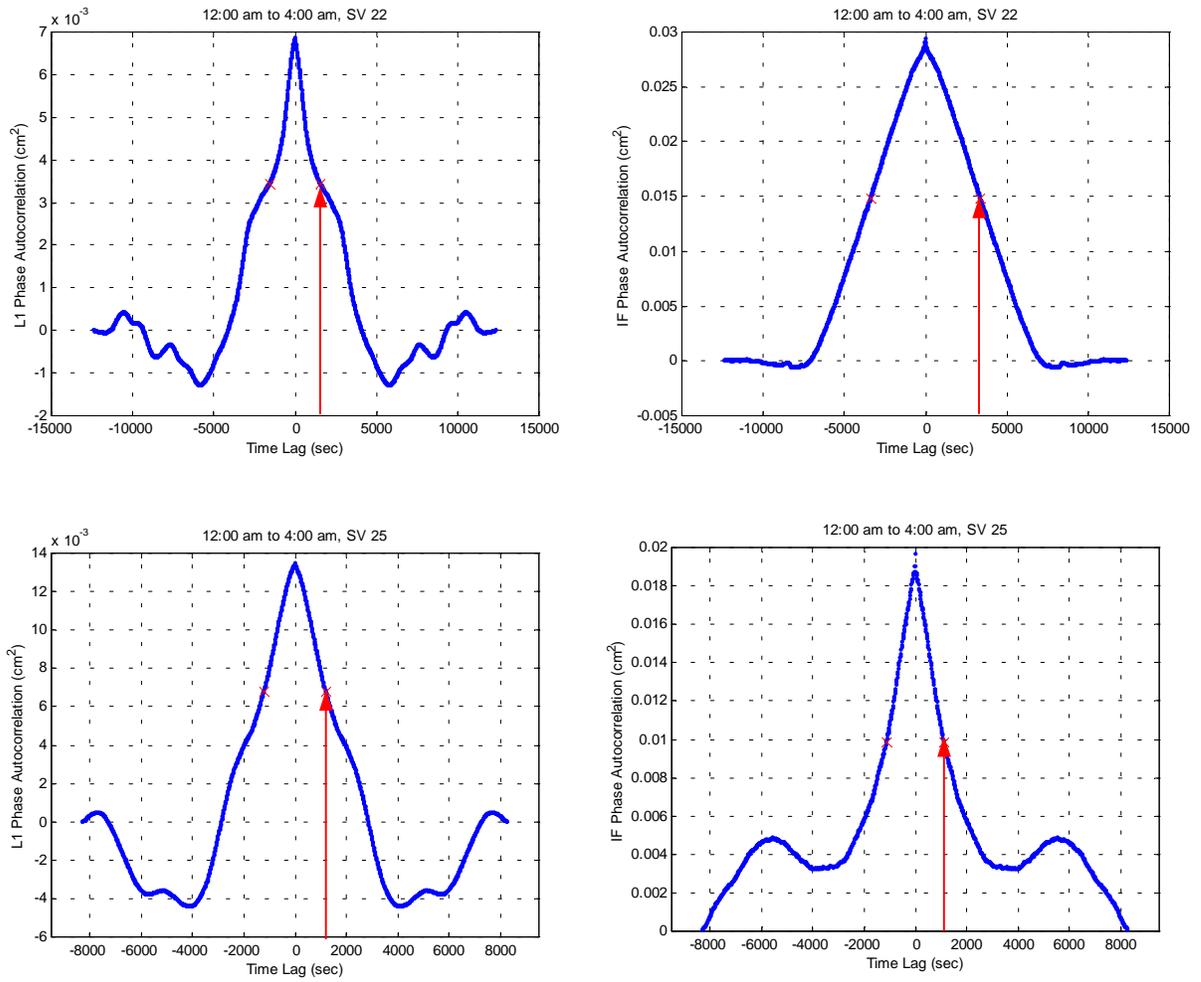
Figure 5.3a: Autocorrelation Functions for L1 and IF Corrections for the
12:00 am to 4:00 am Period
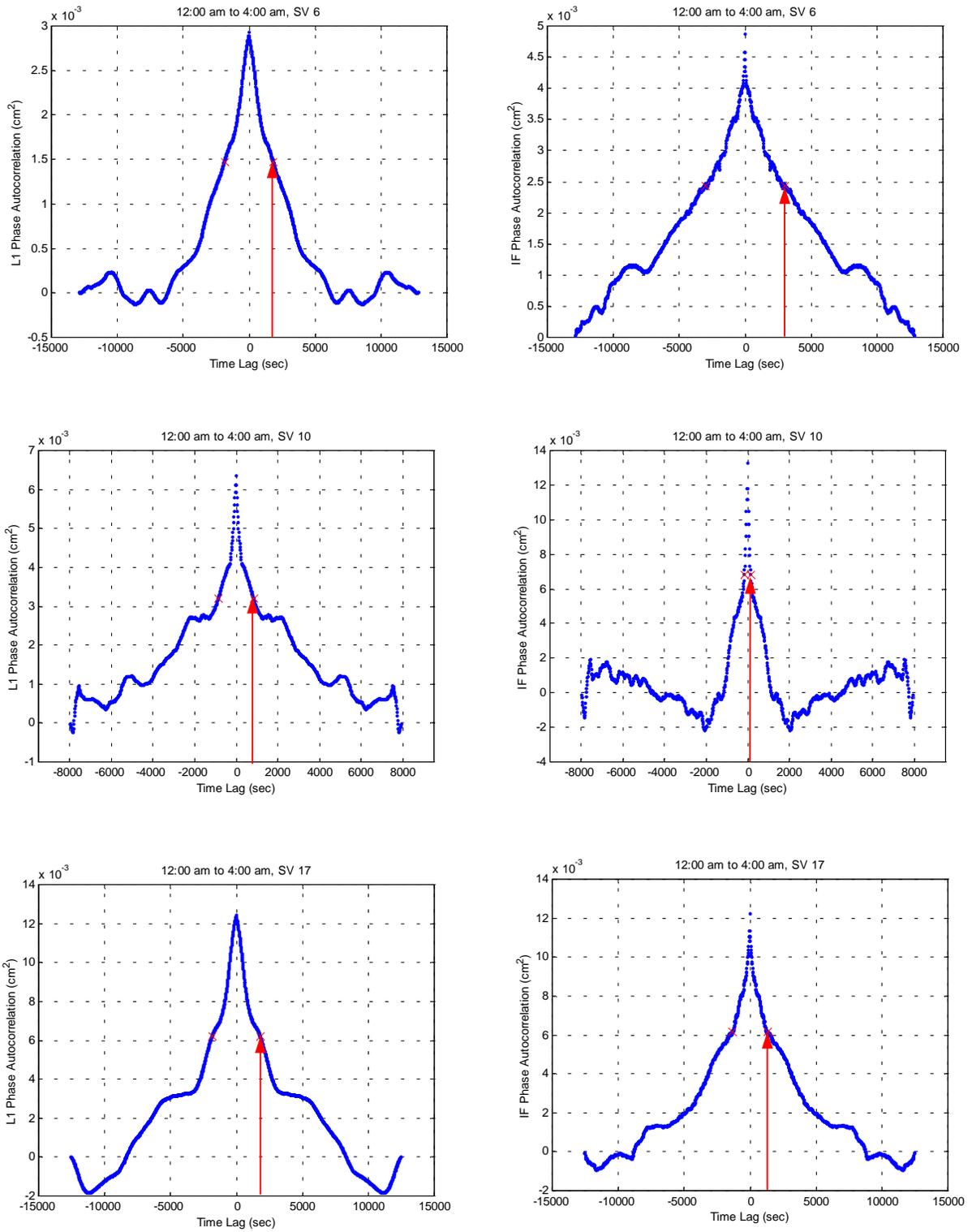
Figure 5.3b: More Autocorrelation Functions for L1 and IF Corrections for the 12:00 am to 4:00 am Period
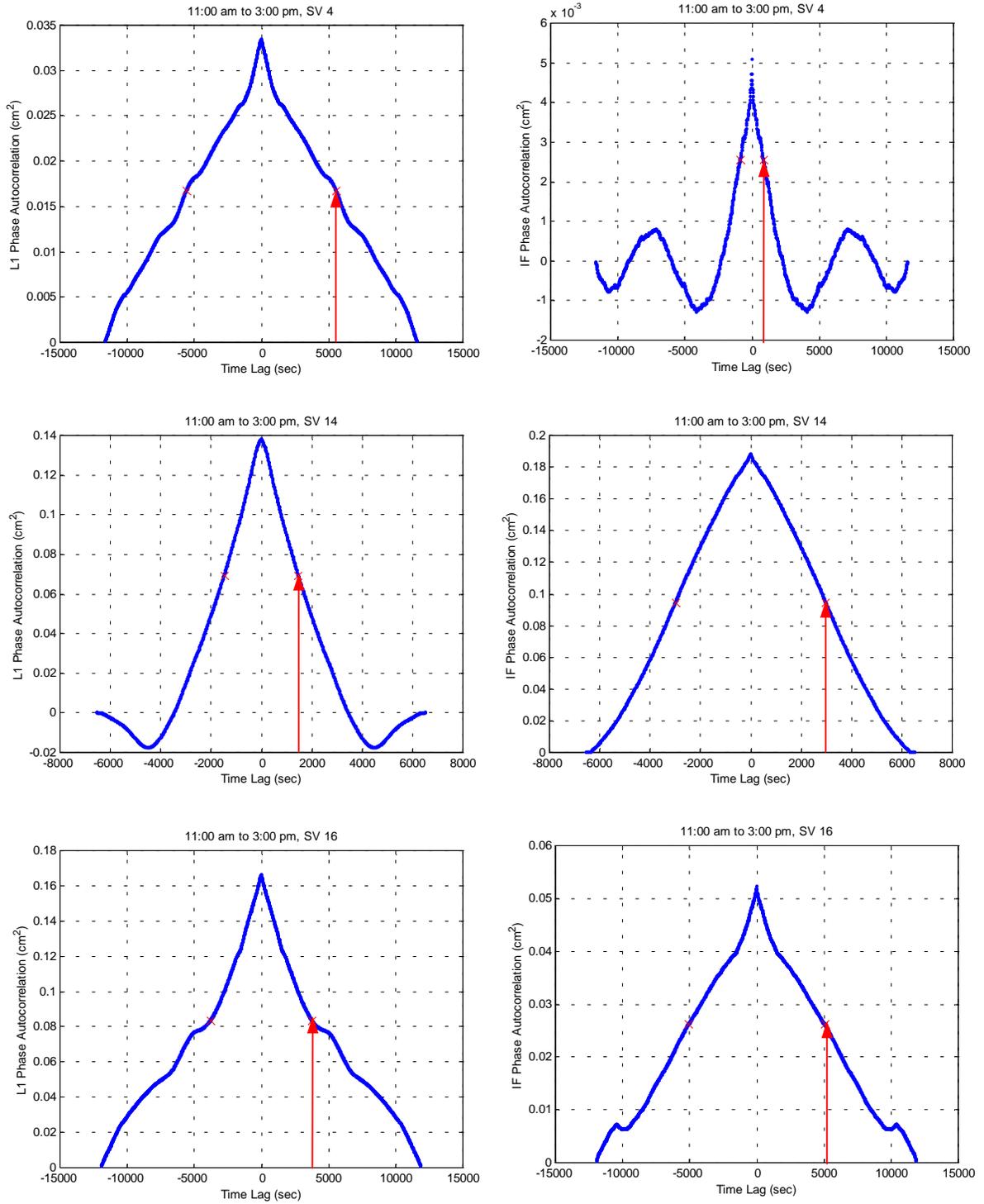
Figure 5.4a: Autocorrelation Functions for L1 and IF Corrections for the
11:00 am to 3:00 pm Period
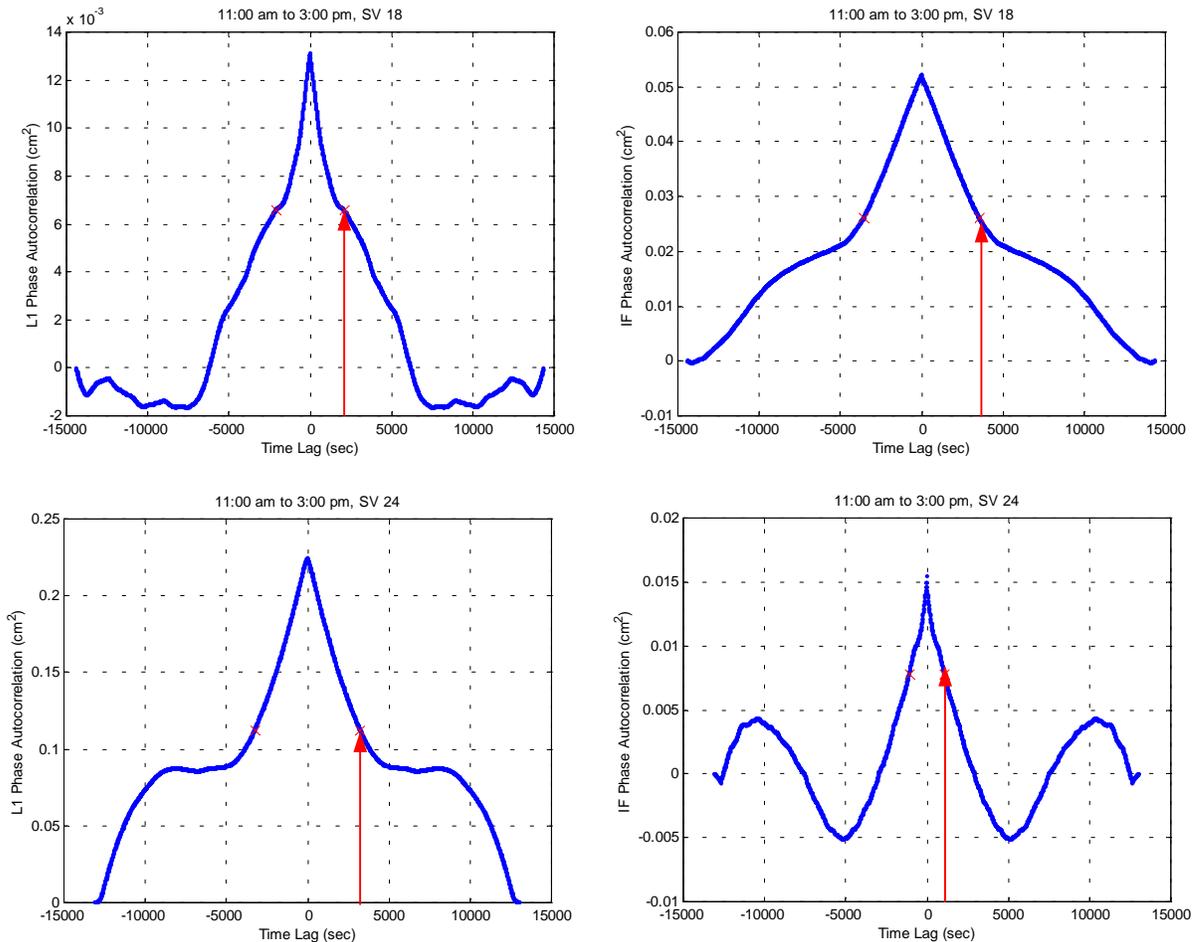
Figure 5.4b: More Autocorrelation Functions for L1 and IF Corrections for the
11:00 am to 3:00 pm Period

Of interest for our analysis is the corresponding mean square value (the value at zero time

lag) where the function takes on its maximum, and the correlation length. The correlation

length is computed from the time lag corresponding to the point where the function takes

on half of the mean square value. Long correlation lengths indicate a much smoother,

more easily modelled signal. The correlation lengths for each of the satellite correction

series is indicated by an arrow on Figures 5.3 and 5.4 (a and b).

The mean square values of the same satellites, associated with the combined and ionospheric-free correlation functions, during the 12:00 am to 4:00 am period, show that the IF values are greater in most cases. In fact, for the one case where the combined corrections mean square value is bigger, it is only by a slight margin (satellite 17). This means that the ionospheric-free signal is stronger than the combined L1 signal. A possible explanation for this is the randomness of the effects of the residual errors observed thus far, which may cancel to some degree. Also the decorrelating behaviour of the most dominant effect in the combined corrections, namely the ionosphere may be responsible for reducing the strength of the signal according to its activity.

The corresponding correlation lengths for the same case range from a minimum of 840 seconds to a maximum of 1860 seconds for the combined corrections and a low of 150 seconds to a high of 3375 seconds for the IF case. Definitely, the range in values for the latter case is much larger than the combined corrections, with considerably large correlation lengths in excess of 3000 seconds seen for the IF case. Recall that long correlation lengths indicate 'smoother' signals. Intuitively, this makes sense since from previous studies, the ionospheric effects have shown to be the cause of the majority of the variations or 'roughness' of the combined corrections signal. Tables 5.3 and 5.4 summarize the autocorrelation function parameters for L1 and IF respectively, discussed for the 12:00 am to 4:00 am period (which can also be obtained from the autocorrelation illustrations in Figure 5.3).

Table 5.3: Autocorrelation Function Parameters for L1 Phase Corrections
(Time Period - 12:00 am to 4:00 am)

Satellite PRN	Mean Square Value (cm^2)	Correlation Length (sec)
6	0.0029	1860
10	0.0063	840
17	0.0124	1815
22	0.0069	1560
25	0.0134	1200

Table 5.4: Autocorrelation Function Parameters for IF Phase Corrections
(Time Period - 12:00 am to 4:00 am)

Satellite PRN	Mean Square Value (cm^2)	Correlation Length (sec)
6	0.0049	3045
10	0.0132	150
17	0.0122	1365
22	0.0294	3375
25	0.0197	1125

Investigations conducted for the day side period (11:00 am to 3:00 pm) should provide a more amplified explanation of the results. During this second period, the mean square values for the combined L1 signal are greater for three satellite samples than the corresponding ionospheric-free cases. The more active ionosphere is also likely to be the cause of this. For the satellites where the IF signal is stronger (SV 14 and SV 18), it may be that residual or 'unmodelled' ionospheric effects remain or perhaps the remaining tropospheric and orbit effects are significant for these satellites. Looking back at the satellite time series for these two cases (Figure 5.2) reveals that satellites 14 and 18 exhibit divergence of the L1 and IF series near the end of their visibility. It is possible that unmodelled systematic behaviour is the cause of these strong autocorrelation

functions. However, in general the results support the fact that the combined signal is much stronger than the ionospheric-free signal during this period of a relatively active ionosphere. Tables 5.5 and 5.6 summarize the autocorrelation function parameters shown in Figure 5.4 (a and b). Also included in these tables are the correlation lengths. It is evident from the tables and Figure 5.4 that the lengths vary from 1470 seconds to 5565 seconds for the combined L1 signals to a minimum of 855 seconds to 5070 seconds for the ionospheric-free signals.

Table 5.5: Autocorrelation Function Parameters for L1 Phase Corrections
(Time Period - 11:00 am to 3:00 pm)

Satellite PRN	Mean Square Value (cm^2)	Correlation Length (sec)
4	0.0335	5565
14	0.1383	1470
16	0.1664	3765
18	0.0131	2100
24	0.2243	3240

Table 5.6: Autocorrelation Function Parameters for IF Phase Corrections
(Time Period - 11:00 am to 3:00 pm)

Satellite PRN	Mean Square Value (cm^2)	Correlation Length (sec)
4	0.0051	855
14	0.1885	2985
16	0.0522	5070
18	0.0523	3540
24	0.0154	1080

Unfortunately, a direct comparison between time periods cannot be made due to the different satellites involved. However, of relevance for the temporal analysis is the

overall large correlation lengths on the order of several thousands of seconds, which translates to tens of minutes. This is encouraging as it reflects the relatively slow variability in parameterized corrections computed for the user location. This is consistent with the temporal analysis presented in the previous chapter, where update rates, based on the position domain results, of up to a few minutes were found to be suitable without jeopardizing the final user position accuracy.

Another point worth mentioning are the differences between the corrections for the combined error effects versus the ionospheric-free values. In many cases the spectral parameters for the same satellites corresponding to combined (three error sources) and separated (two error sources) signals were dissimilar. These differences further support the concept of separating the error sources (at least the ionospheric effects), in order to formulate more consistent information for the user. The combined corrections are often masked by the predominant error source (ionosphere) as well as the remaining errors may combine in a random way, leading to mixed results, as evidenced in some of the satellites' autocorrelation functions.

5.2   Spectral Analysis

In this section, examining the spectral properties of the corrections over the entire coverage area continues the spatial analysis conducted thus far. More specifically, the power spectral density (PSD) functions are generated for samples of correction fields. In

the interests of examining the spatial aspects of the correlated error effects, the satellite-based corrections computed for single epochs in time are evaluated. The satellites chosen for study were visible at midday at approximately 1:00 pm and 3:00 am. The relative level of atmospheric activity varies for the two periods as discussed in the temporal analysis of the previous chapter (Section 4.1.2) and the correlation analysis.

The power spectral density function and the autocorrelation function (Section 5.1.3) are defined as a Fourier transform pair:

$$\phi(\tau) = \int_{-\infty}^{\infty} \Phi(f) e^{j 2\pi f \tau} d\tau \iff \Phi(f) = \int_{-\infty}^{\infty} \phi(\tau) e^{-j 2\pi f \tau} d\tau \tag{5.4}$$

where $\phi(\tau)$ is the autocorrelation function and $\Phi(f)$ is referred to as the power spectral density, which is a function of frequency f (see Brigham, 1988 for a complete explanation). As in the correlation analysis, the empirically derived signal under investigation is only known over a finite interval. Thus, the PSD must be estimated based on the finite duration of the data.

An alternative method for obtaining the PSD, other than the autocorrelation function, is via the Fourier transform over a finite interval. As the spatial dimensionality of the problem has been defined as 2D with horizontal components in the latitudinal (y) and

longitudinal (x) directions, it is necessary to compute the two-dimensional discrete Fourier transform (DFT), as per Dudgeon and Mersereau (1984) and Sideris (1984):

$$H\left(mf_x, nf_y\right) = \frac{T_x T_y}{MN} \sum_{k=0}^{M-1} \sum_{\ell=0}^{N-1} h(k\Delta x, \ell\Delta y) e^{-j2\pi\left(\frac{mk}{M} + \frac{n\ell}{N}\right)}$$

(5.5)

where the grid spacings in the x and y directions are Δx and Δy, respectively and are given by the periods T_x and T_y for $M \cdot N$ discrete values, as follows:

$$\Delta x = \frac{T_x}{M}, \quad \Delta y = \frac{T_y}{N}$$

(5.6)

such that,

$$x = k\Delta x, \qquad k = 0,\ 1,\ 2,\ldots, M-1$$

$$y = \ell\Delta y, \qquad \ell = 0,\ 1,\ 2,\ldots, N-1$$

(5.7)

Using eq. (5.5) the corresponding two-dimensional PSD is easily derived from the squared magnitude of the Fourier transform, as shown below:

$$P_h = \left| H\left(mf_x, nf_y\right) \right|^2$$

(5.8)

The evaluation of the PSD function through eq. (5.8) permits the detection of dominant frequencies contributing to the data. In this case, the data is a discrete number of grid node values containing formulated corrections for the correlated error sources based on a regional network. The corresponding grid spacings, Δx and Δy, vary depending on the resolution of the parameterization scheme (i.e. 0.5°, 1.0°, 1.5°, or 2.0°). This resolution in space is also related to the extent of frequencies over which the PSD is defined, as will be observed in the results that follow.

It is important to make certain that the proper 'scaling' of the correlation and PSD functions is applied when *ready made* algorithms for computing the Fourier transform are used. Some algorithms may include normalization factors, affecting the interpretation of the output (see Sideris, 1984 for a discussion on computer algorithms). In this case, the Fourier transform algorithms available in *MATLAB*® (Math Works, 1998) were implemented to compute the associated PSD. The relationship between the computer algorithms and the values expressed in eq. (5.5) is as follows:

$$H_{ml}(mf_x, nf_y) = \frac{MN}{T_x T_y} H(mf_x, nf_y) \tag{5.9}$$

such that the *MATLAB*® generated DFT $H_{ml}(mf_x, nf_y)$ must be appropriately scaled according to the duration and period of the data signal in both horizontal component directions.

5.2.1   Power Spectral Density Functions

The first case investigated was the correction field of a very high elevation satellite (approximately 85°, also used as the base satellite for the double difference computations at this epoch) at 1:00 pm. Figure 5.5 provides two different views of the same PSD function for this scenario. Here the correction field was computed using the grid-based parameterization scheme with the highest resolution, namely $\Delta x = \Delta y = 0.5°$ (see Section 3.4.3). The higher values near the origin of the PSD shows that the lower frequencies are dominant contributing the most to the data.   These lower frequency components correspond to long wavelength errors, which typically change more slowly and are thus easily modelled. More specifically, the frequency range along the *x*-direction are less than 0.2 cycles/degree longitude, whereas along the *y*-direction the frequencies nearly double, reaching 0.4 cycles/degree latitude. This indicates relatively higher variations in the north-south component than the east-west. A possible explanation for this may be the occurrence of higher gradients in the correlated errors translating to the north-south direction at the grid surface. Overall the frequencies are concentrated near the origin approaching zero at higher frequencies and appear to be uncontaminated by aliasing effects resulting from poorly sampled data.

Results for the same correction field described above were generated for a sparser grid resolution where $\Delta x = \Delta y = 1.0°$. The formulated PSD function is shown in Figure 5.6. Note that the higher spacing also affects the realizable frequency range, limiting our scope to $\pm$ 0.5 cycles/degree. Since the majority of the frequency components are

contained within this range, it is evident that the low frequency information is still maintained with the sparser grid parameterization. However any higher frequency information seen in Figure 5.5 beyond $\pm$ 0.5 cycles/degree is lost. This effect is amplified when the grid resolution is further decreased to $\Delta x = \Delta y = 1.5^{\circ}$. In this case the corresponding recoverable frequency range is decreased to $\pm$ 0.33 cycles/degree, which means that there is a loss of information, especially for the upper limits of the predominant frequencies in the *y*-direction (shown in Figure 5.5). These results echo the analysis performed in the position domain in Chapter 3, which showed the degradation in position accuracy as a function of various grid resolutions.

A second correction field was investigated for the same epoch as satellite 18 above, but in this case the satellite was at a much lower elevation, averaging approximately 26° over the network coverage area. The corresponding PSD function for the high density grid parameterization (0.5° $\times$ 0.5°) is included in Figure 5.7. It is interesting here to note the significantly large amplitudes compared to the higher elevation case. This can be explained from the elevation of the satellite, which is much lower and therefore incurs a higher level of atmospheric activity that is accounted for by higher correction values. Again the amplitudes associated with the lower frequency components are much larger corresponding to longer wavelength trends in the data. The ranges in this case vary from $\pm$ 0.2 cycles/degree in both horizontal component directions. The limited larger frequency values are concentrated in more of the East-West direction this time. Thus, it

is evident from these plots that the directional behaviour of the predominant frequencies vary depending on the satellite correction field.

In terms of the spatial analysis, the aforementioned examples are useful for identifying the dominant frequency distributions for the entire coverage area. The relative strength of the signal is magnified according to the selected epoch in time when the correction field is investigated. Since the examples discussed thus far are obviously contaminated with stronger ionospheric activity, it is useful to examine some cases where the ionospheric activity is relatively lower. To accomplish this, two correction fields were generated at high and low elevations for an epoch in the morning hours at approximately 3:00 am. The corresponding PSD function for the high elevation satellite case (~71°) is shown in Figure 5.8.

Figure 5.8 was generated for the higher resolution grid based parameterization and depicts essentially the same spectral behaviour associated with the complementing scenario visible during the midday period. Additional tests were conducted in the same manner for other correction fields, however the plots are not included here as they would be redundant. In all cases it was found that the majority of the frequency information is retained up to a grid resolution where $\Delta x = \Delta y \leq 1.0^{\circ}$. Once the resolution is decreased by implementing sparser grid-based parameterizations, some of the dominant low frequency information is distorted and some higher frequency information is lost.
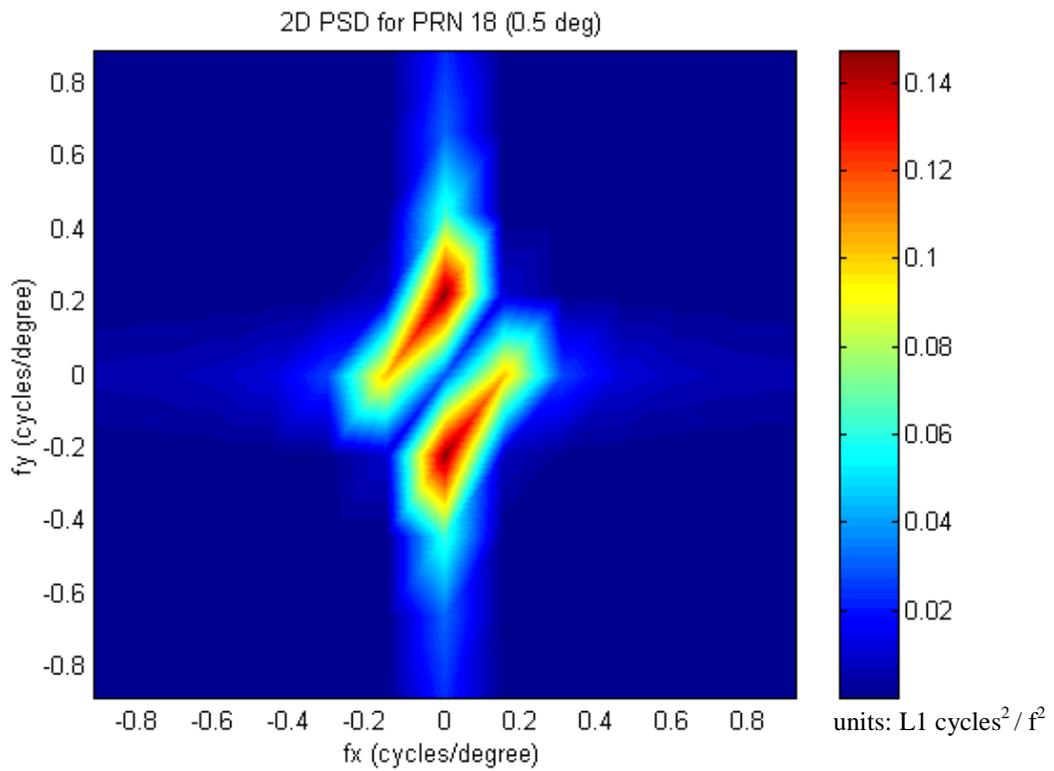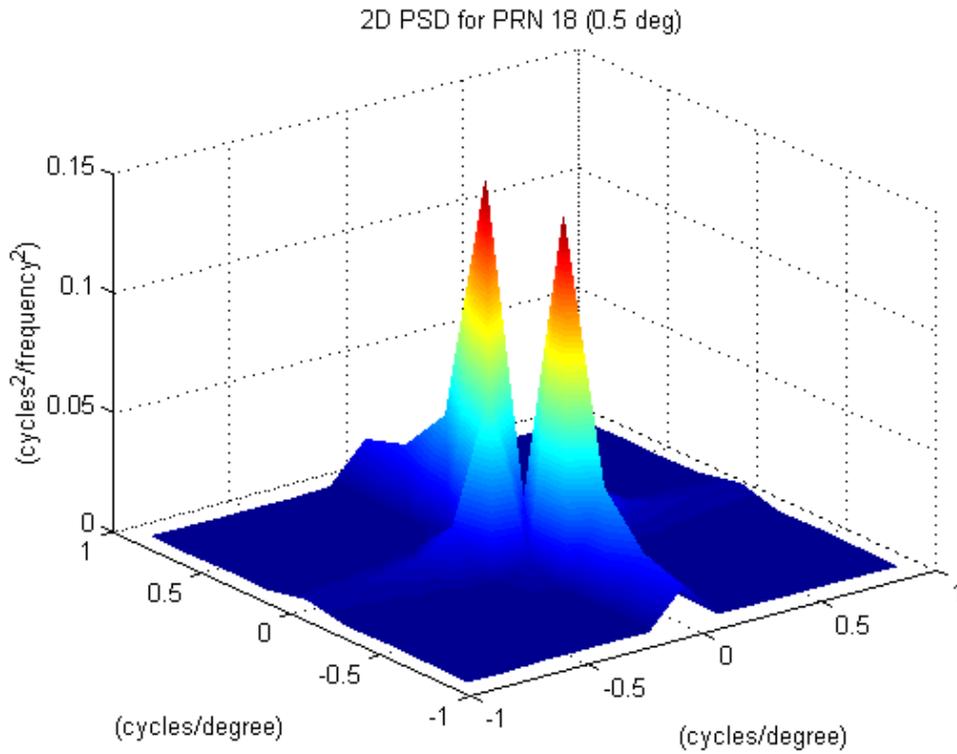
Figure 5.5: 2D PSD Functions for High Elevation Satellite Correction Field (0.5°)
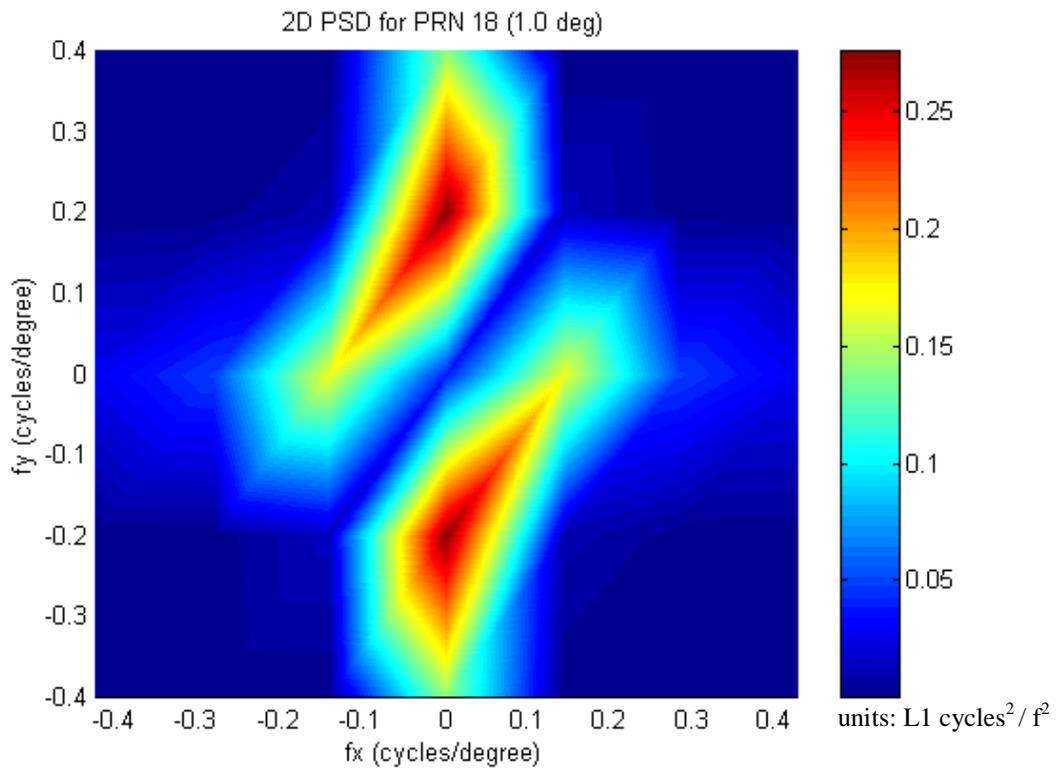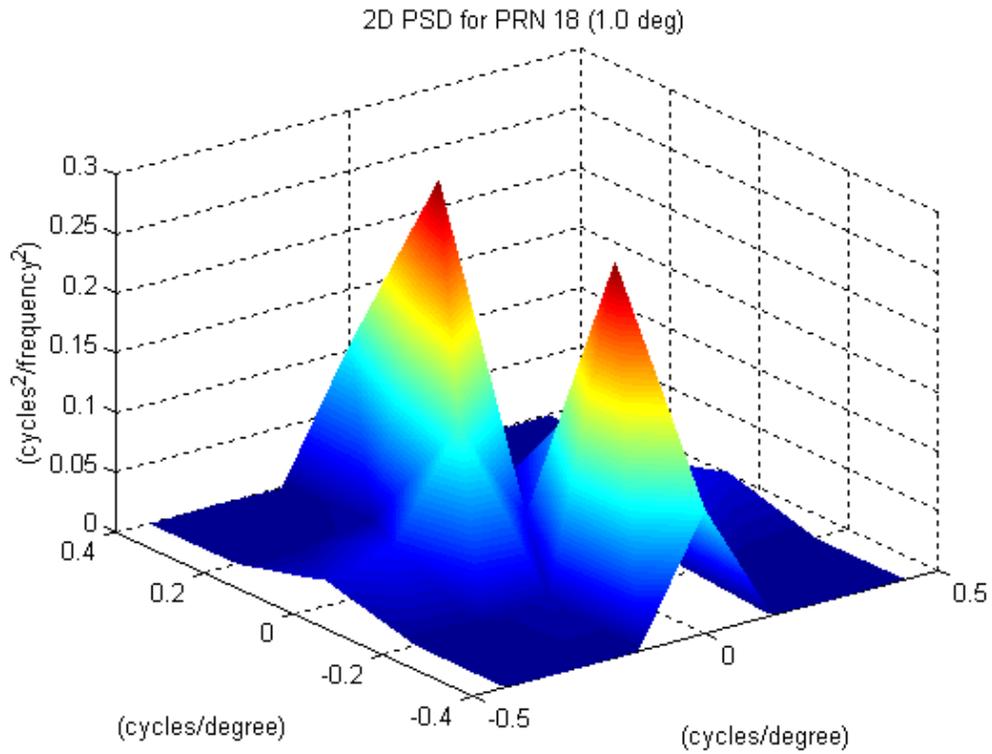
Figure 5.6: 2D PSD Functions for High Elevation Satellite Correction Field (1.0°)
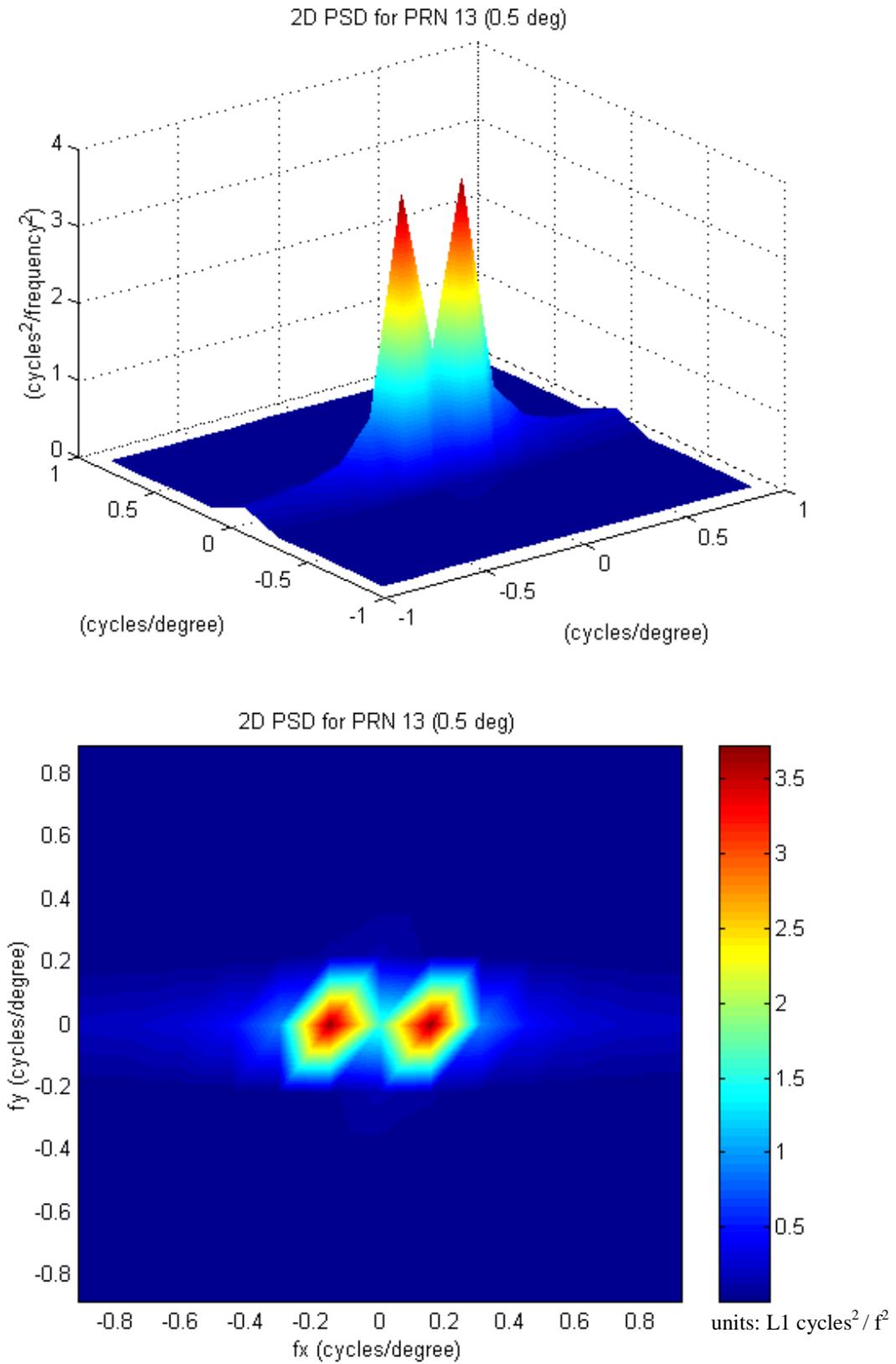
Figure 5.7: 2D PSD Functions for Low Elevation Satellite Correction Field (0.5°)
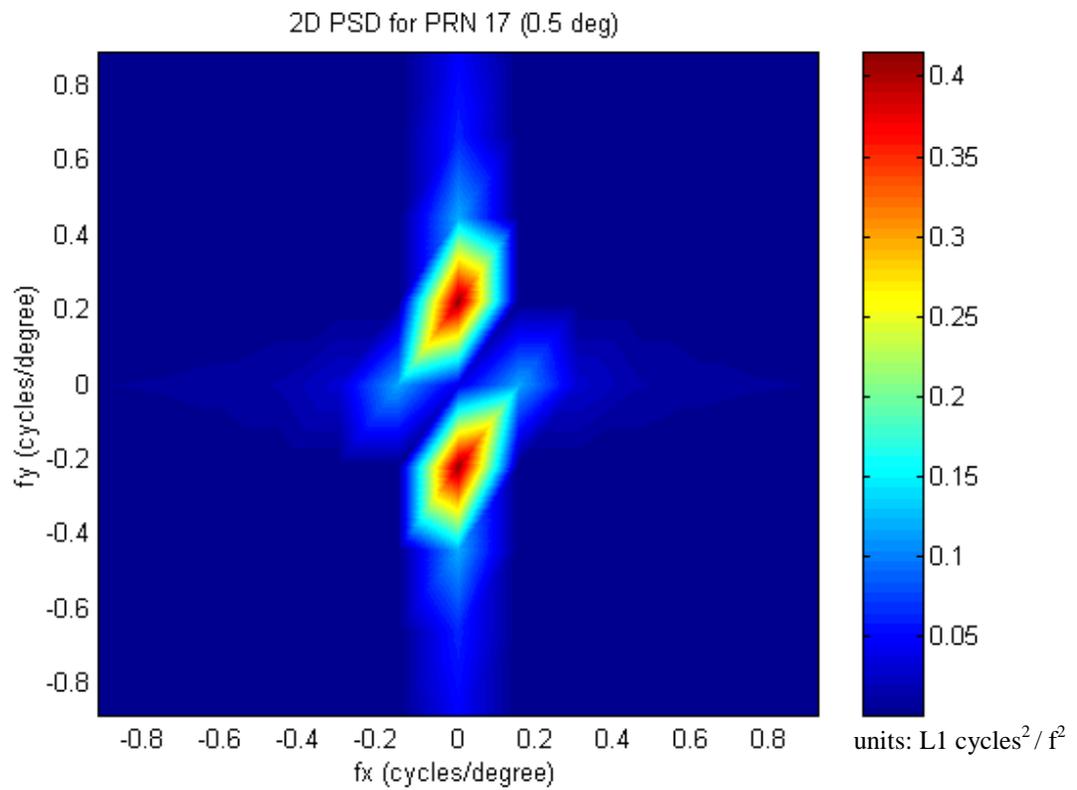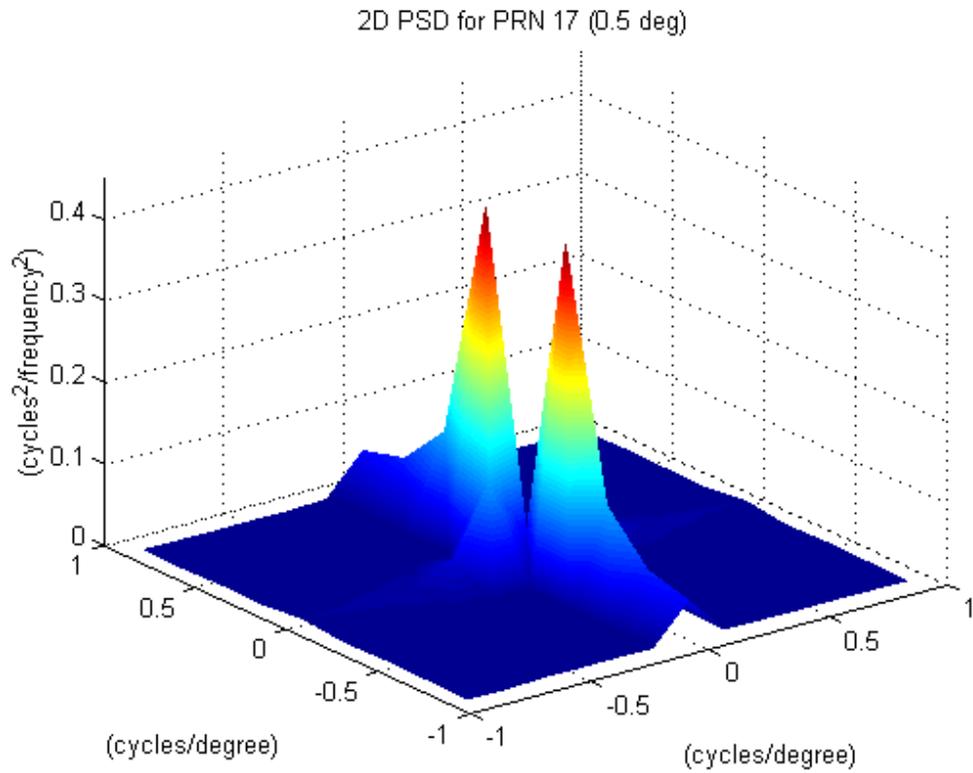
Figure 5.8: 2D PSD Functions for High Elevation Satellite During Morning Period (0.5°)

5.3   Summary

The analysis conducted in this chapter provided valuable insight into the spectral behaviour of the satellite-based correction fields. From a temporal perspective, large correlation lengths on the order of several tens of minutes indicate slowly changing corrections. This confirms the conclusions made in Chapter 4 stating that it is feasible to update model parameters over several minutes, which significantly reduces the transmission bandwidth load. Also, differences between the combined L1 phase correlation functions and the ionospheric-free correlation functions were noticed, which emphasized the need for modelling each error source separately (to some degree).

From a spatial perspective, the spectral analysis of various correction fields showed that grid resolutions of $0.5° \times 0.5°$ and $1° \times 1°$ were adequate for maintaining the frequency information contributing the most to the Swedish network data. This was also seen in Chapter 3, where the position domain results of various parameterization schemes indicated the superiority of grid-based parameterizations with resolutions of $\Delta x = \Delta y \leq 1.0°$ (see Table 3.8 for a summary).

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS

This final chapter summarizes the results presented in this thesis in terms of the three main objectives identified in the first chapter. Recommendations for future work on this area and related issues are also included.

6.1 Conclusions

A comprehensive discussion on a few of the major correction generation algorithms was presented, with an emphasis placed on the correction dissemination alternatives. Although not all methods available were discussed, a sample representing a variety of different approaches was chosen for comparison purposes. It was found that all methods provide some insight into the problems associated with the implementation of a carrier phase based multiple reference station network approach in real-time. Given this brief overview, it can be said that in general the focus of most of the research has been on modelling the spatial behaviour of the distance-dependent errors (especially ionospheric errors). However, in order to successfully implement any one of the algorithms

presented, it is equally important to determine the temporal correlation of these errors over the network coverage area.

The results presented in this thesis revealed several interesting aspects of the spatial and temporal behaviour of the correlated errors over a regional network in southern Sweden. Various methods for parameterizing carrier phase corrections for the combined correlated errors (ionospheric, tropospheric and satellite orbit) over a regional network of reference stations were analyzed in detail. The grid-based method employed the formulation of corrections at discrete points, with resolutions ranging from $0.5° \times 0.5°$ to $2° \times 2°$. A two-dimensional distance-based bilinear interpolation algorithm was found to be the most suitable for representing the formulated satellite-based correction fields. The results showed that the grid-based approach provided the greatest accuracy in terms of user position, showing improvements of 20% (latitude) 54% (longitude) and 27% (height) over the standard *uncorrected* single baseline approach. The achievable accuracies in position decreased with lower grid resolution, to the point where a $2° \times 2°$ spacing was found to be worse than a single baseline approach. These findings are quite significant because they identify the limitations of the grid-based model for RTK applications.

On the other hand, the volume of data required for transmitting the necessary information of a grid-based parameterization to the user was found to be quite significant and options for reducing the transmission load were investigated. In the interest of meeting data transmission bandwidth requirements, a second parameterization approach, namely low-

order surface modelling, was investigated. In this case four different first and second order polynomial surfaces were used to model the corrections over the regional network. The results in the position domain showed up to 5%, 36% and 18% improvements in RMS for latitude, longitude and height respectively, over the single baseline approach, using a second order polynomial fit.

Of the most important findings, were the investigations conducted on the update rates of all spatial model parameters and their effect in the position domain. For instance, it was found that some parameterization schemes provide relatively low RMS position errors (with respect to the corrected single baseline approach) even with corrections as old as 60 seconds. Temporal trends in the corrections were identified based on the time of day. In general, the midday period was found to produce a stronger signal corresponding to the greater ionospheric activity associated with this time period. This was also exhibited in the correlation and spectral analysis of the satellite-based correction fields.

Comparing the two parameterization schemes, it is evident that both models are capable of producing results better than the single baseline solution. Also the temporal characteristics of the corrections were found to be concentrated in the lower frequencies, which indicate smooth and slowly changing behaviour. Results showed that it is possible to update the parameterization parameters (correction grid node values and polynomial coefficients) at intervals of several tens of seconds. Therefore, numerous transmission options were identified as feasible solutions to disseminating the parameterized

information. For the southern Swedish network, a single data transmitter option was used to derive the correction message information for both the grid-based and function-based schemes. The final decision on which option to implement would depend on the data link and bandwidth capacity available. Therefore, depending on the user's accuracy requirements several options for the dissemination and communication of multiple reference station data are available and feasible for use in real-time.

6.1.1 A Note on Higher Residual Correlated Errors

The objectives of this thesis have been addressed and supported with results obtained using the southern Swedish network data. As stated in Section 3.2, the level of ionospheric activity associated with this data was a maximum of approximately 5 ppm and the remaining errors were on the level of approximately 3 ppm. It is appreciated that these levels of residual correlated error effects may vary from day to day and between data sets. In view of this inevitable variability, it is beneficial to discern the effects of such changes (i.e. higher atmospheric activity) with respect to the research conducted thus far.

As described previously (in Chapters 1 and 2) numerous studies have been conducted on developing correction generation algorithms with limited work on the parameterization and subsequent dissemination of these corrections. Thus, the focus of this work was to lay a foundation for future investigations, which will continue and expand the results to

different and perhaps more extreme cases. Therefore, as a first case a moderate data set with relatively quiet atmospheric activity was used for the analysis. It was found throughout the analyses that the majority of the corrections were compensating for the ionospheric effects ranging from levels of 3 ppm to 5 ppm. It was observed that the effects at this level could successfully be parameterized using grid resolutions of $0.5° \times 0.5°$ to $1.5° \times 1.5°$ and five to six coefficient bivariate polynomial surfaces, depending on the accuracy requirements.

From the experience obtained in conducting the investigations, it is evident that higher atmospheric effects (especially ionospheric) may lead to a need for higher resolution parameterization models. In such cases, it is recommended that denser grid-based parameterizations are used (as opposed to function-based models), as the former was found to be the most accurate approach. The actual grid resolution will depend on the correction values and the geometry of the regional network of reference stations. Thus, it must be determined on an individual basis and tests should be conducted in the manner described throughout this thesis.

As it was shown in Chapters 4 and 5, the temporal behaviour of the investigated corrections was slowly changing and therefore the parameterization scheme update rates of several minutes can be used. For higher amplitude and frequency variations in corrections this update rate may also increase. However, the advantage of a satellite-

based correction scheme allows for each correction field to be transmitted at staggered intervals.

From a data transmission point of view, it is clear that if higher amplitude and frequency variations exist (as in the case of ionospheric scintillation or an ionospheric storm, Skone, 1998) a single data transmitter configuration may not be sufficient. Although for the SSN, this option was assumed for the discussion, two additional alternatives were also presented, namely (i) multiple data transmitters and (ii) combined single and multiple data transmitters. These options are more easily adaptable for communicating greater variations in corrections over localized areas.

6.2   Recommendations

Based on the results presented in this thesis, several areas for future work have been identified in order to improve the overall performance of the multi-reference station concept. One of the most prevalent recommendations, which was also mentioned several times during the analysis, was the generation of 'error-specific' corrections. This would involve isolating the ionospheric, tropospheric and satellite orbit effects and formulating separate corrections for each source. It is evident that all of the correlated error sources are independent of each other and they exhibit random behaviour. Since all of the investigations conducted thus far were based on combined corrections of the correlated errors, it was often difficult to identify the source of any disturbances. By separating the

correlated errors, corrections for specific sources can be obtained which will result in more accurate parameterizations. Also, since transmission capacity is a concern for real-time users, it may be possible to transmit the separated corrections at staggered intervals resulting in a more efficient dissemination process.

Another area recommended for future work is testing the parameterization schemes under different conditions. These conditions should involve a number of different networks with varying ground station geometry and inter-station separations. It would also be of benefit to test the two-dimensional options provided with networks of greater vertical separations between reference stations. This would test the limits of the two-dimensional model and identify when (if ever) a third dimension is required.

Finally, it is recommended that real-time tests are conducted which integrate on-the-fly ambiguity resolution techniques, correction formulation algorithms, spatial parameterization schemes, and transmission of multi-reference data to mobile users. Given the continuously increasing needs of GPS users for higher positioning accuracy and the explosion of precise positioning applications, it is evident that further research in this area is warranted.

REFERENCES

Abousalem, M. and Bogle, A. (1997): LandStar: The Differential GPS Service. In *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*, Banff, Canada, pp. 357-365.

Ashkenazi, V., Hill, C. and Nagel, J. (1992): Wide Area Differential GPS: A Performance Study. In *Proceedings of the 5th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-92)*, Albuquerque, New Mexico, pp. 589-598.

Blewitt, G. (1989): Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km. *Journal of Geophysical Research*, vol. 94, no. B8, pp. 10187-10203.

Borre, K. and Strang, G. (1997): Autocorrelation Functions in GPS Data Processing: Modeling Aspects. In *Proceedings of the 10th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-97)*, Kansas City, Missouri, pp.1143-1150.

Braasch, M.S. (1996): Chapter 14: Multipath Effects. In *Global Positioning System: Theory and Applications*, Volume I. Eds. Parkinson, B. and Spilker, J., American Institude of Aeronautics and Astronautics, Inc., pp. 547-566.

Brigham, E. (1988): *The Fast Fourier Transform and its Applications*. Prentice Hall Inc., Englewood Cliffs, New Jersey.

Challstrom, C. (1999): The Benefits of GPS Modernization. In *The Quarterly Newsletter of the Institute of Navigation*, vol. 9, no. 2, Summer 1999.

Chen D., and Lachapelle, G. (1995): A Comparison of the FASF and Least-Squares Search Algorithms for On-the-Fly Ambiguity Resolution. *Navigation: Journal of the Institute of Navigation*, vol. 42, no. 2, pp. 371-390.

Dermanis, A. (1988): Geodetic Applications of Interpolation and Prediction. *Eratosthenes*, vol. 22, pp. 229-262.

Divis, D.A. (2000): SA: Going the Way of the Dinosaur. *GPS World Magazine*, vol. 11, no. 6, June 2000, pp. 16-19.

Dudgeon, D.E. and Mersereau, R. M. (1984): *Multidimensional Digital Signal Processing*. Prentice-Hall Inc., New Jersey.

El-Arini, M.B., Conker, R., Albertson, T., Reegan, J.K., Klobuchar, J.A., and Doherty, P. (1995): Comparison of real-time ionospheric algorithms for a GPS wide area augmentation system (WASS). *Navigation: Journal of the Institute of Navigation*, vol. 41, no. 4, Winter 1994-95.

El-Rabbany, A.E. (1994): The Effect of Physical Correlations on the Ambiguity Resolution and Accuracy Estimation in GPS Differential Positioning. *PhD Thesis*, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada.

FAA. (199): U.S. Department of Transportation Federal Aviation Administration Specification Wide Area Augmentation System (WAAS). Draft, September 1999.

Fortes, L.P.S., Lachapelle, G., Cannon, M.E., Marceau, G., Ryan, S., Wee, S., and Raquet, J. (1999): Testing of a Multi-Reference GPS Station Network for Precise 3D Positioning in the St. Lawrence Seaway. In *Proceedings of the 12th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-99)*, Nashville, Tennessee, pp. 1259-1269.

Fotopoulos, G. (2000): Parameterization of Carrier Phase Corrections Based on a Regional Network of Reference Stations. Accepted for publication in *Proceedings of the 13th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-00)*, Salt Lake City, Utah, 2000.

Fotopoulos, G., and Cannon, M.E. (2000a): Spatial and Temporal Characteristics of DGPS Carrier Phase Errors Over a Regional Network. In *Proceedings of the International Associations of the Institute of Navigation (IAIN)*, San Diego, California, June 2000.

Fotopoulos, G., and Cannon, M.E. (2000b): An Overview of Multi-Reference Station Methods for Cm-Level Positioning. *GPS Solutions*, vol. 4, John Wiley & Sons, Inc.

Frei, E. and Beutler, G. (1990): Rapid static positioning based on the fast ambiguity resolution approach 'FARA': theory and first results. *Manuscripta Geodaetica*, vol. 15, pp. 325-356.

Gao, Y., Li, Z., and McLellan, J.F. (1997): Carrier Phase Based Regional Area Differential GPS for Decimeter-Level Positioning and Navigation. In *Proceedings of the 10th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-97)*, Kansas City, Missouri, pp.1305-1313.

Gao, Y., and Li. Z. (1998): Ionosphere Effect and Modelling for Regional Area Differential GPS Network. In *Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-98)*, Long Beach, California, pp. 91-97.

Goad, C. C. and Goodman, L. (1974): A Modified Hopfield Tropospheric Refraction Correction Model. In *Proceedings of the Fall Annual Meeting of the American Geophysical Union*. San Francisco, California.

Han, S. and Rizos, C. (1996): GPS Network Design and Error Mitigation for Real-Time Continuous Array Monitoring Systems. In *Proceedings of the 9th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-96)*, Kansas City, Missouri, pp. 1827-1836.

Han, S. and Rizos, C. (1997): Instantaneous Ambiguity Resolution for Medium-Range GPS Kinematic Positioning Using Multiple Reference Stations. In *Proceedings of the International Association of Geodesy Symposia*, vol. 118, Advances in Positioning and Reference Frames, Rio de Janeiro, Brazil, pp. 283-288.

Hatch, R. (1990): Instantaneous ambiguity resolution, In *Proceedings of Kinematic Systems in Geodesy, Surveying, and Remote Sensing (IAG Symposia 107)*, Banff, Canada, Sept. 1990, pp. 299-308.

Hedling, G. and Jonsson, B. (1996): New Developments in the SWEPOS Network. In *Proceedings of the 9th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-96),* Kansas City, Missouri, pp. 1803-1808.

Hegarty, C.J. (1993): Optimizing Differential GPS for a Data Rate Constrained Broadcast Channel. In *Proceedings of the 6th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-93)*, Palm Springs, California, pp. 1527-1535.

Hofmann-Wellenhof, B., Licktenegger, H. and Collins, J. (1994): *GPS Theory and Practice*, Third revised edition, Springer-Verlag.

Hopfield, H.S. (1969): Two-quartic Tropospheric Refractivity Profile for Correcting Satellite Data. *Journal of Geophysical Research*, vol. 74, no. 17, pp. 4487-4499.

Howind, J., Kutterer, H., and Heck, B. (1999): Impact of temporal correlations on GPS-derived relative point positions. *Journal of Geodesy*, vol. 73, pp. 246-258.

ICD-GPS-200C. (1993): GPS Signal Specification Interface Control Document, Revision C.

ION Newsletter (2000): The Decision to Discontinue SA: A White House Press Release. *The Quarterly Newsletter of the Institute of Navigation*, vol. 10, no. 1, Spring 2000, page 2.

Jensen, A.B.O. and Cannon, M.E. (2000): Performance of Network RTK Using Fixed and Float Ambiguities. In *Proceedings of the 2000 National Technical Meeting of the Institute of Navigation*, January 26-28, 2000, Anaheim, California.

Jong de, P.J. (1997): GPS Ambiguity Resolution for Navigation, Rapid Static Surveying, and Regional Networks. In *Proceedings of the International Association of Geodesy Symposia*, vol. 118, Advances in Positioning and Reference Frames, Rio de Janeiro, Brazil, pp. 223-228.

Junkins, J. L., Miller, G.W., Jancaitis, J,R. (1973): A Weighting Function Approach to Modeling of Irregular Surfaces. *Journal of Geophysical Research*, vol. 78, no. 11, pp. 1794 – 1803.

Kaplan, D. (1996): *Understanding GPS Principles and Applications*. Artech House Publishers, Boston.

Kee, C. and Parkinson, B. (1992): Algorithms and Implementation of Wide Area Differential GPS. In *Proceedings of the 5th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-92)*, Albuquerque, New Mexico, pp. 565-572.

Klobuchar, J.A. (1996): Chapter 12: Ionospheric Effects on GPS. In *Global Positioning System: Theory and Applications*, Volume I, Eds. Parkinson, B. and Spilker, J., American Institute of Aeronautics and Astronautics, Inc., pp. 485-514.

Lachapelle, G., Cannon, M.E., Fortes, L.P., Stephen, J. (1999): Feasibility Study on the Use of a Multi-Reference Station Approach for OTF Ambiguity Resolution in a Constricted Waterway (St. Lawrence). Report prepared for Electronic Engineering Group, Canadian Coast Guard, Ottawa, 1999.

Lachapelle, G., Cannon, M.E., Fortes, L.P. (2000): Testing of a Multi-Reference Station Approach for OTF Ambiguity Resolution in the St. Lawrence River. Report prepared for Electronic Engineering Group, Canadian Coast Guard, Ottawa, 2000.

Lancaster, P. and Salkauskas, K. (1986): *Curve and Surface Fitting: An Introduction*. Academic Press Inc., London, England.

Leick, A. (1995): *GPS Satellite Surveying*. Second Edition. John Wiley & Sons, Inc., New York.

Lu, G., Cannon, M.E., Chen, D., and G. Lachapelle (1994): Flykin^{TM} Operator's Manual, version 2.0. Department of Geomatics Engineering, The University of Calgary.

Marel, H. van der (1998): Virtual GPS reference stations in the Netherlands. In *Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-98)*, Nashville, Tennessee, pp. 49-58.

Math Works, Inc. (1998): Signal Processing Toolbox User's Guide for Use with MATLAB®, see http://www.mathworks.com.

Mendes, V.B. (1999): Modeling the neural-atmosphere propagation delay in radiometric space techniques. *Ph.D. Thesis*, Department of Geodesy and Geomatics Engineering Technical Report No. 199, University of New Brunswick, Fredericton, New Brunswick, Canada.

Mikhail, E. M. (1976): *Observations and Least-Squares*. Thomas Y. Crowell Company, Inc., New York.

Mueller, T. (1994): Minimum Variance Network DGPS Algorithm. In *Proceedings of the IEEE 1994 Position Location and Navigation Symposium (PLANS)*, Las Vegas, Nevada, pp. 418-425.

Neumann, J. B., Van Dierendonck, K.J., Manz, A., and T.J. Ford. (1997): Real-Time Carrier Phase Positioning Using the RTCM Standard Message Types 20/21 and

18/19. In *Proceedings of the 10th International Technical Meeting of the Satellite Division of the Institute of Navigaiton (ION GPS-97)*, Kansas City, Missouri, pp. 856-866.

Odijk, D., Marel, H. van der, and Song, I. (2000): Precise GPS Positioning by Applying Ionospheric Corrections from an Active Control Network. *GPS Solutions*, vol. 3, John Wiley & Sons, Inc., pp. 49-57.

Parkinson, B. (1996): Chapter 1: Introduction and Heritage of NAVSTAR, the Global Positioning System. In *Global Positioning System: Theory and Applications*, Volume I, Eds. Parkinson, B. and Spilker, J., American Institute of Aeronautics and Astronautics, Inc., pp. 3-28.

Parkinson, B., and Spilker, J. (1996a): *Global Positioning System: Theory and Applications*, Volume I. American Institute of Aeronautics and Astronautics, Inc.

Parkinson, B., and Spilker, J. (1996b): *Global Positioning System: Theory and Applications*, Volume II. American Institute of Aeronautics and Astronautics, Inc.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1988*): Numerical Recipes in C*, Cambridge University Press.

Raquet, J. (1997): Multiple User Network Carrier-Phase Ambiguity Resolution. In *Proceedings of the 1997 International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS-97)*, Banff, Canada, pp. 45-55.

Raquet, J. (1998): Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers. *Ph.D. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20116, Calgary 1998.

Raquet, J., Lachapelle, G., Melgård, T. (1998): Test of a 400 km × 600 km Network of Reference Receivers for Precise Kinematic Carrier-Phase Positioning in Norway. In

Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation, Kansas City, Missouri, pp. 407-416.

Ray, J.K. (2000): Mitigation of GPS Code and Carrier Phase Multipath Effects Using a Multi-Antenna System. *Ph.D. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20136, Calgary 2000.

Rothacher, M. (1997): GPS Satellite Orbits, Orbit Determination, and the IGS. Lecture Notes of the Nordic Autumn School on Geodetic Applications of GPS, Bastad, Sweden, August 26-31, 1996, Nordic Geodetic Commission (NGK): Jonsson, Bo (ed.), Trycksam, Gävle, pp. 55-108.

RTCM (1994): RTCM Recommended Standards for Differential Navstar GPS Service, Version 2.1, RTCM Paper 194-93/SC104, January 3, 1994.

Russell, W.S. (1995): Polynomial Interpolation Schemes for Internal Derivative Distributions on Structured Grids. *Applied Numerical Mathematics*, vol. 17, pp. 129-171.

Saastamoinen, J. (1972): Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. In *The Use of Artificial Satellites for Geodesy, Geophys. Monogr. Ser., American Geophysical Union*, Ser., vol. 15, Washington, D.C. pp. 247-251.

Schaer, S., Beutler, G., Rothacher, M., Brockmann, A., Wiget, A. and Wild, U. (1999): The Impact of the Atmosphere and Other Systematic Errors on Permanent GPS Networks. Presented at the *International Association of Geodesy Symposium of Positioning*, Birmingham, UK, July 21, 1999.

Seeber, G. (1993): *Satellite Geodesy: Foundations, Methods and Applications*. Walter de Gruyter, New York.

Sideris, M. G. (1984): Computation of Gravimetric Terrain Corrections Using Fast Fourier Transform Techniques. *MSc. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20007, Calgary, 1994.

Skone, S. (1998): Wide Area Ionospheric Grid Modelling in the Auroral Region. *Ph.D. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20123, Calgary, 1998.

Spilker, J.J. (1996): Chapter 13: Tropospheric Effects on GPS. In *Global Positioning System: Theory and Applications*, Volume I, Eds. Parkinson, B. and Spilker, J., American Institute of Aeronautics and Astronautics, Inc., pp. 517-545.

Stein, M.L. (1999): *Interpolation for Spatial Data - Some Theory for Kriging*. Springer-Verlag, New York, Inc.

Sun, H., Cannon, M.E., and Melgård, T.E. (1999): Real-Time GPS Reference Network Carrier Phase Ambiguity Resolution. In *Proceedings of the National Technical Meeting of the Institute of Navigation (ION NTM-99)*, San Diego, California, pp. 193-199

Talbot, N.C. (1996): Compact Data Transmission Standard for High-Precision GPS. In *Proceedings of the 9th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-96)*, pp. 861-871.

Teunissen, P. (1998): Some Remarks on GPS Ambiguity Resolution. *LGR Series, Publications of the Delft Geodetic Computing Centre*, no. 20. Artificial Satellites, vol. 32, no. 3., pp. 13-23.

Teunissen, P.J.G. and Kleusberg, A. (1996): *GPS for Geodesy*, 2nd Edition. Springer-Verlag, Berlin.

Townsend, B., Lachapelle, G., Fortes, L.P., Melgard, T., Norbech,T. and Raquet, J. (1999): New Concepts for a Carrier Phase Based GPS Positioning Using a National
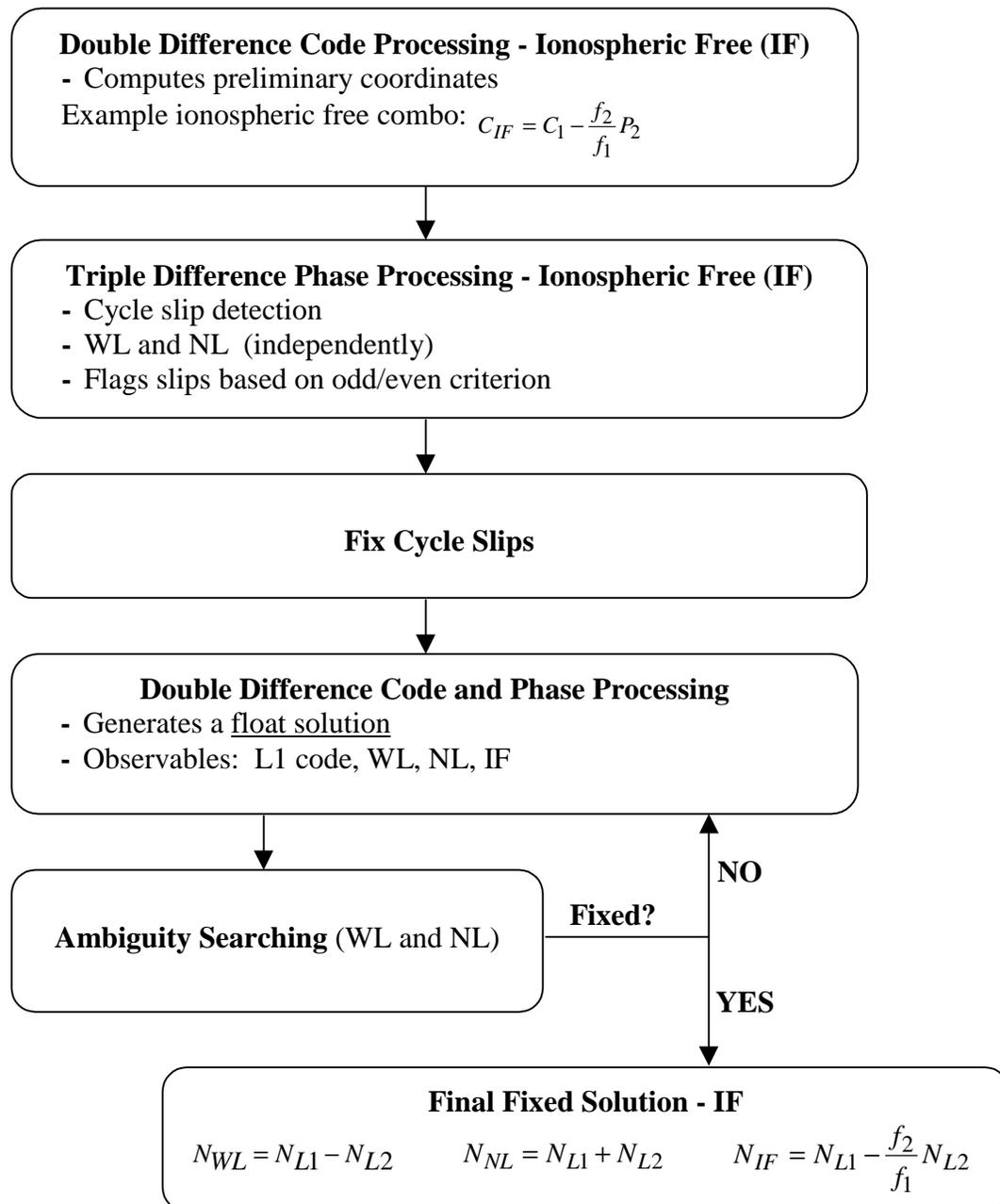
Reference Station Network. In *Proceedings of the National Technical Meeting of the Institute of Navigation (ION NTM-99)*, San Diego, California, pp. 319-326.

Trimble Navigation Limited (1996): *GPSurvey - Software User's Guide*. Trimble Navigation Limited, Surveying and Mapping Division, Sunnyvale, California.

UNAVCO (1994): UNAVCO QC v3 User's Guide. University NAVSTAR Consortium, Boulder, 1994.

Van Dierendonck, A.J., Fenton, P., and Ford, T. (1992): Theory and Performance of Narrow Correlator Spacing in a GPS Receiver. *Navigation: Journal of the Institute of Navigation*, vol. 39, no. 3, pp. 265-283.

Varner, C. (2000): DGPS Carrier Phase Networks and Partial Derivative Algorithms. *Ph.D. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20129, Calgary, 2000.

Varner, C. and Cannon, M.E. (1997): The Application of Multiple Reference Stations and the Determination of Multipath and Spatially Decorrelating Errors. In *Proceedings of the National Technical Meeting of the Institute of Navigation (ION NTM-97)*, Santa Monica, California, pp. 323-333.

Wang, J. (1998): Stochastic Assessment of the GPS Measurements for Precise Positioning. In *Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-98),* Nashville, Tennessee, pp. 81 - 89.

Wanninger, L. (1995): Improved ambiguity resolution by regional differential modelling of the ionosphere. In *Proceedings of the 8th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-95)*, Palm Springs, California, pp. 55-62.

Wanninger, L. (1997): Real-Time Differential GPS Error Modelling in Regional Reference Station Networks. *In Proceedings of the International Association of*

Geodesy Symposia, vol. 118, Advances in Positioning and Reference Frames. Rio de Janeiro, Brazil, pp. 86-92.

Watson, D. (1992): *Contouring A Guide to the Analysis and Display of Spatial Data*. Pergamon Press, Oxford, England.

Weisenburger, S. (1997): Effect of Constraints and Multiple Receivers for On-The-Fly Ambiguity Resolution. *MSc. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20109, Calgary, 1997.

Wübbena, G., Bagge, A., Seeber, G., Volker, B. and Hankemeier, P. (1996): Reducing Distance Dependent Errors for Real-Time Precise DGPS Applications by Establishing Reference Station Networks. In *Proceedings of the 9th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-96)*, Kansas City, Missouri, pp. 1845-1852.

Zhang, J. (1999): Investigations into the Estimation of Residual Tropospheric Delays in a GPS Network. *MSc. Thesis*, University of Calgary, Department of Geomatics Engineering Report Number 20132, Calgary, 2000.

Appendix A

FLOWCHART OF CARRIER PHASE PROCESSING PROCEDURE

The following flowchart outlines the major steps involved in determining the double difference ambiguities for independent baselines in the network.

Double Difference Code Processing - Ionospheric Free (IF)
- Computes preliminary coordinates
Example ionospheric free combo: $C_{IF} = C_1 - \dfrac{f_2}{f_1} P_2$

↓

Triple Difference Phase Processing - Ionospheric Free (IF)
- Cycle slip detection
- WL and NL (independently)
- Flags slips based on odd/even criterion

↓

Fix Cycle Slips

↓

Double Difference Code and Phase Processing
- Generates a <u>float solution</u>
- Observables: L1 code, WL, NL, IF

↓

Ambiguity Searching (WL and NL)   **Fixed?**   **NO**

YES

↓

Final Fixed Solution - IF

$$N_{WL} = N_{L1} - N_{L2} \qquad N_{NL} = N_{L1} + N_{L2} \qquad N_{IF} = N_{L1} - \dfrac{f_2}{f_1} N_{L2}$$

Appendix B

DERIVATION OF ERROR VARIANCES FOR PARAMETERIZED CORRECTIONS

To provide an indication of the quality of the estimated corrections evaluated for the corresponding horizontal user position, error propagation of the parameterized form can be performed.

In general, all parameterization schemes can be represented in the following form (vectors in **bold**):

$$z_i = \mathbf{a_i^T x} \tag{B.1}$$

where, z_i is the parameterized user correction, $\mathbf{a}_i$ is a 'design' vector, which depends on the estimated horizontal user positions (usually the single point position), and $\mathbf{x}$ is the vector of parameters. The error variances of z_i is given by:

$$\sigma_{z_i}^2 = \mathbf{a}_i^T C_x \mathbf{a}_i \tag{B.2}$$

where, C_x is the error covariance matrix of the parameters. This can be computed based on the correction generation scheme implemented.

As an example, consider the six-coefficient bivariate polynomial parameterization of eq. (3.10c), repeated below:

$$z_i = ax_i + by_i + cx_iy_i + dx_i^2 + ey_i^2 + f \tag{B.3}$$

Equation (B.3) can be represented in vector notation as,

$$z_i = \begin{bmatrix} 1 & x_i & y_i & x_iy_i & x_i^2 & y_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \tag{B.4}$$

$$= \mathbf{a}_i^T \mathbf{x}$$

By applying eq. (B.2) to the above equation, the error variance of the derived user corrections is computed.

Appendix C

DERIVATION OF POLYNOMIAL SURFACE COEFFICIENTS

The general equation of a plane is provided by,

$$z = ax + by + c \tag{C.1}$$

where, the unknown vector of parameters (coefficients in this case) is as follows:

$$\mathbf{d}^a = \begin{bmatrix} a & b & c \end{bmatrix}^T \tag{C.2}$$

The design matrix is formed by computing the partial derivatives with respect to the unknowns $\partial z / \partial \mathbf{d}^a$:

$$A = \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{pmatrix} \tag{C.3}$$

The normal matrix is generally computed by

$$N = A^T P A \tag{C.4}$$

where, P is the classical weight matrix of the observations (e.g. inverse of the covariance matrix). Assuming all observations to be of equal weight ($P = I$), a closed form of the normal matrix is given as follows:

$$N = \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i x_i & \sum_{i=1}^{n} y_i^2 & \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} y_i & n \end{pmatrix} \tag{C.5}$$

Note that the matrix is symmetric, which means that the full matrix does not have to be computed on an element-by-element basis. This allows for some computational advantages in terms of computing the matrix inverse.

Now we can compute the closed form,

$$\mathbf{u} = A^T \mathbf{b}$$

$$= \begin{bmatrix} \sum_{i=1}^{n} x_i z_i \\ \sum_{i=1}^{n} y_i z_i \\ \sum_{i=1}^{n} z_i \end{bmatrix} \tag{C.6}$$

where the vector of observations b is represented by:

$$\mathbf{b} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ \vdots \\ z_n \end{bmatrix} \tag{C.7}$$

and therefore, we can finally solve for the unknown coefficients of the best-fit plane (in a least-squares sense) as follows:

$$\hat{\mathbf{x}} = N^{-1}\mathbf{u} \tag{C.8}$$

where $\hat{\mathbf{x}}$ contains the estimated adjusted parameters.

The best-fit low order surfaces can be computed in a similar manner. In fact, the computational burden is lightened somewhat by utilizing the closed forms of the matrices, as shown above. These values are computed for the most *complicated* surface with six unknown parameters. Essentially all surface 'fits' are derived from a sub-set of this form:

$$z = ax + by + cxy + dx^2 + ey^2 + f \tag{C.9}$$

The unknown parameter vector is:

$$\mathbf{d^a}_{(1\times u)} = \begin{bmatrix} a & b & c & d & e & f \end{bmatrix}^T \tag{C.10}$$

and the corresponding design matrix is:

$$
A_{(n \times u)} = \begin{pmatrix}
x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 & 1 \\
x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_n & y_n & x_n y_n & x_n^2 & y_n^2 & 1
\end{pmatrix}
\tag{C.11}
$$

and the normal matrix is given by:

$$
N_{(u \times u)} = \begin{pmatrix}
\sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} x_i^2 y_i & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i y_i^2 & \sum_{i=1}^{n} x_i \\
\sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} y_i^2 & \sum_{i=1}^{n} x_i y_i^2 & \sum_{i=1}^{n} x_i^2 y_i & \sum_{i=1}^{n} y_i^3 & \sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_i^2 y_i & \sum_{i=1}^{n} x_i y_i^2 & \sum_{i=1}^{n} x_i^2 y_i^2 & \sum_{i=1}^{n} x_i^3 y_i & \sum_{i=1}^{n} x_i y_i^3 & \sum_{i=1}^{n} x_i y_i \\
\sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^2 y_i & \sum_{i=1}^{n} x_i^3 y_i & \sum_{i=1}^{n} x_i^4 & \sum_{i=1}^{n} x_i^2 y_i^2 & \sum_{i=1}^{n} x_i^2 \\
\sum_{i=1}^{n} x_i y_i^2 & \sum_{i=1}^{n} y_i^3 & \sum_{i=1}^{n} x_i y_i^3 & \sum_{i=1}^{n} x_i^2 y_i^2 & \sum_{i=1}^{n} y_i^4 & \sum_{i=1}^{n} y_i^2 \\
\sum_{i=1}^{n} x_i & \sum_{i=1}^{n} y_i & \sum_{i=1}^{n} x_i y_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} y_i^2 & n
\end{pmatrix}
\tag{C.12}
$$